Durchgeführt am Max-Planck-Institut für Herz- und Lungenforschung,

W.G. Kerckhoff-Institut Bad Nauheim

Spurious transcripts in *Tet3*-deficient smooth muscle cells provoke innate immune responses and lung pathogenesis

INAUGURAL-DISSERTATION

zur Erlangung des Doktorgrades der Naturwissenschaften
- Doctor rerum naturalium -
(Dr. rer. nat.)

vorgelegt dem
Fachbereich für Biologie und Chemie (FB 08)
der Justus-Liebig-Universität Gießen

eingereicht von

Fan Wu

Gießen, April 2021

Die vorliegende Arbeit wurde am Max-Planck-Institut für Herz- und Lungenforschung, W.G. Kerckhoff-Institut in Bad Nauheim angefertigt.

Erstgutachter:
Prof. Dr. Dr. Thomas Braun
Abteilung Entwicklung und Umbau des Herzens
Max-Planck-Institut für Herz- und Lungenforschung
Ludwigstraße 43, 61231 Bad Nauheim

Zweitgutachter:
Prof. Dr. Reinhard Dammann
Institut für Genetik
Justus-Liebig-Universität Giessen
Heinrich-Buff-Ring 58-62
35392 Giessen

Datum der Disputation: 15th, June 2021

EIDESSTATTLICHE ERKLÄRUNG

„Ich erkläre: Ich habe die vorgelegte Dissertation selbständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind, eingehalten.“

Bad Nauheim, den

ZUSAMMENFASSUNG

Glatte Muskelzellen (GMZ) sind eine wichtige Komponente der Atemwege und der Blutgefäße, die für die Aufrechterhaltung der Lungenfunktion und des Blutdrucks notwendig sind. Im Gegensatz zu anderen differenzierten Zellen zeigen glatte Muskelzellen eine beeindruckende Plastizität hinsichtlich ihrer Fähigkeit unterschiedliche phänotypische Zustände abhängig von den funktionalen Anforderungen anzunehmen. Als Reaktion auf Verletzungen und schädliche Umweltfaktoren durchlaufen kontraktile glatte Muskelzellen einen phänotypischen Wechsel zu unreifen, proliferierenden und stärker sekretierenden GMZ, wie man sie während der Entwicklung findet. Phänotypische Veränderungen und Dysfunktion von glatten Muskelzellen aufgrund von Immunreaktionen spielen eine wichtige Rolle bei Veränderungen der Struktur und Funktion der Atemwege und Blutgefäße. Die Funktion von GMZ hat eine zentrale Rolle in der Pathogenese verschiedener Erkrankungen des Lungenparenchyms inne.

In meiner Dissertation konnte ich zeigen, dass die zehn-elf-Translokation-Dioxygenase 3 (*engl. ten to eleven translocation dioxygenase 3*, TET3) stark in glatten Muskelzellen exprimiert ist und eine wichtige Rolle in der Regulierung der Kontraktilität von den glatten Muskelzellen und der Homöostase des Lungengewebes spielt. TET3 ist ein Enzym der TET-Familie, welche das 5-Methylcytosin (5mC) zum 5-Hydroxymethylcytosin, 5-Formylcytosin (5fC) und 5-Carboxylcytosin (5caC) oxidiert. Die Deletion von *Tet3* in glatten Muskelzellen führt nicht nur zu phänotypischen Transition von kontraktilen zu synthetisch aktiven GMZ, sondern auch löst auch einen auffallenden Atemwegumbau aus, der sich durch massive Becherzellenhyperplasie, verdickte Basalmembranen und Epithelzellenabbau auszeichnet. Ich habe festgestellt, dass TET3 während der transkriptionellen Elongation direkt an RNA-Polymerase II (Pol II) bindet und die intragene Akkumulation von 5hmC in hoch exprimierten Gene vermittelt. Außerdem verstärkt TET3 die Bindung des SETD2-Proteins (*engl. SET Domain Containing 2*) zu Pol II und fördert dadurch die Bildung der repressiven H3K36me3-abhängigen Chromatin-Struktur. Hierdurch wird ein ektopischer intragener Eintritt der RNA-Polymerase II verhindert. Der Verlust von *Tet3* führt zu einem abnormalen Start der intragenen Transkription in stark exprimierten Genen und zur verringerten Produktion von funktionsfähigen vollständigen mRNAs von Genen, die an der Regulation der Kontraktilität von den glatten Muskelzellen beteiligt sind. Zudem kommt es zur Ansammlung von abnormalen Transkripten. Solch abnormale Transkripte stimulieren den TLR7-Signalweg und die Th2-basierte Immunreaktion, welche wiederum zum pathologischen Umbau der Atemwege führen.

Meine Ergebnisse weisen auf einen kritischen epigenetischen Mechanismus hin, welcher DNA-Demethylierung, Suppression von abnormaler Transkription und Immunreaktionen mit der

Homöostase der glatten Muskelzellen verbindet. Eine Störung dieses kritischen Regulationsmechanismus führt zu Lungenerkrankungen. Meine Befunde zeigen, dass die gründliche Analyse der dynamischen DNA-Methylierung/Demethylierung zur Identifizierung neuer Zielmoleküle führen kann, welche zur Behandlung von Funktionsstörungen glatter Muskelzellen eingesetzt werden können.

SUMMARY

The smooth muscle cell (SMC) is a crucial component of airways and vessels, required to maintain lung function and blood pressure. Unlike other differentiated cells, SMCs exhibit remarkable plasticity, allowing them acquire different phenotypic states depending on functional demands. In response to various injuries and harmful environmental factors, contractile SMCs undergo a substantial phenotypic switch to immature, proliferative and more secretory SMCs, similar to early developmental stages. Phenotypic modulation and associated dysfunctions of SMCs, such as alterations of contractility, innate and adaptive immune responses and altered structure of airways and vessels, play central roles in the pathogenesis of diverse lung parenchyma diseases.

In this study, I demonstrate that Ten to Eleven Translocation dioxygenase 3 (TET3), a TET family enzyme that converts 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) successively, is highly expressed in SMCs and plays essential roles in regulating SMC contractility and lung tissue homeostasis. Induced deletion of *Tet3* in SMCs not only leads to a phenotypic switch of contractile SMC to synthetic SMC, but also provokes airway remodeling, characterized by massive goblet cell hyperplasia, thickened basal membranes and epithelial shedding. Mechanistically, I discover that TET3 binds directly to the elongating form of RNA Polymerase II (Pol II) and mediates intragenic accumulation of 5hmC within highly expressed genes. In addition, TET3 enhances recruitment of the SET Domain Containing 2 (SETD2) to Pol II, thereby promoting H3K36me3-dependant formation of repressive chromatin, which prevents aberrant intragenic entry of Pol II. Loss of *Tet3* results in aberrant initiation of intragenic transcription in highly expressed genes, accumulation of spurious transcripts and reduced production of functional full-length mRNAs of genes involved in regulation of SMC contractility. Eventually, spurious transcripts activate the TLR7-signaling pathway and Th2-based immune responses, which in turn causes airway inflammation and pathological airway remodeling.

Taken together, my findings reveal a vital epigenetic mechanism that link DNA demethylation, spurious transcription and innate immune responses to SMC homeostasis and lung pathogenesis. The identification of a novel target molecule pivotal for maintaining the contractile state of SMCs and the prevention of innate immune responses in the lung might facilitate development of new therapies to treat lung diseases associated with smooth muscle dysfunctions.

Table of Contents

INTRODUCTION

I Structure & resident cellular components of the lung

I-1 Structure of the lung

The lung, a complex organ containing more than 40 different cell types (Tata and Rajagopal, 2017), mediates gas exchange through extracting oxygen from the atmosphere, transferring it into the bloodstream, and releasing carbon dioxide from the bloodstream into the atmosphere. A healthy adult breathes every day more than 7000 liters of air containing diverse microbes. Thus, the mammalian lung is considered as a non-sterile organ (O'Dwyer et al., 2016). Albeit the human and murine lungs show some differences since a human lung has two lobes on the left side and three lobes of the right lung, while mouse has one lobe of the left lung and four lobes of the right lung (Travillian et al., 2005), the lung structure in both species is very similar. The lung can be divided into two major compartments: the airway overlaid by mucosal tissue to conduct gases, and the thin-walled alveoli where gas exchange occurs (Holt et al., 2008; Tata and Rajagopal, 2017). The trachea and proximal airways in human and mouse are lined with epithelial cells including ciliated cells, basal cells (progenitor cells), secretory cells (club cells & goblet cells), which are underlaid by smooth muscle cells and interstitial fibroblasts (Zepp and Morrisey, 2019). The distal bronchiolar region of the mouse lung is lined with a monolayered epithelium consisting of secretory, ciliated and neuroendocrine epithelial linages (Kotton and Morrisey, 2014) (Figure 1). The airway epithelium functions as the first barrier against incoming environmental pathogens. It senses the environment, repels infectious agents, processes toxins and removes debris (Holt et al., 2008; Tata and Rajagopal, 2017), thereby delivering almost completely sterile, hydrated gases to the peripheral alveoli for gas exchange. Alveoli are tiny air sacs located at the end of bronchial tubes. Two cell layers form the inter-alveolar septum, an epithelium facing the alveolar lumen and an endothelium facing the capillary lumen, constituting a gas diffusible interface (Knudsen and Ochs, 2018). Under homeostatic condition the adult lung is rather static, but can quickly activate damage repair mechanisms through activation of progenitor populations (Kotton and Morrisey, 2014), cell dedifferentiation/transdifferentiation (Tata and Rajagopal, 2017) and immune responses (Holt et al., 2008) upon insult or injury.

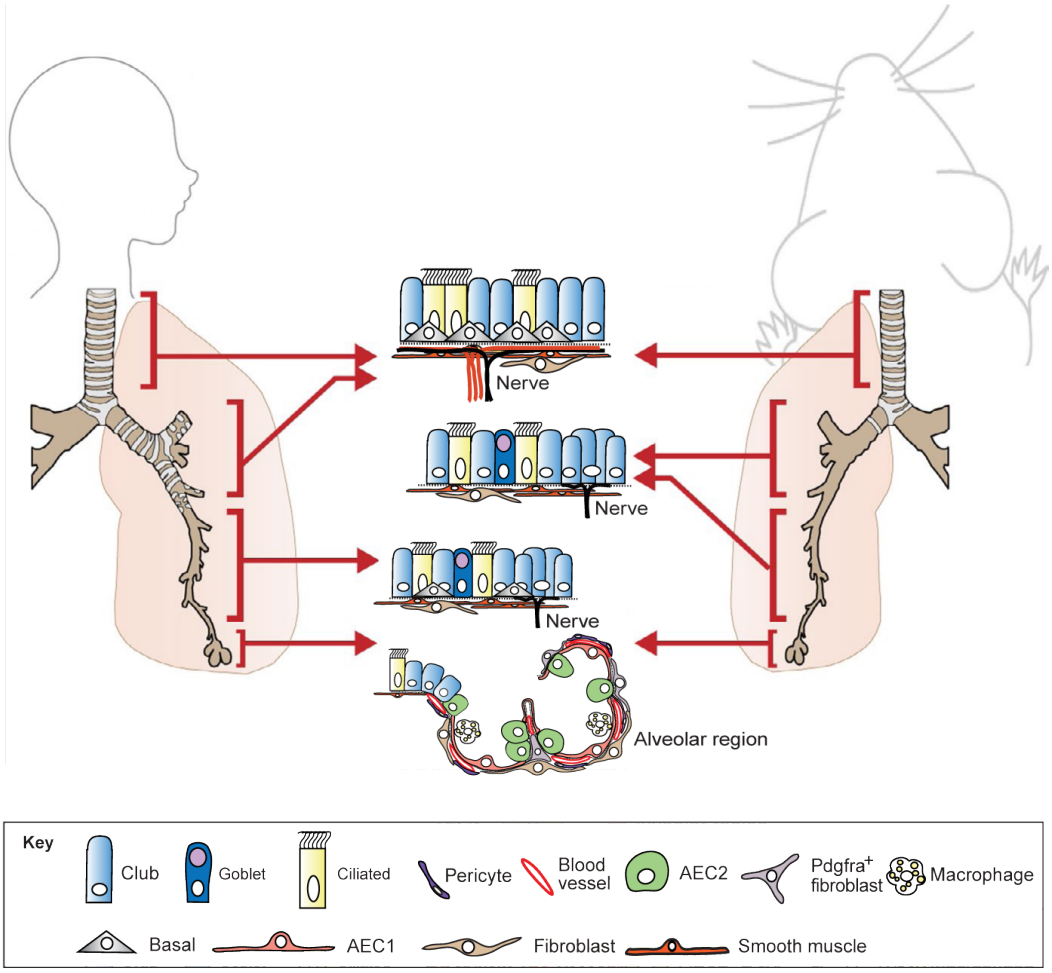


Figure 1. Schematic of the major cell types in different regions of the human and mouse lung. Modified from *Kiyokawa et al., 2019* and *Barkauskas et al., 2017* (Barkauskas et al., 2017; Kiyokawa and Morimoto, 2020).

I-2 Variety of cell types in lung

Distinct epithelial stem cell and progenitor cell populations in the lung have been identified. These cells, together with their differentiated progeny, maintain a stable homeostatic steady state, but show remarkable lineage plasticity following injury.

I-2-1 Ciliated cells

Ciliated cells characterized by the expression of nuclear transcription factor Forkhead box protein J1 (FoxJ1)(Tata and Rajagopal, 2017) are terminally differentiated epithelial cells (Bustamante-Marin and Ostrowski, 2017) and are present throughout the proximal and distal airways. Centrioles formed by ciliated cells can dock to the apical membrane and become basal bodies to facilitate cilia growth (Bustamante-Marin and Ostrowski, 2017). Cilia are specialized organelles marked by α-tubulin (Shah et al., 2008), and provide the necessary force to transport foreign materials out of the respiratory tract (Bustamante-Marin and Ostrowski, 2017). Furthermore, ciliated cells constitute a

crucial part of the primary defense mechanism – mucociliary clearance (MCC) by ciliary beating to assist mucus layer in lubricating and clearing the airway surface efficiently (Bustamante-Marin and Ostrowski, 2017). Under steady state in large airways, ciliated cells are mostly replenished by secretory cells, and only few by basal cells (Tata and Rajagopal, 2017). In response to lung injury, ciliated cells can transiently alter their morphology, whereas they do not proliferate or transdifferentiate as part of the repair process (Rawlins et al., 2007).

I-2-2 Basal cells

Basal cells are found in the trachea and the proximal airway in human and mice, but they extend to distal airways only in human (Tata and Rajagopal, 2017). Basal cells have extensive proliferative potential, a self-renewal capacity and the ability to differentiate into secretory and ciliated lung epithelial cells after injury (Tata and Rajagopal, 2017). In addition, basal cells secret IL33 to recruit immune cells to the airways in humans with severe chronic obstructive pulmonary disease (COPD) and in mice with chronic obstructive lung disease (Byers et al., 2013). Basal cells are characterized by the expression of p63, cytokeratin 5 (KRT5), nerve growth factor receptor (NGFR) and podoplanin (PDPN) (Tata and Rajagopal, 2017).

I-2-3 Secretory cells

Secretory or club cells (formerly known as Clara cells) are dome shape cells filled with secretory granules (Tata and Rajagopal, 2017). Club cells, marked by the expression of Club cell secrete protein (CCSP) (also called SCGB1A1), can self-renew and differentiate into ciliated cells or goblet cells in response of injury or allergen exposure, albeit they represent functionally differentiated cells (Kotton and Morrisey, 2014; Tata and Rajagopal, 2017). While basal cells primarily serve as a stem cell population to replenish club cells after injury, club cells serve as a primary progenitor for both ciliated cells and another sort of secretory cells named goblet cells (Rawlins et al., 2009). In response to naphthalene administration in mice, a small population of club cells termed 'variant club cells', which reside at the branch points of small airways and show low expression of CCSP, replicate and repopulate the damaged airways epithelium (Tata and Rajagopal, 2017). The remarkable resilience and heterogeneity of club cell are crucial factors contributing to plasticity of the lung epithelium.
Goblet cells are situated in the epithelium of the conducting airways, often with their apical surfaces protruding into the lumen. The distribution of goblet cells varies with species, airway level and disease status (Bustamante-Marin and Ostrowski, 2017). The well-appreciated roles of goblet cells are secretion of mucin, anti-microbial proteins, chemokines and cytokines for innate immunity and maintenance of the mucosal barrier (Knoop and Newberry, 2018). In the upper respiratory tract, goblet cells secrete Mucin 5AC (MUC5AC) and Mucin 5B (MUC5B) (Knoop and Newberry, 2018),

which are gel-forming mucins encoded by MUC genes. Goblet cells can be stimulated by paracrine and autocrine mediators, especially ATP (Bustamante-Marin and Ostrowski, 2017). The increasing production of secreted mucins is a typical symptom of asthma, COPD and cystic fibrosis (Bustamante-Marin and Ostrowski, 2017). Goblet cells can be derived from conventional club cells after injury or disease states (Kotton and Morrisey, 2014). Goblet cell hyperplasia is a hallmark of T help type 2 cells (Th2) responses, provoked by excessive production of interleukin 13 (IL-13) in the lung (Knoop and Newberry, 2018).

I-2-4 Resident immune cells

Distinct populations of innate immune cells including macrophages (i.e. alveolar and interstitial macrophages) and innate lymphoid cells (e.g. Natural Killer cells, dendritic cells) reside in adjacent areas of conducting airways and alveoli and orchestrate the adaptive immune system via an extensive crosstalk (Holt et al., 2008). For instance, dendritic cells (DCs) reside in the airway mucosa and secret CC-chemokine ligand 21 (CCL21) to increase infiltration of DCs, $CD4^+$ and $CD8^+$ T cells to interstitial areas of the lung (Altorki et al., 2019). Macrophages, T cells, B cells as well as DCs in alveolar duct and interstitium are involved in local immune responses following injury or infection. Resident immune cells within the lung not only provide local immunity and protection from infections but also cause disease when dysregulated. For example, exposure to allergens may activate innate immune cells such as dendritic cells and provoke exaggerated Th2 immune responses including excessive production of Interleukin-4 (IL-4) and IL-13, thereby contributing to the development and progression of asthma (Kudo et al., 2012).

I-2-5 Smooth Muscle Cells

Smooth muscle cells (SMCs) are major parts of airways and vessels, required to maintain the structure and contractile function of the lung. Mature SMCs contain myofilaments and express a unique repertoire of proteins required for their contractile function including smooth muscle myosin heavy chain 11 (MYH11), α-SMA (smooth muscle α-actin, ACTA2), and SM22α (TAGLN) (Rensen et al., 2007). Albeit morphologically similar, SMCs in the vasculature (VSMC) and airway (ASMC) exhibit marked differences in signaling and transcriptional programs during development and disease. A subpopulation of VSMCs is derived from mesenchymal cells (Moiseenko et al., 2017), and another population from the mesothelium (Que et al., 2008). The development of ASMCs occurs at an earlier timepoint than VSMCs. ASMCs appear to develop from the local mesenchymal cells around the tips of epithelial bud in the distal lung. These mesenchymal cells express fibroblast growth factor 10 (*Fgf10*), elongate, and start to express smooth muscle (SM)-

specific proteins at gestation day 11 in the mouse and week 5 in humans in response to mechanical tension and epithelial cell derived morphogens (Badri et al., 2008; Mailleux et al., 2005) (Figure 2). Mature airway smooth muscle (ASM) cells in adult lung do not only represent a contractile tissue required for mechanical stability, but also participate in bronchial remodeling and airway inflammation, which has consequences for regulation of the bronchomoter tone (Pelaia et al., 2008). Similar to immune cells, ASMC may secrete a broad range of cytokines/chemokines such as interleukin 1β (IL-1β), tumor necrosis factor alpha (TNF-α) (Stamatiou et al., 2012), interferon-γ-inducible protein-10 (CXCL10) (Clarke et al., 2010) and transforming growth factor β1(TGF-β1) (Pan et al., 2018) to elicit autocrine and paracrine effects, which affects proliferation and migration of ASMC or neighboring cells (Figure 2). On the other hand, interleukin-17A (IL-17A) secreted by T help type 7 cells (Th17) and IL-13 secreted by Th2 cells inhibit proliferation of ASMC but enhance its contractility leading to hyper-responsiveness (Kudo et al., 2012; Risse et al., 2011). High plasticity is a hallmark of ASMCs during lung development and during pathogenic events at adult stages.

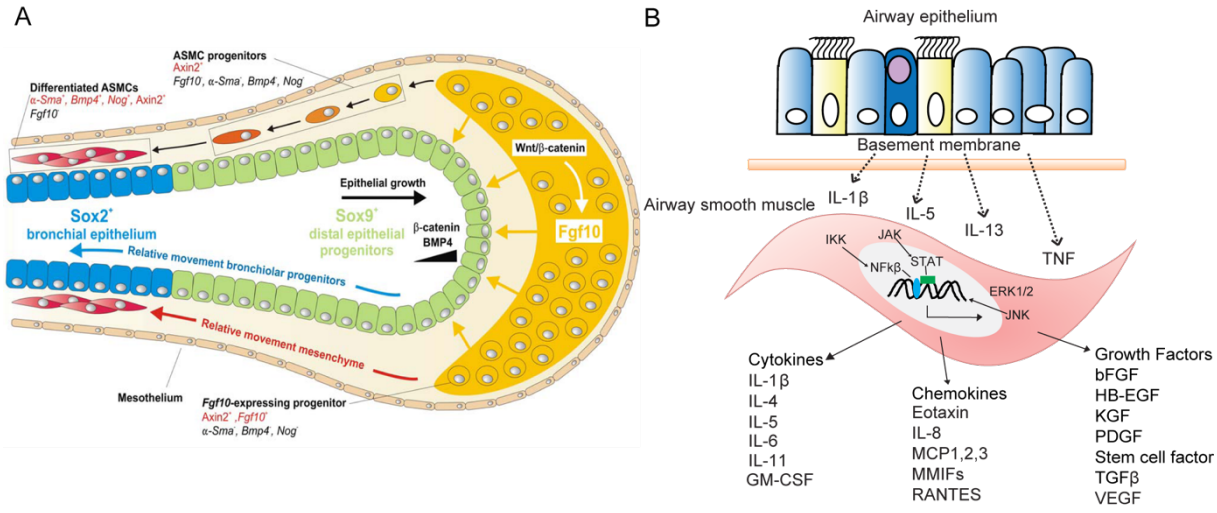


Figure 2. (A) Model of coordinated differentiation of conducting airway epithelium and airway smooth muscle. A subpopulation of *Fgf10*-positive cells located in the distal submesothelial mesenchyme supply progenitor cells for ASMCs. Their proliferation is dependent on mesenchymal Wnt signaling. Adapted from *Volckaert et al., 2014* (Volckaert and De Langhe, 2015). **(B)** Paracrine and autocrine function of ASMCs. Modified from *Volckaert et al., 2015*, *Barkauskas et al., 2017*, *Faiz et al., 2018* (Barkauskas et al., 2017; Faiz et al., 2018; Volckaert and De Langhe, 2015).

I-2-5-1 SMC plasticity

Unlike other differentiated cells, SMCs exhibit remarkable plasticity, meaning they are able to acquire different phenotypic states depending on functional demands. In response to various injuries and harmful environmental factors, contractile SMCs undergo a substantial phenotypical switch to immature and proliferative SMCs, reminiscent of SMCs that are seen during embryonic developmental. In general, this phenotypic switch is characterized by markedly reduced expression

of SMC-specific contractile marker genes and gain of large amounts of synthetic organelles i.e. rough endoplasmic reticulum, Golgi, mitochondria and free ribosomes. De-differentiated SMCs elicit autocrine or paracrine function by secreting various extracellular matrix, cytokines/chemokines, and were therefore named synthetic SMCs (Alexander and Owens, 2012) (Figure 3). The phenotypical modulation and subsequent dysfunction of SMC play a central role in the pathogenesis of diverse lung parenchymal diseases due to alterations in contractility, innate and adaptive immune responses and altered structure of airways and vessels.

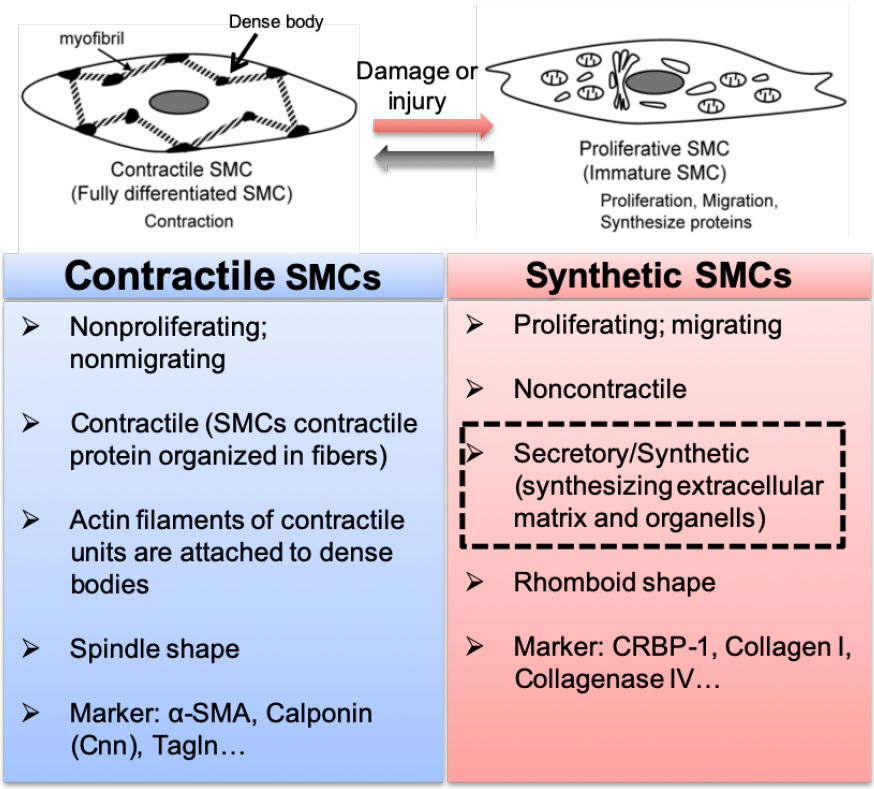


Figure 3. Characteristics of contractile and synthetic SMCs. Modified from *Milewicz et al., 2010* (Milewicz et al., 2010).

I-2-5-2 SMC related diseases

Phenotype switching of SMCs plays an important function in vascular diseased but also outside the vessels, e.g. in bronchial musculature of the lung. In vessels, the prototypical disorder that involves phenotype switching of SMCs is atherosclerosis, a chronic disease of the arterial wall characterized by accumulation of lipids together with infiltration of immune cells (Hansson and Hermansson, 2011). A recent study revealed that the majority of SMCs within atherosclerotic lesions is derived from resident SMCs in the medial layer that underwent phenotypic modulation and re-entered the cell cycle. A significant fraction of these SMCs expresses both SMC and macrophage markers and were therefore termed macrophage-like cell (Allahverdian et al., 2018).

In lungs, SMCs exert a pivotal role in the pathogenesis of asthma. Asthma is a chronic inflammatory airway disorder. Acute asthma attacks are provoked by stimuli such as cold air and exercise, inhaled allergens such as house dust mites (HDM) and pollen, leading to bronchial hyperreactivity, airway remodeling and ultimately airway constriction (Lambrecht and Hammad, 2012). Airway remodeling that occurs in nearly all forms of asthma includes increased airway smooth muscle (ASM) mass (James et al., 2012), mucus-producing goblet cell hyperplasia, subepithelial fibrosis and airway epithelial alterations (Lambrecht and Hammad, 2012). In the pathogenesis of asthma, airway smooth muscle cells not only function as the primary effector cells by regulation of the bronchomotor tone, but also secret chemokines and cytokines thereby modulating inflammatory responses in an autocrine or paracrine manner (Damera et al., 2009). In addition, ASMCs express numerous cell adhesion molecules (CAMs) that promote interactions among ASM and inflammatory cells like T cells and Mast cells in response to cytokines such as IL-1β, TNFα and Interferon γ (IFNγ) (Damera et al., 2009; Ramos-Barbon et al., 2010). Thus, ASMCs and immune cells form an intricate intercellular network that has a decisive function for airway remodeling in asthma.

II TLR mediated innate immune response in lung pathogenesis

Since the lungs are continuously exposed to a diverse array of microbes and organic or inorganic particulate materials, various defense strategies are needed to maintain proper tissue homeostasis. Vertebrates have developed two interactive protective systems: innate and adaptive immunity. The innate immune response represents and immediate but non-specific line of defense against pathogens. The immune cells involved in innate immunity include macrophages, neutrophils, dendritic cells (DCs) and natural killer (NK) cells, which sense the presence of pathogens within minutes. The adaptive immune response is a secondary but long-lasting, highly antigen specific defense line with a long memory, characterized by the rapid increase of T and B lymphocytes through clonal expansion.

II-1 Pathogen recognition in innate immunity

Innate immunity is critical for host defense against pathogens and also involved in pathogenesis of diverse human inflammatory disease such as asthma and atopy, as well as a variety of autoimmune disorders (Turvey and Broide, 2010). The molecular mechanism of innate immunity mainly is based on a limited repertoire of germline-encoded receptors termed 'pattern recognition receptors' (PRRs), which detect conserved and unique molecular patterns in microbes called pathogen-associated molecular patterns (PAMPs), host molecules in the form of 'damage-associated

molecular patterns' (DAMPs) or "missing self" molecules expressed by normal healthy cells but not in infected cells (Turvey and Broide, 2010).

II-2 TLR signaling in innate immune response

Recognition of extracellular or endosomal PAMPs is mediated by transmembrane proteins known as toll-like receptors (TLRs), which are the best characterized class of PRRs in mammalians (Kumar et al., 2009). There are 10 TLR receptors in human but 12 in murine that are expressed in various immune and non-immune cells such as B cells, NK cells, DCs, macrophages, epithelial cells, smooth muscle cells and endothelial cells (Kumar et al., 2009). TLR family members are differentially localized in the cell membrane and in intracellular endosomal membranes. TLR1, TLR2, TLR4, TLR5 and TLR6 localize to the cell surface and recognize mainly microbial membrane components such as lipids, lipoproteins and other related proteins (Kawai and Akira, 2010; Kumar et al., 2009). The intracellular TLRs consisting of TLR3, TLR7, TLR8 and TLR9 reside on the endosomal membrane where they recognize microbial nucleic acids (Kawai and Akira, 2010; Kumar et al., 2009). Among the endosomal TLRs, TLR3 recognizes double-stranded RNA (dsRNA) after viral infection and induces antiviral immune responses by promoting the production of type I interferon and inflammatory cytokines (Lavelle et al., 2010). TLR7 and TLR8 detect single-stranded RNA (ssRNA) derived from RNA viruses such as influenza virus and certain small interfering RNAs (Lavelle et al., 2010). TLR9 recognizes unmethylated 2'-CpG DNA motifs that are frequently present in bacteria and viruses (Lavelle et al., 2010). Upon recognition of potentially dangerous foreign nucleic acids, individual endosomal TLRs trigger specific biological responses by recruiting distinct TIR (Toll-interleukin-1 receptor) domain-containing adaptor molecules and activate distinct signaling pathways (Kumar et al., 2009). For instance, TLR7/8 and TLR9 recruit myeloid differentiation primary response 88 (MYD88), activate MAPKs (mitogen-activated protein kinases) and transcription factors such as NF-κB (nuclear factor κB) and IRF7 [IFN (interferon) regulatory factors 7] to induce the production of inflammatory cytokines (Kumar et al., 2009) (Figure 4). TLR3 recruits TRIF [TIR-containing adaptor-inducing IFNβ] (Kumar et al., 2009) and stimulates downstream signaling pathways that lead to activation of the transcription factors IRF3 [IFN (interferon) regulatory factors 3] and NF-κB, which leads to subsequent induction of type I interferon and inflammatory cytokines. To discriminate self from non-self in nucleic acid sensing, a multilevel, fail-safe system that combines cellular localization of nucleic acids, nucleic acid secondary structure, nucleic acid sequence and chemical modification with nuclease-mediated degradation has evolved in vertebrates to selective recognize foreign nucleic acids. Nucleases rapidly degrade most self-nucleic acids before they can be sensed by nucleic acid receptors.

However, if potential immune-stimulatory self-nucleic acids are accessible by their detectors, which may happen when nucleic acids show up at aberrant locations or are inappropriately cleared, autoimmune diseases will occur (Kawai and Akira, 2010; Schlee and Hartmann, 2016).

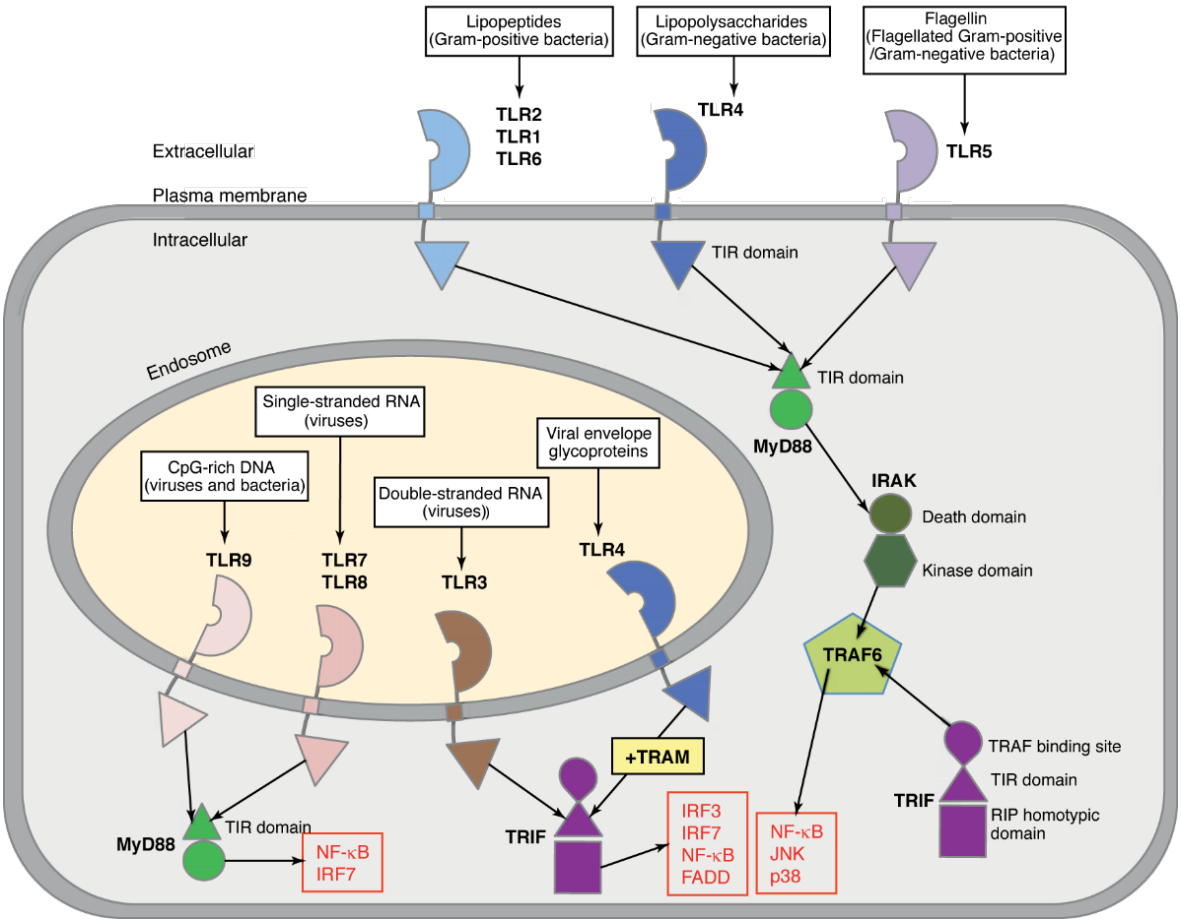


Figure 4. TLRs signaling pathways. Modified from *Gangloff ef al., 2012* (Gangloff, 2012).

II-3 Endosomal TLR signaling pathway in lung diseases

The lung expresses the full array of TLRs that play crucial roles in host defense against bacteria, fungi and respiratory viruses. In particular, endosomal RNA-sensing TLRs including TLR7 and TLR8, which are not only expressed in the resident and recruited immune cells but also in the pulmonary epithelial cells, airway SMCs and innate lymphoid cells, play a prominent role in both viral clearance and disease pathogenesis during infection (Kovach and Standiford, 2011). TLR7/8 recognize ssRNA present in the influenza A virus (IAV), while TLR3, TLR7 and TLR8 all contribute to cytokine and interferon production in respiratory syncytial virus (RSV) induced innate immune response (Jensen and Thomsen, 2012). In addition to their essential function in infection, TLRs also have the capacity to recognize ligands from nonpathogenic organisms as well as various endogenous ligands in non-infectious disease processes. For instance, TLR4 signaling is important in chronic obstructive pulmonary disease (COPD) and cystic fibrosis (CF), TLR2/4 signaling in

acute lung injury and interstitial lung disease, and the endosomal TLRs in the induction (TLR3) and exacerbation (TLR7, TLR8 and TLR9) of asthmatic inflammation (Kovach and Standiford, 2011). The functions of endosomal TLRs in asthma include: 1) activation of residential mast cells during acute responses (Papaioannou et al., 2017); 2) activation of Th2 cells and recruitment of eosinophils during late responses (Papaioannou et al., 2017); 3) activation of lung DCs and resistance to immunosuppression by TREGS during induction and maintenance of asthmatic phases (Akbari et al., 2003; Barnes, 2008). PRR-mediated sensing directs specific adaptive immune responses by presenting antigens from pathogens to antigen receptor-carrying T cells and B cells, according to the origin of the antigens and the type of infection encountered. TLRs have been suggested to play an important role in the initiation and modulation of adaptive immune responses by regulating dendritic cell maturation, proliferation/maturation of B cells for antibody production (IgM, IgA, IgG), T helper subset differentiation (Th1, Th2, Th17) and immune tolerance (Tipping, 2006). Therefore, TLR represent important targets for development of new therapeutic strategies in inflammatory disease like asthma, where activation at the interface between innate and adaptive immune responses is critical (Papaioannou et al., 2017).

III Epigenetic mechanisms regulating lung cell/tissue homeostasis

Epigenetics is the study of heritable changes in gene activity or expression that occur without any changes in the DNA sequences. Such effects on the transcription pattern often result from external or environmental factors. Many types of epigenetic processes including DNA methylation, histone modifications, chromatin remodeling, insertion of histone variants and noncoding RNAs (ncRNAs) regulation have been identified (Figure 5).

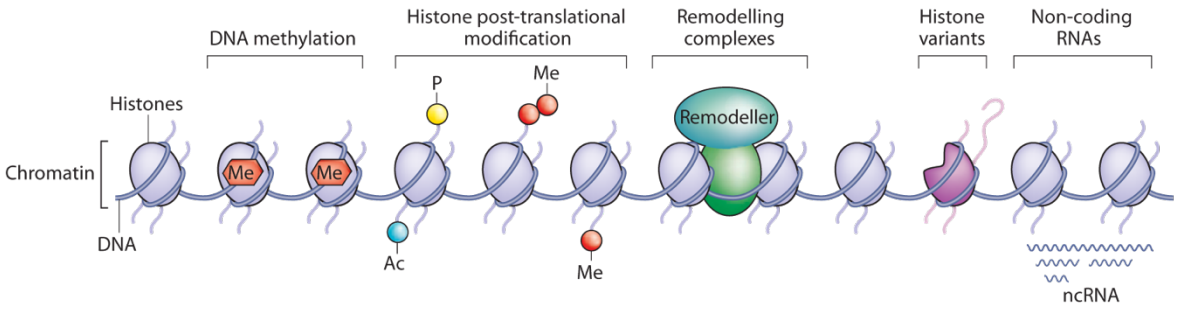


Figure 5. Epigenetic processes that affect chromatin structure: DNA methylation, histone modification, chromatin-remodeling, histone variant deposition, non-coding RNA (ncRNAs) guided chromatin structure alternation. Ac, acetyl; Me, methyl; P, phosphate. Adapted from *Dulac et al., 2010* (Dulac, 2010).

III-1 Major Mechanisms of Epigenetic Regulation

III-1-1 DNA methylation

DNA methylation is a biological process involving the covalent transfer of a methyl group to the C-5 position of the cytosine ring by DNA methyltransferases (DNMTs) (Jin et al., 2011). DNA cytosine methylation does not occur in several animal models such as Drosophila melanogaster, Caenorhabditis elegans, fission yeasts and bakers' yeasts (Greenberg and Bourc'his, 2019). In contrast, about 60–80% of the CpG sites in the mammalian genome is modified by 5-methylcytosine (5mC) (Smith and Meissner, 2013). Most of the CpG methylation is involved in genome-wide silencing of retrotransposons and repetitive elements. DNA methylations are also found at gene promoter sites where the methylation status of CpGs plays decisive roles to modulate cell-type specific transcription required for establishing and maintaining cell identity during development and tissue homeostasis. Unmethylated CpGs within promoters are normally associated with transcriptionally active genes, while hyper-methylation of CpGs within promoters usually blocks binding of transcription factors to these motifs, thereby impeding activation of transcription (Greenberg and Bourc'his, 2019). Paradoxically, strong DNA methylation occurs within transcribed gene bodies of several tissue-specific and imprinted genes. The degree of methylation is often positively correlated with transcriptional strength, indicating that DNA methylation in gene bodies might have a complex role for gene transcription (Jones, 1999).

DNA methylation errors have been discovered in several diseases including Prader-Willi, and Angelman syndromes, Fragile X syndrome and Facial anomalies (ICF) syndrome. In addition, genome-wide DNA hypomethylation (20-60% loss of overall 5mC) is implicated to play an important role in carcinogenesis (Portela and Esteller, 2010), while global hypomethylation occurs mainly at repetitive sequences and leads to chromosomal instability. Hypomethylation is also found at promoters of oncogenes in several cancer types, which results in aberrant cell growth and survival of cancer cells (Feinberg et al., 2006; Portela and Esteller, 2010). Likewise, hypermethylation is observed at specific CpG islands of tumor suppressor genes, thereby causing transcriptional inactivation of genes involved in main cellular pathways such as DNA repair and apoptosis (Portela and Esteller, 2010). DNA methylation is also involved in the aging-related epigenetic drift. 5mC levels gradually decline over the lifetime of an organism, most likely due to aging-related changes in DNMT expression and availability of methyl donors resulting from cumulative environmental exposures and lifestyle factors (Jung and Pfeifer, 2015).

III-1-2 Histone modifications

The nucleosome is the fundamental unit of chromatin encompassing 146 base pairs of DNA wrapped around an octamer of histone proteins (Li and Reinberg, 2011). The histone octamer is composed of two H2A–H2B dimers and two H3–H4 dimers, wherein the N-terminal tails of the four core histones and the C-terminal tail of histone H2A protrude from the nucleosome core particle. The histone tails are highly basic and contain residues targeted by post-translational modifications including acetylation, methylation, phosphorylation, ubiquitination, and ADP-ribosylation (Liu et al., 2015). These histone–tail modifications modulate the histone–DNA and nucleosome–nucleosome interactions, and play important roles during transcription, replication, recombination, and DNA repair. Histone acetylation is catalyzed by histone acetyltransferases (HATs) that transfer an acetyl group from Acetyl Coenzyme-A (Acetyl-CoA) to the NH3+ group on lysine. The acetyl group can be erased by histone deacetylases (HDACs) (Beerman and Rossi, 2015). Acetylation of nucleosomes surrounding transcription starts sites (TSSs) may lead to decreased nucleosome occupancy by neutralizing the positive charge of lysine residues, weakening charge-dependent interactions between a histone and nucleosomal DNA, linker DNA or adjacent histones. This process increases the accessibility of DNA to the transcriptional machinery, thus activating gene transcription (Zentner and Henikoff, 2013). Methylation of histones involves the transfer of a methyl group from S-adenosyl methionine (SAM) to nitrogen atoms in lysine (Lys or K) and arginine (Arg or R) residues. Unlike histone acetylation, histone lysine methylation does not change the charge at the histone tails and these methylated lysine moieties function as docking sites for binding of downstream reader proteins that are involved in gene activation or repression. Histone methylation has varying effects on gene transcription activity depending on the specific residues that are modified, the degree and pattern of methylation, and the genomic context in which the methylation occurs. Histone H3 is the primary site of histone methylation, although other core histones can be methylated as well. Among the known histone lysine methylations on Histone H3, di- and tri-methylation at H3K4, H3K36 and H3K79 are largely associated with gene transcription activation. H3K4 tri-methylation (H3K4me3) is enriched at promoters, while H3K36 and H3K79 methylations are primarily enriched within gene bodies. Mono-methylation of H3K4 is an activating mark unique to enhancers. In contrast, H3K27me3 is generally linked to long-term transcriptional repression. H3K27me3 also marks the bivalent domains of the differentiation related genes together with H3K4me3 in embryonic stem cells, which is critical for proper cell differentiation. H3K9me3 is the hallmark of heterochromatin, whereas H3K9me2 is found more commonly at silent or lowly expressed genes in euchromatin (Jambhekar et al., 2019).

III-1-3 Epigenetic mechanisms safeguarding transcription fidelity

During transcriptional elongation, nucleosomes within gene bodies must be temporarily evicted from the DNA, thereby allowing efficient transit of the RNA polymerase II machinery. After one round of transcription histone-DNA interactions need to be reestablished to prevent initiation of transcription at undesired intragenic regions (Zentner and Henikoff, 2013). So far, several epigenetic mechanisms including histone deacetylation, histone methylation at H3K4, H3K36 and DNMT3b mediated DNA methylation have been reported to participate in inhibition of aberrant intragenic transcription initiation. Deregulation of these epigenetic events will lead to excessive production of cryptic transcripts that originate from the non-canonical promoter regions from both the sense and antisense direction (McDaniel and Strahl, 2017).

III-1-3-1 Inhibition of spurious transcription in yeast

Histone H3K36me3, which is highly enriched in the gene bodies of actively transcribed genes, is considered as a hallmark of active transcription and plays key roles in preclusion of cryptic transcription in yeast. SET2 is the only histone methyltransferases that acts on H3K36 in yeast, and is recruited by Ser-2-phosphorylated Pol II during elongation. H3K36me3 functions as a docking site for the binding of Rpd3S histone deacetylase complex, which makes the chromatin inaccessible for undesired intragenic transcription initiation by transforming gene bodies into a hypoacetylated state after passage of the elongating Pol II (Venkatesh and Workman, 2013) (Figure 6A). Set2 and H3K36me3 are also required for recruitment of the Isw1b chromatin-remodeling complex to establish proper nucleosome spacing across the gene body (McDaniel and Strahl, 2017). Nucleosomes marked with Set2-mediated H3K36me3 are refractory to binding of Asf1, Spt6, and FACT, the histone chaperons contributing to nucleosome disassembly and reassembly during transcription elongation, thus ensuring the stability of reassembled chromatin structure in the wake of Pol II transcription. Loss of genome wide H3K36me3 or Rpd3 leads to a significant increase in cryptic transcripts (McDaniel and Strahl, 2017).

III-1-3-2 Inhibition of spurious transcription in mammalian cells

Eight different mammalian enzymes, which all contain the catalytic SET domain, have been identified to methylate H3K36 *in vitro* and *in vivo* (Wagner and Carpenter, 2012). However, only the SET domain containing 2 (SETD2), the homolog of Set2 in yeast, is able to generate H3K36me3. Similar to yeast Set2, SETD2 interacts with elongating Pol II, deposits H3K36me3 in the gene body and recruits a Rpd3S-related HDAC complex via the association with the Eaf3 homolog MRG15 (Xie et al., 2011). H3K4me3, an active transcriptional histone mark enriched at promoters

near TSSs, is deposited by mixed lineage leukemia (MLL) family methyltransferases within intragenic regions of genes involved in cell identity maintenance or cell type-specific tumor suppression during transcriptional elongation (Chen et al., 2015; Smith et al., 2011).

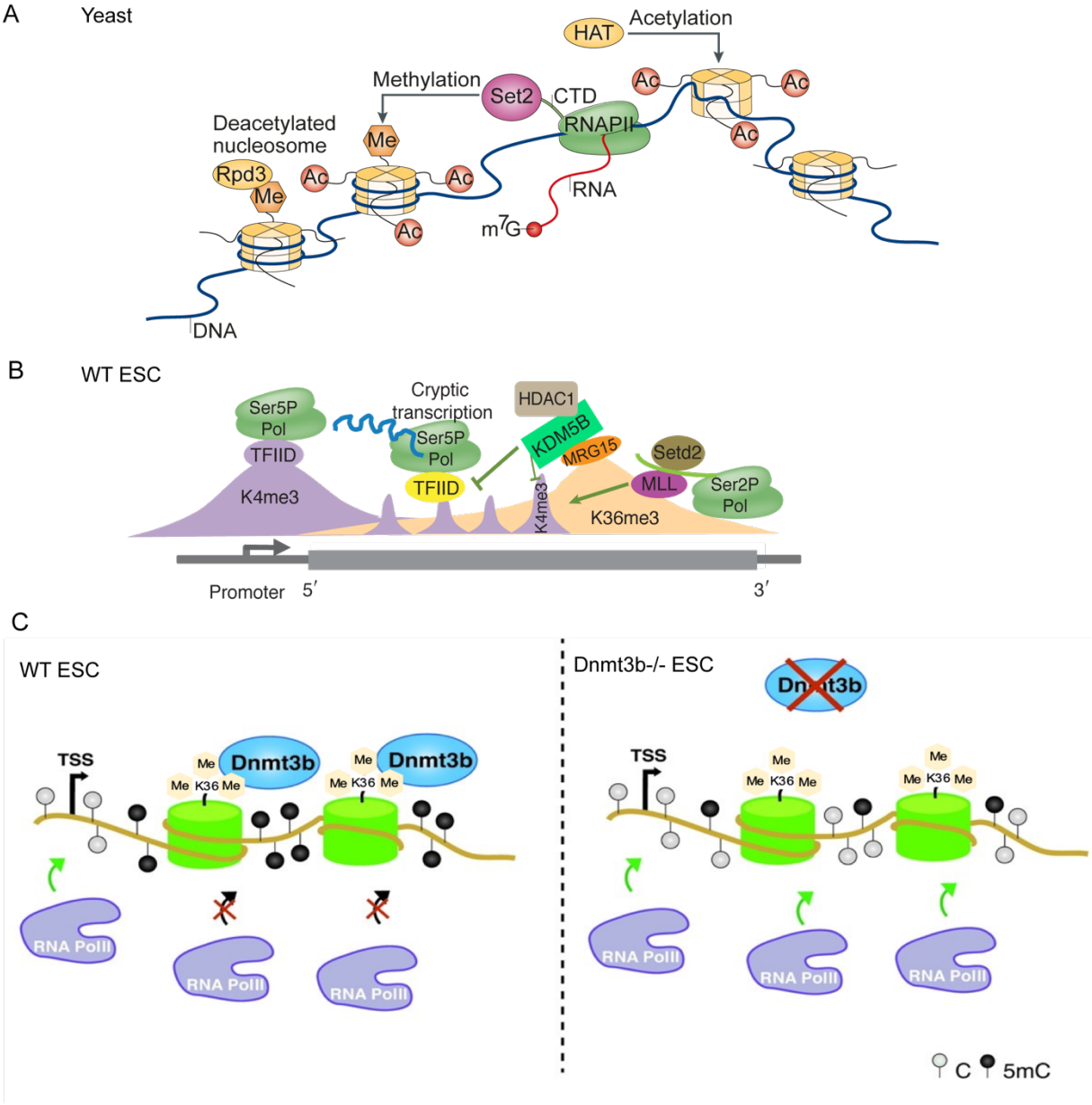


Figure 6. Models of cryptic transcription inhibition. **(A)** H3K36me3-dependent prevention of aberrant transcription in yeast. Adapted from *Wagner et al., 2012* (Teissandier and Bourc'his, 2017). **(B)** KDM5B mediated intragenic H3K4me3 demethylation represses cryptic initiation and safeguards mRNA elongation. Modified from *Xie et al., 2011* (Xie et al., 2011). **(C)** Schematic model depicting the role of Dnmt3b-dependent intragenic DNA methylation in ES cells. In wild type ESC, DNMT3b is able to methylate gene body and contributes to the formation of a repressive chromatin environment that inhibits spurious entry of Pol II. After ablation of *Dnmt3b*, gene body hypomethylation leads to Pol II intragenic entries that generate cryptic transcripts (Neri et al., 2017). Modified from *Neri et al., 2017* (Neri et al., 2017).

KDM5B, a H3K4me2/3 demethylase serving as a transcriptional repressor by removing promoter-associated H3K4me3, is recruited to H3K36me3 at least in part through interaction with the chromodomain of MRG15, primarily targeting genes that are highly expressed and involved in self-

renewal of embryonic stem cell (ESC). KDM5B-dependent demethylation of intragenic H3K4me3 is essential to safeguard transcriptional elongation via repressing aberrant intragenic entry of transcription-initiated Pol II machinery (Xie et al., 2011) (Figure 6B). H3K36me3 is also necessary for establishing DNA methylation in gene bodies. In the absence of SETD2, intragenic regions of actively transcribed genes cannot be properly methylated, primarily due to the loss of DNMT3b recruitment through its binding with H3K36me3 (Baubec et al., 2015; Morselli et al., 2017). A more recent study has demonstrated that DNMT3b mediated hypermethylation of cryptic TSSs prevents intragenic entry of Pol II initiation complex, therefore inhibiting spurious transcription (Figure 6C). In *Dnmt3b* deficient ESCs, spurious transcription is activated and generates capped (5'-cap) and partially polyadenylated spurious transcripts that might be transcribed into aberrant proteins (Neri et al., 2017). However, the cellular effects or roles of these aberrant transcripts and proteins remain enigmatic.

III-2 Epigenetic regulation of SMCs and lung tissue homeostasis

Similar to other differentiated cells, SMC display a cell-specific epigenetic signature made up by a unique pattern of histone and DNA modifications. In response to extracellular signals during injury repair and disease, SMC can undergo phenotypic modulation, de-differentiate into and acquire a more synthetic phenotype. Phenotypical switch of SMC is always accompanied with epigenetic reprogramming that dynamically control the chromatin structure to coordinate gene activation or repression. Changes in the epigenome has been observed in SMC-based lung diseases including COPD, pulmonary hypertension (PH) and asthma, indicating the crucial role of epigenetic modifiers in the regulation of SMCs and lung tissue homeostasis.

III-2-1 DNA methylation and SMC phenotype modulation

Global hypomethylation with decreasing DNA methyltransferase activity has been observed during the shift of SMCs from the contractile to the synthetic state and in proliferating SMC cultured *in vitro* (Liu et al., 2015). During the initiation or progression of atherosclerosis, in which VSMCs play a major role (Wierda et al., 2010), aberrant DNA methylation patterns can be observed in a set of genes involved in regulating proliferation, (de)differentiation and migration of SMC such as the *Srf*, platelet-derived growth factor (*Pdgf*), estrogen receptor α (*Erα*), estrogen receptor β (*Erβ*), and SMC-specific *Sm22α* (Gomez et al., 2015). DNA methylation also regulates the SMC phenotype by modulating express of extracellular matrix genes. For instance, DNA methylation inhibitor 5-Aza can partially repress collagen type I and III genes in SMCs isolated from neurogenic bladders,

attenuate PDGF-induced airway SMC migration and proliferation, and increase cellular contractility (Hodges et al., 2010).

III-2-2 Histone modification and SMC homeostasis

Histone 3 lysine 4 dimethylation (H3K4me2) has been found to promote expression of SMC contractile genes by serving as a docking site for MYOCD binding within CArG box motifs during early stages of SMC differentiation (Alexander and Owens, 2012), (Liu et al., 2015). Interestingly, H3K4me2 levels are retained when contractile gene expression is reduced or undetectable after phenotypic switching of SMC towards a less differentiated state (Gomez et al., 2013). H3K4me2 is also persistently enriched at *Myh11* promoter in SMCs that have undergone a transit to a macrophage-like state within mouse and human atherosclerotic lesions (Gomez et al., 2015). It remains unclear whether the persistence of H3K4me2 deposition within SMC contractile gene loci is essential for the reversibility of phenotypic switching during repair of vascular injuries. H3K9me2 has been implicated in regulating the pro-inflammatory phenotype of vascular SMCs by enhancing binding of NFκB and activator protein-1 (AP-1) transcription factors to inflammation-responsive genes to activate gene expression (Harman et al., 2019). P300 mediated histone acetylation is necessary for SMC differentiation and probably essential for SMC phenotypic switching (Spin et al., 2012). In vascular SMCs, P300 interacts with phosphorylated Kruppel-like factor 4 (KLF4) after TGF-β stimulation and subsequently translocates into the nucleus, where it acetylates H3 on target genes leading to gene activation (He et al., 2015). Moreover, P300 mediated hyperacetylation of H3K18 at *Cxcl8* promoter enhances CXCL8 production and secretion in airway SMCs from patients with asthma (Clifford et al., 2015). On the other hand, HDAC functions as a critical regulator for SMC proliferation. For instance, the class II HDAC 4 and 5 suppress transcription of the contractile genes *Acta2* and *Tagln* by inhibiting MYOCD activation (Spin et al., 2012). Furthermore, pharmacological inhibition of HDACs prevents mitogen-induced SMC proliferation by negatively regulating the activity of CDK/cyclin complexes and repressing cyclin D1 transcription, thus reducing the neointima formation following vascular injury (Findeisen et al., 2011).

IV TETs function in cell/tissue homeostasis control and disease

IV-1 TETs mediated 5mC oxidation and DNA demethylation

5mC is chemically and genetically stable. Despite its stability, DNA demethylation occurs by passive replication-dependent dilution or active enzymatic removal mediated by members of the Ten to eleven translocation dioxygenase (TET) family, or a combination of both. The TET family

is comprised of TET1-3, which are α-ketoglutarate (α-KG) dependent methyl cytosine dioxygenases, converting successively 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (Rasmussen and Helin, 2016). Since the oxidized forms of 5mC cannot be bound by the DNA methylation maintenance machinery, 5mC can be converted back to unmodified cytosine via replication-dependent dilution of 5hmC, 5fC or 5caC, a process known as passive demethylation. During the cycle of active DNA demethylation, 5fC and 5caC can be further excised by thymine DNA glycosylase (TDG) coupled with base excision repair (BER) to regenerate unmodified cytosine (C) (Kohli and Zhang, 2013) (Figure 7).

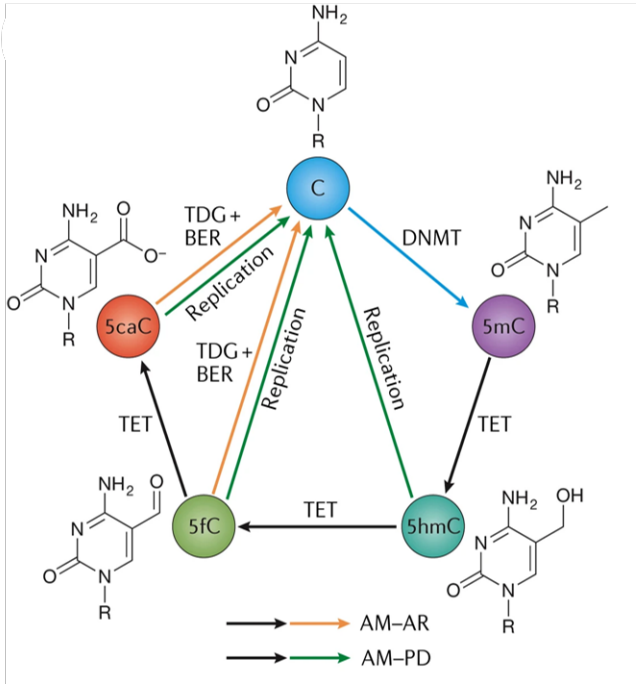


Figure 7. The cycle of active DNA demethylation. DNMTs convert unmodified cytosine to 5-methylcytosine (5mC). 5mC can be converted back to unmodified cytosine by TET-mediated oxidation to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), followed by excision of 5fC or 5caC mediated by thymine DNA glycosylase (TDG) coupled with base excision repair (BER) (the process of active modification-active removal (AM-AR)) or by replication-dependent dilution of 5hmC, 5fC or 5caC [the process of active modification-passive dilution (AM-PD)]. Adapted from Wu et al., 2017 (Wu and Zhang, 2017).

IV-2 Genomic distribution and potential function of oxidized 5mC

5mC is found in all tissues, involving up to ~4%–5% of all cytosines. Among the 5mC oxidation derivatives, 5hmC is more stable and most abundant presenting at a range of 0.3%–0.7% in various cell types or tissues. By comparison, the products of further TET-mediated oxidation, 5fC and 5caC, show ~10-fold to 100-fold lower levels than 5hmC and are often close to or under the detection limit, likely due to the low affinity of TETs to 5hmC or rapid excision by TDG-BER pathway or both. Studies of genome wide distribution of oxidized derivatives of 5mC show that

vast majority of 5hmC and 5fC/5caC are found on distinct cytosine bases in the genome, indicating that the oxidized derivatives of 5mC might have additional regulatory functions aside from being intermediates within the DNA demethylation process.

Among oxidized derivatives of 5mC by TETs proteins, 5hmC is abundant and enriched at promoters or enhancers in ESCs, PGCs (Primordial germ cells) and neurons. Although 5hmC has been implicated in the regulation gene expression, the relationship between 5hmC level and gene transcription activity is complex. 5hmC peaks are depleted around the TSS region in actively transcribed gene of human and mouse ES cells, mouse NPCs (neural progenitor cells), neurons and the cerebellum, although they usually can be found 500-2000 bases before and after TSSs (Greco et al., 2016; Pastor et al., 2013; Tan et al., 2013). However, genes with low transcriptional activity are characterized by accumulation of 5hmC near the TSSs in ESCs, NPCs and neurons (Colquitt et al., 2013; Hahn et al., 2013; Tan et al., 2013). 5hmC is most abundant at promoter-distal regulatory elements, particularly enriched in distal regulatory elements like enhancers in ESCs, but globally depleted at enhancers (p300 sites) in NPC and neurons (Hahn et al., 2013; Yu et al., 2012). The enrichment of 5hmC in enhancers is also found in other cells/tissues although its impact on enhancer activity is controversial. For instance, 5hmC is enriched within poised enhancers but depleted within active enhancers in the placenta. In contrast, 5hmC is accumulated in enhancers of tissue specific genes and showed greater activity in precursor T-cells and cardiomyocytes (Greco et al., 2016; Green et al., 2016; Tsagaratou and Rao, 2013). 5hmC not only displays unique regulatory functions in these tissues via demethylation of cytosine but also directly through altering the chromatin environment or through the recruitment or exclusion of binding proteins such as methyl-CpG binding domain protein 1 (MBD1) or methyl CpG binding protein 2 (MECP2) (Wu and Zhang, 2017), thereby regulating chromatin structure and gene expression.

Besides the characteristic accumulation at promoters or distal regulatory elements, high levels of 5hmC in gene bodies have been also detected in various cells or tissues including human and mouse ESCs, mouse cardiomyocytes, mouse liver, brain and human melanomas, in which intragenic 5hmC level is positively correlated with gene expression (Khare et al., 2012)·(Greco et al., 2016; Pastor et al., 2013). During neuronal development, 5hmC is enriched preferentially within gene bodies of activated neuronal function-related genes and gets increased during neuronal differentiation (Hahn et al., 2013). Furthermore, DNA hydroxymethylation marks the bodies of highly expressed cardiac specific genes as well as distal regulatory regions with enhanced activity in cardiomyocytes (Greco et al., 2016). In thymic and peripheral T cell subsets, the intragenic 5hmC level shows a remarkable positive correlation with the level of gene expression, Pol II occupancy and the level of H3K36me3 (Tsagaratou et al., 2014). Mapping the genomic distribution of 5hmC in neurons disclosed that 5hmC peaks at exon-intron boundaries related to RNA splicing in genes involved in synapse-

related functions in both human and mouse brain, suggesting a potential link of 5hmC with splicing (Khare et al., 2012; Wen et al., 2014). In line with this, intragenic 5mC and 5hmC exhibit opposing effects on CCCTC-binding factor (CTCF) mediated pre-mRNA splicing: CTCF promotes inclusion of weak upstream exons through RNA polymerase II pausing, whereas 5-methylcytosine evicts CTCF, leading to exon exclusion. TET-catalyzed oxidation of intragenic 5-methylcytosine promotes recruitment of CTCF, thereby facilitating CTCF-dependent alternative splicing. Reduced TET levels culminate in increased 5-methylcytosine, CTCF eviction and exon exclusion (Marina et al., 2016; Shukla et al., 2011).

5fC and 5caC are much less abundant than 5hmC, likely due to the high efficiency of TDG-BER mediated excision of these two bases or the low efficiency of conversion of 5hmC to 5fC and 5caC, or both. 5fC and 5caC accumulate preferentially at distal regulatory elements, bivalent promoters and gene bodies of *Tdg*-depleted cells, similar to the enrichment of 5hmC at these regions (Neri et al., 2015; Rasmussen and Helin, 2016). In *Tdg*-depleted mouse ESCs, accumulated 5caC and 5fC in gene bodies significantly reduce the Pol II transcription elongation rate, indicating that 5fC/5caCs within a transcribed genomic region may serve as speed brakes for the transcriptional machinery to enable fine-tuning of the elongation rate. The pausing effect might be most significant for the transcription of highly expressed genes, which is relatively rich in 5hmC (Wang et al., 2015). The impact of 5fC/5caCs on the transcription elongation rate might be achieved by compromising NTP incorporation due to the specific hydrogen bonds between the 5-carboxyl group of 5caC and the conserved epi-DNA recognition loop in the polymerase (Wang et al., 2015). In addition, 5fC and 5caC appear to interact with numerous proteins including TDG, p53, DNA repair factors, chromatin remodeling factors, and forkhead box transcription factors (TFs) (Iurlaro et al., 2013; Spruijt et al., 2013). While some of these proteins may be recruited as a result of DNA repair associated with active DNA demethylation process, the biochemical and physiological relevance of these interactions is unclear.

IV-3 TET proteins

TET family (TET1, TET2 and TET3) are evolutionarily conserved iron (II)/α-ketoglutarate (Fe (II)/α-KG)-dependent dioxygenases that contain Cysteine-rich and double-stranded β-helix (DSBH) domains at the carboxyl terminus responsible for the catalytic activity (Solary et al., 2014). Whereas TET1 and TET3 have an amino-terminal CXXC-type zinc-finger domain that can bind DNA directly, TET2 lacks the putative CXXC domain and can be recruited to chromatin via interacting with its cofactor IDAX (also known as CXXC4), which originates from ancestral *Tet2* gene duplication (Solary et al., 2014) (Figure 8). The TET proteins, although harboring the same

catalytic activity, are involved in distinct biological processes due to their differential expression in diverse cells/tissues or their substrate preference (Hu et al., 2015; Jin et al., 2016). *Tet1* is highly expressed in ESC and PGCs but maintained at low levels in differentiated cells. *Tet2* and *Tet3* are more widely expressed. *Tet2* is found in ESCs and plays a unique role in the regulation of hematopoietic stem cell differentiation. Although TET1 and TET2 functions are required for ESC lineage specification, they are dispensable for pluripotency and their loss is compatible with embryonic and postnatal development. While both *Tet1* and *Tet2* are highly expressed in mouse ESCs, enrichment of 5hmC at the promoter/TSS is most likely associated with TET1, and the 5hmC level in gene bodies and boundaries of exons in active genes is correlated with TET2 activity (Huang et al., 2014). *Tet2*, the most highly expressed member of the TET family, is vital to create a balance between DNA methylation and hydroxymethylation together with DNA methyltransferase (DNMTs) in hematopoietic stem cells (HSCs), which ensures correct gene expression and cellular function. Ablation of *Tet2* leads to reduced 5hmC levels, increased global 5mC levels and myeloid malignancies due to skewing of differentiation towards the myeloid lineage and expansion of HSC pool (Avgustinova and Benitah, 2016). TET2 enzyme has been shown to regulate the expression of key cardiac genes, such as myosin heavy chain 7 (*Myh7*), through 5hmC deposition on the gene bodies and at enhancers (Greco et al., 2016).

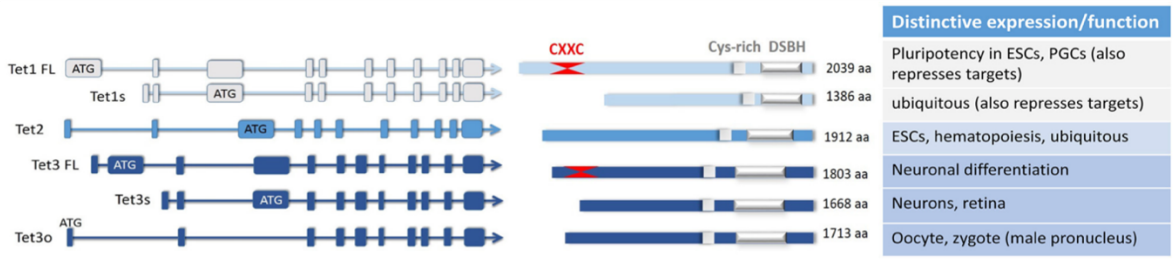


Figure 8. The major isoforms of the three TET enzymes, and their differential expression and functions. Adapted from *Melamed et al., 2018 (Melamed et al., 2018)*. *Tet1* full length isoform (*Tet1* FL) and *Tet3* full length isoform (*Tet3* FL) have at their N-terminus two Cys4-type zinc finger motifs which make up the CXXC domain. Short forms of *Tet1* (*Tet1*s) and *Tet3* (*Tet3*s) are produced as a result of differential promoter usage, giving rise to short transcripts and truncated proteins that completely lack the CXXC domain. In the oocytes, a unique *Tet3* isoform (*Tet3*o) is found. Its expression is initiated at a distinct promoter, presumably as a result of activation by oocyte specific factors (Jin et al., 2016). Catalytic domain of TETs is comprised Cys-rich domain and DSBH domain, which harbors key catalytic motifs and interacts with Fe (II) and α-KG (An et al., 2017).

Recent studies showed that mammalian TET3 exists in three major isoforms with distinct functions and localizations in different cell types (Figure 8). The *Tet3*o isoform is the sole TET protein involved in epigenetic reprograming in zygotes. *Tet3*o is highly expressed in the paternal pronucleus and rapidly oxidizes 5mC to 5hmC shortly after fertilization. DNA replication results in the dilution of 5mC and all its oxidized derivatives, thereby rapidly reducing the 5mC signal from the highly methylated paternal genome (Gu et al., 2011). Blocking TET3 mediated 5hmC affects paternal gene activation, leading to reduced developmental fitness and survival (Gu et al., 2011). The

expression of *Tet3* in ESC is negligible but increases upon differentiation, which is consistent with its abundance in various adult tissues (Gu et al., 2011; Wang et al., 2013; Wossidlo et al., 2011). *Tet3* is expressed at the highest level among 3 enzymes in the brain and has been recently implicated in neural progenitor proliferation and differentiation (Hahn et al., 2013). *Tet3* full length (*Tet3* FL) and *Tet3* short (*Tet3*s) isoforms are upregulated during ESCs cell differentiation towards the neuronal lineage, in which *Tet3* FL exhibits a high affinity for 5caC and localizes precisely at the TSSs of genes involved in lysosome function, mRNA processing and the BER pathway (Jin et al., 2016). *Tet3* FL binding peaks are largely overlapping with H3K4me3 but not with H3K27me3 in embryonic mouse brain (Jin et al., 2016), indicating the positive correlation between active transcription of targeted genes and TET3-mediated 5hmC deposition. *Tet3*s is recruited to the DNA in the retina by the RE1-silencing transcription factor (REST) DNA-binding factor and TET3s mediated 5hmC enrichment within promoter region results in the transcriptional activation of targeted gene (Perera et al., 2015). TET3s is also shown to interact with H3K36 methyl transferases (NSD3) to facilitate neuronal gene activation (Perera et al., 2015). However, the mechanisms regulating gene transcription in specific cell type or tissue are still poorly explored, although the distinct recruitment mechanisms and diverse expression patterns of TET enzyme isoforms have been described.

IV-4 Biological functions of TETs

IV-4-1 The role of TETs in the regulation of stem cell homeostasis

The biological importance of TETs has been widely explored by loss-of-function studies. Depletion of single *Tet1* or *Tet2*, which are highly expressed in mouse ESCs, decreases 5hmC levels and induces transcription changes, albert the deficiency in either gene does not impair pluripotency or prevent development. *Tet1* and *Tet2* double knockout (DKO) ESCs lose most of their 5hmC content without disrupting pluripotency. However, most DKO embryos die perinatally, although some develop normally (Wu and Zhang, 2017). Compared to *Tet1* and *Tet2*, *Tet3* expression levels are very low in ESCs. Nevertheless, *Tet3* knockout ESCs undergo rapid apoptosis and differentiate poorly, although they can be successfully induced to generate neural progenitor cells (Li et al., 2015). *Tet3* is highly expressed in oocytes and zygotes and plays an important role in rapidly demethylating the paternal genome after fertilization thereby producing genome-wide increases in the oxidized 5mC intermediates 5hmC, 5fC, and 5caC. Ablation of *Tet3* in mice leads to embryonic lethality (Shen et al., 2014). Importantly, *Tet3* haploinsufficiency causes neonatal lethality or sub-Mendelian ratios in mice, indicating a requirement for TET3 function in differentiated tissues (Inoue et al., 2015). Besides critical functions in pre-implantation development, ESC pluripotency, and differentiation, TETs also play key roles in the homeostatic control of adult stem cells. After *Tet1*

deletion *in vitro*, the expression of several genes related to adult neural progenitor cell (NPC) proliferation was repressed due to hypermethylation in promoter regions and *Tet1* knockout mice exhibit decreased numbers and reduced proliferation of NPCs in the adult sub granular zone (SGZ) (Yao et al., 2016), indicating that TET1 regulates neurogenesis via modulating 5hmC deposition at promoters (Yao et al., 2016). In adult neural stem cells (aNSCs), TET2 mediated 5hmC is elevated during the differentiation (Li et al., 2017). Abrogating *Tet2* expression within the hippocampal neurogenic niche or adult neural stem cells impairs neurogenesis, learning and memory (Gontier et al., 2018). Along the same line, restoration of TET2 rescues age-related decline in adult neurogenesis and enhances learning and memory in mice (Gontier et al., 2018). *Tet2* and *Tet3* are both highly expressed in murine hematopoietic stem/progenitor cells (HSPCs), and their individual deletion in mice results in aberrant hematopoiesis characterized by the expansion of HSPCs with enhanced self-renewal capacity, augmented hematopoietic repopulation and preferential differentiation towards myeloid lineages.

IV-4-2 The role of TETs in the regulation of somatic cell/tissue homeostasis

Tet2 and *Tet3* double mutant mice develop a non-compaction cardiomyopathy with embryonic lethality (Fang et al., 2019). TET2 and TET3 mediated DNA hydroxymethylation increases chromatin accessibility and subsequent binding of Ying-yang1 (YY1) to their genomic targets, which is required for the formation of proper long-distance interactions (enhancer-promoter looping) at key genes involved in heart development (Fang et al., 2019). Moreover, *Tet2* knock down in embryonic cardiac cells does not affect the global level of 5hmC but erases the 5hmC deposition within the gene bodies or enhancers of essential cardiac genes such as *Myh7*, thereby repressing their expression (Greco et al., 2016). TET2 was described as an essential regulator of SMC differentiation. TET2 binds to the CArG rich region at the promoters of contractile genes such as *Myocd*, *Srf* and *Myh11* in SMC and mediates conversion of 5mC to 5hmC for activation of contractile gene expression (Liu et al., 2013). *Tet2* knockdown in human coronary artery SMCs (hCASMCs) results in impaired SMC differentiation, while localized delivery of *Tet2* overexpressing viruses to the site of femoral arterial wire injury enhances SMC differentiation and improve vascular repair (Liu et al., 2013). Moreover, *Tet2* knockdown significantly diminishes chromatin accessibility and increases the level of histone 3 lysine 27 trimethylation (H3K27me3) at *Myocd*, *Srf* and *Myh11* gene loci, resulting in suppression of SMC contractile genes (Liu et al., 2013). TET3 is the most highly expressed TET enzyme in the brain and plays a critical role in the regulation of brain functions by controlling neuronal differentiation and neural plasticity. Inhibition or depletion of TET3 in mouse differentiated neurons can impact synaptic function (Yu et al., 2015; Zhu et al., 2016). During development of mature olfactory sensory neurons (mOSM), 5hmC accumulation

over gene bodies increases and positively correlates with gene transcription activity. Overexpression of *Tet3* in mOSNs markedly enhances gene-body 5hmC levels and gene expression consistent with the positive role of 5hmC in transcription (Colquitt et al., 2013).

IV-4-3 The role of TETs in diseases

TET loss-of-function is strongly associated with cancer. In humans, recurrent deletions or mutations of all three *Tet* genes were found in a wide range of myeloid malignancies and in T-cell lymphomas (Abdel-Wahab and Levine, 2013; Ko et al., 2015). In mouse models, single deletion *Tet1* or *Tet2* or *Tet1*/*Tet2* and *Tet2*/*Tet3* double deficiency causes hematopoietic malignancies by inducing aberrant self-renewal. Alternatively, differentiation of hematopoietic stem/progenitor cells occurs, which triggers the rapid development of an aggressive, fully-penetrant and cell-autonomous myeloid leukemia (Rasmussen and Helin, 2016). Dysregulation of TET function has also been implicated in lung pathogenesis. Hypomethylation of a CpG site in the *Tet1* promoter and increased global 5hmC levels in nasal mucosa are associated with childhood asthma (Zhu et al., 2020). In addition, marked changes of genome-wide 5mC + 5hmC, 5mC, and 5hmC modifications were found in air pollution–induced lung injury, particularly at genes and pathways associated with oxidative stress responses, epithelial function and immune cell responses, which are linked to altered TET1 activity. TET1 was also suggested to inhibit asthma-related phenotypes including airway hyperresponsiveness (AHR) and lung eosinophilia in mice. *Tet1*, *Tet2* and *Tet3* are downregulated, while *Dnmts* remain unchanged in the dust mites (HDM)-induced allergic airway model (Zhu et al., 2020). These findings indicate that TET1 plays a vital role in responding to pollutants and may serve as a therapeutic target for neutralizing the negative impact of environmental factors. However, the roles of TET2 or TET3 in lung pathogenesis are still elusive.

IV-4-4 Non-enzymatic function of TETs

In addition to the proverbial enzymatic role, TET proteins also possess non-enzymatic activities. TET proteins may either recruit other factors or prevent DNMT-dependent deposition of 5mC especially at methylation-free promoters. For instance, TET1 contributes to transcriptional repression by recruiting the SIN3A histone deacetylase, the PRC2 complex or HDAC1 to target genes (Damal Villivalam et al., 2020; Williams et al., 2011; Wu et al., 2011). Interaction with TET1 activates the enzymatic activity of hMOF and facilitates hMOF-mediated H4K16 acetylation by inhibiting auto-acetylation of hMOF, thereby enhancing transcription of target genes (Zhong et al., 2017). These effects do not seem to require the catalytic activity of TET1. TET2 selectively represses transcription of interleukin-6 (*Il-6*) by facilitating recruitment of HDAC2 in dendritic cells

and macrophages during resolution of inflammation responses (Zhang et al., 2015). TET3 interacts with and recruits HDAC1 and SIN3 transcription regulator family member A (SIN3A) to the IFNβ gene promoter and subsequently represses its expression in a catalytic activity-independent manner (Xue et al., 2016). Additionally, TET2/3 proteins physically interact with OGT (O-linked β-D-N-acetylglucosamine (O-GlcNAc) transferase), an enzyme responsible for adding a GlcNAc sugar to Ser and Thr residues of numerous proteins. TET2/3-OGT co-localize in the chromatin at active promoters and mediate GlcNAcylation of the SET1/COMPASS complex, resulting in enhanced H3K4 trimethylation and subsequent transcriptional activation (Deplus et al., 2013). In mouse ESCs, depletion of TET2 abolishes recruitment of OGT to transcription start sites and reduces the amount of histone 2B Ser112 GlcNAc marks, which is positively associated with transcriptional activation (Chen et al., 2013). Nevertheless, it is often difficult to distinguish between non-catalytic and catalytic functions of TETs.

OBJECTIVES

Emerging evidences demonstrate that epigenetic processes such as histone post-translation modification and DNA methylation/demethylation on cytosine play important functions in modulating SMC plasticity and SM tissue homeostasis. Disrupted epi-genome has often been detected in smooth muscle dysfunction associated diseases characterized by a pathological phenotypical switching from a more differentiated and quiescent "contractile" state to a highly proliferative and migratory "synthetic" phenotype. Deciphering the role of distinct epigenetic modifications regulating SMC state and function during development and disease will allow new strategies for the treatment of smooth muscle dysfunction associated diseases such as allergic or autoimmune diseases. In this study, I wanted to elucidate the role of the TET enzymes, which play essential roles in regulation of DNA demethylation, in modulating SMC plasticity and lung pathogenesis *in vivo*. To achieve this aim, I will address following issues:

(1) Investigation of the function of TET2/3 in regulating SMC and lung tissue homeostasis by using various knockout mouse lines.
(2) Examination of the level and genome-wide distribution of 5hmC in lung SMCs.
(3) Identification of molecules and signaling pathways regulated by TET enzymes in SMCs and lung tissue using multiple deep-sequencing methods.
(4) Elucidation of underlying epigenetic mechanisms accounting for molecular or cellular events provoked by TET enzyme ablation.
(5) Investigation of the biological consequence of TET2/3 inactivation in adult SMCs and lung tissue.

MATERIAL

Drugs

Table 1

Drug	Company	Catalog no.
Tamoxifen ≥ 99%	SIGMA-ALDRICH	T5648-5G

Enzymes

Table 2

Enzyme	Company	Catalog no.
Proteinase K	Carl Roth	75282
SuperScript™ II Reverse	ThermoFisher SCIENTIFIC	18064014
Taq Polymerase	Homemade	
Trypsin EDTA	ThermoFisher SCIENTIFIC	R001100
Trypsin	SIGMA-ALDRICH	T0303

Inhibitors

Table 3

Inhibitor	Company	Catalog no.
Aprotinin	SIGMA-ALDRICH	10820
Leupeptin	SIGMA-ALDRICH	L8511-5MG
Sodium fluoride (NaF)	SIGMA-ALDRICH	S7920-100G
Sodium orthovanadate (Na_3Vo_4)	SIGMA-ALDRICH	450243
Pheylmethylsulfonyl fluoride (PMSF)	SIGMA-ALDRICH	P7626

Chemicals and Reagents

Table 4

Chemicals	Company	Catalog no.
Agarose	Carl Roth	2267
Ammonium persulfate (APS)	Serva	3375
Anti-Biotin Microbeads	Miltenyi Biotec	130-090-485
Anti-HA Affinity Gel	SIGMA-ALDRICH	SLBV2324
Benzonase® Nuclease	MERCK	E1014-25KU
Bis-Tris	AppliChem	A10250500
Bovine serum albumi (BSA)	Fisher Scientific	BP1605
Calcium Chloride x $2H_2O$	Carl Roth	52392
CD31 Microbeads mouse	Miltenyi Biotec	130-097-418
CD326 (EpCAM) Microbeads mouse	Miltenyi Biotec	130-105-958
CD45 MicroBeads mouse	Miltenyi Biotec	5190403681
Chelex® 100 Resin	BioRad	142-1253
Collagen type 2	Worthington Biochemical Corporation	4176
DAPI ((4', 6-diamidino-2-phenylindole)	ThermoFisher SCIENTIFIC	D1306
Deoxycholic acid	Fluka	30970
5,6-dichloro-1-β-D-ribofuranosyl-1H-benzimidazole (DRB)	Cayman	10010302
Dimethyl sulfoxide (DMSO)	SIGMA-ALDRICH	D-4540
Disodium hydrogen phosphate (Na_2HPO_4)	MERCK	65800500
Dispase	BD Biosciences	354235
Dnase I	Roche Diagnostics GmbH	11284932001
Dynabeads™ Protein A	ThermoFisher SCIENTIFIC	10002D

3,3'-DIAMINOBENZIDIN TETRAHYDROCHLORID (DAB)	SIGMA-ALDRICH	D5905
1,4-Dithiothreit (DTT)	Carl Roth	69082
EDTA	Carl Roth	80403
EGTA	Carl Roth	E3889
Elastase	Worthington Biochemical Corporation	9004-06-2
Entellan	MERCK	1079600500
Eosin Y solution, aqueous	WALDECK	2C-140
Ethanol (100%)	Carl Roth	90654
Ethidium bromide solution (1%)	AppliChem	A11520100
FSC 22 Frozen Section Compound	Leica	75806-668
Fetal calf serum (FCS)	SIGMA-ALDRICH	F2442
Gelatin from porcine skin	SIGMA-ALDRICH	G1890
L-Glutamine-pericillin-streptomycin solution	SIGMA-ALDRICH	G6784-100ML
Gardiquimod	InvivoGen	GDQ-39-03
GFP mRNA	OZBIOSCIENCES	MRNA11-20
Glutaraldehyde	SIGMA-ALDRICH	G5882
Glycerin (87%)	MERCK	1040942500
Glycine	SIGMA-ALDRICH	15527
Glycogen for mol.biol.	Roche	10901393001
Hematoxylin solution modified	MERCK	1.05174.0500
HEPES	Carl Roth	91053
Hydrochloric acid (37%)	Carl Roth	46251
InstantBlue Protein Stain	Expedeon	ISB1L
Isopropanol	Carl Roth	67524

Lipofectamine™ 2000 Transfection Reagent	ThermoFisher SCIENTIFIC	11668027
Lipofectamine™ 3000 Transfection Reagent	ThermoFisher SCIENTIFIC	L3000008
Lipofectamine™ MessengerMax™ Transfection Reagent	ThermoFisher SCIENTIFIC	LMRNA001
MACS buffer	Miltenyi Biotec	A9576
Magnesium chloride x $6H_2O$	Carl Roth	21891
Magnesium sulfate hetahydrate ($MgSO_4$ x $7H_2O$)	MERCK	5886
Methanol	Carl Roth	46275
Miglyol	Cacsar + Loretz	3274
Mowiol 4-88	MERCK	3186101
NP-40 (Tergitol-type NP-40)	Fluka	74385
Oligo 15 primer	Promega	C110A
Opti-MEM™ I Reduced Serum Medium	ThermoFisher SCIENTIFIC	31985062
Paraformaldehyde	Carl Roth	03354
Penicillin-Streptomycin. (10,000 U/mL)	ThermoFisher SCIENTIFIC	15140122
Phenylmethylsulfonyl fluoride (PMSF)	SIGMA-ALDRICH	P7626
Potassium chloride (KCl)	Carl Roth	67813
Potassium bicarbonate ($KHCO_3$)	MERCK	1048540500
Potassium dihydrogen phosphate (KH_2PO_4)	Carl Roth	P0181
Protease Inhibitor cocktail tablets	Roche	11836145001
Protein A-Agarose	Roche	11134515001
Protein G-Agarose	Roche	11243233001
Rnase, Dnase - free	Roche	11119915001

RNaseZAP™	SIGMA-ALDRICH	SLBQ7780V
Red Alert™ 10x	MERCK	71078-3
REDTaq® ReadyMix™ PCR Reaction Mix	SIGMA-ALDRICH	R2523-100RXN
Rnase inhibitor	Promega	N2515
Rotiphorese® Gel 40 (29:1)	Carl Roth	A5151
SDS	Carl Roth	23263
Skim milk	Fluka	70166
Sodium azide (NaN_3)	Carl Roth	K3051
Sodium Bicarbonate ($NaHCO_3$)	MERCK	1063290500
Sodium chloride (NaCl)	Carl Roth	39572
Sodium deoxycholate	MERCK	1065040100
Surcose	Carl Roth	90971
2 x Taq Master Mix	Vazyme	P111/P112
TEMED	Carl Roth	23671
Ter-119 MicroBeads mouse	Miltenyi Biotec	5190917111
Tissue - Tek® O.C.T Polyfreeze TM freezing	Leica	3801480
Tris	Carl Roth	54292
Triton-X 100	Carl Roth	66831
TRIZOL® reagent	Invitrogen	15596018
TurboFect transfection reagent	ThermoFisher SCIENTIFIC	R0531
Tween-20	MERCK	8221840500
Xylol	Carl Roth	97133

Culture media

Table 5

Medium	Composition	Company
Growth medium for H293T, H293, Hela cells	DMEM (Dulbecco's Modified Eagle Medium) with 4.5 g / L glucose, 10% fetal calf serum, 100 U / mL penicillin / streptomycin, 20mM glutamine	SIGMA-ALDRICH
Growth medium for primary SMCs	SmBM™ Smooth Muscle Cell Growth Basal Medium (#CC-3181) with SmGM™- 2 Smooth Muscle Cell Growth Medium-2 SingleQuots™ Supplements and Growth Factors (#CC-4149)	LONZA

Buffer solutions

Table 6

Buffer	Composition
10% APS	Dilute 10 g APS in 100 mL Aqua dest
5% BSA	Dilute 5 g BSA in 100 mL 1 x PBS-T
2 N HCl for 5hmC, 5mC immunostaining	Dilute 200 mL Hydrochloric acid (37%) into 800 mL volume of Aqua dest
Sodium borate buffer pH 8.0 for 5hmC, 5mC immunostaining	Dilute 61.83 g boric acid and 10 g NaOH with 1 L aqua dest
Electron Microscopy fixation buffer	1.5% glutaraldehyde; 1.5% paraformaldehyde in 0.15 mol L-1 HEPES
Immunofluorescence blocking buffer (general)	3% BSA in 1x PBS
Immunofluorescence blocking buffer (5mC, 5hmCl)	2% Horse serum, 3% BSA in 1x PBS
5% Immunoblot blocking solution	Dilute 5 g skim milk in 100 mL 1 x PBS-T
10 x PBS, pH 7.4	Dilute 80 g NaCl; 2 g KCL; 14.4 g Na_2HPO_4; 2.4 g KH_2PO_4 in 800 mL Aqua dest.; adjust pH to 7.4; refill to 1000 mL with Aqua dest.

1 x PBS, pH 7.4	100 mL 10 x PBS + 900 mL Aqua dest
1 x PBS-T	1 x PBS + 0.1% Tween 20
SDS running buffer (10x)	250 mM Tris base, 1 % SDS, 1.9 M Glycine
Red blood cell lysis buffer	1.5 M NH_4Cl, 100 mM $KHCO_3$, 10 mM EDTA, add 5 µg/ml Dnase I before use
ChIP dilution buffer	20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, complete protease inhibitor cocktail
High salt washing buffer for ChIP - qPCR	20 mM Tris-HCl, pH 8.0, 500 mM NaCl, 2 mM EDTA, 1% Triton X-100
LiCl washing buffer for ChIP - qPCR & ChIP - seq	10 mM Tris-HCl, pH 8.0, 250 mM LiCl, 1% Na-deoxycholate, 1% NP40
TE washing buffer for ChIP - qPCR & ChIP - seq	10 mM Tris-HCl pH 8.0, 1 mM EDTA
RIPA buffer for ChIP - seq	0.1% SDS, 0.1% Sodium deoxycholate, 1% Triton X-100, 1mM EDTA, 10mM Tris-HCl pH 8.1, 150mM NaCl
RIPA 500 buffer for ChIP - seq	0.1% SDS, 0.1% Sodium deoxycholate, 1% Triton X-100, 1mM EDTA, 10 mM Tris-HCl pH 8.1, 500 mM NaCl
Elution buffer for ChIP - seq	10 mM Tris-HCl pH 7.4, 5 mM EDTA, 300 mM NaCl, 0.5% SDS
Lysis buffer (200mM)	20mM Tris-HCL, pH 8.0, 200 mM NaCl, 1mM EDTA, 1mM EGTA, 1% Triton X-100
10% SDS	Dilute 10 g SDS in 100 mL Aqua dest
50 x TAE	Dilute 242 g Tris Base; 57.1 mL acetic acid; 100 mL 0.5 mol EDTA, pH 8.0 in 1000 mL Aqua dest.
20 x Transfer buffer	Dilute 163.2 g Bicine; 209.6 g Bis-Tris; 12 g EDTA in 1000 mL Aqua dest

1 x Transfer buffer	250 mL 20 x transfer buffer; 1000 mL methanol; 3750 mL Aqua dest
1 M Tris, pH 6.8	Dilute 12.1 g Tris in 80 mL Aqua dest; adjust pH to 6.8; refill to 100 mL with Aqua dest
1.5 M Tris, pH 8.8	Dilute 18.17 g Tris in 80 mL Aqua dest; adjust pH to 8.8; refill to 100 mL with Aqua dest

Antibodies

Table 7

Antibody	Use	Use Supplier	Cat. No.
5hmC	IF, DB, hMeDIP	Active Motif	39769
5mC	DB, MeDIP	Eurogentec	81103
5fC	DB	Active Motif	61225
5caC	DB	Active Motif	61223
Tet3	IF, WB, ChIP, IP	GeneTex	GTX121453
Tet2	WB	BETHYL	A303-604A
Tet1	WB	Active Motif	61443
α-SMA-FITC	IF	Sigma	F3777
α-SMA-Cy3	IF	Sigma	C6198
Anti-Clara cell secretory protein antibody	IF, DAB	Millipore	07-623
α-Tubulin	IF	Sigma	T6074
Mucin-5AC	IF	Abcam	ab3649
CD3	FACS	BD	560590
CD45R	FACS	Invitrogen	4329548
Siglec-F	FACS	Miltenyi Biotec	5161104265
CD11c	FACS	eBioscience	4300054

CD11b	FACS	eBioscience	4289817
F4/80	FACS	Bio-Rad	MCA497FT
Ly6G	FACS	BD	563005
MHCII	FACS	eBioscience	4271684
CD4	FACS	BD	561830
CD3	FACS	BD	561798
CD196	FACS	Biolegend	129803
CD194	FACS	Biolegend	131213
CD183	FACS	Biolegend	126511
Ki67	IF		
Gradient	IF	Abcam	ab76473
CD68	IF	eBioscience	E24592
CD4	IF	life technologies	MA-146
Collagen I	IF		
α-SMA	WB	Sigma	A2547
Myh11	WB	Abcam	ab5319
Tagln	WB	ProteinTec	10493-1-AP
Vimentin	WB	Abcam	ab92547
Tpm4	WB	Millipore	AB5449
Pan actin	WB	Cell signaling	4968
RNA Pol II Pan	WB	Active Motif	39097
RNA Pol II ser2	WB	Abcam	ab5095
RNA Pol II ser5	ChIPseq	Abcam	ab5408
Histone H3	WB, ChIP	Abcam	ab1791

H3K36me3	WB, ChIP	Abcam	ab9050
TLR7	WB, PLA	proteintech	17232-1-AP
Myd88	WB, PLA	Santa Cruz	sc-74532
TLR8	WB	gift by Dr. Yu	

hMeDIP: hydroxymethylcytosine immunoprecipitation qPCR; MeDIP: methylcytosine immunoprecipitation qPCR; IF: Immunofluorescence; WB: Western blot; ChIP: Chromatin immunoprecipitation qPCR; FACS: Fluorescence-activated cell sorting; PLA: proximity ligation assay.

Primers

Table 8

Name	Primer sequence (5'-->3')	Application
Tet3 loxp3 F	CGCCCAGAGCACTTGGCTGC	Tet3 flox Genotyping
Tet3 loxp3 R	GCCTTAGTTTCAGCTGAAGCTC	
Cre F	GACCAGGTTCGTTCACTCATGG	Cre genotyping
Cre R	AGGCTAAGTGCCTTCTCTACAC	
tdTomato	CTGTTCCTGTACGGCATGG	tdTomato genotyping
tdTomato	GGCATTAAAGCAGCGTATCC	
Acta2 Full length F	CCCTTCAGCGTTCAGCCTCCCG	Acta2 full length mRNA
Acta2 Full length R	GGAAGGTAGACAGCGAAGCCAGGATG	
Cnn1 Full length F	GTGTGAGGAGGGAAGAGGGTGCAGACG	Cnn1 full length mRNA
Cnn1 Full length R	CGTCGCAGGAATGGGGCAGGGG	
Vim full length F	TGCCACTCTTGCTCCGGGACCC	Vim full length mRNA

Vim full length R	CTTTTCTACGTCTTTTGGGGTGTCAGTTGTTA	
Tpm4 full length F	CGCCGGAGCCCAGCAGAACG	Tpm4 full length mRNA
Tpm4 full length R	GCAGCCTACAGAGCACTCCACATCCAG	
Tet3 deletion F	CCCCTATTTCTCTTGCCTCTCCCTACC	
Tet3 deletion R	AAGTGGATGGATGAGATGTTGAGGTGAG	
Tet3 exon 10 F	AACGCCAAGGTCAGTAGTGG	qPCR
Tet3 exon 10 R	GCGTGCTCAGTTTCTCCTTC	
IFNg F	ATGAACGCTACACACTGCATC	qPCR
IFNg R	CCATCCTTTTGCCAGTTCCTC	
IL4 F	GGTCTCAACCCCCAGCTAGT	qPCR
IL4 R	GCCGATGATCTCTCTCAAGTGAT	
IL17a F	TTTAACTCCCTTGGCGCAAAA	qPCR
IL17a R	CTTTCCCTCCGCATTGACAC	
IL6 F	TAGTCCTTCCTACCCCAATTTCC	qPCR
Il6 R	TTGGTCCTTAGCCACTCCTTC	
IL13 F	CCTGGCTCTTGCTTGCCTT	qPCR
IL13 R	GGTCTTGTGTGATGTTGCTCA	
TLR7 qpcr F	ATGTGGACACGGAAGAGACAA	qpCR
TLR7 qpcr R	GGTAAGGGTAAGATTGGTGGTG	
TLR8 qpcr F	GAAAACATGCCCCCTCAGTCA	qPCR
TLR8 qpcr R	CGTCACAAGGATAGCTTCTGGAA	
CCL5 qpcr F	GCTGCTTTGCCTACCTCTCC	qpCR
CCL5 qpcr R	TCGAGTGACAAACACGACTGC	

CCR5 qpcr F	TTTTCAAGGGTCAGTTCCGAC	qPCR
CCR5 qpcr R	GGAAGACCATCATGTTACCCAC	
IRF7 qpcr F	GAGACTGGCTATTGGGGGAG	qPCR
IRF7 qpcr R	GACCGAAATGCTTCCAGGG	
MYD88 mouse F	TCATGTTCTCCATACCCTTGGT	qPCR
MYD88 mouse R	AAACTGCGAGTGGGGTCAG	
CXCL9 human F	CCAGTAGTGAGAAAGGGTCGC	qPCR
CXCL9 human R	AGGGCTTGGGGCAAATTGTT	
CCL5 human F	CCAGCAGTCGTCTTTGTCAC	qPCR
CCL5 human R	CTCTGGGTTGGCACACACTT	
CD68 human F	GGAAATGCCACGGTTCATCCA	qPCR
CD68 human R	TGGGGTTCAGTACAGAGATGC	
IRF5 human F	GGGCTTCAATGGGTCAACG	qPCR
IRF5 human R	GCCTTCGGTGTATTTCCCTG	
IL1B human F	ATGATGGCTTATTACAGTGGCAA	qPCR
IL1B human R	GTCGGAGATTCGTAGCTGGA	
TLR7 human F	TCGTGGACTGCACAGACAAG	qPCR
TLR7 human R	GGTATGTGGTTAATGGTGAGGGT	
MYD88 human F	GGCTGCTCTCAACATGCGA	qPCR
Hsa-miR-145	TaqMan probes #000467 (ThermoFisher)	qPCR
Hsa-miR-145 RT	TaqMan probes #000467 (ThermoFisher)	
Acta2 promoter F	CAAGTCCTCAGCTAATGGCC	hMeDIP-qPCR ChIP-qPCR
Acta2 promoter R	GGGGATAAACATCCTAAGCC	

Acta2 exon 2 F	CGGGCTTCTCGCTCGCTGTCTCTGC	hMeDIP-qPCR ChIP-qPCR
Acta2 exon 2 R	CCGCACGATGGATGGGAAAACAGCC	
Acta2 exon 3 to exon 4 F	GTGGGCATGGGCTTTTGGGGGAAC	hMeDIP-qPCR ChIP-qPCR
Acta2 exon 3 to exon 4 R	ACGCTTGCCTGAAAACCTACTTGCCATCTT	
Acta2 exon 6 to exon 7 F	GCCACACCTGAAATGCACACACACACACA	hMeDIP-qPCR ChIP-qPCR
Acta2 exon 6 to exon 7 R	GGCGGCATTGGGGTCAGGCTTTACT	
Myh11 promoter F	TAGGGTCTTAGCACGCATCC	hMeDIP-qPCR ChIP-qPCR
Myh11 promoter R	GTCACCGCATATCCTCCAGT	
Tagln promoter F	AGCACCTGACTACCCACCAC	hMeDIP-qPCR ChIP-qPCR
Tagln promoter R	TTTGGGCCTAACACATAGCC	
Vim promoter F	GAGAAAGAAAGAGAGAAAGGGGGAGGGA	hMeDIP-qPCR ChIP-qPCR
Vim promoter R	TGTGATTGTTGGGGAAGTGTGGAGAGTT	
Tpm4 promoter F	TGCCTCAGCCTCCCAAACTCCG	hMeDIP-qPCR ChIP-qPCR
Tpm4 promoter R	AGTGCCTGTCCAGTTCGTCCTCCA	
Myh10 promoter F	TGAATGAGTGAGTGAACAAATGAAGTCGCT	hMeDIP-qPCR ChIP-qPCR
Myh10 promoter R	GTCCCCCCTTACTGATGCTTCCTTCC	
Vim genebody F	ATGCGTGAGATGGAAGAGAATTTTGCC	hMeDIP-qPCR ChIP-qPCR
Vim genebody R	TTCCTGTAGGTGGCGATCTCAATGTCC	

Tpm4 genebody F	CAAAGAAGCCAAGCACATCACTGACG	hMeDIP-qPCR ChIP-qPCR
Tpm4 genebody R	AGTCAAGTGGCTGTGTCTGGAAACCC	
Myh10 exon 3 F	TGTTTTTGGCTTATGGTTGTATCTGAGTGA	hMeDIP-qPCR ChIP-qPCR
Myh10 exon 3 R	ATCTCATGGCGTTTCTTCCCTCTATACA	
Myh10 exon 30 F	GCCAAAAAGAAACTGCTCAAGGATGTG	hMeDIP-qPCR ChIP-qPCR
Myh10 exon30 R	ACGATCTGGCGCTGGTGGTCC	
Myh11 exon 8 F	AACCCATCACAGGATACACTCA	hMeDIP-qPCR ChIP-qPCR
Myh11 exon 8 R	GCCCCAGAACCCTTAACTTC	
Myh11 exon22 F	AGCCATCTTCTTCCTCTCAGCCTGTAGTT	hMeDIP-qPCR ChIP-qPCR
Myh11 exon22 R	GCTGTATGCTGAGGCTGAGGAGATGC	
Tagln genebody F	TTACTCACCACACCATTCTTCAGCCACA	hMeDIP-qPCR ChIP-qPCR
Tagln genebody R	CATGAGCCGAGAAGTGCAGTCCAAA	
Rbp1 promoter F	ACCAACTTCGCATCATTTCC	hMeDIP-qPCR ChIP-qPCR
Rbp1 promoter R	CAGGTTTGCTAGCGTCATCA	
Rbp1 exon3 F	CAGTGTGTGCAGAAGGGAGA	hMeDIP-qPCR ChIP-qPCR
Rbp1 exon3 R	CCCAGTGCTCTTTTCCATTG	

Plasmid

Table 9

Plasmid	Catalog no.	Company
FH-TET3-pEF	#49446	Addgene (Anjana Rao lab)

Kits

Table 10

Kit	Company	Catalog no.
AllPrep® DAN/RNA Micro Kit	QIAGEN	80284
Diff-Quick stain kit	SIGMA-ALDRICH	
Dneay Blood & Tissue Kit	QIAGEN	69504
Duolink® in Situ Red Starter Kit Mouse/Rabbit	SIGMA-ALDRICH	Duo92101-1KT
Duolink® in Situ Detection Reagents Green	SIGMA-ALDRICH	Duo92014-30RXN
hMeDIP kit x 16 (polyclonal rabbit antibody)	Diagenode	C02010032
MeDIP kit x 10	Active Motif	55009
miRNeasy Mini Kit	QIAGEN	217004
MinElute® PCR Purification Kit	QIAGEN	28004
Periodic Acid-Schiff (PAS) Kit	SIGMA-ALDRICH	395B-1KT
Rneasy® Micro Kit	QIAGEN	74004
truChIP Chromatin Shearing Kit with Formaldehyde	COVARIS	RK000958

Equipment

Table 11

Equipment	Company
Agarose gel electrophoresis chamber	Peqlab
Bio Doc analyzer	Bio RAD
ChemiDoc™ MP Imaging System	Bio RAD

Cold Plate for Tissue Embedding System	Leica
Eppendorftubes, 0.5 mL, 1.0 mL	Eppendorf
Filter tips 10 µL, 20 µL, 100 µL, 1000 µL	Fisher Scientific
Gelsystems (Mini, Maxi)	VWR
Greiner centrifuge tubes, 15 mL	SIGMA-ALDRICH
Greiner centrifuge tubes, 50 mL	SIGMA-ALDRICH
Keyence Fluorescence Microscope	Keyence
Nitrocellulose membranes	Invitrogen
StepOnePlus™ Real-Time PCR System	ThermoFisher SCIENTIFIC
SuperFrost Plus slides	Menzel-Glaeser
Tips 10 µL, 20 µL, 100 µL, 1000 µL	Greiner Bio-One
Transmission electron microscope	Zeiss
UV Transilluminator	INTAS
Xcell SureLock™ Mini-Cell and Xcell 11™ Blot Module	ThermoFisher SCIENTIFIC

Software

Table 12

Software	**Company**
Graph Pad Prism 6.0	Graph Pad Software
Image J / Fiji	NIH
Image Lab	Bio RAD
Adobe Phtoshop CS4	Adobe
Adobe Illustrator CS4	Adobe
Microsoft Office 2018	Microsoft Corporation®

METHODS

Animals. $Tet3^{floxLacZ/+}$ and $Tet3^{fl/fl}$ mice were generated in house by using targeting vector purchased from the European Conditional Mouse Mutagenesis Program (EUCOMM), in which exon 10 of the *Tet3* gene is flanked with two *LoxP* sequences. α-$SMA^{ERT2Cre}$ transgenic mice were provided by Professor Chambon P. from France. $ROSA26^{tdTomato}$ mice were obtained from The Jackson Laboratory. C57bl/6 mice were obtained from Charles River. All mouse strains were backcrossed and maintained on a C57bl/6 genetic background. Primers used for genotyping are listed in Table 8. All animal experiments were done in accordance with the Guide for the Care and Use of Laboratory Animals published by the US National Institute of Health (NIH Publication No. 85-23, revised 1996) and according to the regulations issued by the Committee for Animal Rights Protection of the State of Hessen (Regierungspraesidium Darmstadt, Wilhelminenstr. 1-3, 64283 Darmstadt, Germany) with the project number B2/1137, B2/1125.

Genotyping

To perform mouse genotyping, tail biopsies were first digested in 200 µl tail lysis buffer (NID buffer) supplemented with 3 µl Proteinase K (20 mg mL^{-1}). After incubation overnight at 56 °C with 1000 rpm shaking, digested tail DNA samples were incubated at 95 °C for 10 min to inactivate Proteinase K. Genotyping PCRs were performed with self-designed primers as indicated in Table 8 using appropriate cycling conditions.

The master mix for one PCR reaction was prepared as followed:

Reaction mixture with homemade Taq polymerase:

10x PCR buffer	2.5
$MgCl_2$ (50 mmol L^{-1})	1.0
dNTPs (10 mmol L^{-1})	0.5
Primer 1 (forward)	0.5
Primer 2 (reverse)	0.5
Taq polymerase (5U L-1)	0.2
ddH_2O	18.5
Genomic DNA	1.0

Reaction mixture with commercial Taq polymerase:

Taq mix (Red taq 2 x Taq Master Mix)	5
ddH_2O	3.5
Primer 1 (forward)	0.5
Primer 2 (reverse)	0.5

Genomic DNA	1

PCR products were loaded on 2 % agarose gel, and ran for 30 min at 150-200 V. Double-stranded DNA was visualized under UV light and gels imaging was done by using Bio Doc analyzer.

Tamoxifen injections

Tamoxifen (Sigma) was dissolved in Miglyol® 812 (Caesar & Loretz GmbH) at concentration of 20mg/1mL after incubation at 56 °C for 10 min and with gentle rotation at room temperature for 1 hour, and then was stored at -20 °C. 8–10 weeks old mice were injected intraperitoneally with Tamoxifen at a dose of 3 mg per 40 g body weight daily for continuous 10 days to induce DNA recombination of floxed alleles.

Lung smooth muscle cells (SMCs) isolation

Mice were sacrificed by overexposure to carbon dioxide (CO_2) and blood was removed by perfusion with cold PBS through the RV prior to lung dissection. Lung tissues were dissected and dissociated into single cells by enzyme digestion. Briefly, individual lung tissue was minced to small pieces and incubated with 3 ml digestion buffer (DPBS containing Collagenase type 2 (2mg/ml, Worthington), Elastase (0.04mg/ml, Worthington) and DNase (5U/ml, Roche) with frequent agitation at 37 °C for 10min. Next, 10 times volume of cold DMEM supplemented with 10% FBS was added to single-cell suspensions immediately. Cells were mechanically dissociated by pressing through 30 ml syringe plunger 4-5 times, and consecutively filtered through 100-, 70- and 40-μm cell strainers (BD). The filtrate was centrifuged at 300 g, room temperature (RT) for 10min. Pellets were re-suspended in 1ml pre-cooled MACS buffer (Cat# A9576, Miltenyi Biotec) with 1% BSA. After 5min centrifugation at 300g, 4 °C, cell pellets were re-suspended in 90 μL MACS buffer and directly incubated with 10 μL CD45 MicroBeads (Cat#130-052-301) and anti-Ter-119 MicroBeads (Cat#130-049-901) at 4 °C for 15min to deplete mouse leukocytes and hematopoietic cells. After washing with MACS buffer, cells were loaded into pre-conditioned LS columns on a MACS separator and the flow-through containing unlabeled cells were collected. Cell sorting was done with FACSAria™ III (BD Biosciences). Debris and clumps were excluded using forward and side scatter parameters. $DAPI^+$ dead cells were excluded and sort gate for tdTomato positive population was set based on tdTomato negative control (Figure 4.4 E).

FACS analysis

Single cell suspensions from the lung were stained with antibodies as follows: Inflammatory cell panel: CD3 (APC-Cy7), B220 (PE-Cy5), Siglec-F (APC), CD11c (eFluo450), CD11b (PE-Cy7),

F4/80 (FITC), Ly6G (BV605), MHCII (PE). T helper (Th) cell panel: CD4 (APC-Cy7), CD3 (FITC), CD196 (PE), CD194 (PE-Cy7), CD183 (APC). T cells were defined as $CD3^+$; B cells were defined as $B220^+$; eosinophils were defined as Siglec-F^+CD11c^-); alveolar macrophages (AMs) were defined as Siglec-F^+ $CD11c^+$ $CD11b^-$ $F4/80^+$; interstitial macrophages (IMs) were defined as Siglec-F^- $CD11c^-$ $CD11b^+$ $F4/80^+$; neutrophils were defined as Siglec-F^- $CD11c^-$ $CD11b^+$ $Ly6G^+$; dendritic cells were defined as Siglec-F^- $CD11c^{hi}$ $MHCII^{hi}$ (Han and Ziegler, 2013). Th1 cells were defined as $CD4^+$ $CD183^+$; Th2 cells were defined as $CD4^+$ $CD194^+$ $CD196^-$; Th9 cells were defined as $CD4^+$ $CD194^-$ $CD196^+$; Th17 cells were defined as $CD4^+$ $CD194^+$ $CD196^+$ (https://www.miltenyibiotec.com/DE-en/resources/macs-handbook/mouse-cells-and-organs/mouse-cell-types/cd4-t-cells-mouse.html#gref). Single color controls and fluorescence-minus-one (FMO) controls were used to accurately identify sub-populations within the multicolor-stained samples. Flow cytometry was performed using LSR Fortessa (BD). Data acquisition and analysis were done using BD FACS Diva v6 software.

Bronchoalveolar Lavage fluid (BALF) analysis

Mice were scarified by overexposure to carbon dioxide (CO_2). The trachea was surgically exposed and intubated with a syringe catheter. The lungs underwent lavage with 1 ml pre-warmed PBS. Cells in BALF were pelleted by centrifugation at 300 g for 5 minutes at 4°C, and then re-suspended in 500 µl PBS and spun onto a microscope slide at 700 rpm for 5min with a cytocentrifuge (ThermoScientific/Shandon). Cells on the slides were dried out and stained using Diff-Quick stain kit (Sigma).

Gene expression analysis

Total RNA was extracted using TRIzol reagent (Invitrogen) according to manufacturer instruction. RNA was reverse-transcribed with Superscript II (Invitrogen) following standard procedures. Real-time PCR was performed with 2 technical replicates using StepOne™ Real time PCR system and KAPA SYBR ® FAST qPCR Master Mix (KAPA Biosystems, CH). Relative quantitation of gene expression was performed using the ΔΔCT method. The Ct values of the target genes were normalized to the β-actin gene using the equation $\Delta Ct = Ct_{reference} - Ct_{target}$ and expressed as ΔCt. Relative mRNA expressions are shown with the average from control samples set as 1. Primers and PCR conditions are listed in Table 8.

Immunohistochemistry, Immunofluorescence and histological analysis.

After perfusion with PBS, tissues were dissected and immediately fixed in 4% PFA. For paraffin sections, samples were dehydrated following standard protocols, embedded into paraffin, sectioned at 7 μm for immunofluorescence staining, hematoxylin/eosin (H&E) staining or trichrome staining using established techniques. For cryosections, fixed tissues were equilibrated in 30% sucrose/PBS at 4°C overnight and frozen on dry ice. 7 μm sections were mounted on superfroast slides for immunofluorescence staining or PAS staining with Periodic Acid-Schiff (PAS) Kit (Sigma) following standard protocol. Macroscopic images were acquired with a fluorescent stereo microscope (Leica M205 FA). Immunofluorescence images were acquired with a fluorescent microscope (ZEISS Imager Z1). Histological image acquisition was performed with a light microscope (ZEISS Axioplan2). Antibodies for immunofluorescence staining are listed in Table 7.

Western blot assays

Sorted cells were washed with ice cold PBS and lysed in cell lysis buffer (20 mM Tris (pH 7.5), 400 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1% Triton X-100, 1x Complete Protease Inhibitor Cocktail (Roche Diagnostics) for 10 min on ice, followed with the sonication. The concentration of the isolated proteins was determined using BCA Protein Assay Reagent (Pierce, Rockford, IL). Samples were mixed with 4 x SDS sample buffer and incubated at 95°C of 10 min. 10 μg whole cell lysate of each sample was loaded on a 10-15 % SDS gel, depending on the size of the interested protein. The gel was blotted onto nitrocellulose membrane (Millipore, Billerica, MA) and blocked with 5% milk/PBS-T 1hour at room temperature. The primary antibody was applied over night at 4°C. Blots were washed twice for 10 min with PBS-T and blots were incubated with the secondary antibody for 1 h at room temperature. Blots were washed three times for 10 min with PBS-T. Protein expression was visualized using an enhanced chemiluminescence detection system (GE Healthcare, Little Chalfont, United Kingdom) and quantified using a ChemiDoc gel documentation system (Bio-Rad). Antibodies are listed in Table 7.

Electronic microscope images

Lungs were isolated and fixed in 1.5% glutaraldehyde (v/v), 1.5% paraformaldehyde (v/w) in 0.15 M HEPES (v/w), pH 8.0 at 4 °C for at least 24 h, and subsequently incubated with 1 % osmium tetroxide for 2 hours. Samples were then en bloc stained with 50 % saturated watery uranyl acetate, followed by sequential ethanol dehydration (30%, 50%, 75%, 95%), and embedded in Agar 100. Ultrathin sections were cut using an ultramicrotome and image acquisition was performed with a Philips CM10 electron microscope. All images were captured with a slow-scan 2K CCD camera.

Laser capture microdissection and Pico-array.

7-10 μm cyrosections mounted on glass microscope slides were immersed into the 70% ethanol fixative solution (10s) - dH_2O (10s) - Mayer's hematoxylin (45s) - ddH_2O (10s) – Tap water 10s - 70% ethanol (10s) – 95% ethanol (10s) - 95 % ethanol (10s) - 100% ethanol (60 s) - 100% ethanol (60 s) successively. Air dry slides as quickly as possible. Bronchi/bronchioles or vasculature were micro-dissected using the Laser Microbeam System (P.A.L.M., Bernried, Germany), and collected into a tube containing 200 μl of RNA lysis buffer for RNA extraction using Rneasy® Micro Kit (QIAGEN). 0.8 ng RNA was applied for microarray assay using GeneChip® WT Pico Reagent Kit (P/N 703262 Rev.1), GeneChip® WT Pico Kit, P/N: 902622; GeneChip® Hybridization, Wash, and Stain Kit P/N 900720; and Mouse transcriptome array 1.0 ST (ClariomD, Ref: 520851) according to Affymetrix protocol User Guide.

RNA-seq

RNA was isolated from sorted smooth muscle cells using the miRNeasy micro Kit (Qiagen) combined with on-column DNase digestion (DNase-Free DNase Set, Qiagen) to avoid contamination by genomic DNA. RNA and library preparation integrity were verified with BioAnalyzer 2100 (Agilent) or LabChip Gx Touch 24 (Perkin Elmer). 50 ng of total RNA was used as input for ribosomal depletion with RiboGone-Mammalian (Clontech) followed by library preparation using SMARTer Stranded Total RNA Sample Prep Kit (Clontech). Sequencing was performed on the NextSeq500 instrument (Illumina) using v2 chemistry, resulting in average of 44M reads per library with 1x75bp single end setup. The resulting raw reads were assessed for quality, adapter content and duplication rates with FastQC (Andrews S. 2010, FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc. Trimmomatic version 0.36 was employed to trim reads after a quality drop below a mean of Q20 in a window of 5 nucleotides(Bolger et al., 2014). Only reads between 30 and 150 nucleotides were cleared for further analyses. Trimmed and filtered reads were aligned versus the Ensembl mouse genome version mm10 (GRCm38) using STAR 2.6.0c with the parameter "--outFilterMismatchNoverLmax 0.1" to increase the maximum ratio of mismatches to mapped length to 10% (Dobin et al., 2013). The number of reads aligning to genes was counted with featureCounts tool from the Subread package 1.6.0 (Liao et al., 2014). Only reads mapping at least partially inside exons were admitted and aggregated per gene. Reads overlapping multiple genes or aligning to multiple regions were excluded. The Ensemble annotation was enriched with UniProt data (release 06.06.2014) based on Ensembl gene identifiers (Activities at the Universal Protein Resource (UniProt)).

hMeDIP sequencing.

1.5 μg genomic DNA was sheared to 100-600bp using the EpiSonic 2000 Sonication System (EpiGentek, Cat. #EQC-2000). 1 μg of sheared DNA, 4 μg of 5hmC monoclonal antibody (EpiGentek Cat. #A-1018) were incubated with 8 μl of protein A/G beads and captured DNA eluted in 15 μl H_2O and input DNA were used for DNA end polishing and adaptor ligation, library amplification using indexed primers and library purification. Purified library DNA was eluted in 12 μl of water. Library was verified on a Bioanalyer and KAPA qPCR Library Quantification for quality control. 10 nM of samples libraries were subjected to next generation sequencing on an Illuminar HiSeq 4000. Reads of two replicates were processed utilizing a set of public available bioinformatics tools: Quality assessment was performed via FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc) and reads were trimmed using Reaper (Davis et al., 2013). Mapping was performed by STAR (Dobin et al., 2013) against the mouse genome (version mm10). Deduplication of reads was done by Picard (http://broadinstitute.github.io/picard/; Broad Institute). In order to find broader regions enriched for reads, the peak caller MUSIC (Harmanci et al., 2014) was employed. The annotation of peaks was performed via UROPA (https://pubmed.ncbi.nlm.nih.gov/28572580/) against GENCODE annotation (V.16), annotation was generated in context of gene bodies. Recounting on different regions (unified peaks (all replicates), promoters, and genes) was calculated using bigWigAverageOverBed (UCSC Tools). To compare the peaks and regions from different samples and conditions, read counts for were submitted to DESeq2 for normalization (Anders and Huber, 2010). The RNA-seq counts (as generated above) were joined with the hMeDIP-seq counts based on the Ensembl ID. Profiles for the TSS or gene bodies, were created using deepTools. Genomic regions were defined as indicated on the plots (+- in bp), gene bodies include introns and exons.

hMeDIP – qPCR /MeDIP - qPCR

1 μg genomic DNA was extracted from control and mutant SMCs by using AllPrep® DAN/RNA Micro Kit (Qiagen). hMeDIP and MeDIP were done following manufacturer's instructions for hMeDIP kit (Diagenode) and MeDIP kit (Active Motif). Normal IgG antibodies were used as control. Input and precipitated DNA from hMeDIP or MeDIP were used as templates for quantitative real-time PCR. Relative enrichment of 5hmC or 5mC was using the ΔΔCT method. The Ct values of the target genes were normalized to the input using the equation $\Delta Ct = Ct_{input} - Ct_{hMeDIP}/ Ct_{MeDIP}$ and expressed as ΔCt. Primers used for qPCR are listed in Table 8.

RNA Pol II Ser5 ChIP sequencing and data analysis

Purified ChIP DNA samples were quantified by Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific). 0.5-10ng of DNA was used for TruSeq ChIP Library Preparation Kit (Illumina) with

modifications. Briefly, libraries were size selected by SPRI-bead based approach after final PCR with 18 cycles: samples were firstly cleaned up by 1x bead: DNA ratio to eliminate residuals from PCR reaction, followed by 2-sided-bead cleanup step with initially 0.6x bead: DNA ratio to exclude larger fragments. Supernatant was transferred to new tube and incubated with additional beads in 0.2x bead: DNA ratio for eliminating smaller fragments, like adapter and primer dimers. Bound DNA samples were washed with 80% ethanol, dried and re-suspended in TE buffer. Library integrity was verified with LabChip Gx Touch 24 (Perkin Elmer). Sequencing was performed on the NextSeq500 instrument (Illumina) using v2 chemistry with 1x75bp single end setup. Quality assessment and trimming was performed according to RNAseq methods. Reads were further deduplicated using Picard 2.17.10 (Picard: A set of tools (in Java) for working with next generation sequencing data in the BAM format) to mitigate PCR artefacts leading to multiple copies of the same original fragment. The Macs2 peak caller version 2.1.0 was employed to accommodate for the range of peak widths as typically expected (Zhang et al., 2008). Minimum q value was set to -1.5 and FDR was changed to 0.001. Peaks overlapping ENCODE blacklisted regions (known misassemblies, satellite repeats) were excluded. In order to determine thresholds for significant peaks per IP, the data was manually inspected in IGV 2.3.52 (Robinson et al., 2011). To be able to compare peaks in different samples to assess reproducibility, the resulting lists of significant peaks were overlapped and unified to represent identical regions and recounted. A background-correction was performed to correct read counts on various regions (unified peaks, promoters, genes). Treatment and Input samples were normalized for sequencing depth, before subtracting reads of the Input sample from reads of the respective Treatment sample in windows of 50nt length(Ramirez et al., 2014). All windows with negative values (Input > Treatment) were set to zero(Zerbino et al., 2014). Background-corrected counts for regions were calculated using bigWigAverageOverBed (UCSC Tools) and normalized with DESeq2(Anders and Huber, 2010). Peaks were annotated with the promoter (TSS +- 5000 nt) of the gene most closely located to the centre of the peak based on reference data from GENCODE v25. Spearman correlations were produced to assess the degree of reproducibility between samples. To permit comparative display of samples in IGV, raw BAM files were scaled with DESeq2 size factors based on all unified peaks using bedtools genomecov resulting in normalized BigWig files(Quinlan and Hall, 2010). Finally, DESeq2 was used to identify significantly differentially modified peaks based on background-corrected read counts from recounted unified peak regions. Reads of three replicates were processed by an in-house pipeline, utilizing a set of public available bioinformatics tools: Quality assessment was performed via FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc) and reads where trimmed using Reaper (Davis et al., 2013) .

ChIP – qPCR

ChIP were performed according to published protocols(Dahl and Collas, 2008). Briefly, FACS-purified SMCs (300,000) were first cross-linked with 1% formaldehyde for 10 min and then quenched by quenching buffer for 10 min at RT. Chromatin was prepared with truChIP Chromatin Shearing Kit (COVARIS), and sheared to an average size of 200–500 bp by sonication (Diagnode Biorupter). Protein-DNA complexes were immunoprecipitated with IgG or antibodies shown in Table 7, followed by incubation with Protein A/G magnetic beads (Dynabeads, Invitrogen). For ChIP-qPCR, beads were washed and protein-DNA complexes were eluted and purified using 10% chelex-100 (w:v, Bio-Rad Laboratories) in Tris-EDTA (Nelson et al., 2006). Immunoprecipitated chromatin was analyzed by qPCR using SYBR Green quantitative real time analysis with primers listed in Table 8. For RNA Pol II pSer5 ChIP-seq, $3x10^6$ freshly isolated SMCs were treated by DRB (100 μmol) for 1 hour. 50 μg sheared genomic DNA was subjected to immunoprecipitation with 4 μg RNA Pol II Ser5 antibody according to the established protocol. Protein-DNA complexes were eluted from beads by incubation with 50 μl elution buffer (10 mM Tris-HCl pH 7.4, 5 mM EDTA, 300 mM NaCl, 0.5% SDS) at RT for 5 min and treated with 1 μg DNase free-RNase (Roche) at 37 °C for 30 min. After incubation with 25 μg of proteinase K (10 mg/ml), 1 μg Glycogen at 37 °C for 2 hours, samples were further heated at 65 °C with constant shaking at 1350 rpm overnight, and DNA was purified with PCR purification Kit (MinElute® PCR purification Kit).

CAGE sequencing

Total RNA was isolated using the miRNeasy micro Kit (Qiagen) combined with on-column DNase digestion (DNase-Free DNase Set, Qiagen) to avoid contamination of genomic DNA. CAGE library preparation, sequencing, mapping and motif discovery analysis were performed by DNAFORM (Life Science Research Center 402, 1-1-43, Japan). In brief, RNA quality was assessed by Bioanalyzer (Agilent) to ensure that RIN (RNA integrity number) is over 7.0, and A260/280 and 260/230 ratios are over 1.7. First strand cDNA was transcribed to the 5' end of capped RNAs, attached to CAGE "bar code" tags and the sequenced CAGE tags were mapped to the mouse mm9 genomes using BWA software (v0.5.9) after discarding ribosomal or non-A/C/G/T base-containing RNAs. CAGE tag start site (CTSSs) were acquired as described previously (Takahashi et al., 2012). CTSS number at gene body was calculated by excluding exon 1 and normalized to total tags. Genes with 1.3-fold higher intragenic CTSS signal in Tet3 deficient cells compared to control were selected. The average CTSS coverage at single-base nucleotide were calculated, and those with CTSS < 8 were determined as lowly expressed TSS. For visualization, only CTSS with

tag higher than 8 were preserved, and regions of 25 bp around TSSs specific to TET3 KO were used for motif enrichment analysis by HOMER (Heinz et al., 2010).

Cell culture, plasmid transfection and Co-immunoprecipitation (Co-IP)

HEK293, HEK293T and Hela cells were grown in DMEM (Sigma) supplemented with 10% FCS (Sigma), 2mM L-Glutamine, 100U Penicillin and 100 µg /ml Streptomycin at 37°C, 5% CO_2. HEK293T cells ($2x10^6$/10 cm dish) were transfected with 8 µg Flag-HA-TET3-pEF (#49446, Addgene) using calcium phosphate precipitation method. 48 hours after transfection, HEK293T cells were collected and washed twice with ice-cold PBS. Cells were re-suspended in 300 µl lysis buffer (20mM Tris-HCL, pH 8.0, 200 mM NaCl, 1mM EDTA, 1mM EGTA, 1% Triton X-100) and sonicated with bioruptor for 15 min. The lysates were supplemented with 500 □l lysis buffer and incubated by wheeling on rotor at 4°C for 30min. Cell debris were removed by centrifugation at 12,000 x g for 20min, 4°C, and 800 µg protein lysate was incubated with HA or SETD2 antibody overnight at 4°C followed by incubation with Protein A-Agarose beads (Roche) antibodies on rotor at 4°C for 4 hrs. After 3 time washing with lysis buffer, precipitated proteins were eluted from beads in 2 x SDS loading buffer and analyzed by western blots.

RNA transfection

RNA from 1 million freshly sorted SMCs were extracted by TRIZOL according to manufacturer instruction. Hela or HEK293 cells with 70% confluence in 24 well plate was mock transfected or transfected with 1 µg total RNA using Lipofectamin™ MessengerMAX™ (LipoMAX) Transfection Reagent (Thermo Fisher) according to manufacturer instruction. 20 hours after transfection, cells were collected and RNA was extracted using TRZOL/Chloroform protocol for gene expression analysis.

Proximity ligation assay (PLA)

Proximity ligation assay (PLA). Cryosections of lung tissue or FACS-sorted SMCs after 6 days cultivation in 6 well plate (approx. 200,000 SMCs per well) with SmBM™ Smooth Muscle Cell Growth Basal Medium (LONZA) were fixed with 4 % PFA for 10 min, permeabilized with 0.3% Triton-100x in PBS for 15 min and washed twice with PBS. PLA were performed with Duolink PLA Fluorescence Protocol (MERRK) and antibodies listed in Table 7 according to manufacturer instruction. Tissue sections and cells were finally mounted by Duolink In Situ Mounting Media with DAPI (sigma). Image acquisition were performed by confocal (Leica SP8).

Statistical analysis.

For all quantitative analyses, a minimum of three biological replicates was analyzed. Statistical tests were selected based on the assumption that sample data comes from a population following a probability distribution based on a fixed set of parameters. Student's t-tests were used to determine statistical significance of differences between two groups. One-way AVOVA was used for multiple comparison tests. The following values were considered to be statistically significant: *P<0.05, **P<0.01, P***<0.001, P****<0.0001. Calculations were done using the GraphPad Prism 6 software package. Error bars always represent standard error of the mean. No statistical method was used to predetermine sample size.

RESULTS

1. Inactivation of *Tet2* has no impact on global 5hmC levels in the heart and lung

Tet-genes plays a crucial role in different cell types during embryonic development and in adults but limited information is available about the function of TET-proteins and their product 5hmC in the heart and lung. *Tet2* and *Tet3* but not *Tet1* are highly expressed in the neonatal heart resulting in high level of 5hmC in adult cardiomyocytes (Greco et al., 2016), Similarly, the expression of *Tet1* in the adult lung is virtually undetectable, which is why I focused on *Tet2* and *Tet3* during the course of this study. To investigate the impact of TET2 inactivation on 5hmC deposition in heart and lung, I compared 5hmC levels in wild-type and *Tet2* germ line knockout mice ($Tet2^{KO}$) generated by the CRISPR-Cas9 technology (Wang et al., 2013) (Figure 1 A & B). Unexpectedly, no pronounced change of 5hmC levels was observed in *Tet2* deficient hearts and lungs, despite high expression of TET2 (Figure 1 C & D). Similarly, loss of *Tet2* did not affect heart morphology and the ratio between heart weight and body weight (Figure 1 E). To corroborate these findings, a second mouse line ($Tet2^{cKO}$) was employed, in which the catalytic domain of *Tet2* gene was excised by CMV-cre (Figure 1 F). Again, no obvious morphological defects were observed in the heart and the lung. No signs of airway remodeling such as mucus-producing ($Mucin5AC^{+}$) cell hyperplasia and increased expression of the synthetic SMC marker Vimentin was visible (Figure 1 G & H). Of note, no evident change of 5hmC level in airway smooth muscle layer were observed in *Tet2*-deficient lung (Figure 1 I). These findings are not consistent with a previous study showing that knockdown of *Tet2* in human coronary artery SMCs (hCASMCs) results in reduced 5hmC level. However, the previous study did mostly rely on knockdown approaches and did not use genetic models, which leaves the possibility for potential off-target effects of siRNAs. It is also possible that the biological consequences of a loss of *Tet2* might are highly context-dependent, or that the function of TET2 is compensated by TET3 in lung SMCs *in vivo*.

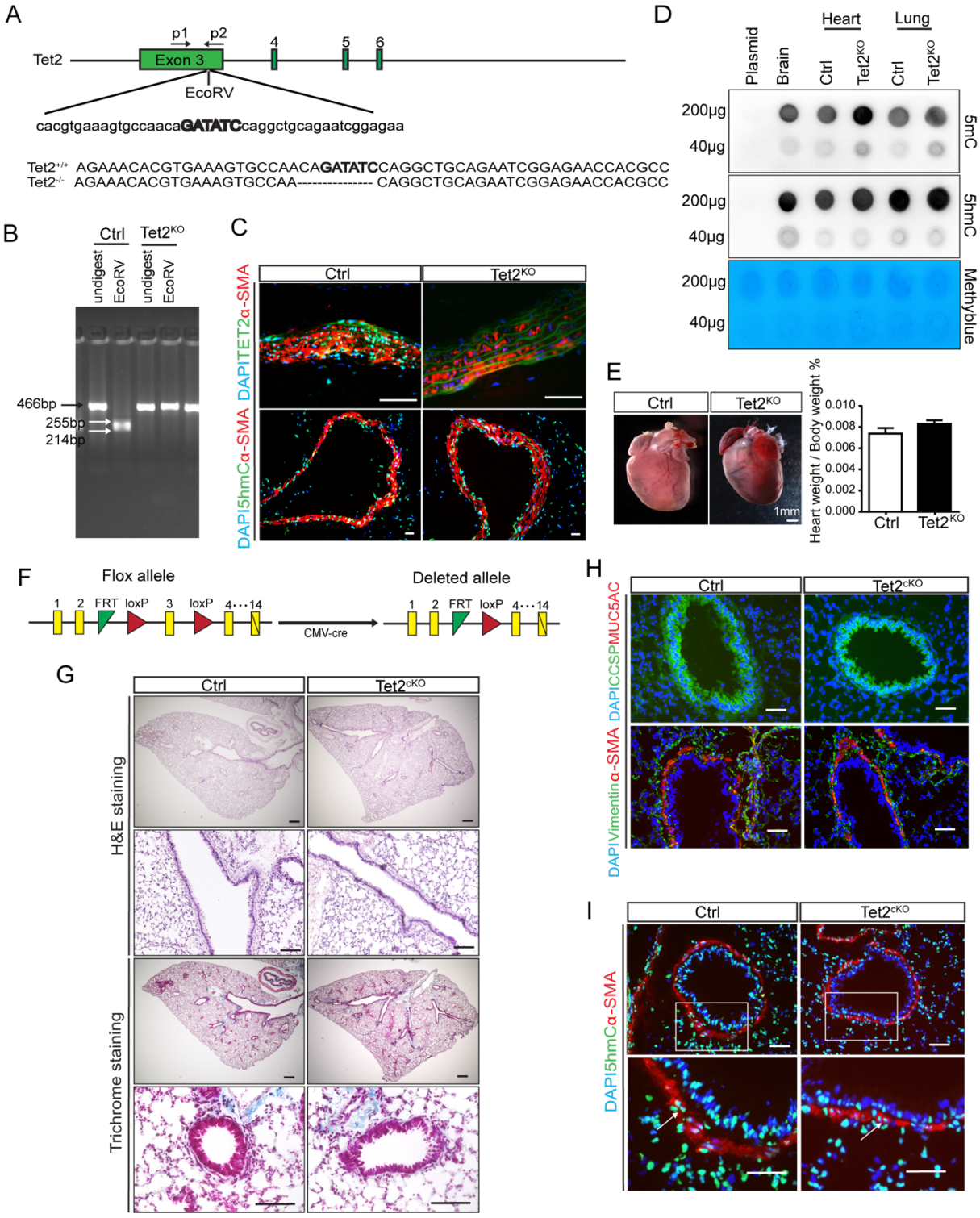


Figure 1 Inactivation of *Tet2* neither reduces global levels of 5-hydroxymethylation nor causes a pathological phenotype in the heart and lung.

(A) Schematic of the Cas9/sgRNA-targeting sites in *Tet2* gene. The sgRNA-targeting sequence is underlined, the protospacer-adjacent motif (PAM) sequence is labeled in green. The restriction site at the targeted region is in bold and capitalized (Wang et al., 2013). **(B)** Genotyping of wild type and *Tet2* knockout (*Tet2*^{KO}) mice by EcoRV digestion. **(C)** Representative immunofluorescence images of co-staining for TET2 & α-SMA or 5hmC & α-SMA with cryosections prepared from aortas of 8-week-old wild type and *Tet2*^{KO} mice (n=3). DNA was stained with DAPI. Scale bar: 50 μm. **(D)** Dot blots to detect global 5mC, 5hmC levels of genomic DNA extracted from heart and lung dissected from 8-week-old wild type and *Tet2*^{KO} mice (n=3). Plasmid and brain DNA were used as

negative or positive control respectively. **(E)** Representative macroscopic images of wild type and *Tet2*^{KO} hearts (n=3). Ratios between heart weight and body weight of 8-week-old wild type or *Tet2*^{KO} mice were calculated and shown in the right panel (n=3). Data are represented as mean ± SEM. **(F)** Schematic depicting the floxed *Tet2* allele and deleted *Tet2* allele after CMV-cre-mediated excision of exon 3. **(G)** Representative images of H&E staining and Trichrome staining with paraffin sections of lung dissected from 8-week-old control and *Tet2*^{cKO} mice (n=3). Scale bar: 200μm (upper panels of H&E staining and Trichrome staining); 50 μm (lower panels of H&E staining and Trichrome staining). **(H)** Representative immunofluorescence images of co-staining for CCSP & Mucin-5AC (MUC5AC) or Vimentin & α-SMA with cryosections prepared from the lungs of 8-week-old wild type and *Tet2*^{cKO} mice (n=3). DNA was stained with DAPI. Scale bar: 50 μm. **(I)** Representative immunofluorescence images of co-staining for 5hmC & α-SMA with cryosections prepared from the lungs of 8-week-old wild type and *Tet2*^{cKO} mice (n=3). DNA was stained with DAPI. Scale bar: 50 μm.

2. Cardiomyocyte specific *Tet3* inactivation does not have an obvious impact on 5hmC levels in the heart and does not affect cardiac morphology

TET3 is the major TET isoform highly expressed in oocytes, zygotes and neuron, but its function in differentiated cells or adult tissues remains enigmatic. To investigate the function of TET3 in the heart, I generated the *Tet3*$^{fl/fl}$α-MHC-MerCreMer^{pos} mouse line (thereafter referred as *Tet3*^{cKO}) in which floxed exon 10 of *Tet3* gene, an exon encoding the conserved catalytic domain of TET3, can be excised by Cre recombinase specifically in cardiomyocytes (CMs) after tamoxifen (TAM) injection (Figure 2 A-C). Two months after TAM injection, *Tet3* expression was efficiently abolished in CMs isolated from *Tet3*^{cKO} (Figure 2 D). However, no obvious changes of 5mC and 5hmC levels were detected in mutant CMs compared to control CMs (Figure 2 D & E). Loss of TET3 also did not affect CM size, heart morphology, body weight and the ratio between heart weight and body weight (Figure 2 E-G). Taken together, separate inactivation of *Tet2* or *Tet3* does not affect 5hmC levels in the heart and does not cause any gross cardiac abnormality, indicating possible functional redundancy of these two enzymes in regulating 5hmC deposition in the heart.

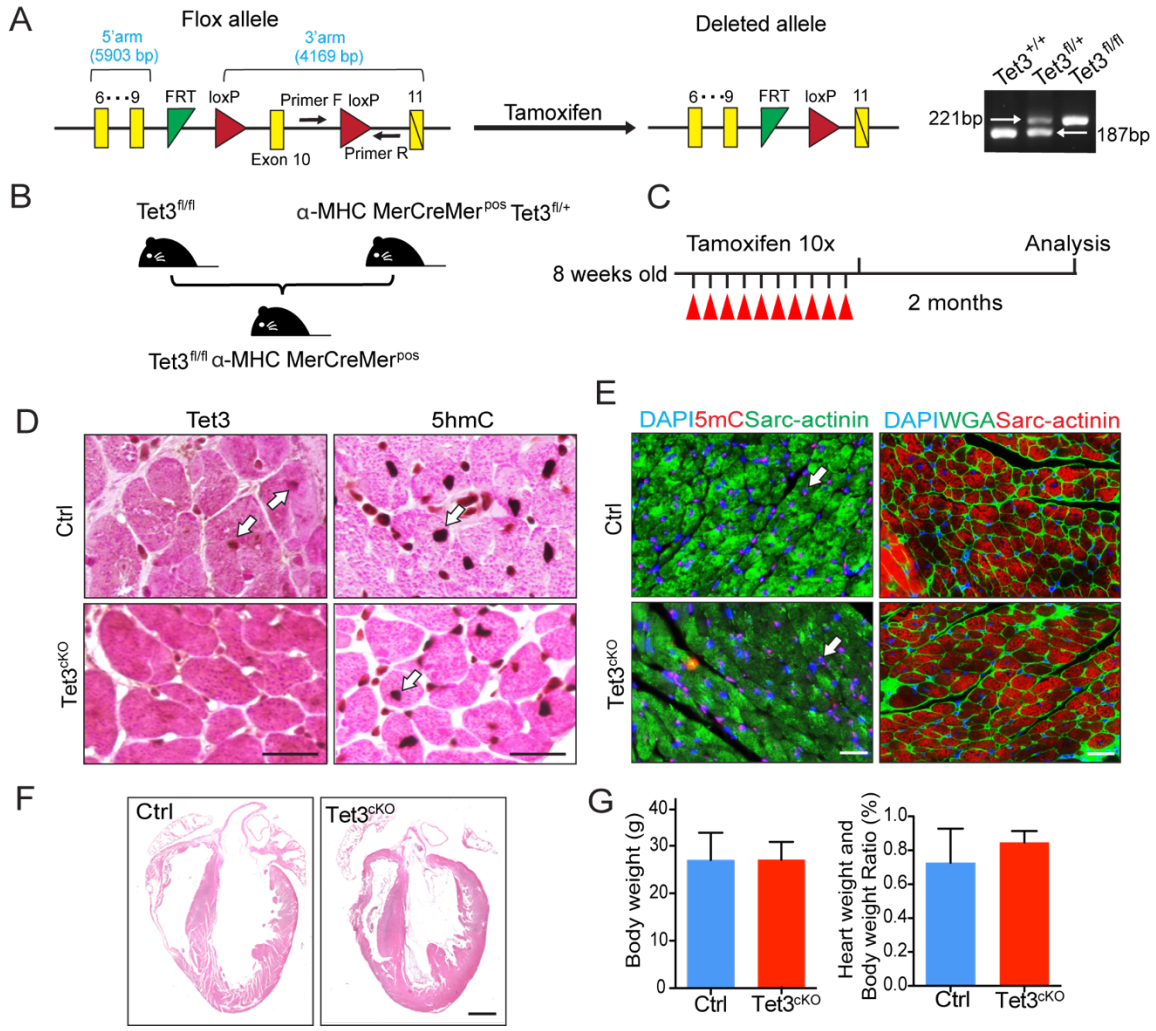


Figure 2 5hmC level in CMs and cardiac morphology are not altered after *Tet3* inactivation.
(A) Schematic depicting the floxed and deleted *Tet3* alleles. The primers used for genotyping (Primer F & Primer R) are indicated by arrows. The PCR results for genotyping are shown in the right panel. **(B)** Outline of the strategy to generate inducible *Tet3* knockout mouse lines in CMs. **(C)** Experimental design to induce CM-specific *Tet3* inactivation with TAM. **(D)** Representative immunohistochemistry images of TET3 and 5hmC co-staining with cryosections of the heart dissected from control (Ctrl) and *Tet3*^{cKO} mice two months after tamoxifen injection (n=3). Scale Bar: 50 µm. **(E)** Representative immunofluorescence images of co-staining for 5mC and Sarc-actinin or WGA and Sarc-actinin with cryosections of the hearts dissected from Ctrl and *Tet3*^{cKO} mice two months after tamoxifen injection (n=3). Scale Bar: 50 µm. **(F)** Representative images of H&E staining with cryosections of the hearts dissected from Ctrl and *Tet3*^{cKO} mice (n=3) two months after tamoxifen injection. Scale Bar: 1 mm. **(G)** Body weight (left panel) and ratio between heart weight and body weight (right panel) of 8-week-old Ctrl and *Tet3*^{cKO} mice (n=3). Data are represented as mean ± SEM. Control mice used in D-G: *Tet3*$^{fl/fl}$ α-MHC MerCreMer^{neg}.

3. *Tet3* is highly expressed in smooth muscle cells in various adult tissues

Next, I generated the transgenic mouse line *Tet3*$^{floxlacZ/+}$, in which *LacZ* gene expression is driven by the endogenous *Tet3* promoter, to monitor the expression of *Tet3* in different cell types (Figure 3 A). X-gal staining revealed that *LacZ* gene was ubiquitously expressed in E9.5 embryo, while high levels of *LacZ* expression were specifically observed in smooth muscle cells (SMCs) in various

adult tissues including heart, lung, skeletal muscle, brain and esophagus (Figure 3 B & C). These findings suggest that *Tet3* is highly expressed in SMCs in adult tissues.

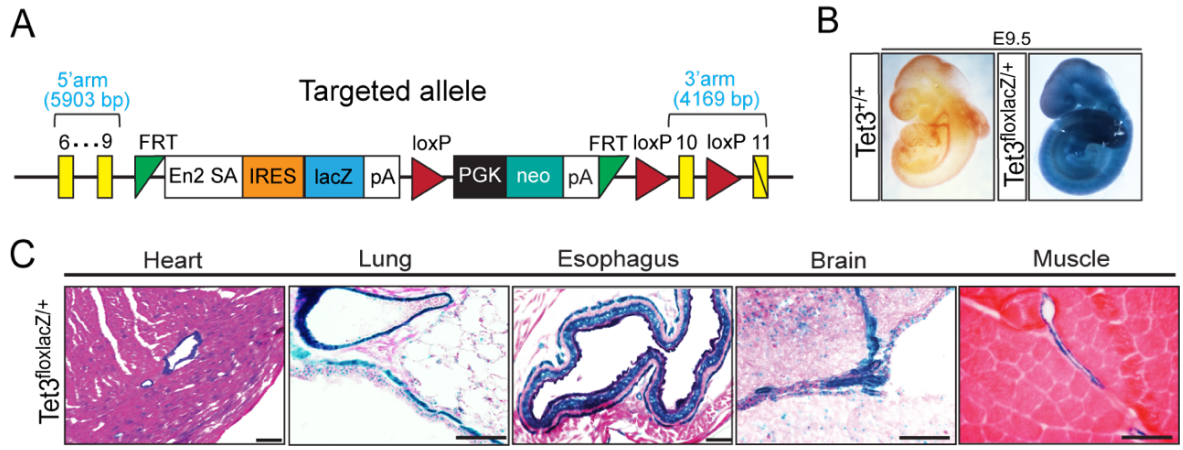


Figure 3 TET3 is highly expressed in SMCs in various adult tissues.
(A) Schematic diagram of targeted *Tet3* allele containing LacZ reporter. **(B)** Whole-mount lacZ staining of *Tet3*$^{+/+}$, *Tet3*$^{floxlacZ/+}$ embryos on E9.5 (n=10). **(C)** Representative images of LacZ staining with cryosections of the heart, lung, esophagus, brain and skeletal muscle dissected from 8-week-old *Tet3*$^{floxlacZ/+}$ mice. Sections were counterstained for and Eosin (n=3). Scale bar: 50 µm.

4. *Tet3* inactivation in SMCs results in pathological changes in the lung

4.1 Generation of an inducible, SMC-specific *Tet3* deficient mouse line

To address the function of TET3 in SMCs in vivo, I generated a SMC-specific *Tet3* knockout mouse line using an inducible α-SMA-creERT2 allele (LeBleu et al., 2013). In the resulting *Tet3*$^{fl/fl}$α-SMA-creERT2^{pos} (thereafter referred as *Tet3*^{smKO}) mice, the floxed exon 10 of *Tet3* gene can be specifically excised by Cre recombinase in SMCs after tamoxifen (TAM) injection (Figure 4.1 A & B). Immunofluorescence staining revealed that *Tet3* expression in smooth muscle cells of aorta dissected from *Tet3*^{smKO} was substantially diminished (Figure 4.1 C). RT-qPCR further confirmed the high efficiency of *Tet3* depletion in aortas isolated from *Tet3*^{smKO} mice (Figure 4.1 D). *Tet3*^{smKO} mice were viable and fertile but exhibited significant reduction of body weight, indicating an important function of TET3 in regulating smooth muscle cell/tissue homeostasis (Figure 4.1 E).

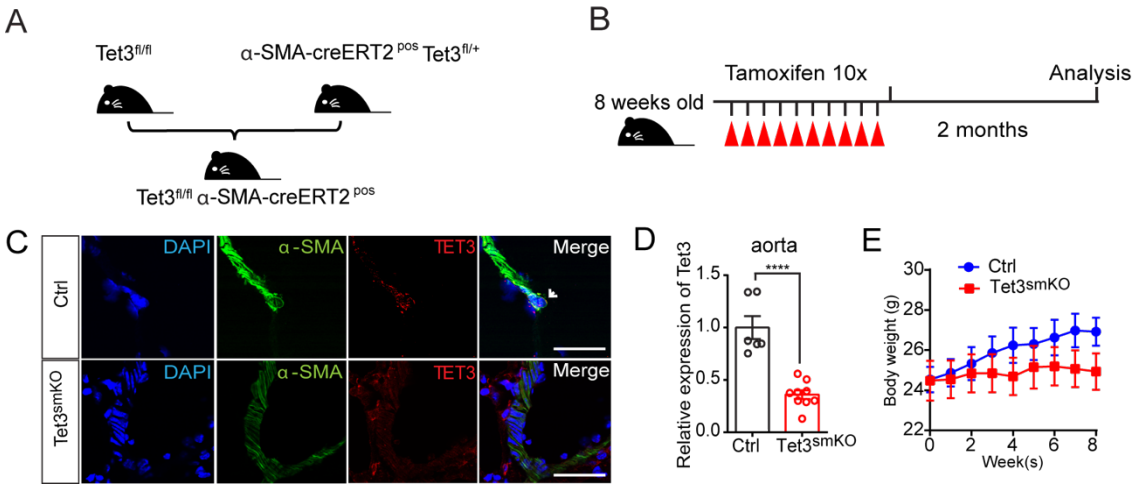


Figure 4.1 Generation of the inducible SMC-specific *Tet3* knockout mouse line.

(A) Outline of the strategy to generate inducible SMC-specific *Tet3* knockout mouse model. **(B)** Experimental design to induce SMC-specific *Tet3* inactivation with tamoxifen injection. **(C)** Representative immunofluorescence images of co-staining for α-smooth muscle actin (α-SMA) and TET3 with cryosections prepared from aortas of control (Ctrl) or *Tet3*^{smKO} mice 8 weeks after tamoxifen injection (n=3). DNA was stained by DAPI. Scale bar: 50 µm. **(D)** RT-qPCR analysis of *Tet3* expression in aortas from control (Ctrl) and *Tet3*^{smKO} mice 8 weeks after Tamoxifen injection. Quantitative PCR was performed with the primers to detect exon 10 flanked by *LoxP* sites. The *β-actin* gene was used as reference for normalization. The values represent mean ± SEM (n=6). (Unpaired t-test: ****p<0.0001). **(E)** Body weight of control (Ctrl) or *Tet3*^{smKO} mice at different time points after tamoxifen injection (n=6). Data are represented as mean ± SEM. Tamoxifen injection was performed with 8-week-old mice. Control mice used in C-E: *Tet3*$^{+/+}$α-SMA-creERT2^{pos}.

4.2 *Tet3* ablation in SMC induces airway remodeling but does not cause pathological changes in the heart, aorta and intestine

Smooth muscle cells are highly enriched in multiple tissues/organs including the heart, aorta, lung and intestine. To analyze the impact of SMC-specific inactivation of *Tet3* on tissue homeostasis in vivo, I dissected various tissues/organs from control and *Tet3*^{smKO} mice and performed morphological analysis by H&E staining. *Tet3* inactivation did not cause obvious morphological defects in the heart, aorta and intestine (Figure 4.2 A & B). However, a shift from a simple columnar to a cuboidal epithelium and thickening of the basal membrane in bronchi of *Tet3*^{smKO} lung were observed, indicating that *Tet3* ablation in SMCs induces airway remodeling (Figure 4.2 C).

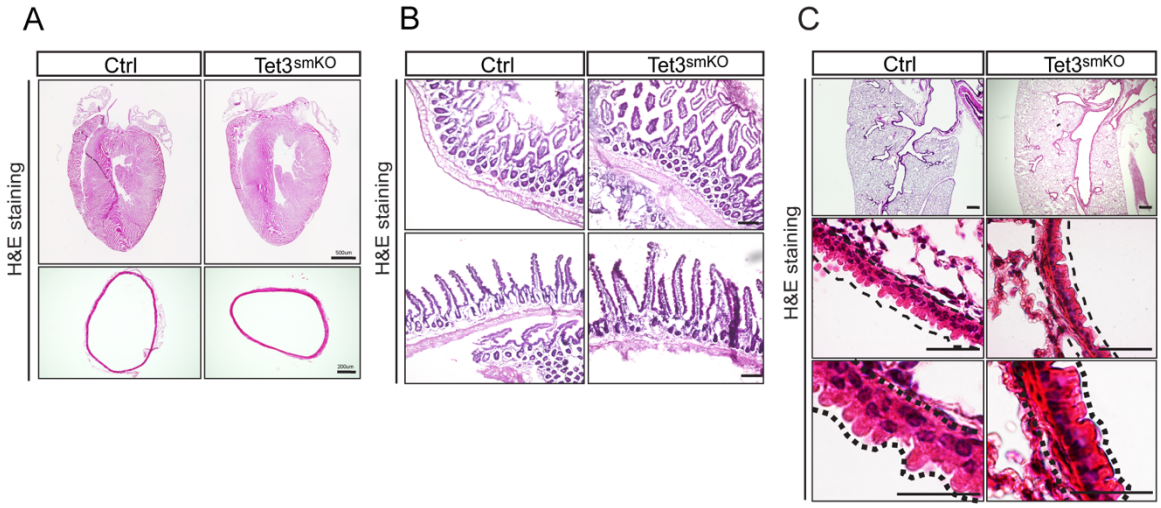


Figure 4.2 Inactivation of *Tet3* in SMCs induces airway remodeling but does not cause pathological phenotype in the heart, aorta and intestine.

(A) Representative images of H&E staining with cryosections from the hearts and aortas of control (Ctrl) and *Tet3*^{smKO} mice 8 weeks after tamoxifen injection (n=3). Scale bar: 50 µm. **(B)** Representative images of hematoxylin & eosin (H&E) staining with cryosections of the intestines dissected from control (Ctrl) and *Tet3*^{smKO} mice 8 weeks after tamoxifen injection (n=3). Scale bar: 50 µm. **(C)** Representative images of hematoxylin & eosin (H&E) staining with paraffin sections of the lungs dissected from control (Ctrl) and *Tet3*^{smKO} mice 8 weeks after tamoxifen injection (n=4). Scale bar: 200µm (upper panel); 50 µm (middle panel); 10 µm (lower panel). Note the morphological change of epithelia in mutant lung indicated by dash lines. Control mice used in this figure: *Tet3*$^{+/+}$α-SMA-creERT2^{pos}.

4.3 Loss of *Tet3* in SMCs provokes metaplasia of club to goblet cells

Lung bronchial epithelium is lined with different cell types including mucus-producing goblet cells, ciliated cells and secretory progenitor club cells. Various injuries promote airway remodeling by stimulating metaplasia of club cells to mucus-producing goblet cells during allergic lung pathogenesis. To test whether *Tet3* inactivation induces epithelial cell metaplasia, I analyzed the expression of club cell secreted protein (CCSP, a marker of club cells) and α-tubulin (a marker of ciliated cells) in control and *Tet3*^{smKO} bronchiole by immunohistochemistry staining. Depletion of *Tet3* resulted in dramatically attenuated expression of CCSP, while no change of α-tubulin expression was detected in bronchiolar epithelia, implying metaplasia of club but not ciliated cells (Figure 4.3 A & B). Concurrently, *Tet3* inactivation led to impaired expression of α-SMA but elevated expression of Mucin-5AC, the most prominent secreted mucin, in the lung but not in the intestine epithelium, suggesting that the metaplasia of club cells to mucus-producing goblet cells following *Tet3* ablation is tissue specific (Figure 4.3 C). The increase of mucus-producing goblet cells in *Tet3*-deficient lungs was further demonstrated by elevated expression of another prominent secreted mucin protein, Anterior Gradient 2, markedly augmented number of Periodic Acid-Schiff (PAS)-staining positive cells and highly enriched secreting granules in the respiratory tract of *Tet3*

mutant lung (Figure 4.3 D-F). Taken together, these findings indicate that *Tet3* inactivation provokes metaplasia of club cells to mucus producing goblet cells, indicating extensive airway remodeling.

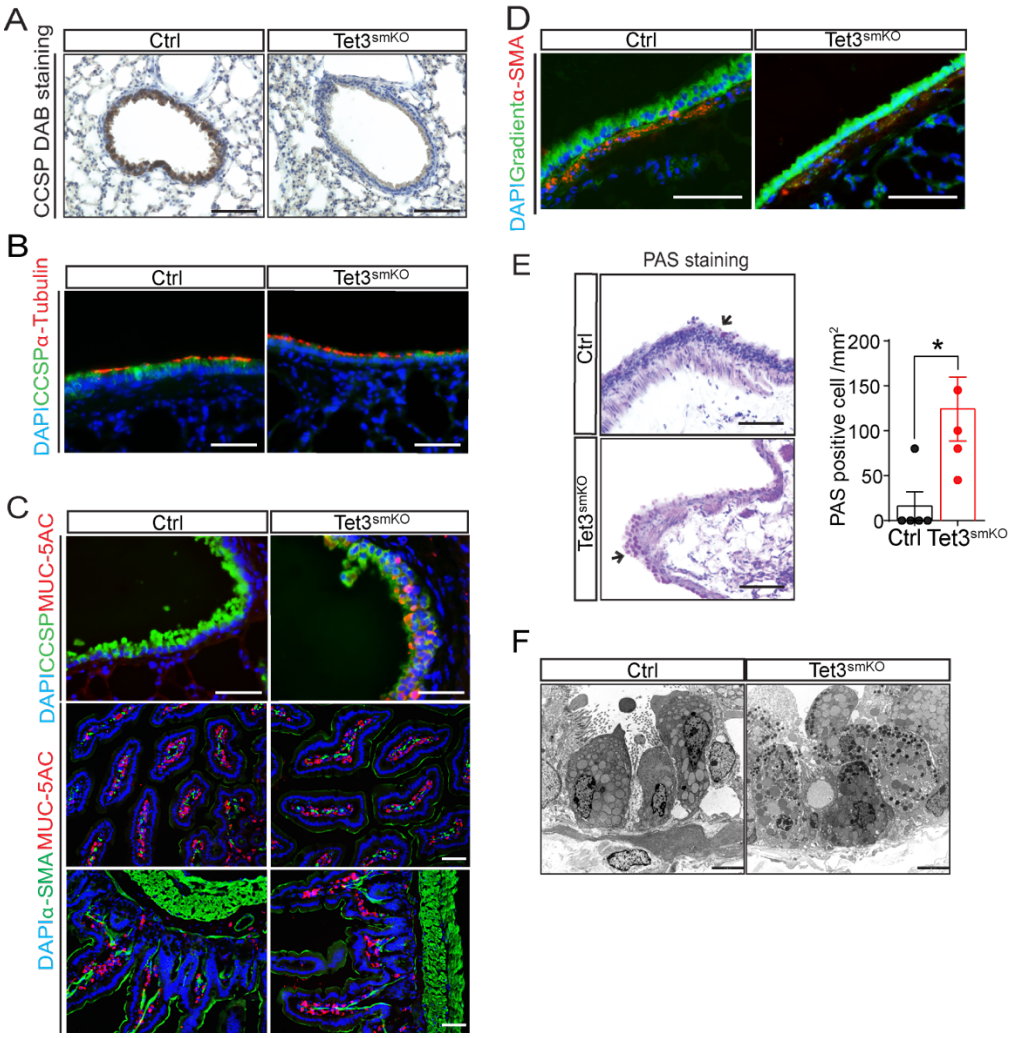


Figure 4.3 Inactivation of *Tet3* in SMCs provokes metaplasia of club cells to mucus-producing goblet cells
(A) Representative images of immunohistochemistry staining for CCSP with paraffin sections prepared from control (Ctrl) and *Tet3*^{smKO} lungs 8 weeks after Tamoxifen injection (n=5). Scale bar: 50 µm. **(B)** Representative immunofluorescence images of co-stainings for α-Tubulin and CCSP with paraffin sections prepared from control (Ctrl) and *Tet3*^{smKO} lungs 8 weeks after tamoxifen injection (n=5). DNA was counterstained by DAPI. Scale bar: 50 µm. **(C)** Representative immunofluorescence images of co-staining for α-SMA and Mucin5AC with cryosections from the lungs (upper panel) or intestines (middle and lower panels) of control (Ctrl) and *Tet3*^{smKO} mice 8 weeks after tamoxifen injection (n=3). Scale bar: 50 µm. **(D)** Representative immunofluorescence images of co-staining for Gradient and α-SMA with cryosections prepared from the lungs of control (Ctrl) and *Tet3*^{smKO} mice 8 weeks after tamoxifen injection (n=3). DNA was stained by DAPI. Scale bar: 50 µm. **(E)** Representative images of Periodic Acid-Schiff staining (PAS) with cryosections prepared from control (Ctrl) and *Tet3*^{smKO} lungs 8 weeks after tamoxifen injection. Quantification of mucus-producing cell (PAS staining positive) was performed by Image J software and shown in the right panel (Unpaired t-test: *p<0.05; n=5). Data are represented as mean ± SEM. Scale bar: 50 µm. **(F)** Representative electronic microscope images of cryosections of lungs dissected from control (Ctrl) and *Tet3*^{smKO} mice (n=4). Scale bar: 5000nm. Control mice used in this figure: *Tet3*$^{+/+}$α-SMA-creERT2^{pos}.

4.4 Inactivation of *Tet3* reduces global 5-hydroxymethylcytosine levels in SMCs

To address the role of TET3 in regulating 5hmC deposition in lung SMCs, I compared the expression levels of *Tet3* to the two other *Tets* (i.e *Tet1, Tet2*) in adult lung SMCs.

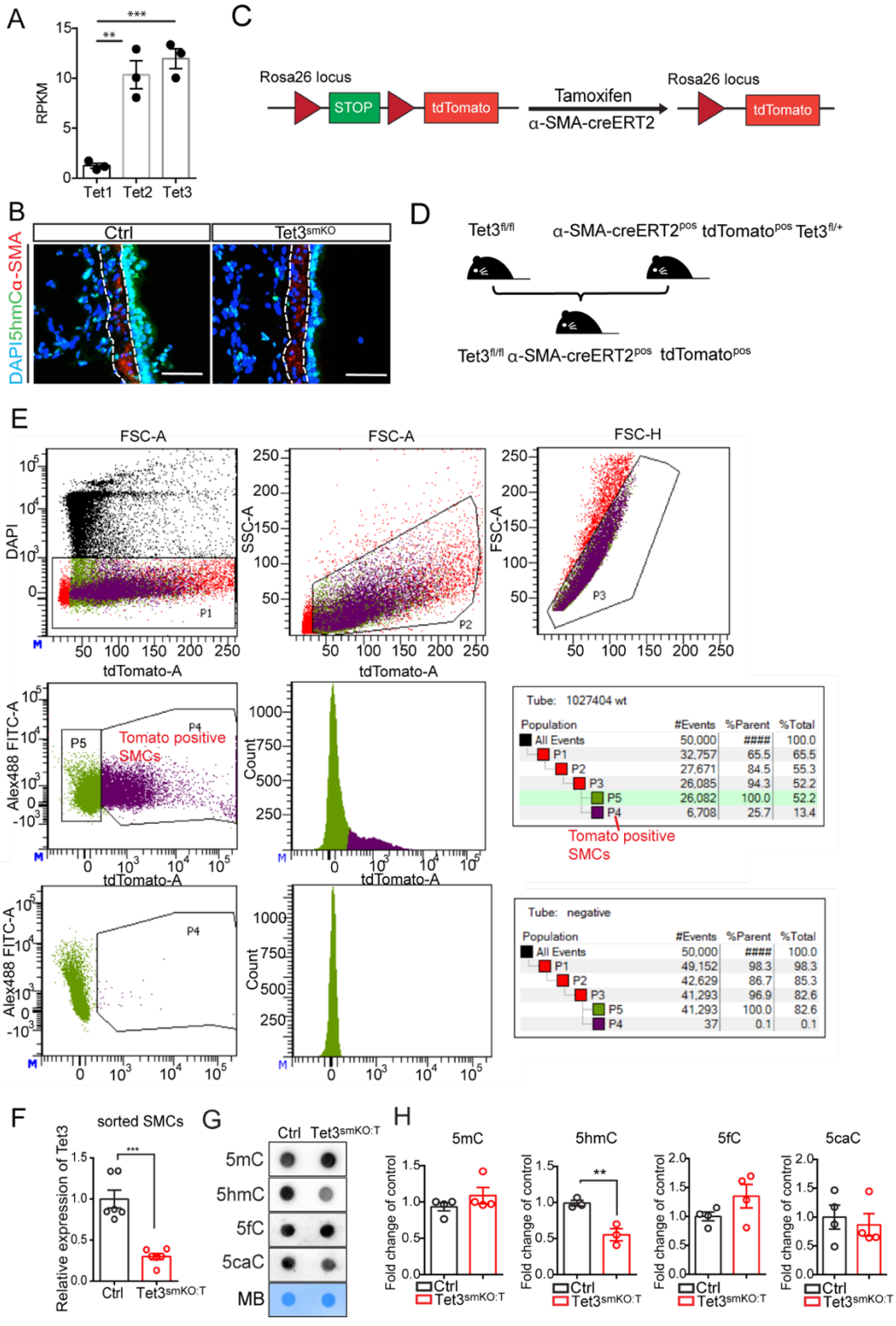


Figure 4.4 Inactivation of *Tet3* in SMCs significantly reduces global levels of 5-hydroxymethylation.
(A) Bar graph showing the normalized RNA-seq read counts (RPKM) of *Tet1, Tet2* or *Tet3* gene in FACS-sorted lung SMCs isolated from *Tet3*$^{+/+}$α-SMA-creERT2^{pos} tdTomato^{pos} mice (n=3). Data are represented as mean ± SEM.

One-way ANOVA: **p<0.01, ***p<0.001. **(B)** Representative immunofluorescence images of co-staining for α-SMA and 5hmC with cryosections from the lungs of control (Ctrl) and *Tet3*^{smKO} mice 8 weeks after Tamoxifen injection (n=3). DNA was stained by DAPI. Scale bar: 50 µm. **(C)** Outline of the strategy to induce SMC-specific expression of *tdTomato* reporter gene. **(D)** Outline of the strategy to generate inducible *Tet3* knockout mouse line with conditional expression of *tdTomato* reporter gene in SMCs. **(E)** FACS strategy to isolate SMCs from the lungs dissected from control (Ctrl) and *Tet3*$^{smKO:T}$ mice 8 weeks after tamoxifen injection. **(F)** RT-qPCR analysis of *Tet3* expression in sorted lung SMCs from control (Ctrl) and *Tet3*$^{smKO:T}$ mice 8 weeks after tamoxifen injection with primers specifically detect *LoxP*-flanked exon 10 (n=6). The *β-actin* gene was used as reference for normalization. The values represent mean ± SEM. Unpaired t-test: ***p<0.001. **(G)** Dot blots to monitor global 5-methylcytosine (5mC), 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) levels with genomic DNA of sorted lung SMCs from control (Ctrl) or *Tet3*$^{smKO:T}$ mice 8 weeks after tamoxifen injection (n=4). Genomic DNA stained by methylene blue (MB) were used as loading control for normalization. **(H)** Quantifications of 5mC (n=4), 5hmC (n=3), 5fC (n=4), 5caC (n=4) levels monitored by dot blots in **G** using image lab software from ChemiDoc gel imaging system. Data are represented as mean ± SEM. Unpaired t-test: **p<0.01. Control mice used in F-H: *Tet3*$^{+/+}$α-SMA-creERT2^{pos} tdTomato^{pos}.

RNA-seq revealed negligible expression of *Tet1* while both *Tet2* and *Tet3* were highly expressed in FACS-sorted lung SMCs (Figure 4.4 A). Interestingly, 5hmC levels were dramatically reduced in *Tet3*-deficient airway SM layer, implying that the function of TET3 in lung SMCs cannot be compensated by TET2 (Figure 4.4 B). In order to investigate the molecular events that occurred following *Tet3* inactivation, *Tet3*^{smKO} mouse was crossed with a *tdTomato* reporter mouse to generate *Tet3*$^{fl/fl}$α-SMAcreERT2^{pos}tdTomato^{pos} mouse (thereafter referred as *Tet3*$^{smKO:T}$). *Tet3*$^{smKO:T}$ mice allow tracing and isolation of SMCs, since the CAG promoter-driven *tdTomato* can be specifically expressed in SMCs after tamoxifen induced removal of *LoxP*-flanked STOP cassette inserted into the *ROSA26* locus (Figure 4.4 C & D). By using this mouse line, I isolated tdTomato positive lung SMCs from *Tet3*$^{smKO:T}$ mice by FACS-sorting and validated the high efficiency of *Tet3* depletion by RT-qPCR (Figure 4.4 E & F). Next, I performed dot blots to determine the global levels of 5mC and its oxidized derivatives in lung SMCs sorted from control and *Tet3*$^{smKO:T}$ mice. Depletion of *Tet3* in SMCs resulted in a substantial decrease of 5hmC levels, while no obvious influence on global level of 5mC, 5fC and 5caC were detected (Figure 4.4 G & H).

5. TET3 is a key regulator of SMC plasticity in the lung

5.1 *Tet3*-deficient lung SMCs shift from a contractile to a synthetic phenotype

Next, I determined the primary cellular events occurring in lung SMCs after *Tet3* inactivation. The smooth muscle cell is a unique cell type that exhibits remarkable plasticity and can acquire different phenotypes in response to different genetic or environmental cues ranging from a contractile to a synthetic state.

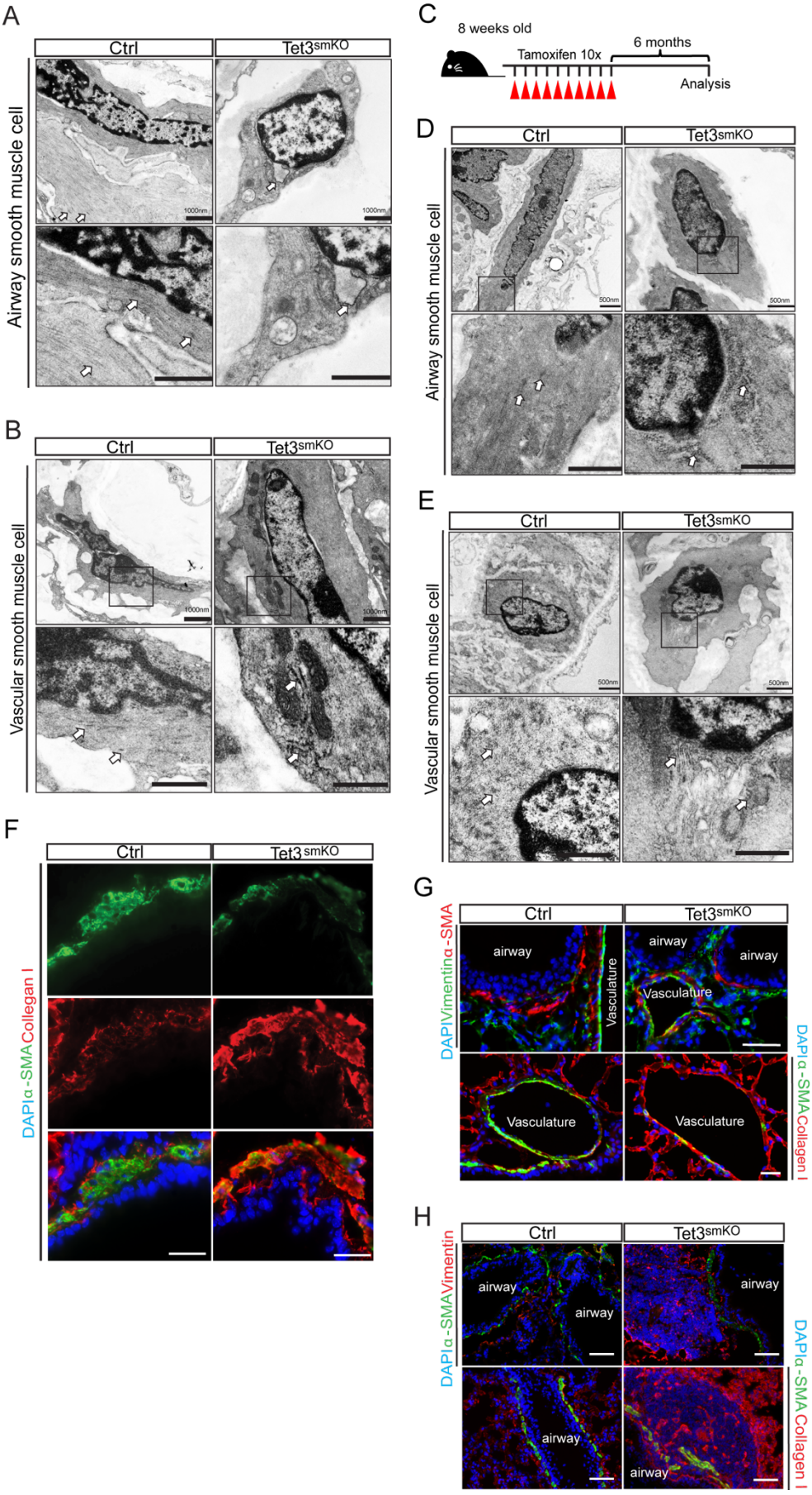

A
Ctrl
Tet3smKO
Airway smooth muscle cell
1000nm
B
Vascular smooth muscle cell
C
8 weeks old
Tamoxifen 10x
6 months
Analysis
D
500nm
E
F
DAPIα-SMACollegan I
G
DAPIVimentinα-SMA
airway
Vasculature
H
DAPIα-SMAVimentin

Figure 5.1 *Tet3*-deficient SMCs shift from a contractile to a synthetic state.
(A&B) Representative electronic microscope images of ASMC (A) or lung VSMC (B) on sections prepared from control (Ctrl) and *Tet3*^{smKO} mice 8 weeks after tamoxifen injection (n=4). Scale bar: 1000nm. **(C)** Experimental design to induce SMC-specific *Tet3* inactivation with tamoxifen injection in a long-term experiment. **(D&E)** Representative electronic microscope images of ASMC (D) or lung VSMC (E) on sections prepared from control (Ctrl) and *Tet3*^{smKO} mice 24 weeks after tamoxifen injection (n=4). Scale bar: 1000nm. **(F)** Representative immunofluorescence images of co-staining for α-SMA and Collagen I with crysections of the lung ASMC dissected from control (Ctrl) and *Tet3*^{smKO} mice 8 weeks after tamoxifen injection (n=5). DNA was stained by DAPI. Scale bar: 50 µm. **(G&H)** Representative immunofluorescence images of co-stainings for α-SMA and Vimentin (upper panel) or α-SMA and Collagen I (lower panel) with cryosections of the lungs dissected from control (Ctrl) and *Tet3*^{smKO} mice 8 weeks (G) or 24 weeks (H) after tamoxifen injection (n=3). DNA was stained by DAPI. Scale bar: 50 µm. DNA was stained by DAPI. Scale bar: 50 µm. Control mice used in this figure: Tet3$^{+/+}$α-SMA-creERT2^{pos}.

Assuming that dynamic DNA methylation/demethylation on cytosine might be a major epigenetic mechanism involved in the regulation of cell plasticity, I asked whether *Tet3* deficiency has an impact on the phenotype of SMC. To answer this question, I compared the ultrastructure of vascular and airway SMCs in control and *Tet3*^{smKO} lung 2 months after TAM injection. Electron microscopy (EM) showed that the majority of control SMCs maintained a defined spindle-shape and contained actin filaments and dense bodies, while most of *Tet3*^{smKO} SMCs showed rhomboid-shape with rough endoplasmic reticulum (rER) in the cytoplasm (Figure 5.1 A & B). Similar ultrastructural changes were observed in *Tet3*-deficient SMCs 6 months after TAM injection (Figure 5.1 C-E). Consistently, immunofluorescence staining revealed attenuated expression of α-SMA, a classic contractile marker gene, whereas levels of synthetic markers such as Collagen I and Vimentin (Vim) were elevated in *Tet3*-deficient smooth muscle layers in both short term and long-term experiments (Figure 5.1 F-H). In summary, loss of *Tet3* in SMCs leads to phenotypic switching from a contractile to a synthetic state.

5.2 TET3 is indispensable to maintain SMC contractility

To further prove the loss of contractile SMC after *Tet3* inactivation, I performed laser capture microdissection and microarray analysis using bronchioles or vasculature dissected from control and *Tet3*^{smKO} lungs. Expression profiling analysis revealed that *Mir145a*, a master regulator of SMC contractility (Boettger et al., 2009), was substantially reduced in both bronchioles and lung vasculature (Figure 5.2 A). RT-qPCR using FACS-sorted lung SMCs further confirmed the downregulation of *Mir145a* in *Tet3*-deficient SMCs (Figure 5.2 B). Interestingly, very few deregulated genes overlapped in bronchioles and vasculature after *Tet3* inactivation, which might imply tissue-specific paracrine effect of *Tet3*-deficient SMCs on neighboring cells. The phenotypic shift of SMCs was further validated by western blot analysis showing reduced protein levels of contractile markers like α-SMA, Myosin heavy chain 11 (MYH11), Transgelin (TAGLN) but

increased levels of synthetic markers like Vimentin (VIM) and Tropomyosin alph-4 chain (TPM4) (Figure 5.2 C & D). These findings indicate that TET3 plays a vital role in maintaining SMC contractility.

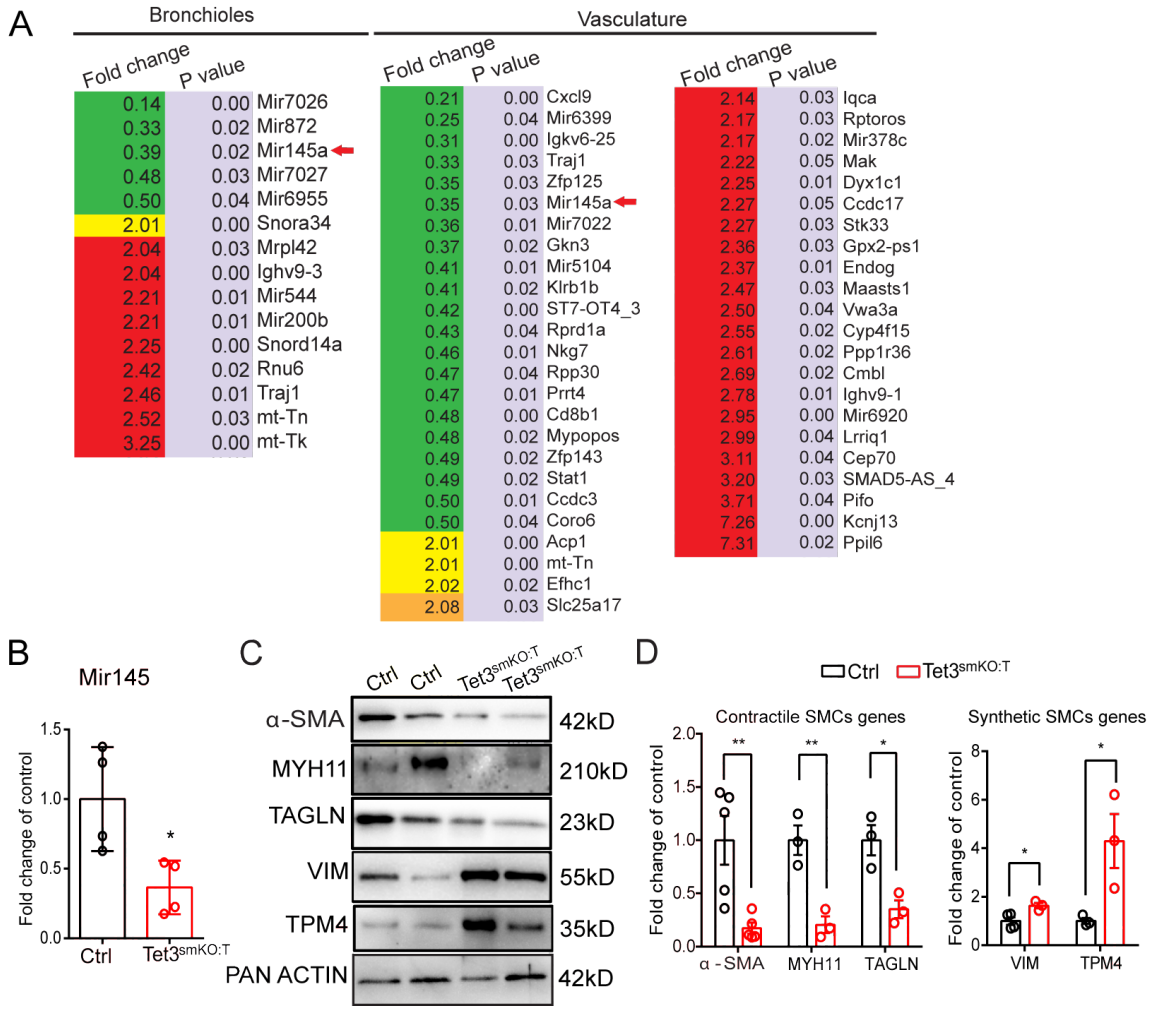


Figure 5.2 Contractile genes are downregulated but synthetic genes are upregulated in *Tet3*-deficient SMCs. **(A)** Deregulated genes (fold change (Tet3$^{smKO:T}$/Ctrl)>2 or <0.5, *P* value=<0.05) identified by microarray analysis of bronchiole (left panel) or vasculature (middle and right panels) tissues isolated by laser capture dissection of cryosections prepared from control (Ctrl) and *Tet3*^{smKO} lungs 8 weeks after tamoxifen injection (n=3). **(B)** RT-qPCR analysis of *Mir145* expression in sorted lung SMCs from control (Ctrl) and *Tet3*$^{smKO:T}$ mice 8 weeks after tamoxifen injection (n=4). Data are represented as mean ± SEM. Unpaired t-test: *p<0.05. **(C)** Western blots to detect protein levels of indicated contractile and synthetic genes in sorted lung SMCs from control (Ctrl) and *Tet3*$^{smKO:T}$ mice 8 weeks after tamoxifen injection. The protein level of pan-actin was used as loading control. **(D)** Quantification of protein levels of indicated contractile and synthetic genes in sorted lung SMCs isolated from control (Ctrl) and *Tet3*$^{smKO:T}$ mice 8 weeks after tamoxifen injection using image lab software from ChemiDoc gel imaging system. Data are represented as mean ± SEM. (n=3) Unpaired t-test: *p<0.05; **p<0.01. Control mice used in this figure: *Tet3*$^{+/+}$α-SMA-creERT2^{pos}tdTomato^{pos}. (microarray data analysis was assisted by Dr. Thomas Böttger)

6. 5hmC is enriched in intragenic regions of highly expressed genes and is dramatically reduced after *Tet3* ablation in SMCs

The genome-wide distribution of 5hmC depends on the cell type and is tissue-specific. To analyze the genomic distribution of 5hmC in lung SMCs, I performed 5hmC antibody-based hydroxymethylated DNA immunoprecipitation sequencing (hMeDIP-seq).

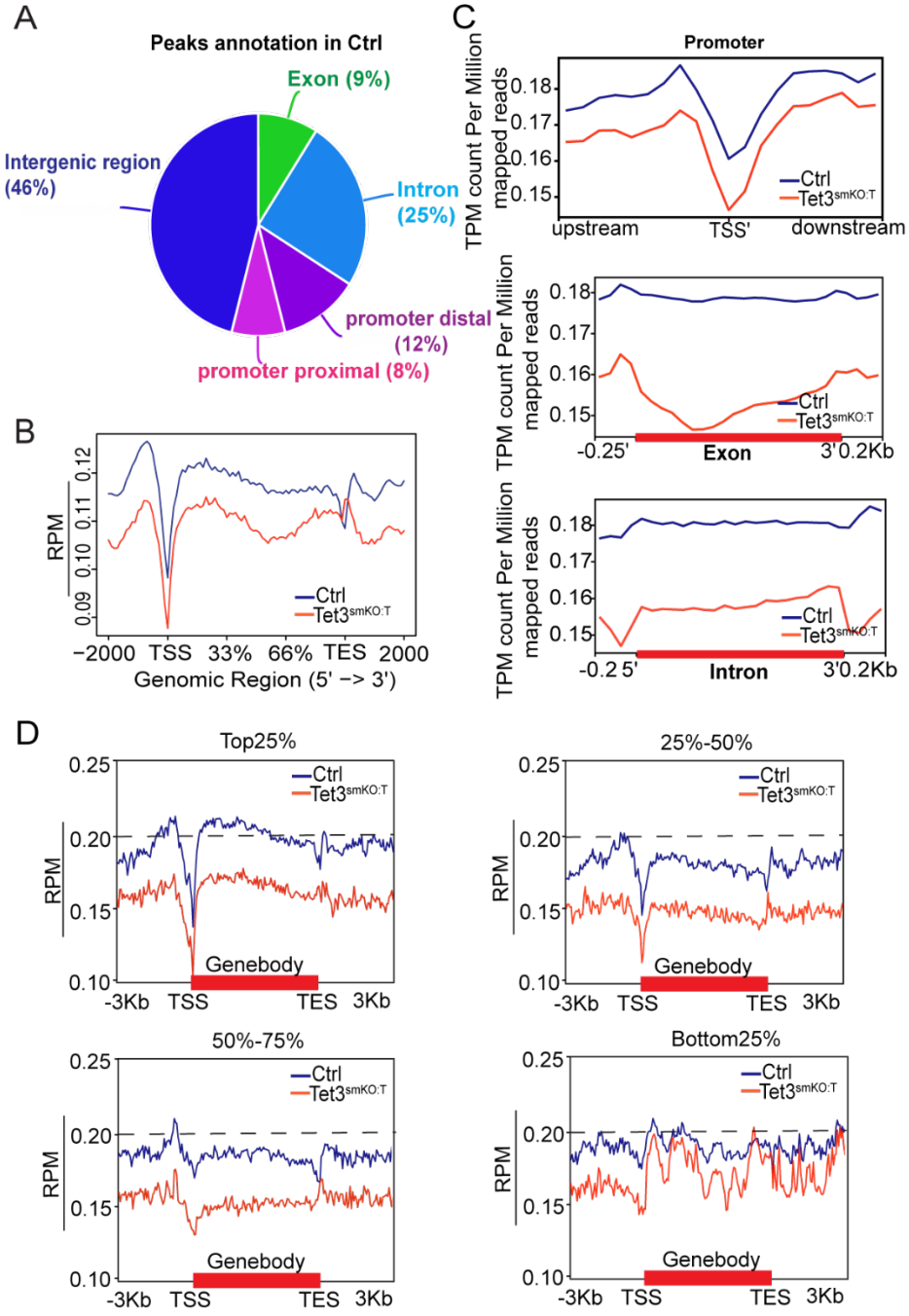


Figure 6. 5hmC is enriched in intragenic region of highly expressed genes in lung SMCs.
(A) Pie-chart of hMeDIP-seq data showing genomic distribution of 5hmC peaks in control lung SMCs (n=2). **(B&C)** Distribution of 5hmC within gene bodies (B) or promoter, exon and intron (C) in sorted lung SMCs prepared from control (Ctrl) and *Tet3*$^{smKO:T}$ mice 8 weeks after tamoxifen injection (n=2). **(D)** Integrative analysis of hMeDIP-seq and RNA-seq datasets of sorted lung SMCs prepared from control (Ctrl) and *Tet3*$^{smKO:T}$ mice 8 weeks after tamoxifen injection. Genes were divided into quartiles based on expression levels measured by RNA-seq with control lung SMCs and the relevant distributions of 5hmC were shown (n=2). Control mice used in this figure: *Tet3*$^{+/+}$α-SMA-creERT2^{pos}tdTomato^{pos}. (hMeDIP-seq data analysis was assisted by Dr. Mario Looso)

Bioinformatic analysis showed that around 34% of 5hmC peaks were located within gene bodies (9% in exons and 25% in introns, respectively), 46% of peaks were located in intergenic regions and 20% were found in promoters of wild-type SMCs (Figure 6A). The genome-wide 5hmC profile across averaged Refseq genes revealed that 5hmC was highly enriched in proximal 5'-upstream regulatory regions but depleted at transcription initiation sites (TSSs). In addition, 5hmC was detected at high levels in gene bodies (Figure 6B). *Tet3* inactivation led to a dramatic global reduction of 5hmC level with more pronounced drop within exons and introns compared to the vicinity of TSSs (Figure 6C).

Next, I characterized the correlation of 5hmC levels with transcriptional activity. Based on RNA-seq data in control SMCs, genes were sub-grouped into quartiles according to their expression levels. Impressively, 5hmC was strongly enriched within gene bodies of highly transcribed genes, but was reduced in genes with low transcriptional activity (Figure 6D). Intragenic 5hmC in actively transcribed genes was strongly erased after *Tet3* inactivation. In contrast, repressed genes (RPKM 0-1) exhibited a similar genome-wide 5hmC abundance in both control and *Tet3*-deficient SMCs (Figure 6D). These findings indicate that TET3 mediated accumulation of 5hmC within gene bodies is positively correlated with transcriptional activity.

7. TET3 mediated 5hmC is required to prevent intragenic entry of RNA Pol II pSer5 in highly transcribed genes

7.1 TET3 interacts with the RNA Pol II elongation machinery

To analyze whether TET3 associates with transcription elongation machinery and thereby deposites intragenic 5hmC within actively transcribed genes, I preformed co-immunoprecipitation (co-IP) with HEK293T cells overexpressing HA-tagged TET3 ($TET3^{OE}$). I discovered that TET3 co-precipitated with elongating Pol II modified by phosphorylation of serine 2 on its carboxy terminal domain (Pol II pSer2) (Figure 7.1 A). Proximity ligation assay (PLA) further confirmed the co-localization of TET3 and Pol II (Figure 7.1 B). Interestingly, I also detected interaction of TET3 with SETD2, a crucial methyltransferase for H3K36me3 (Figure 7.1 A). Transcriptional elongation is coupled with H3K36me3 deposition which is vital to prevent aberrant intragenic transcription initiation through facilitating local nucleosome deacetylation or DNA methylation (Wagner and Carpenter, 2012). Since SETD2 can bind to the RNA Pol II elongation complex by directly interacting with phosphorylated CTD at Ser2 (Hsin and Manley, 2012), I wondered whether TET3 or TET3 mediated 5hmC has an impact on the interaction between RNA Pol II and SETD2. Co-IP experiments in HEK293T cells demonstrated that overexpression of *Tet3* enhanced binding of SETD2 with Pol II pSer2, implying that TET3 or TET3 mediated 5hmC

formation might improve the recruitment of SETD2 to the RNA Pol II-containing elongating complex (Figure 7.1 C). Indeed, PLA demonstrated co-localization of 5hmC with SETD2 but not with H3K36 dimethyltransferase NSD3, suggesting that TET3 mediated 5hmC might be directly involved in the recruitment of SETD2 to the RNA Pol II-containing elongating complex (Figure 7.1 B). To evaluate the impact of *Tet3* inactivation on H3K36me3 deposition, I performed ChIP-qPCR and monitored H3K36me3 levels within a set of highly transcribed contractile genes or synthetic genes that are normally transcribed at very low levels in control or *Tet3*^{smKO} SMCs. *Tet3* inactivation significantly reduced H3K36me3 level within intragenic region of contractile genes such as *Acta2*, *Tagln* and *Cnn1* but not within promoters (Figure 7.1 D). In contrast, H3K36me3 levels within both promoters and gene bodies of synthetic genes like *Vim*, *Tpm4* and *Myh10* were not changed after *Tet3* inactivation (Figure 7.1 D). Thus, my findings indicate that intragenic H3K36me3 deposition is largely dependent on TET3 and/or 5hmC by stabilizing recruitment of SETD2 to the elongating RNA Pol II machinery.

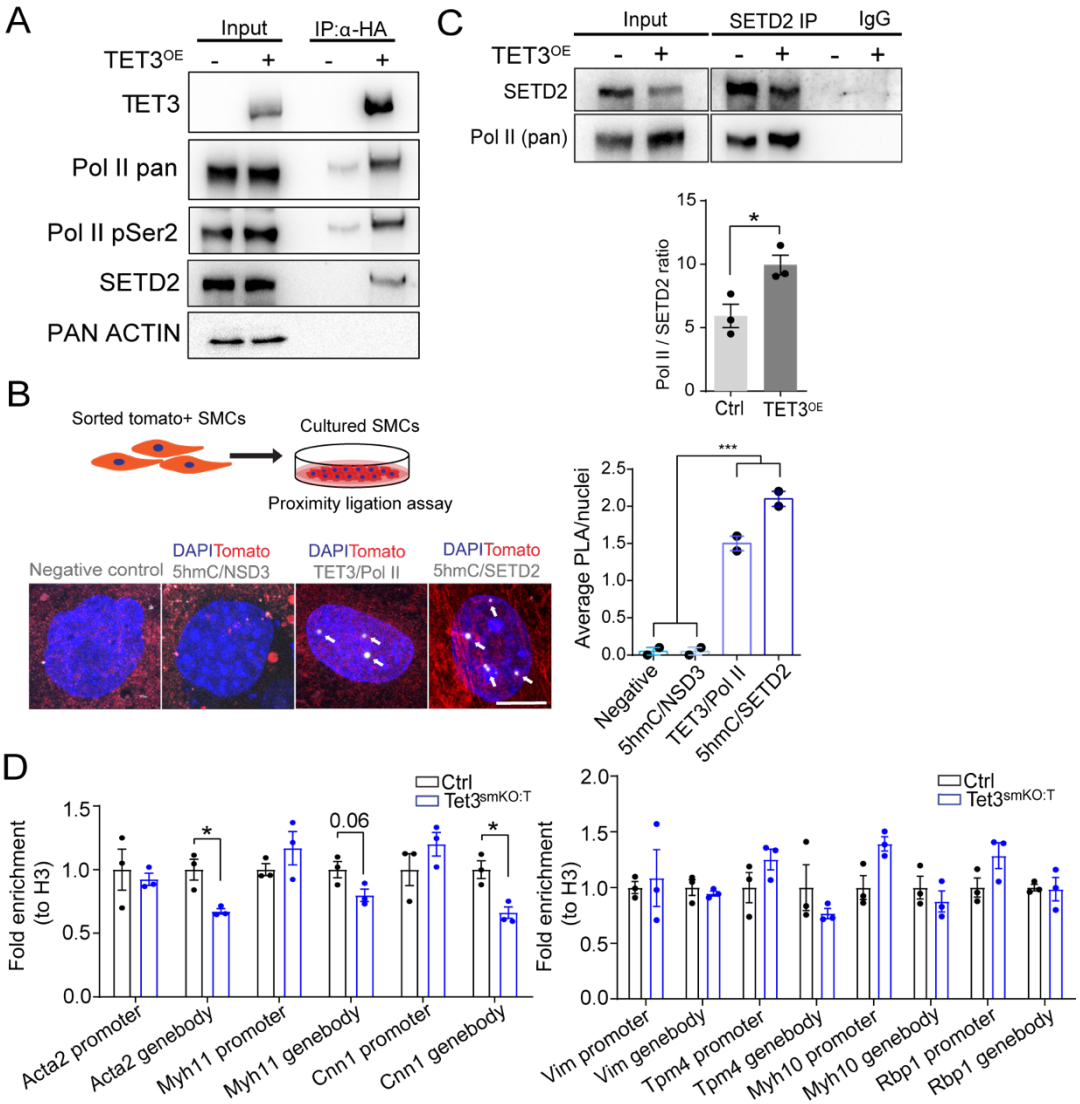


Figure 7.1 TET3 interacts with RNA Pol II pSer2 and SETD2 and plays an important role for intragenic H3K36me3 deposition within highly transcribed genes.
(A) Co-Immunoprecipitation (Co-IP) with HEK293T cells mock-transfected (Ctrl) or overexpressing HA-tagged TET3 (TET3^{OE}). TET3 was precipitated with HA antibody and co-precipitated proteins were monitored by western blots with indicated antibodies (n=3). **(B)** Left panel: representative images of PLA in FACS-sorted SMCs (tdTomato+) with antibodies against 5hmC (negative control), 5hmC & NSD3, TET3 & RNA Pol II, and 5hmC & SETD2. Arrows indicate PLA positive signals. Right panel: quantification of PLA signals in each nuclei. Data are represented as mean ± SEM. (100 cells each sample have been quantified, n=3), One-way ANOVA: ***p<0.001. **(C**) Upper panel: Co-IP with HEK293T cells mock-transfected (Ctrl) or overexpressing HA-tagged TET3 (TET3^{OE}). SETD2 was precipitated with anti-SETD2 antibody and co-precipitated RNA Pol II was monitored by western blots with indicated antibodies. The quantification of co-precipitated RNA Pol II in mock transfected and TET3^{OE} HEK293T cells using image lab software from ChemiDoc gel imaging system was shown in the right panel. n=3, Data are represented as mean ± SEM. Unpaired t-test: *p<0.05. **(D)** Chromatin immunoprecipitation to monitor H3K36me3 enrichment within promoters and exons of *Acta2, Myh11, Cnn1, Vim, Tpm4, Myh10* genes in lung SMCs isolated from control (Ctrl) and *Tet3*$^{smKO:T}$ 8 weeks after tamoxifen injection (n=3). ChIP with H3 antibody was used as reference for normalization. Data are represented as mean ± SEM. Unpaired t-test: *p<0.05. Control mice used in D: *Tet3*$^{+/+}$α-SMA-creERT2^{pos}tdTomato^{pos}

7.2 Loss of *Tet3* promotes ectopic intragenic entry of RNA Pol II pSer5 into highly transcribed genes of SMCs

Next, I performed ChIP-seq after blockage of transcription elongation with 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole (DRB) to determine the binding profiles of Pol II pSer5, a specific form of Pol II required for transcriptional initiation. Binding of Pol II pSer5 within gene bodies was significantly increased in *Tet3*-deficient SMCs (Figure 7.2 A). To further investigate the correlation of Pol II pSer5 binding, intragenic 5hmC levels and transcriptional activity, genes were sub-grouped into quartiles based on Pol II pSer5 ChIP-seq data ($\log_2$*Tet3*$^{smKO:T}$/control), followed by analysis of 5hmC levels and transcriptional activities. I observed that in Group a, a subset of genes with substantial increase of intragenic Pol II pSer5 binding after *Tet3* inactivation, both transcriptional activity and intragenic 5hmC accumulation were at the highest levels (Figure 7.2 B-E). Pol II pSer5 binding in intragenic regions of *Tet3*-deficient SMCs was also weakly but significantly increased in Group b. However, the transcription activities and intragenic 5hmC accumulations decreased together with reduced intragenic entry of Pol II in groups b to d (Fig. 7.2 B-E). Hence, these results indicate that intragenic accumulation of *Tet3*-mediated 5hmC in highly transcribed genes is essential to prevent spurious entry of Pol II in adult SMCs.

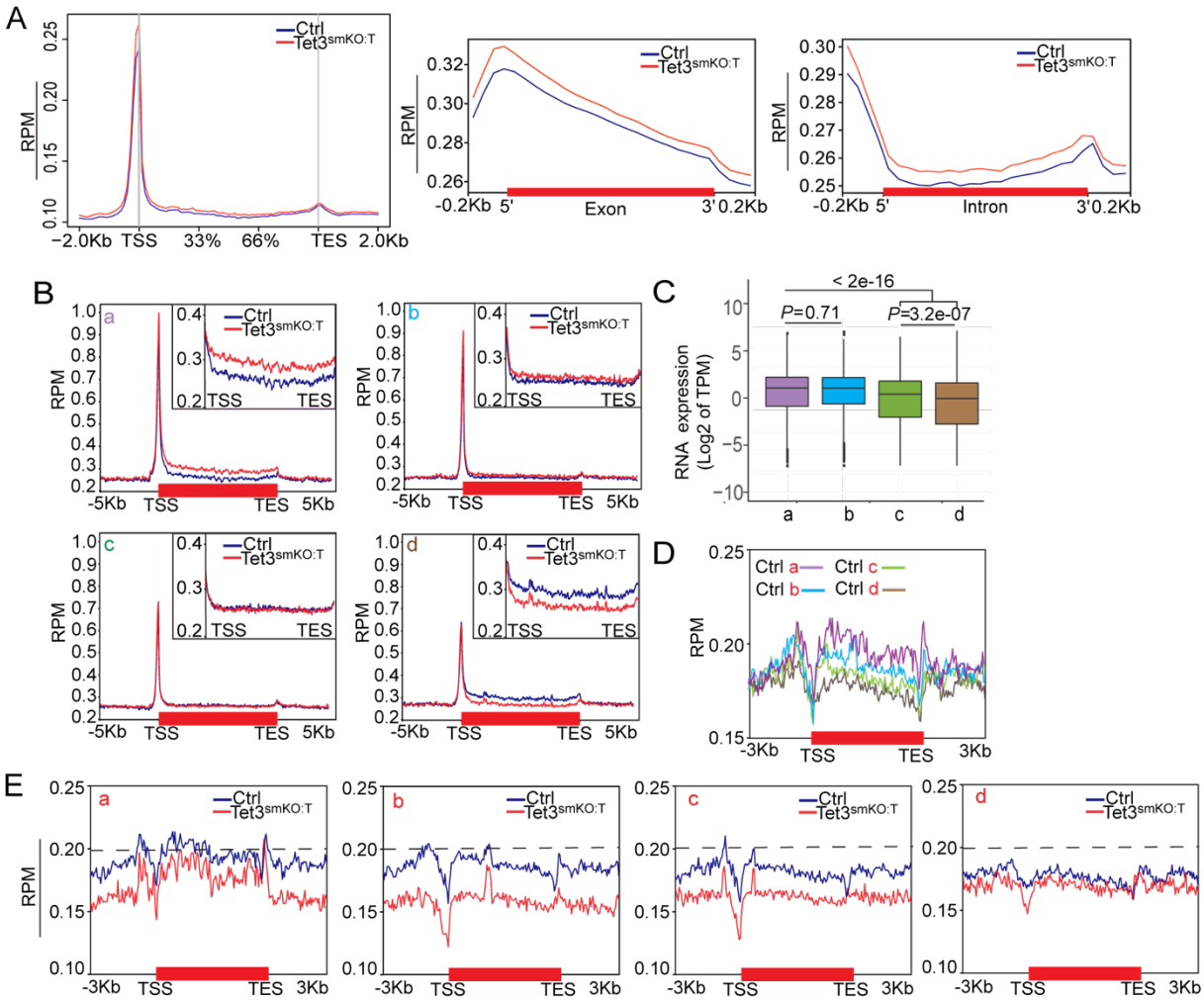


Figure 7.2 *Tet3* inactivation increases intragenic entry of RNA Pol II into actively transcribed genes with high levels of intragenic 5hmC accumulation.

(A) Plot of the RNA Pol II pSer5 distribution within gene body, exon and intron in sorted lung SMCs from control and *Tet3*$^{smKO:T}$ mice after DRB treatment (n=3). **(B-D)** Integrative analysis of Pol II pSer5 ChIP-seq, RNA-seq and hMeDIP-seq datasets showing the correlation of *Tet3* deficient specific intragenic Pol II pSer5 entry **(B)** with transcription activity **(C)** and 5hmC accumulation **(D)** in genes grouped into four quartiles according to differential intragenic Pol II pSer5 binding signals. **(E)** Enrichment and distribution of 5hmC in subgroups of genes indicated in **(D)** in control and *Tet3*-deficient lung SMCs isolated from control (Ctrl) and *Tet3*$^{smKO:T}$ 8 weeks after tamoxifen injection (n=2). *P*-values shown in C were determined by Kruskal-Wallis test. Control mice used in this figure: *Tet3*$^{+/+}$α-SMA-creERT2^{pos}tdTomato^{pos}. (Pol II pSer5 ChIP-seq data analysis was assisted by Dr. Mario Looso)

8. *Tet3* inactivation promotes aberrant intragenic initiation of transcription in highly expressed genes

8.1 *Tet3* inactivation provokes aberrant intragenic transcription of highly expressed genes

To identify genes with TET3-dependent intragenic spurious transcription, I firstly analyzed RNA-seq data by calculating the ratio between the RPKM (reads per kilobase per million mapped reads) of all intermediate exons from second exons onwards and the first exon. To exclude the impact of alternative transcription, the occurrence of intragenic transcription was determined by the fold

change of log_2 ratio between control and *Tet3*-deficient SMCs (Figure 8.1 A & B). Of all genes containing more than 4 exons, 7761 genes had a log_2 ratio of all intermediate exons from second exon onwards versus first exons >1 in *Tet3*-deficient SMCs. To detect bona fide cryptic transcription initiation events in *Tet3*-deficient SMCs, I performed Cap analysis gene expression sequencing (CAGE-seq), a method detecting transcription start sites (TSSs) at single-base resolution (Shiraki et al., 2003) (Kodzius et al., 2006). Consistently, the *Tet3*-deficiant SMC specific binding of Pol II pSer5 within the intragenic regions occurred in the actively transcribed genes and was positively correlated with transcriptional activity reflected by the CAGE signals at canonical TSSs (Figure 8.1 C). Moreover, the number of intragenic CTSS (defined as TSS with CAGE-tag >8, the average value of each single-base TSS on annotated TSSs) in *Tet3*-deficient SMCs was significantly increased (Figure 8.1 D). Among 2114 genes containing single-base intragenic CTSS specific to *Tet3*-deficient SMC and producing spurious transcripts from the sense strand, 515 genes (26%, thereafter referred as spurious genes) exhibited increased ratios of RNA-seq reads between downstream and first exons as well as enhanced Pol II intragenic entry (Figure 8.1 E & F). Importantly these genes showed significantly higher transcription activity than non-spurious genes (Figure 8.1 G).

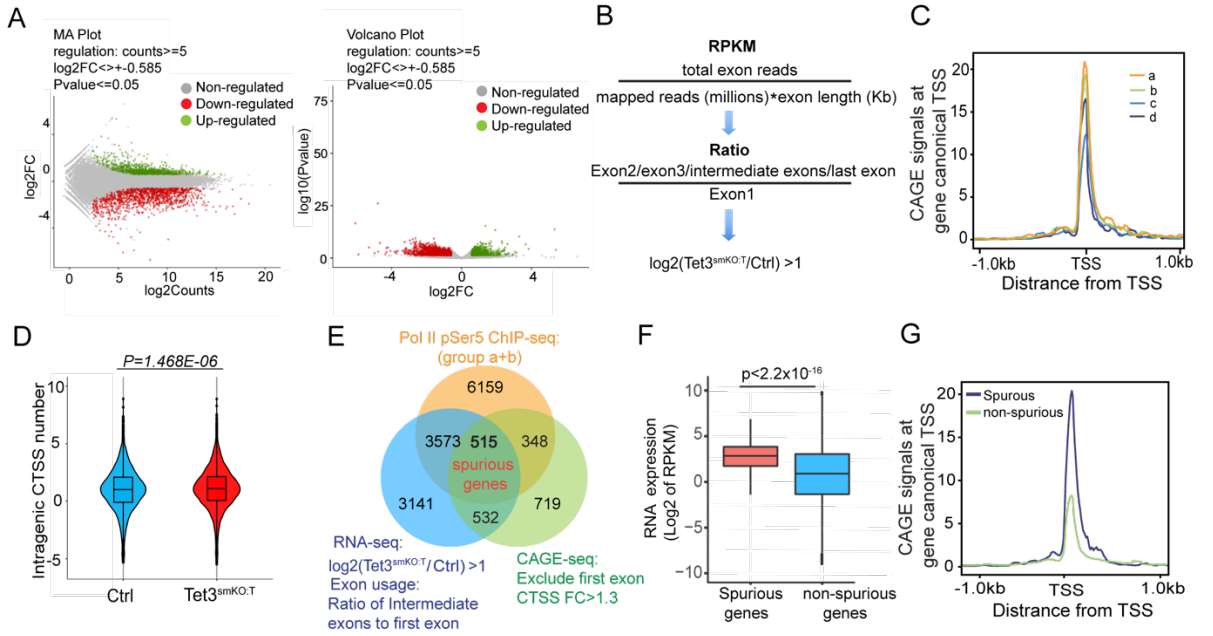


Figure 8.1 *Tet3* inactivation provokes aberrant intragenic transcription in highly expressed genes.
(A) MA and Volcano plots showing the deregulated genes identified by RNA-seq with sorted lung SMCs from control and *Tet3*$^{smKO:T}$ lungs 8 weeks after tamoxifen injection (n=3). **(B)** Formulas for calculating the ratio of normalized RNA-seq RPKM of downstream exon versus first exon between control and *Tet3*$^{smKO:T}$ SMCs. **(C)** Distribution of CAGE-signal at canonical TSSs in four quantiles (a-d) based on differential intragenic binding signals of Pol II pSer5 as in Figure 7.2B. **(D)** Violin plot showing the CTSS number at gene body (without exon1) in sorted lung SMCs from control and *Tet3*$^{smKO:T}$ lungs. n=2, *P* value was calculated by likelihood-ratio test. **(E)** Venn Diagram of spurious genes based on integrative analysis of Pol II pSer5 ChIP-seq, RNA-seq and CAGE-seq datasets. **(F&G)** Transcription activities of spurious and non-spurious genes based on RNA-seq (**F**) or CAGE-seq

signals at canonical TSSs (**G**). The data analysis of RNA-seq and CAGE-seq was performed with help from Hang Liu and Dr. Stefan Grünther.

8.2 *Tet3* inactivation induces aberrant intragenic transcriptional initiation within pro-contractile genes

To elucidate the biological functions of spurious genes, I performed GO term enrichment assays.

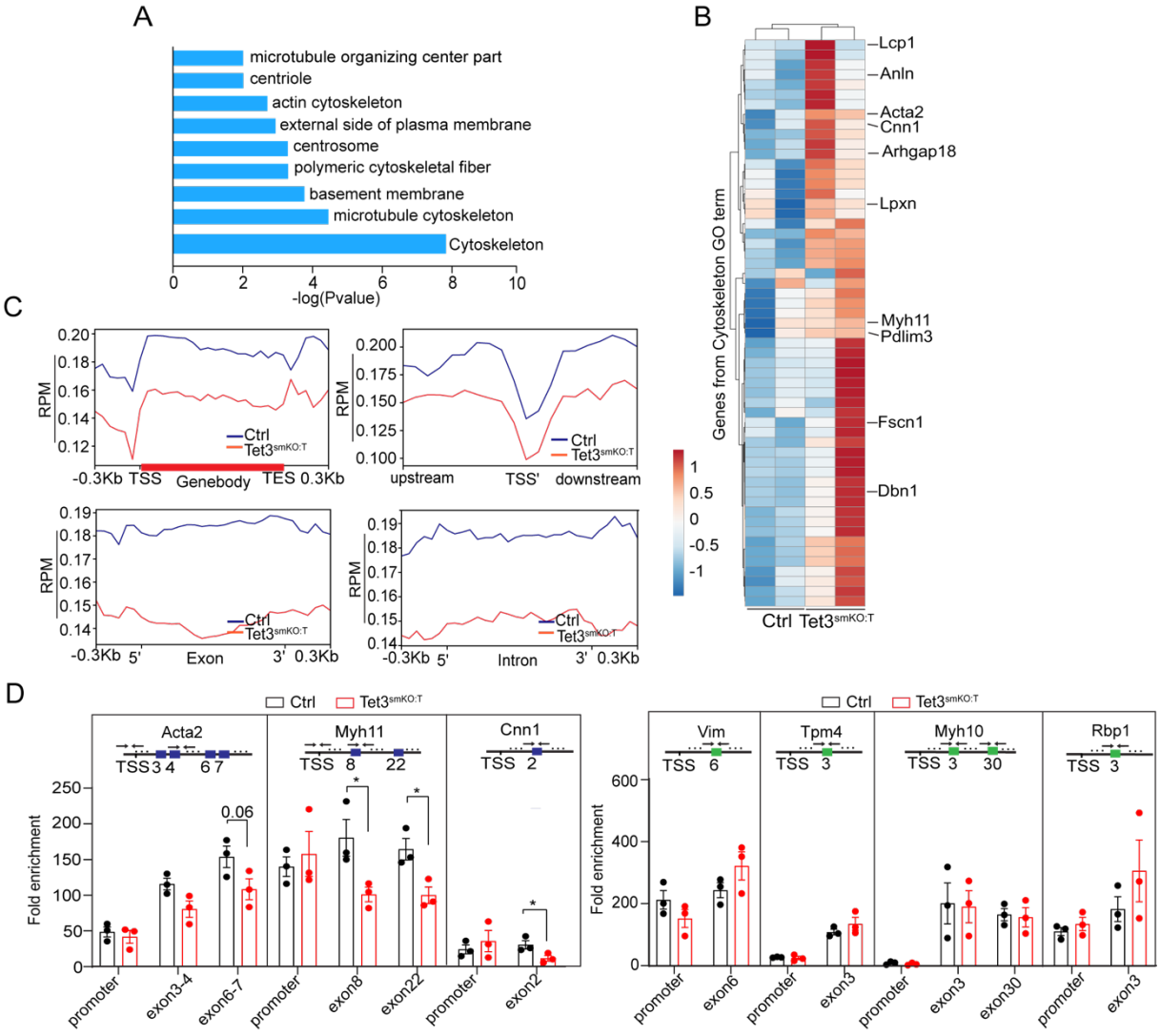


Figure 8.2 *Tet3* inactivation promotes aberrant intragenic transcription within pro-contractile genes.
(A) Gene Ontology enrichment analysis of spurious genes in Figure 8.1 (E). **(B)** Heat map of genes involved in regulation of cytoskeleton based on absolute reads of RNA-seq with sorted lung SMCs ($p<0.05$; n=2). Genes related to contractile actin filament bundle and actomyosin structure organization were indicated. **(C)** Distribution profiles of 5hmC within gene body, promoter, exon and intron of spurious genes (indicated in Fig. 8.1 C) in SMCs isolated from control and *Tet3*$^{smKO:T}$ lungs (n=2). **(D)** Quantitative PCR to monitor 5hmC levels at promoter and exon of indicated gene in lung SMCs isolated from control and *Tet3*$^{smKO:T}$ mice 8 weeks after tamoxifen injection. n=3, unpaired t-test: *$p<0.05$. The graphic data in **D** are represented as mean ± SEM. Control mice used in **B-D**: *Tet3*$^{+/+}$α-SMA-creERT2^{pos} tdTomato^{pos}. hMeDIP-seq data analysis was performed with help from Dr. Mario Looso.

Interestingly, genes with spurious transcription (e.g *Acta2*, *Cnn1*, *Myh11*, *Synpo2*) are mainly involved in the regulation of smooth muscle contraction, e.g. contractile actin filament bundle and actomyosin structure organization (Figure 8.2 A & B). In addition, 5hmC accumulation was more constant and pronounced the gene bodies of spurious genes compared to the genome-wide 5hmC profile, and no enrichment at proximal 5'-upstream regulatory regions was observed, indicating predominant role of TET3 in regulating transcription elongation rather than initiation in this subset of genes (Figure 8.2 C). hMeDIP-qPCR validated this assumption by demonstrating that *Tet3* depletion resulted in a marked reduction of 5hmC level specifically at intragenic regions but not in promoters of contractile genes like *Cnn1*, *Myh11* and *Acta2* (spurious genes), while 5hmC levels in both promoters and gene bodies of synthetic genes such as *Vim*, *Myh10*, *Tpm4* and *Rbp1* (non-spurious) were not affected (Figure 8.2 D).

8.3 Spurious transcription from intragenic cryptic TSSs reduces generation of functional full-length mRNA of pro-contractile genes

To characterize the intragenic cryptic TSSs in spurious genes, the DNA regions of 50 base pairs centered at the intragenic CTSSs specific to *Tet3*-deficient SMCs were analyzed.

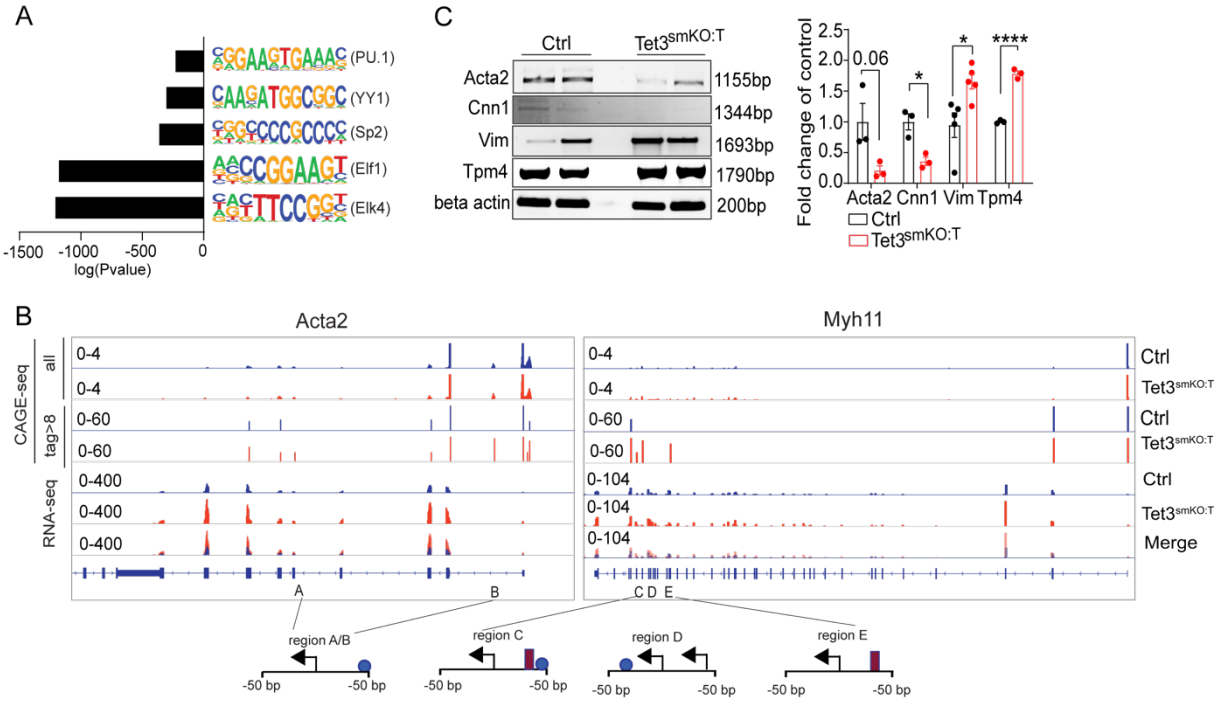


Figure 8.3 *Tet3* inactivation reduces generation of functional full-length mRNA of pro-contractile genes due to aberrant intragenic transcription initiation.

(A) Histogram plot showing the enrichment of CpG-containing motifs in a region of 50 nucleotides centered at intragenic TSSs (CAGE-tag >8) specific to *Tet3*-deficient SMCs. **(B)** Integrated Genome Viewer (IGV) tracks showing first single nucleotide of CAGE-seq capture sequences and RNA-seq peaks within *Acta2* and *Myh11* gene in control or *Tet3*-deficient lung SMCs (n=2). Schematic representation of putative transcription factor binding elements within cryptic TSS regions are shown in the bottom. **(C)** Semi-quantitative RT-PCR analysis to detect

full-length mRNA of *Acta2, Cnn1, Vim, Tpm4* genes in sorted SMCs from control and *Tet3*$^{smKO:T}$ lungs. The *β-actin* was used as reference gene for normalization. Quantification of expression levels was performed by image J software and shown in the right panel (n=3). Unpaired t-test: *p<0.05; ****p<0.0001. The graphic data in **C** are represented as mean ± SEM. Control mice used in **B & C**: *Tet3*$^{+/+}$α-SMA-creERT2^{pos}tdTomato^{pos}. CAGE-seq data analysis was performed with help from Hang Liu.

I found a significant enrichment of the CpG dinucleotide and several transcription factor binding motifs containing CpG sequences in the vicinity of CTSSs, such as binding motifs for Sp2 and members of the Ets family (Figure 8.3 A). The activation of intragenic binding motifs specifically in *Tet3*-deficient SMCs within contractile genes such as *Acta2* and *Myh11* was further verified by CAGE-seq analysis (Figure 8.3 B). The spurious transcription subsequently resulted in enhanced RNA-seq reads of contractile SMCs genes e.g *Acta2*, *Cnn1* after *Tet3* ablation, which were unable to produce functional full-length mRNA. This phenomenon was not observed for synthetic genes like *Vim* and *Tpm4* (Figure 8.2 B and 8.3 C). Altogether, these data support the conclusion that TET3 mediated 5hmC is indispensable for the definition of functional transcription units by suppressing aberrant intragenic transcription initiation in lung SMCs.

9. Spurious transcripts in *Tet3*-deficient SMCs induce innate immune responses by activating the endosomal TLR signaling pathway

9.1 Genes related to innate immune response are upregulated in *Tet3*-deficient SMCs

To investigate biological consequences induced by spurious transcription, I performed GO term enrichment analysis of significantly upregulated genes in *Tet3*-deficient SMCs. Impressively, all top 10 upregulated GO term pathways are associated with inflammation responses including innate immune response signaling pathways (Figure 9.1 A). The top 10 downregulated GO term pathways related to fatty acid or lipid metabolism, implying that the phenotypic and functional changes of *Tet3*-deficient SMCs are accompanied by a switch in the metabolic profile (Figure 9.1 B). Of note, a set of macrophage specific genes such as *Cd68*, *Adgre1* and *Lgals3* were significantly upregulated, suggesting that *Tet3*-deficient SMCs might gain macrophage-like functions similar to SMCs during atherogenesis (Feil et al., 2014) (Figure 9.1 C). The Toll-like receptor (TLR) signaling pathway, which is highly active to direct production of cytokine or chemokine in macrophages, plays a pivotal role in innate immune responses. KEGG pathway analysis of significantly upregulated genes revealed that the Toll-like receptor (TLR) signaling pathway was among the top 10 upregulated KEGG pathways in *Tet3*-deficient SMCs (Figure 9.1 D). Four of the 13 mammalian TLRs, i.e. TLRs 3, 7, 8 and 9, localize within the intracellular endosome and can recognize self and foreign double-stranded RNA, single-stranded RNA and DNA to activate downstream signaling. Intriguingly, transcriptome analysis and RT-qPCR revealed that expression of *Tlr7/8*, *Myd88* and

their downstream targets such as *Ccl5* and *Il1b* were significantly upregulated after *Tet3* inactivation (Figure 9.1 E & F). These results indicate that activation of TLR7/8, which specifically recognize single-stranded RNA, might contribute to the massive inflammation observed in *Tet3*-deficient lungs.

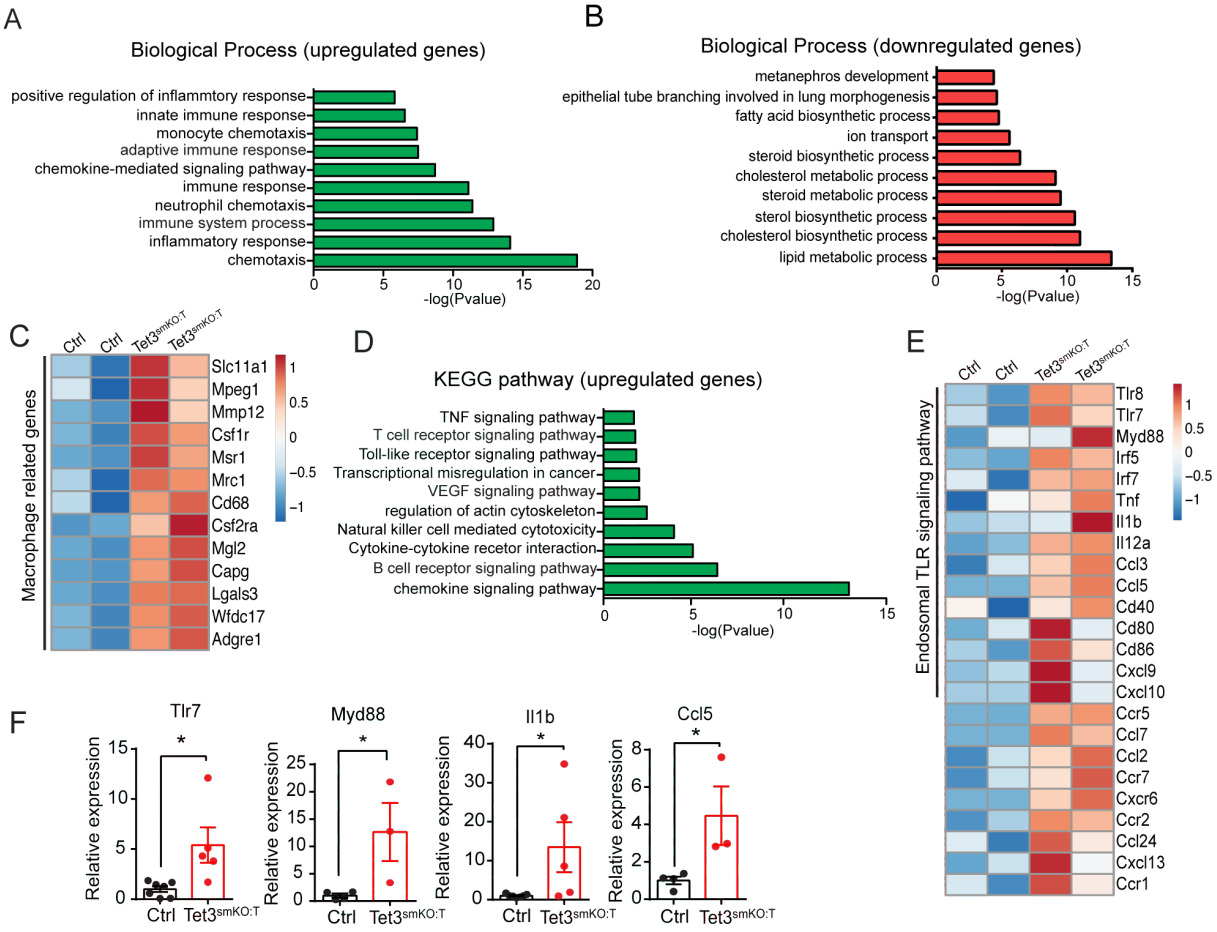


Figure 9.1 Loss of *Tet3* activates the endosomal TLR7/8 signaling pathway.
(A&B) Biological process analysis of significantly upregulated (A) or downregulated (B) genes in sorted lung SMCs isolated from *Tet3*$^{smKO:T}$ mice (n=3). **(C)** Heat map showing the macrophage related genes based on absolute reads of RNA-seq with sorted lung SMCs isolated from control (Ctrl) and *Tet3*$^{smKO:T}$ mice. **(D)** KEGG pathway analysis of significantly upregulated genes in sorted lung SMCs isolated from *Tet3*$^{smKO:T}$ mice (n=3). **(E)** Heat map of differentially regulated genes involved in TLR signaling pathway based on absolute reads of RNA-seq with sorted lung SMCs isolated from control (Ctrl) and *Tet3*$^{smKO:T}$ mice. **(F)** RT-qPCR analysis to monitor expression levels of *Tlr7, Myd88, Il1b, Ccl5* genes in sorted lung SMCs from control (Ctrl) and *Tet3*$^{smKO:T}$ mice 8 weeks after Tamoxifen injection (n=3). Data are represented as mean ± SEM. Unpaired t-test: *p<0.05; Control mice used in this figure: *Tet3*$^{+/+}$α-SMA-creERT2^{pos}tdTomato^{pos}.

9.2 Spurious transcripts in *Tet3*-deficient SMCs induce activation of TLR7-MYD88 signaling cascade

Western blot analysis indicated that in addition to the enhanced expression of *Tlr7* and *Myd88*, protein levels of EEA1 and RAB7 were strongly increased in *Tet3*-deficient lung SMCs. EEA1 and

RAB7 are involved in regulating endosome trafficking and co-localize with TLR7 in stimulated macrophages (Luo et al., 2017), (Figure 9.2 A). Since the recruitment of specific adaptor molecules MYD88 (myeloid differentiation factor-88) by TLR7/8 upon engagement of ligands is crucial to propagate downstream signaling, I asked whether the association of TLR7 and MYD88 is increased after deletion of *Tet3*. In situ proximity ligation assays (PLA) with TLR7 and MYD88 antibodies in control and *Tet3* mutant lung sections indicated substantially increased PLA signals in *Tet3*-deficient compared to controls SMCs (Figure 9.2 B). These results suggest that aberrant single stranded RNA transcripts in *Tet3*-deficient SMCs might be recognized by endosomal TLR7/8 and stimulate signal transduction via the TLR7/8-MYD88 axis.

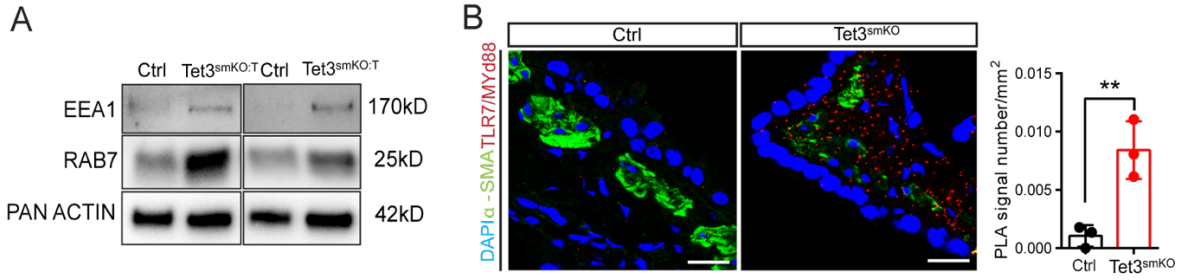


Figure 9.2 *Tet3* inactivation promotes the interaction between TLR7 and MYD88.
(A) Western blots to monitor protein levels of EEA1 (early endosomal marker) and RAB7 (late endosomal marker) in sorted lung SMCs isolated from control (Ctrl) and *Tet3*$^{smKO:T}$ mice 8 weeks after tamoxifen injection. Protein level of pan-actin was used as loading control (n=3). **(B)** In situ proximity ligation assay (PLA) to monitor interaction of TLR7 and MYD88 with cryosections prepared from control (Ctrl) and *Tet3*^{smKO} lungs 8 weeks after tamoxifen injection (n=3). The PLA signals in airway smooth muscle layers stained by α-SMA antibody in control (Ctrl) and *Tet3*^{smKO} lungs were quantified and shown in the right panel (n=3). Data are represented as mean ± SEM. Unpaired t-test: **p<0.01. The nuclei were counterstained with DAPI. Control mice used in this figure: *Tet3*$^{+/+}$α-SMA-creERT2^{pos}.

9.3 Spurious transcripts from *Tet3*-deficient SMCs display high immunogenicity

To further test whether spurious transcripts in *Tet3*-deficient SMCs stimulate the TLR7/8-MYD88 axis and thereby innate immune responses, I transfected HeLa cells with total RNAs extracted from control and *Tet3*-deficient SMCs and monitored the expression levels genes downstream of endosomal TLR7/8 (Figure 9.3 A). Since the RT-qPCR assays were done with human-specific primers, the possibility was excluded that mouse specific mRNA molecules used for transfection were detected. The expression of a set of endosomal *Tlr7/8* downstream genes such as *Il1b, Cd86, Ccl5, Cd80, Irf7, Cxcl9, Ifnb* were significantly higher in cells transfected with RNAs extracted from *Tet3*-deficient SMCs compared to cells transfected with RNAs extracted from control SMCs, suggesting a strong induction of *Tlr7/8* signaling by spurious RNAs (Figure 9.3 B). Expression of *Tlr7/Myd88* and the non-endosomal TLR downstream target *Ccr5* were not altered after transfection with RNA extracted from *Tet3*-deficient SMCs, suggesting that activation of TLR7/8

signaling rather than induction of *Tlr7/8* expression is decisive for increased expression of TLR7/8 targets (Figure 9.3 B). These findings indicate that spurious transcripts in *Tet3*-deficient SMCs provoke excessive production of cytokine/chemokine via activation of endosomal TLR7/8 signaling.

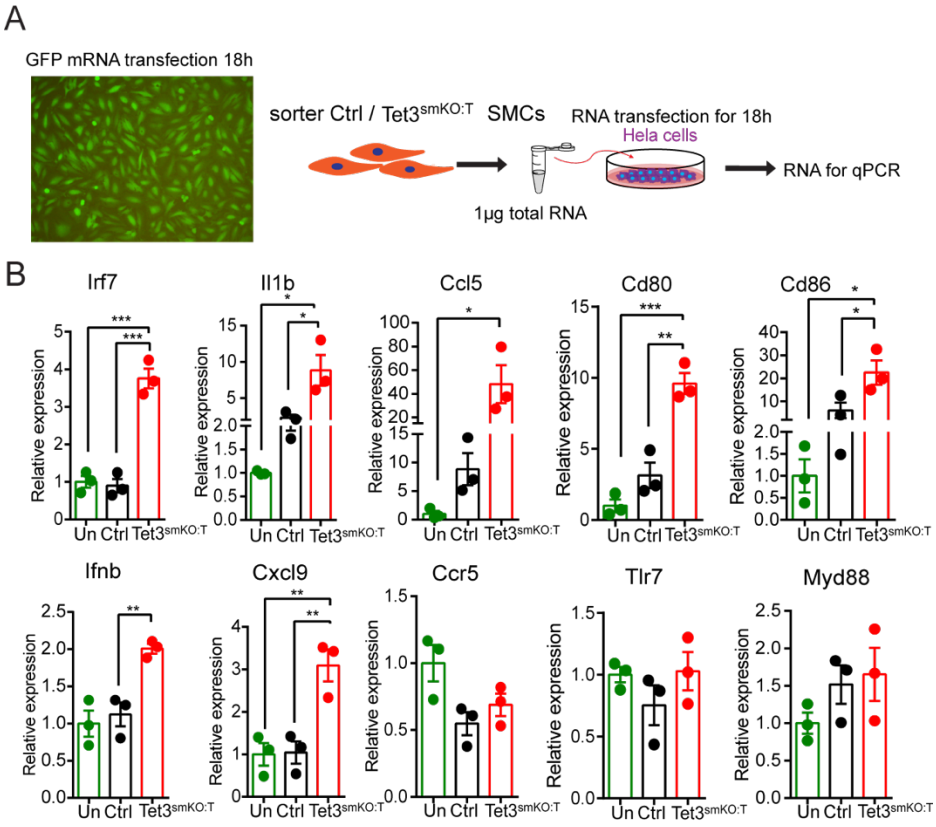


Figure 9.3 Spurious transcripts from *Tet3*-deficient lung SMCs activate endosomal TLR7 signaling pathway in HeLa cells.

(A) Experimental strategy for RNA transfection into HeLa cells. The transfection efficiency was monitored using commercial GFP mRNA (0.3µg/200,000 cells) and representative image of GFP^{+} cells after GFP mRNA transfection is shown in the left panel. **(B)** RT-qPCR expression analysis of *Irf7, Il1b, Ccl5, Cd80, Cd86, Ifnb, Cxcl9, Ccr5, Tlr7* and *Myd88* genes in HeLa cells mock-transfected (Un) or transfected with total RNAs isolated from lung SMCs of control (Ctrl) and *Tet3*$^{smKO:T}$ mice 8 weeks after tamoxifen injection (n=3). Data are represented as mean ± SEM. One-way ANOVA: *p<0.05, **p<0.01, ***p<0.001. Control mice used in B: *Tet3*$^{+/+}$α-SMA-creERT2^{pos}tdTomato^{pos}.

9.4 Immunostimulatory effect of spurious transcripts from *Tet3*-deficient SMCs is TLR7/8-dependent

To further verify that innate immune responses trigged by spurious transcripts in *Tet3*-deficeint SMCs are TLR7/8-dependent, I performed RNA transfection experiments in HEK293 cells, in which TLR7/8 is not expressed (Figure 9.4 A). In contrast to the strong effects in HeLa cells, transfection of RNAs from *Tet3*-deficient SMCs did not stimulate expression of endosomal TLR7/8 downstream cytokines or chemokines in HEK293 cells (Figure 9.4 B & C). Thus, I

conclude that TLR7/8 signaling pathway is indispensable for spurious transcript-induced production and secretion of chemokines and cytokines in *Tet3*-deficient cells.

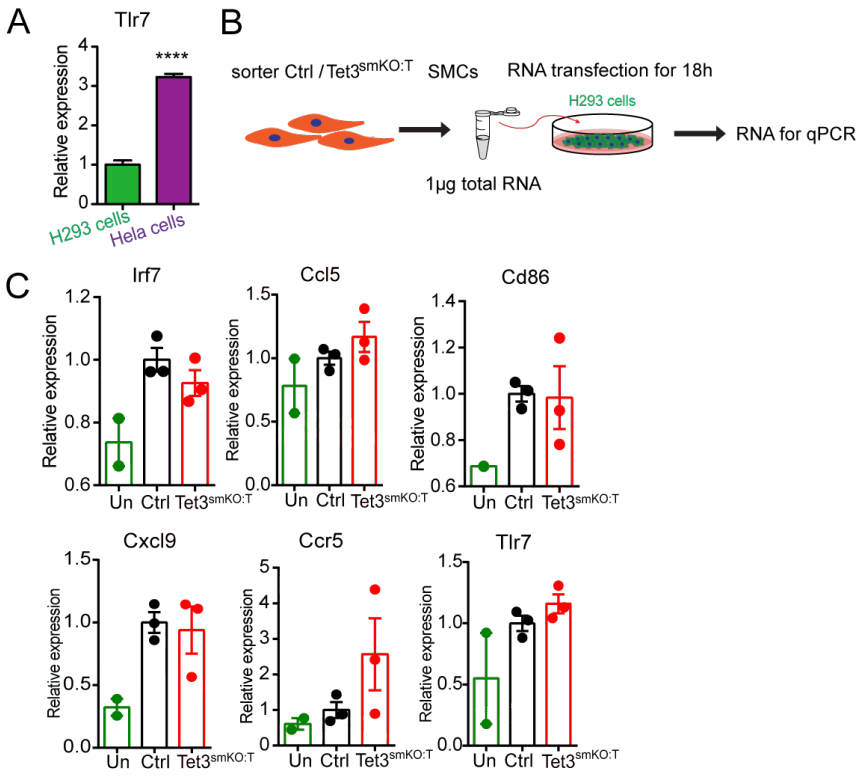


Figure 9.4 Spurious transcripts from *Tet3*-deficient lung SMCs are unable to activate endosomal TLR7/8 signaling pathway in HEK293 cells.

(A) RT-qPCR analysis to monitor expression levels of *Tlr7* in HEK293 and HeLa cells (n=3). The *β-actin* was used as reference gene for normalization. Data are represented as mean ± SEM. Unpaired t-test: ****p<0.0001. **(B)** Experimental strategy for RNA transfection on HEK293 cells. **(C)** RT-qPCR analysis of *Irf7, Ccl5, Cd86, Cxcl9, Ccr5, Tlr7* mRNA in HEK293 cells mock-transfected (Un, n=2) or transfected with total RNAs isolated from lung SMCs of control (Ctrl, n=3) and *Tet3*$^{smKO:T}$ (n=3) mice 8 weeks after tamoxifen injection. Data are represented as mean ± SEM. One-way ANOVA: no statistical significance (p>0.05). Control mice used in C: *Tet3*$^{+/+}$α-SMA-creERT2^{pos}tdTomato^{pos}.

9.5 *Tet3* in activation in SMCs induces Th2 cell based immune responses in the lung and enhances lung fibrosis

Activation of the innate immune system in response to different pathogens will eventually also result in activation of the adaptive immune system and recruitment of T and B lymphocytes. To characterize the inflammatory responses in mutant lungs, I first analyzed cell in the bronchoalveolar lavage (BAL) of control and *Tet3*^{smKO} lungs, a method typically performed to diagnose pathogenic infections of the airways. I did not observe pronounced change in the number of immune cells, but detected an increased number of epithelial cells in the BAL of *Tet3*^{smKO} lung (Figure 9.5 A & B). Consistently, H&E staining revealed shedding of the epithelium in *Tet3*^{smKO} lung, a typical pathological trait in asthmatic airways (Figure 9.5 C). The failure to detect enhanced number of immune cells in BAL fluids from *Tet3* mutant lungs might due to the dilution of the lavage fluid,

since the major sites of inflammation in *Tet3* mutant lung are conducting airway while the lavage fluid was mainly sampled from alveoli. Next, I determined the inflammatory response in *Tet3*^{smKO} lung tissue. Immunofluorescence and FACS analysis revealed that the numbers of interstitial macrophages (IMs), T cells and neutrophils were upregulated in *Tet3* mutant lungs 2 months after tamoxifen injection (Figure 9.5 D & E). Since innate immune system provides key signals for differentiation of naive CD4^{+} T cells into functionally distinct Th subtypes, I performed FACS analysis of cells dissociated from whole lungs to characterize potential effector T cells that might get activated in response to *Tet3* inactivation. The number of Th2 cells but no other subtypes was significantly increased *Tet3* mutant lungs, which coincides with a strong increase of Interleukin 4 (IL4) and Interleukin 13 (IL13) production. IL4 and IL13 are cytokines primarily secreted by Th2 cells and promote club cell metaplasia and mucin secretion (Figure 9.5 F & G, Figure 4.3 A, C, E & F).

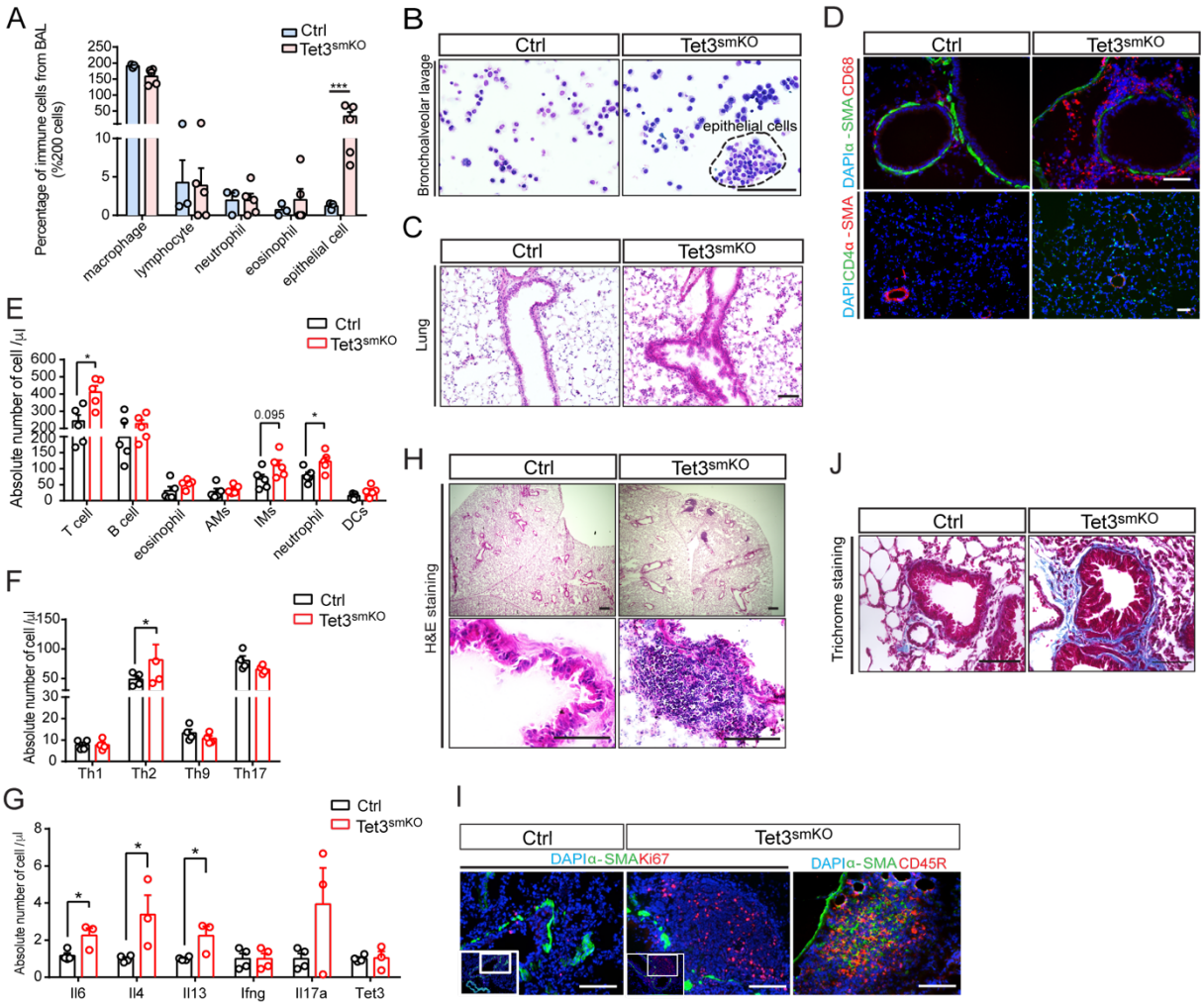


Figure 9.5 Inactivation of *Tet3* in SMCs induces persistent immune responses, epithelial shedding and fibrosis. **(A)** Quantification of macrophage, lymphocyte, neutrophil, eosinophil and epithelial cells in BAL cells in bronchoalveolar lavages (BAL) prepared from control (Ctrl) and *Tet3*^{smKO} mice 8 weeks after tamoxifen injection using Diff-Quick stain kit (n=3). Data are represented as mean ± SEM. Unpaired t-test: ***p<0.001. **(B)** Representative images of BAL cells stained with Diff-Quick stain kit (n=3). Scale bar: 50 µm. **(C)** Representative

images of hematoxylin & eosin (H&E) staining with cryosections from the lungs of control (Ctrl) and *Tet3*^{smKO} mice 8 weeks after Tamoxifen injection) (n=3). Scale bar: 50 µm. **(D)** Representative immunofluorescence images of co-staining for α-SMA & CD68 (upper panel), or CD4 & α-SMA (lower panel) with paraffin sections prepared from the lungs of control and *Tet3*^{smKO} mice 8 weeks after tamoxifen injection (n=3). DNA was stained by DAPI. Scale bar: 50 µm. **(E)** FACS analysis to monitor absolute number of different immune cells in the whole lungs dissected from control (Ctrl) and *Tet3*^{smKO} mice 8 weeks after tamoxifen injection (n=5). Data are represented as mean ± SEM. Unpaired t-test: *p<0.05. **(F)** FACS analysis to monitor absolute number of different T help Cell (Th cell) with the dissociated lung tissues dissected from control (Ctrl) and *Tet3*^{smKO} mice 6 months after tamoxifen injection (n=4). Data are represented as mean ± SEM. Unpaired t-test: *p<0.05. **(G)** RT-qPCR analysis of *Il6, Il13, Il4, Ifng, Il17a and Tet3* expression in sorted CD3^{+} T cells isolated from the lungs of control (Ctrl) and *Tet3*^{smKO} mice 6 months after tamoxifen injection. The *β-actin* gene was used as reference for normalization (n=4). Data are represented as mean ± SEM. Unpaired t-test: *p<0.05. **(H)** Representative images of H&E with cryosections prepared from the lungs of control (Ctrl) and *Tet3*^{smKO} mice 6 months after tamoxifen injection (n=5). Scale bar: 50 µm. **(I)** Representative immunofluorescence images of co-staining for α-SMA and Ki67 or α-SMA and CD45R (mature B cell marker) with cryosections prepared from the lungs of control (Ctrl) and *Tet3*^{smKO} mice 6 months after Tamoxifen injection (n=3). Scale bar: 50 µm. (J) Representative images of Trichrome staining with cryosections prepared from the lungs of control and *Tet3*^{smKO} mice 6 months after Tamoxifen injection (n=5). Scale bar: 50 µm. Control mice used in this figure: *Tet3*$^{+/+}$α-SMA-creERT2^{pos}.

To clarify whether inflammatory responses persistent and/or progress, I conduced histo-pathological analysis with control and *Tet3*^{smKO} lungs 6 months after TAM injection. Interestingly, I observed peri-bronchiolar lesions in mutant lung 6 months after tamoxifen injection that were primarily composed of CD45R^{+} B cells (Figure 9.5 H & I). In addition, Trichrome staining revealed enhanced fibrosis in *Tet3*^{smKO} lung tissue 6 months after tamoxifen injection, indicating an exacerbated inflammatory response due to persistent loss of *Tet3* in mutant lung (Figure 9.5 J). Taken together, the findings indicate that *Tet3* inactivation in SMCs provokes excessive Th2 cells-mediated immune responses that promotes pathological airway remodeling.

Discussion

Dysregulation of SMCs plays central role in most lung parenchymal diseases, contributing to changes in the contractility of airways, inflammatory responses, and immune reactions amongst others. By using a combination of genetic, biochemical and cell biology approaches, I explored the function of TET3 in modulating smooth muscle cells and lung tissue homeostasis. I discovered a pivotal role of TET3 in regulating the fidelity of gene transcription, maintaining the identity of SMC and balancing immune response (Figure 10 A & B). My study suggests that modulation of TET3 activity might represent a novel approach to treat numerous lung parenchymal diseases.

1. *Tet3* deficiency in SMC specifically promotes pathogenic processes in the lung

SMCs are non-striated muscle cells that act as essential components of several hollow organs such as artery, lung, bladder, intestine, and stomach. In this study, I discovered that loss of *Tet3* in SMCs induces pathological airway remodeling and massive inflammatory response in lung, but has negligible impact on tissue homeostasis of the intestine and the aorta. In addition, massive airway remodeling but not vascular remodeling was observed in mutant lungs. This tissue specific effect might be explained by three reasons: 1) distinct effects of SMCs on neighboring cells; 2) different developmental origins of SMCs; 3) distinctive compensatory effects of other TETs in different SMC cell types due to different expression levels or activities (Hirschi and Majesky, 2004).

1-1. Loss of *Tet3* leads to airway remodeling but not pulmonary vascular remodeling.

ASMC and VSMC in the lung exhibited a synthetic phenotype characterized by decreased actin filament and increased rER in cytoplasm after *Tet3* inactivation (Figure 5.1 A & B, D & E), suggesting a crucial role of TET3 in maintaining the contractile phenotype of ASMC and VSMC. In addition, enhanced interactions between TLR7 and MYD88 were detected in both ASMCs and VSMCs after *Tet3* inactivation (Figure 9.2 B, data of VSMC not shown), suggesting that activation of endosomal TLR7 signaling pathway and downstream chemokine and cytokine occurs in both cell types. Signaling by *Tet3*-deficeint SMC enhanced recruitment of peripheral lymphocyte and triggered massive inflammatory responses in lung. Interestingly, *Tet3*-deficiency also led to substantial airway remodeling (Figure 4.2 C), whereas no detectable histological changes of endothelium in detected in the aorta and lung vessels (Figure 4.2 A & lung vessel data not shown). This observation suggests that loss of TET3 in ASMCs and VSMCs influences differentially affect neighboring cells. Further evidence for this hypothesis comes from the comparison of transcriptome profiles of bronchioles and the vasculature. After *Tet3* ablation, very few overlapping

DEGs (differentially regulated genes) were identified in bronchioles and the vasculature, suggesting that different signaling pathways were induced in neighboring cells. This argues for different susceptibilities of cells in the lung and the vasculature to cytokines and chemokines secreted by *Tet3*-deificient SMCs (Figure 5.2 A). For example, excessive production of IL-13 by Th2 cells provokes metaplasia of club cells to goblet cells, which is a hallmark of Th2 cells responses (Knoop and Newberry, 2018), but did not affect endothelial cells. Due to the high heterogeneity of lung epithelial cells and complex cell-cell interaction in the airway in response to various extrinsic or intrinsic cues, it seems likely that ASMCs, club cells and inflammatory cells form a complex network of positive and negative feedback loops. Such a network might transmit pathological signals in the lung resulting in pathological airway remodeling after *Tet3* deficiency but not in other tissues such as the vasculature, which own a different molecular circuitry of cell-cell communications.

1-2 The absence of *Tet3* in SMCs elicits distinct effect on tissue homeostasis in a context-dependent manner

SMCs interact with epithelial cells in the gastrointestinal tract (Gays et al., 2017). However, I observed neither gross morphological abnormalities nor mucus-producing cell hyperplasia in the intestine of *Tet3* mutants (Figure 4.2 B, Figure 4.3 C). It is still unclear whether 5hmC levels are lowered in intestinal SMCs of *Tet3* mutants and whether loss of *Tet3* in intestinal SMC leads to a phenotype switch as in the lung. Nevertheless, the low vulnerability of intestine to the loss of *Tet3* in SMC might be attributed to a lower susceptibility of intestinal epithelium to secreted cytokines or chemokines. Since SMCs develop from local mesenchymal cells with different developmental origins, it seems possible that different types of SMCs in the intestine and the lung own distinct intrinsic mechanisms resulting in proprietary responses after loss of TET3 SMCs. Since the expression of *Tet1*, *Tet2* and *Tet3* is cell type and/or tissue specific at adult stages, it also possible that *Tet3* expression is lower in intestinal SMCs and that *Tet1* and *Tet2* are the major enzymes converting 5mC to its oxidized derivatives. On the other hand, I did not detect substantial differences in the intensity of lacZ staining in $Tet3^{floxlacZ/+}$ mice, when analyzing different SMC-containing organs. Additional experiments will show in the future whether TLR7/8 signaling, which senses single-strand nucleic acids, is lower in intestinal compared to lung SMCs.

2. TET3 displays unique function to modulate gene expression in SMC

So far, a potential functional overlap among TET family members for embryonic development or adult tissue homeostasis has not been examined systematically. In mouse ESCs, *Tet1* and *Tet2* but not *Tet3* are highly expressed, while both *Tet2* and *Tet3* are expressed at a higher level than *Tet1* in

several cell types such as mouse $CD4^+$ T cells (Lio and Rao, 2019), B cells (Lio et al., 2016) and human $CD34^+$ bone marrow cells in adult stage (Rampal et al., 2014). The RNA-seq data performed in this study indicated that both *Tet2* and *Tet3* are highly expressed at similar levels in primary lung SMCs while the expression of *Tet1* was negligible. In contrast, previous reports described higher expression of *Tet2* than *Tet1* and *Tet3* in human coronary artery smooth muscle cells (Liu et al., 2013), (Figure 4.4 A). The co-expression of *Tet*s in the same cell or tissue may either suggest a functional redundancy or different targets preferences at distinct genomic regions. In support of the latter notion, TET1 was found to be enriched at active promoters and bivalent promoters, while TET2 prefers to target gene bodies of highly expressed genes and enhancers (Hon et al., 2014; Huang et al., 2014). Furthermore, a recent study revealed that TET2 activates the expression of key pro-contractile genes via increasing chromatin accessibility at the promoters in human coronary SMCs (Liu et al., 2013). In contrast, my study revealed that 5hmC in lung SMCs was highly enriched in gene bodies of highly transcribed genes, particularly in genes related to SMC contractility (Figure 8.2 D). Thus, the findings from us and others suggest that TET2 might mainly target promoter region, whereas TET3 deposits 5hmC mainly on gene body in SMCs. Due to the lack of antibodies suitable for ChIP-seq, the precise binding targets of TET2 and TET3 in SMCs are not known. Development of non-antibodies-based approach (e.g. DamID-seq) or generation knock-in mice expressing epitope-tagged TET2/3 via CRISPR-Cas9 technology for ChIP-seq will be helpful to solve this problem.

Several studies suggested functional redundancy of TETs in cells committed to differentiation or terminally differentiated cells/tissue. For instance, *Tet1* and *Tet2* double knockout mice are viable, fertile and overtly normal, albeit abnormal methylation at various imprinted loci was observed. The finding that loss of both enzymes is compatible with development suggests a compensatory function of TET3, which is up-regulated in *Tet1* and *Tet2* double knockout mice during the development. Earlier studies have shown that separate inactivation of *Tet1* and *Tet2* or cardiac-specific *Tet3* knockout mice display no overt cardiac phenotype, while *Tet2*/*Tet3* double KO led to no-compaction cardiomyopathy and early embryonic lethality, suggesting a functional redundancy of TET2 and TET3 during cardiac development (Fang et al., 2019). In human vascular SMC (HVSMC), TET2 binds within CArG-rich regions of SMC contractile promoters (*Myocd*, *Srf*, and *Myh11*) and increased 5hmC concentrations, which results in strong gene activation (Liu et al., 2013). Knockdown of *Tet2* in HVSMC inhibited expression of key pro-contractile genes including *Myocd* and *Srf* with concomitant transcriptional upregulation of *Klf4*. In this study, however, I found that deletion of *Tet2* in lung SMCs *in vivo* does not reduce 5hmC levels and does not cause any obvious lung abnormality (Figure 1 C & H-I). Furthermore, no apparent phenotype was discernible in the vasculature of SMC-specific *Tet2* mutants, which does not support a role of *Tet2* as a "master

regular" of SMCs as claimed previously (Liu et al., 2013). However, it is possible that the role of *Tet2* becomes only apparent under disease condition, which were not interrogated during the course of this study.

The conflicting results might also be explained by different expression levels of *Tet3* in the two type of SMCs: TET3 may compensate for the absence of TET2 in lung SMCs, which show high levels of *Tet3* expression. Human vascular SMC have extremely low levels of TET3, which may prevent compensation for the absence of *Tet2* (Figure 4.4 A)(Liu et al., 2013). In stark contrast, *Tet3* deficiency in lung SMCs led to a dramatic drop of 5hmC levels, impaired expression of contractile genes such as *Acta2*, *Cnn1* and *Myh11* and induced a switch from the contractile to the synthetic state. In vascular SMCs expression of the *Myocd* gene, which is an upstream activator of contractile genes including *Acta2*, *Cnn1* and *Myh11*, was repressed after loss of *Tet2*. Yet, expression of *Myocd* was not significantly altered in *Tet3*-deficient lung SMCs (data not shown), implying that TET2 and TET3 have distinct targets in lung SMCs. Of note, TET2 might preferably target promoters in lung SMC, while TET3 prefers targets at gene bodies of actively transcribed genes. Importantly, TET2 appears unable to compensate for the loss of TET3 function in lung SMCs, probably due to the lack of DNA binding domain and/or absence of key factors bridging interaction of TET2 with the Pol II-SETD2 complex. With other words, structural differences may prevent recruitment of TET2 to gene bodies, which is a typical feature of TET3 (Figure 7.1 A). Overall, my findings indicate that TET3 might compensate for the absence of TET2 but not *vice versa* (Figure 1 I, Figure 4.4 G). This hypothesis in agreement with finding in T cells that TET3 exhibits broader regulatory function than TET2 (Tsagaratou et al., 2017).

3. TET3 plays indispensable roles in orchestrating proper transcription elongation

3.1 The positive correlation of intragenic 5hmC enrichment with transcription activity appears to be conventional

In this study, I examined genome-wide distribution of 5hmC in post-mitotic SMCs *in vivo*, which allowed me to exclude the influence of passive demethylation processes on 5hmC. Instead I was able to directly monitor the dynamics of 5mC oxidation during gene transcription. Similar to terminally differentiated neurons that contain higher levels of 5hmC than most other tissues, I found that 5hmC in SMCs was globally enriched in gene bodies with a preference to the 5' end of intragenic regions (Figure 6 B). This intragenic distribution of 5hmC differs from other cell types such as ESC, in which 5hmC is enriched at the 3' end of intragenic regions, and in mouse cerebellum tissue where 5hmC distribution is almost equal in the intragenic region from the 5' to the 3' end (Xu et al., 2011). I also observed a positive correlation in SMCs of intragenic 5hmC accumulation with transcription level. This positive correlation was also observed in ESC and

mOSN (mature olfactory sensory neurons) but not in NPCs, in which 5hmC levels are low and not particularly enriched in gene body regions (Tan et al., 2013). This finding suggests that the impact of intragenic 5hmC enrichment on gene transcription is the rule and not cell/tissue specific. Of note, 5hmC level was markedly reduced in actively transcribed genes but not in repressed genes, in which 5hmC was evenly distributed within the gene after *Tet3* depletion, indicating that TET3 in SMCs targets specifically genes with a high turnover rate of transcription elongation but not in transcriptionally silent genes (Figure 6 D). In addition, 5hmC in SMCs is generally depleted from TSSs regardless of gene expression levels but was accumulated at around 1 kb upstream of TSSs. This situation is reminiscent of what has been observed in mOSN but different from the finding in ESCs, in which 5hmC accumulates at promoters of poised or repressed genes (Colquitt et al., 2013; Tan et al., 2013).

3.2 TET3–dependent 5hmC accumulation is required for maintenance of cell identity of primary lung SMCs by defining cell-specific transcription units

Loss of *Tet3* in SMC led to persistent switch from a contractile towards a synthetic state. The paradox of unexpectedly increased RNA-seq reads of contractile genes but decreased full-length mRNA and protein levels in *Tet3*-deficient SMC implies cryptic transcription in these genes. Integrative bioinformatics analysis of CAGE-seq, RNA-seq and RNA Pol II perS5 ChIP-seq revealed that loss of 5hmC after *Tet3* ablation resulted in aberrant initiation of intragenic transcription in genes that are highly expressed and involved in cytoskeleton structure organization. Genes related to SMC contractility such as *Acta2*, *Cnn1*, *Myh11*, *Pdlim3*, *Arhgap18* etc. showed increased numbers of RNA-reads but failed to produce functional transcripts and proteins, which is mainly caused by increased intragenic entry of Pol II pSer5 (Figure 8.2 B). This result suggests that 5hmC enrichment in intragenic region of highly expressed genes, which coincides in SMCs with genes required to maintain contractility, reflects dynamic transcriptional elongation and defines the correct transcription units.

Chromatin organization is evolutionarily highly conserved in transcription units from yeast to humans. Nucleosomes at proximal promoter regions show a high turnover rate characterized by continuous eviction and reincorporation. In contrast, gene coding regions show a more closed chromatin structure with nucleosomes tightly organized with roughly uniform spacing, avoiding the random appearance of larger nucleosome-free DNA segments, which prevents cryptic promoter activation. H3K36me3 is a characteristic histone modification of gene coding regions and plays pivotal role to maintain the repressive chromatin structure (Hennig and Fischer, 2013). Although the precise roles of H3K36me3 in preventing aberrant intragenic transcription remain elusive, it has been demonstrated that H3K36me3 serves as docking site for other repressive

chromatin remodelers including KDM5b, Rpd3 and DNMT3b. A recent study demonstrated that intragenic DNA methylation mediated by DNMT3b is critical to inhibit Pol II intragenic entry thereby precluding cryptic transcription in ESCs. In this model, DNMT3b is recruited to and methylates intragenic DNA by interacting with SETD2-mediated H3K36me3. Loss of DNMT3b results in cryptic entry of Pol II without interfering with the deposition of H3K36me3 in gene bodies (Neri et al., 2017). To date, no DNA methyltransferase was found and no DNA methylation was detected in yeast, indicating that intragenic DNA methylation is an additional molecular mechanism in mammals to prevent activation of cryptic promoters in gene bodies. Interestingly, around 91.8% of 5mC show a symmetrically distribution on sense and anti-sense strands, while this percentage drops to 21% for 5hmC (Yu et al., 2012). Another study revealed that 5hmC is slightly but significantly higher on the sense strand, linking 5hmC accumulation with the orientation of transcription (Wen et al., 2014). This asymmetric distribution of 5hmC might also explain the preferential occurrence of cryptic transcription on the sense strand in mammalian cells, In line with this hypothesis cryptic transcripts were found to occur both at the forward and the reverse DNA strand in yeast (Hennig and Fischer, 2013). In this study, the CAGE-seq analysis identified a lower level of antisense transcripts (average 96 tags within 904 genes) in *Tet3*-deficient SMCs than of sense transcripts in gene bodies (average 101 tags within 2114 genes). Moreover, GO term enrichment analysis revealed that genes with a significant increase in sense transcripts were largely correlated to SMC cell identity, while no SMCs cell identity correlated signaling pathways were discovered in genes with antisense transcripts. Thus, my results suggest that TET3 mediated 5hmC might preferentially locate at the sense strands of highly expressed genes to ensure transcriptional fidelity.

3.3. The role of TET3 and TET3-dependent 5hmC in transcriptional elongation

Studies with mESCs and in yeast disclosed that intragenic 5-cytosine methylation is a controlled process carried out by the *de novo* DNA methyltransferase DNMT3b and is positive correlated with transcription. A recent study from Neri et al. revealed that DNMT3b-mediated intragenic DNA methylation after the passage of elongating Pol II is required for a closed chromatin structure formation to prevent cryptic transcription. This model actually raises the question of the fate of DNA methylation for the next round of transcription elongation, since elongation occurs continuously within highly transcribed genes even before full-length RNA is synthesized. This is an important question, particularly for post-mitotic cells in which replication-dependent passive DNA demethylation does not occur. In addition, the function of intragenic DNA methylation during transcription elongation needs to be critically questioned, since the readout of the broadly used 5mC mapping method, bisulfite sequencing (BS-seq), is the sum of 5mC and 5hmC levels.

Hence, the precise distribution of 5mC versus 5hmC and the relationship between these two modifications remain unclear. Indeed, a recent study revealed that loss of 5hmC in gene body regions of tumor-related genes in kidney tumors was linked to reduced transcription. Such investigations are only feasible by careful simultaneous profiling at single-nucleotide resolution of 5mC (BS-seq corrected by Tet-assisted BS-seq, TAB-seq) and 5hmC (TAB-seq) (Chen et al., 2016). These findings suggest that a regulatory layer beyond intragenic DNA methylation exist for the activation of transcription.

Notably, it has been proposed that intragenic DNA methylation collaborates with H3K36me3 to prevent aberrant transcription. However, several studies in various cell types have demonstrated that H3K36me3 deposition is established in gene bodies independent of DNA methylation, indicating that DNA methylation occurs after SETD2-mediated H3K36me3 formation (Hahn et al., 2011; Neri et al., 2017). In this study, I demonstrated that TET3-mediated 5hmC in gene bodies plays a decisive role to maintain fidelity of transcription in SMCs. Importantly, I found that inactivation of *Tet3* leads to decreased enrichment of H3K36me3 in intragenic regions of highly transcribed genes, which demonstrates the involvement of 5hmC in renewed formation of repressive chromatin required to prevent cryptic intragenic transcription initiation. In addition, CAGE-seq IGV tracks (Figure 8.3 B) revealed that spurious transcription initiation sites are specific but not randomly distributed, indicating that TET3-mediated generation of 5hmC only occurs at specific genomic loci. It is tempting to speculate that TET3 interacting with Pol II preferentially demethylates 5mC at sites with strong CpG methylation (at least 90%), which is known to slow down the elongation rate (Veloso et al., 2014), giving TET3 more time to exert its enzymatic activity. My results support the hypothesis that TET3-mediated 5mC oxidation allows the elongating Pol II complex to overcome obstacles installed by DNA methylation (Neri et al., 2017) by transiently removing 5mC while subsequently re-establishing a closed H3K36me3-dependent state, followed by replenishment of 5mC.

So far, all three TETs have been reported to physically interact with TDG in various cell types, allowing TET-TDG-BER to efficiently excise 5mC (Spallotta et al., 2018; Weber et al., 2016)·(Zhang et al., 2017). Additionally, a previous study by Wang *et al.* demonstrated the existence of oxidized derivatives of 5hmC, i.e 5caC and 5fC, in gene bodies, which blocks incorporation of NTPs during elongation. In *Tdg*-depleted mouse ES cells, 5caC and 5fC accumulated in gene bodies and the rate of Pol II transcription elongation is significantly reduced, indicating that 5fC/5caCs within a transcribed genomic region serve as a speed brake for the transcriptional machinery. This may enable fine-tuning of the elongation rate for highly expressed genes that are relatively rich in 5hmC (Wang et al., 2015). Furthermore, the 5fC/5caC glycosylase TDG is involved in transcription through its association with the co-activators CBP/p300 in euchromatin (Cortellino et al., 2011).

Since my study revealed that TET3 interacted not only with elongating Pol II but also with TDG in primary SMCs (data not shown), I therefore hypothesize that TET3 allows transcriptional elongation to proceed in primary SMCs by enabling DNA demethylation and transcription-coupled processing of 5fC/5caC at a few critical 5mC-rich regions.

I found that TET3 does not only oxidize 5mC, but has also non-catalytic function, e.g. enhancement or stabilization of SETD2 binding to Pol II, although I cannot be fully excluded that the catalytic activity of TET3 plays a role in this process. This finding is consistent with the impaired enrichment of H3K36me6 in gene bodies of TET3 target genes. Based on these findings, I postulate that TET3 is recruited to Pol II and mediates 5hmC formation when the RNA transcription machinery encounters methylated cryptic starts sites (TSSs) or transposons. TET3 and/or TET3-dependent 5hmC deposition stabilize the recruitment of SETD2, which subsequently results in H3K36 tri-methylation and restores DNA methylation thereby preventing aberrant activation of non-canonical promoters in gene bodies. In this process, 5hmC might serve as an intermediate towards DNA demethylation. The resulting de-methylated cytosines are quickly re-methylated and then again oxidized, perhaps during each round of transcription as part of mechanisms to ensure correct initiation of transcription at canonical TSSs (Figure 10 B). Based on the high prevalence of 5hmC in the genome compared to its oxidized derivatives, the low overlap between 5hmC and 5fC/5caC at individual CpGs, and the unchanged global levels of 5fC/5caC levels after *Tet3* inactivation I reason that the observed 5fC/5caC modifications are not the results of dynamic demethylation but represent oxidation events in the past, probably during SMC differentiation. We also cannot exclude the possibility that due to rapid excision of 5fC/5caC mediated by TDG, subtle differences caused by the presence of absence of TET3 are not detectable (Wu et al., 2014).

4. Activation of innate immune responses by spurious transcripts

Eukaryotes from yeast to humans have developed evolutionarily conserved mechanisms to repress cryptic transcription of coding genes to ensure generation of functional proteins for proper cell function. Excessive cryptic transcription and accumulation of aberrant non-coding RNAs has been reported to impair genomic stability in yeast (Wagner and Carpenter, 2012). However, the biological consequence of inappropriate cryptic transcription in mammals were not adequately determined. In this study, I discovered a link between the production of spurious transcripts and the massive inflammation in *Tet3*-deficient lung *in vivo*. Furthermore, I demonstrated the ability of aberrant spurious transcripts from *Tet3*-deficient SMCs to trigger TLR7/8-dependent innate immune responses.

4.1 Features of spurious transcripts in *Tet3*-deficient SMCs

The recognition of nucleic acids by TLRs represents a potential risk for inappropriate responses to self-derived nucleic acids, which can lead to autoimmunity or autoinflammatory diseases (Sharma et al., 2015). Thus, the activation of nucleic acid sensing TLR7/8 is tightly controlled by various security mechanisms to prevent unwanted activation by host derived ssRNAs. This includes rapid degradation of endogenous RNAs before encountering the endosomal compartment (Chi and Flavell, 2008). Furthermore, the high content of modified nucleotides such as 5-methylcytidine or pseudouridine makes endogenous RNA non-stimulatory (Kariko et al., 2005). In my study, the nature of spurious transcripts in *Tet3*-deficient SMCs is still unclear but they are detectable in RNA-seq, suggesting that aberrant RNAs are sufficiently stable and protected from rapid degradation by RNases, probably due to proper 5' end m^7G Capping and 3'-terminal polyadenylation, similar to full-length RNAs. Cryptic transcripts that escape nuclease-mediated degradation, enabling them to activate the endosomal TLR7/8 signaling pathway. It is still unclear how these self-originated spurious transcripts are recognized by the intracellular nucleic acid sensing machinery in *Tet3*-deficient SMCs. One possibility is that incorporation of modified nucleosides m5C, m6A, m5U, s2U, or pseudouridine are missing in spurious transcripts in *Tet3*-deficient SMCs, which would disguise them as "foreign" and allow TLR7/8 activation (Kariko et al., 2005). Modifications might be missing due to incorrect transcription or splicing that might reduce the affinity for RNA modifiers. This idea is compatible with previous reports showing that transfected RNAs containing nucleoside modifications are not able to activate TLR signaling in human 293 cell lines stably overexpressing exogenous human TLR3, TLR7, or TLR8. For instance, m6A modifications prevent activation of TLR3-signaling by RNAs, while m6A and 5mC modifications block stimulation of TLR7 and TLR8. This observation clearly demonstrates that modified RNAs have a much lower potential to activate innate immune responses than non-modified bacterial mRNAs (Kariko et al., 2005). However, I cannot completely exclude the possibility that spurious transcripts are modified and contain 5mC and m6A, but are recognized via an additional RNA sensing machinery to activate TLR7 signaling. In fact, RNAs with m6A and 5mC modifications are almost as potent as unmodified foreign RNAs to induce cytokine production in human primary blood dendritic cells, in which only TLR7 but not TLR3 or TLR8 are expressed (Kariko et al., 2005). Another explanation for the strong activation of innate immune responses by spurious transcripts is the deregulation of intracellular RNA sorting due to disrupted delivery of RNA granules containing spurious transcripts. Accumulation of spurious transcripts in such granules, which encompass stress granules and P-bodies, might interfere with processing to late endosome/lysosomes or lead to defects in autophagy (direct uptake of aberrant nucleic acid by lysosomes (Fujiwara et al., 2014). Furthermore, it is noteworthy that TET enzymes are able to

convert 5mC to 5hmC in RNA, which accelerates mRNA translation (Delatte et al., 2016; Fu et al., 2014). Aberrant transcripts will most likely lack hydroxymethylated CpG(5hmrC) in TET3-deficient SMCs. The absence of 5hmrC might be sufficient to activate TLR7/8 and innate immunity. Thus, it will be interesting to further study the immune-stimulatory capacity of synthetic RNA with or without 5hmrC modifications.

4.2. Innate immune responses evoked by loss of TET3 and its correlation with chronic allergic inflammation

Activation of endosomal TLR signaling plays a role in several lung diseases including allergic asthma. However, the function of TLR7/8-signaling in SMCs and the association with lung diseases is elusive. In this study, I demonstrated that TLR7/8-signaling was activated after *Tet3* ablation in SMCs. I observed enhanced T helper type 2 cell (Th2) differentiation and excessive production of cytokine IL4 and IL13 in *Tet3* mutant lung 6 months after TAM injection ("long-term experiment"). Th2 cells play critical roles in orchestrating, perpetuating and amplifying the inflammatory response. Th2 cells are characterized by secretion of IL-4, IL-5, IL-9 and IL-13, prototypical cytokines that play vital roles in the pathogenesis of asthma (Lambrecht and Hammad, 2012). For instance, IL-4 promotes Th2 cell differentiation and IgE synthesis, while IL-13 signaling plays a central role in airway hyperreactivity in the effector phase of the disease by enhancing ASMC sensitivity to a number of contractile agonists (Gour and Wills-Karp, 2015). IL-13 also induces metaplasia of club cells to goblet cells by regulating transcription factors such as SAM Pointed Domain Containing ETS Transcription Factor (SPDEF) and Forkhead Box A2 (FOXA2). This process will promote expression of MUC5AC and anterior gradient 2 and impair homeostasis of airway epithelium (Gour and Wills-Karp, 2015). Excessive goblet cell differentiation and mucus hypersecretion exacerbate airway hyperresponsiveness and airway remodeling in asthma patients and mouse models. In addition, excessive production of IL-4 and IL-13 promotes airway SMCs proliferation (Khan, 2013). The occurrence of excessive Th2 cell based immune responses (Figure 9.5 F & G), airway remodeling with massive goblet cell hyperplasia (Figure 4.3 A & C), thickened basal membrane (Figure 4.2 C), B cell accumulation (Figure 9.5 I) and epithelial shedding (Figure 9.5 A) in *Tet3*^{smKO} mice indicates that loss of *Tet3* in SMCs induces an asthma-like phenotype. However, it is still unclear what are the causative signals for Th2 cell activation, since activation of TLR7 signaling was reported to inhibit allergic airways responses and Th2-mediated airway inflammation (Zakeri et al., 2018). Based on the observation that TLR7 agonists relax the contraction of airway and that stimulation of TLR7 in airway nerves results in bronchodilatation, stimulation of TLR7 was proposed as a therapeutic approach to forestall asthma (Drake et al., 2013). However, stimulation of TLRs, especially endosomal TLRs on eosinophils, increases

cytokine release, suggesting that TLR7/8 and TLR9 in eosinophils are responsible for exacerbation of asthma during viral infections (Hatchwell et al., 2015). Finally, the asthma-like phenotype observed in *Tet3*^{smKO} mice differs substantially from classical allergic asthma. A potential beneficial role of TLR7/8 activation in allergic asthma does not rule out that TLR7/8 activation is disease-promoting and detrimental in another context. Activation of Th2 cell-based immune response in *Tet3* mutant lung might be a cell-specific consequence of TLR7 signaling in *Tet3*-deficient SMC or occur indirectly by other means. Furthermore, TLR7-induced production of pro-inflammatory cytokine and chemokine may differ between SMCs and different types of immune cells.

Sterile airway inflammation is caused by a combination of complex and incompletely understood environmental and genetic interactions, which is reflected by epigenetics change in airway SMCs (Busse et al., 2007). Changes in the metabolic activity of cells has profound consequences for DNA methylation/demethylation, since most substrates used to modify DNA are intermediate metabolites like SAM and α-KG, whose intracellular level is affected by nutrient availability and the activity of metabolic pathways. It will be interesting to evaluate TET3 activity and 5hmC levels in obese asthma patients, who have a distinct phenotype with a unique pathophysiology (low eosinophilic inflammation, low allergen sensitization, and showed limited responsiveness to inhaled corticosteroids).

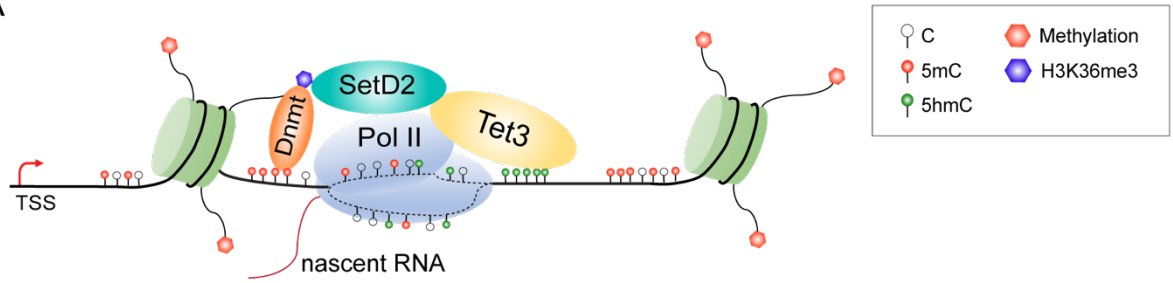
During aging the dynamic balance between methylation and demethylation critical for cell/tissue homeostasis maintenance is disrupted, resulting in reduced global methylation (Avrahami et al., 2015; Jung and Pfeifer, 2015). Asthma frequently develops in children, but as many as 25% of asthma patients develop symptoms after the age of 40 (Sood et al., 2013). Patients with late onset asthma may have more severe airway disease than patients with earlier onset of asthma. It will be interesting to study the activity of TETs and 5hmC levels in patients with late onset of asthma.

5. Conclusions and proposed models

In this study, I revealed a critical dual role of TET3 for transcriptional elongation and for defining canonical transcription units in SMCs (Figure 10 A). In actively transcribed genes, TET3 is recruited to the RNA Pol II elongation machinery and converts 5mC to its different oxidized forms when encountering methylated cryptic starts sites (TSSs) or transposons. TET3 and/or 5hmC stabilize the recruitment of SETD2 that subsequently mediates H3K36me3 deposition followed by DNA methylation, thereby enforcing the usage of canonical upstream promoters. According to this model, 5hmC serves as intermediate towards DNA demethylation. Demethylated cytosines will be quickly re-methylated when the elongation complex has passed until the next round of transcription, which will cause another oxidation event. The model explains why a "stable" presence of 5hmC in gene bodies of highly transcribed genes is seen, although in reality a rapid

succession of 5mC, 5hmC, 5fC, 5caC and C occurs. The dynamic changes in chromatin accessibility during transcriptional elongation prevent ectopic initiation and ensure the usage of regular TSSs. Disruption of this process by inactivation of *Tet3* in SMCs, a cell type with remarkable plasticity and vulnerability in response to intrinsic or extrinsic cues, results in accumulation of aberrant intragenic transcripts that are detected by nucleic acid sensing system and subsequently provokes massive immune responses and lung pathogenesis (Figure 10 B). My study suggests that modulation of TET3 activity might be an attractive approach for therapeutic interventions to treat several lung parenchymal diseases.

A



B

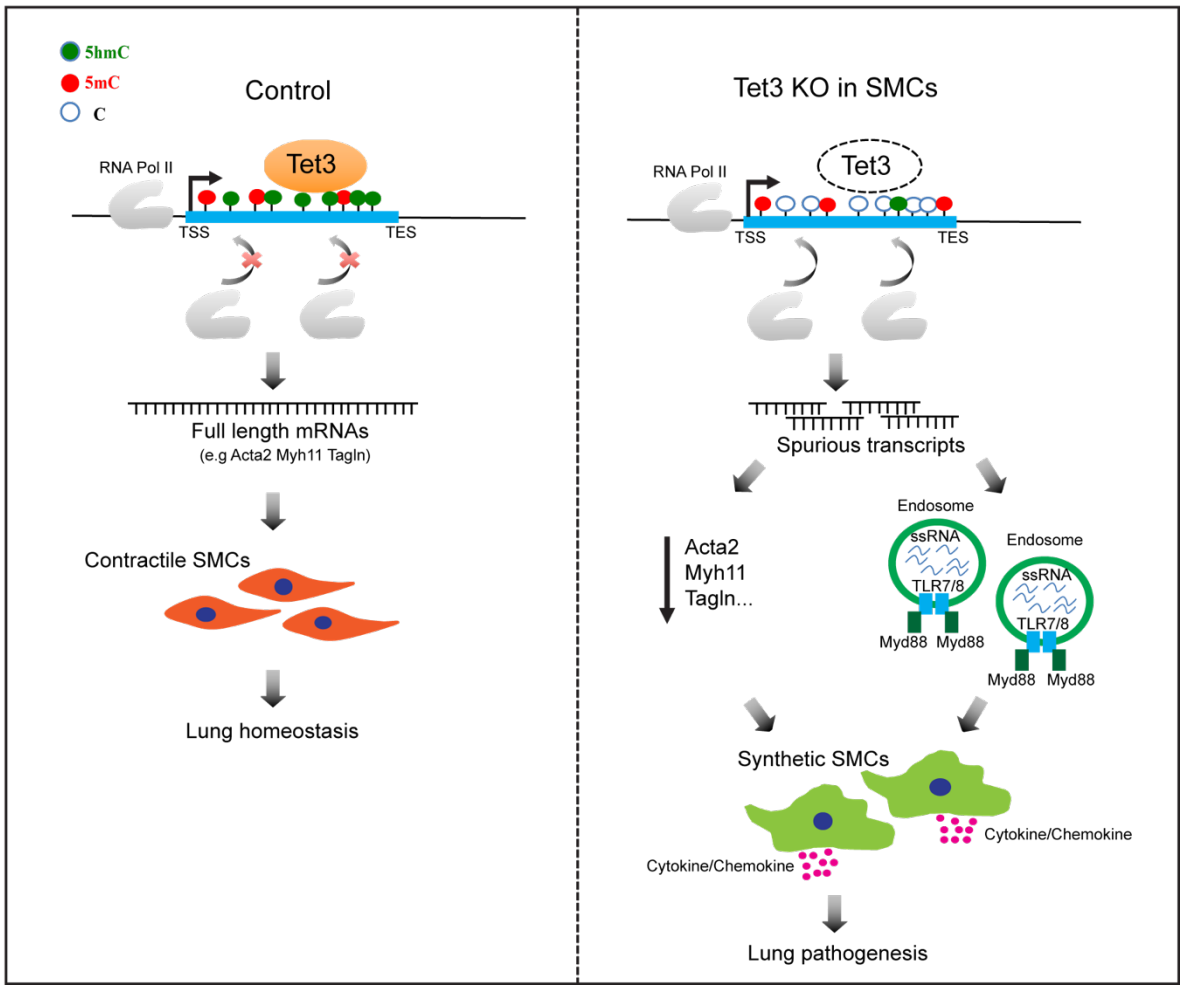


Figure 10. Summary models. (A) Model showing the potential functions of TET3 during transcription elongation. **(B)** Schematic model showing the epigenetic mechanisms regulating lung SMC identity and lung tissue homeostasis in wild-type and *Tet3*^{smKO} mice.

Reference

Abdel-Wahab, O., and Levine, R.L. (2013). Mutations in epigenetic modifiers in the pathogenesis and therapy of acute myeloid leukemia. Blood *121*, 3563-3572.
Akbari, O., Stock, P., DeKruyff, R.H., and Umetsu, D.T. (2003). Mucosal tolerance and immunity: regulating the development of allergic disease and asthma. Int Arch Allergy Immunol *130*, 108-118.
Alexander, M.R., and Owens, G.K. (2012). Epigenetic control of smooth muscle cell differentiation and phenotypic switching in vascular development and disease. Annu Rev Physiol *74*, 13-40.
Allahverdian, S., Chaabane, C., Boukais, K., Francis, G.A., and Bochaton-Piallat, M.L. (2018). Smooth muscle cell fate and plasticity in atherosclerosis. Cardiovasc Res *114*, 540-550.
Altorki, N.K., Markowitz, G.J., Gao, D., Port, J.L., Saxena, A., Stiles, B., McGraw, T., and Mittal, V. (2019). The lung microenvironment: an important regulator of tumour growth and metastasis. Nat Rev Cancer *19*, 9-31.
An, J., Rao, A., and Ko, M. (2017). TET family dioxygenases and DNA demethylation in stem cells and cancers. Exp Mol Med *49*, e323.
Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol *11*, R106.
Avgustinova, A., and Benitah, S.A. (2016). Epigenetic control of adult stem cell function. Nat Rev Mol Cell Biol *17*, 643-658.
Avrahami, D., Li, C., Zhang, J., Schug, J., Avrahami, R., Rao, S., Stadler, M.B., Burger, L., Schubeler, D., Glaser, B.*, et al.* (2015). Aging-Dependent Demethylation of Regulatory Elements Correlates with Chromatin State and Improved beta Cell Function. Cell Metab *22*, 619-632.
Badri, K.R., Zhou, Y., and Schuger, L. (2008). Embryological origin of airway smooth muscle. Proc Am Thorac Soc *5*, 4-10.
Barkauskas, C.E., Chung, M.I., Fioret, B., Gao, X., Katsura, H., and Hogan, B.L. (2017). Lung organoids: current uses and future promise. Development *144*, 986-997.
Barnes, P.J. (2008). The cytokine network in asthma and chronic obstructive pulmonary disease. J Clin Invest *118*, 3546-3556.
Baubec, T., Colombo, D.F., Wirbelauer, C., Schmidt, J., Burger, L., Krebs, A.R., Akalin, A., and Schubeler, D. (2015). Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. Nature *520*, 243-247.
Beerman, I., and Rossi, D.J. (2015). Epigenetic Control of Stem Cell Potential during Homeostasis, Aging, and Disease. Cell stem cell *16*, 613-625.
Boettger, T., Beetz, N., Kostin, S., Schneider, J., Kruger, M., Hein, L., and Braun, T. (2009). Acquisition of the contractile phenotype by murine arterial smooth muscle cells depends on the Mir143/145 gene cluster. J Clin Invest *119*, 2634-2647.
Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114-2120.
Busse, P.J., Zhang, T.F., Srivastava, K., Schofield, B., and Li, X.M. (2007). Effect of ageing on pulmonary inflammation, airway hyperresponsiveness and T and B cell responses in antigen-sensitized and -challenged mice. Clin Exp Allergy *37*, 1392-1403.
Bustamante-Marin, X.M., and Ostrowski, L.E. (2017). Cilia and Mucociliary Clearance. Cold Spring Harb Perspect Biol *9*.
Byers, D.E., Alexander-Brett, J., Patel, A.C., Agapov, E., Dang-Vu, G., Jin, X., Wu, K., You, Y., Alevy, Y., Girard, J.P.*, et al.* (2013). Long-term IL-33-producing epithelial progenitor cells in chronic obstructive lung disease. J Clin Invest *123*, 3967-3982.
Chen, K., Chen, Z., Wu, D., Zhang, L., Lin, X., Su, J., Rodriguez, B., Xi, Y., Xia, Z., Chen, X.*, et al.* (2015). Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes. Nat Genet *47*, 1149-1157.

Chen, K., Zhang, J., Guo, Z., Ma, Q., Xu, Z., Zhou, Y., Xu, Z., Li, Z., Liu, Y., Ye, X.*, et al.* (2016). Loss of 5-hydroxymethylcytosine is linked to gene body hypermethylation in kidney cancer. Cell Res *26*, 103-118.
Chen, Q., Chen, Y., Bian, C., Fujiki, R., and Yu, X. (2013). TET2 promotes histone O-GlcNAcylation during gene transcription. Nature *493*, 561-564.
Chi, H., and Flavell, R.A. (2008). Innate recognition of non-self nucleic acids. Genome Biol *9*, 211.
Clarke, D.L., Clifford, R.L., Jindarat, S., Proud, D., Pang, L., Belvisi, M., and Knox, A.J. (2010). TNFalpha and IFNgamma synergistically enhance transcriptional activation of CXCL10 in human airway smooth muscle cells via STAT-1, NF-kappaB, and the transcriptional coactivator CREB-binding protein. J Biol Chem *285*, 29101-29110.
Clifford, R.L., Patel, J.K., John, A.E., Tatler, A.L., Mazengarb, L., Brightling, C.E., and Knox, A.J. (2015). CXCL8 histone H3 acetylation is dysfunctional in airway smooth muscle in asthma: regulation by BET. Am J Physiol Lung Cell Mol Physiol *308*, L962-972.
Colquitt, B.M., Allen, W.E., Barnea, G., and Lomvardas, S. (2013). Alteration of genic 5-hydroxymethylcytosine patterning in olfactory neurons correlates with changes in gene expression and cell identity. Proc Natl Acad Sci U S A *110*, 14682-14687.
Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., Cigliano, A., Le Coz, M., Devarajan, K., Wessels, A., Soprano, D.*, et al.* (2011). Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. Cell *146*, 67-79.
Dahl, J.A., and Collas, P. (2008). MicroChIP--a rapid micro chromatin immunoprecipitation assay for small cell samples and biopsies. Nucleic Acids Res *36*, e15.
Damal Villivalam, S., You, D., Kim, J., Lim, H.W., Xiao, H., Zushin, P.H., Oguri, Y., Amin, P., and Kang, S. (2020). TET1 is a beige adipocyte-selective epigenetic suppressor of thermogenesis. Nat Commun *11*, 4313.
Damera, G., Tliba, O., and Panettieri, R.A., Jr. (2009). Airway smooth muscle as an immunomodulatory cell. Pulm Pharmacol Ther *22*, 353-359.
Davis, M.P., van Dongen, S., Abreu-Goodger, C., Bartonicek, N., and Enright, A.J. (2013). Kraken: a set of tools for quality control and analysis of high-throughput sequence data. Methods *63*, 41-49.
Delatte, B., Wang, F., Ngoc, L.V., Collignon, E., Bonvin, E., Deplus, R., Calonne, E., Hassabi, B., Putmans, P., Awe, S.*, et al.* (2016). RNA biochemistry. Transcriptome-wide distribution and function of RNA hydroxymethylcytosine. Science *351*, 282-285.
Deplus, R., Delatte, B., Schwinn, M.K., Defrance, M., Mendez, J., Murphy, N., Dawson, M.A., Volkmar, M., Putmans, P., Calonne, E.*, et al.* (2013). TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. EMBO J *32*, 645-655.
Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.
Drake, M.G., Scott, G.D., Proskocil, B.J., Fryer, A.D., Jacoby, D.B., and Kaufman, E.H. (2013). Toll-like receptor 7 rapidly relaxes human airways. Am J Respir Crit Care Med *188*, 664-672.
Dulac, C. (2010). Brain function and chromatin plasticity. Nature *465*, 728-735.
Faiz, A., Weckmann, M., Tasena, H., Vermeulen, C.J., Van den Berge, M., Ten Hacken, N.H.T., Halayko, A.J., Ward, J.P.T., Lee, T.H., Tjin, G.*, et al.* (2018). Profiling of healthy and asthmatic airway smooth muscle cells following interleukin-1beta treatment: a novel role for CCL20 in chronic mucus hypersecretion. Eur Respir J *52*.
Fang, S., Li, J., Xiao, Y., Lee, M., Guo, L., Han, W., Li, T., Hill, M.C., Hong, T., Mo, W.*, et al.* (2019). Tet inactivation disrupts YY1 binding and long-range chromatin interactions during embryonic heart development. Nat Commun *10*, 4297.
Feil, S., Fehrenbacher, B., Lukowski, R., Essmann, F., Schulze-Osthoff, K., Schaller, M., and Feil, R. (2014). Transdifferentiation of vascular smooth muscle cells to macrophage-like cells during atherogenesis. Circ Res *115*, 662-667.
Feinberg, A.P., Ohlsson, R., and Henikoff, S. (2006). The epigenetic progenitor origin of human cancer. Nat Rev Genet *7*, 21-33.

Findeisen, H.M., Gizard, F., Zhao, Y., Qing, H., Heywood, E.B., Jones, K.L., Cohn, D., and Bruemmer, D. (2011). Epigenetic regulation of vascular smooth muscle cell proliferation and neointima formation by histone deacetylase inhibition. Arterioscler Thromb Vasc Biol *31*, 851-860.
Fu, L., Guerrero, C.R., Zhong, N., Amato, N.J., Liu, Y., Liu, S., Cai, Q., Ji, D., Jin, S.G., Niedernhofer, L.J.*, et al.* (2014). Tet-mediated formation of 5-hydroxymethylcytosine in RNA. J Am Chem Soc *136*, 11582-11585.
Fujiwara, Y., Furuta, A., Kikuchi, H., Aizawa, S., Hatanaka, Y., Konya, C., Uchida, K., Yoshimura, A., Tamai, Y., Wada, K.*, et al.* (2014). Discovery of a novel type of autophagy targeting RNA. Autophagy *9*, 403-409.
Gangloff, M. (2012). Different dimerisation mode for TLR4 upon endosomal acidification? Trends Biochem Sci *37*, 92-98.
Gays, D., Hess, C., Camporeale, A., Ala, U., Provero, P., Mosimann, C., and Santoro, M.M. (2017). An exclusive cellular and molecular network governs intestinal smooth muscle cell differentiation in vertebrates. Development *144*, 464-478.
Gomez, D., Shankman, L.S., Nguyen, A.T., and Owens, G.K. (2013). Detection of histone modifications at specific gene loci in single cells in histological sections. Nat Methods *10*, 171-177.
Gomez, D., Swiatlowska, P., and Owens, G.K. (2015). Epigenetic Control of Smooth Muscle Cell Identity and Lineage Memory. Arterioscler Thromb Vasc Biol *35*, 2508-2516.
Gontier, G., Iyer, M., Shea, J.M., Bieri, G., Wheatley, E.G., Ramalho-Santos, M., and Villeda, S.A. (2018). Tet2 Rescues Age-Related Regenerative Decline and Enhances Cognitive Function in the Adult Mouse Brain. Cell Reports *22*, 1974-1981.
Gour, N., and Wills-Karp, M. (2015). IL-4 and IL-13 signaling in allergic airway disease. Cytokine *75*, 68-78.
Greco, C.M., Kunderfranco, P., Rubino, M., Larcher, V., Carullo, P., Anselmo, A., Kurz, K., Carell, T., Angius, A., Latronico, M.V.*, et al.* (2016). DNA hydroxymethylation controls cardiomyocyte gene expression in development and hypertrophy. Nat Commun *7*, 12418.
Green, B.B., Houseman, E.A., Johnson, K.C., Guerin, D.J., Armstrong, D.A., Christensen, B.C., and Marsit, C.J. (2016). Hydroxymethylation is uniquely distributed within term placenta, and is associated with gene expression. FASEB J *30*, 2874-2884.
Greenberg, M.V.C., and Bourc'his, D. (2019). The diverse roles of DNA methylation in mammalian development and disease. Nat Rev Mol Cell Biol *20*, 590-607.
Gu, T.P., Guo, F., Yang, H., Wu, H.P., Xu, G.F., Liu, W., Xie, Z.G., Shi, L., He, X., Jin, S.G.*, et al.* (2011). The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature *477*, 606-610.
Hahn, M.A., Qiu, R., Wu, X., Li, A.X., Zhang, H., Wang, J., Jui, J., Jin, S.G., Jiang, Y., Pfeifer, G.P.*, et al.* (2013). Dynamics of 5-hydroxymethylcytosine and chromatin marks in Mammalian neurogenesis. Cell Rep *3*, 291-300.
Hahn, M.A., Wu, X., Li, A.X., Hahn, T., and Pfeifer, G.P. (2011). Relationship between gene body DNA methylation and intragenic H3K9me3 and H3K36me3 chromatin marks. PLoS One *6*, e18844.
Han, H., and Ziegler, S.F. (2013). Bronchoalveolar Lavage and Lung Tissue Digestion. Bio Protoc *3*.
Hansson, G.K., and Hermansson, A. (2011). The immune system in atherosclerosis. Nat Immunol *12*, 204-212.
Harman, J.L., Dobnikar, L., Chappell, J., Stokell, B.G., Dalby, A., Foote, K., Finigan, A., Freire-Pritchett, P., Taylor, A.L., Worssam, M.D.*, et al.* (2019). Epigenetic Regulation of Vascular Smooth Muscle Cells by Histone H3 Lysine 9 Dimethylation Attenuates Target Gene-Induction by Inflammatory Signaling. Arterioscler Thromb Vasc Biol *39*, 2289-2302.
Harmanci, A., Rozowsky, J., and Gerstein, M. (2014). MUSIC: identification of enriched regions in ChIP-Seq experiments using a mappability-corrected multiscale signal processing framework. Genome Biol *15*, 474.

Hatchwell, L., Collison, A., Girkin, J., Parsons, K., Li, J., Zhang, J., Phipps, S., Knight, D., Bartlett, N.W., Johnston, S.L.*, et al.* (2015). Toll-like receptor 7 governs interferon and inflammatory responses to rhinovirus and is suppressed by IL-5-induced lung eosinophilia. Thorax *70*, 854-861.
He, M., Zheng, B., Zhang, Y., Zhang, X.H., Wang, C., Yang, Z., Sun, Y., Wu, X.L., and Wen, J.K. (2015). KLF4 mediates the link between TGF-beta1-induced gene transcription and H3 acetylation in vascular smooth muscle cells. FASEB J *29*, 4059-4070.
Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.
Hennig, B.P., and Fischer, T. (2013). The great repression: chromatin and cryptic transcription. Transcription *4*, 97-101.
Hirschi, K.K., and Majesky, M.W. (2004). Smooth muscle stem cells. Anat Rec A Discov Mol Cell Evol Biol *276*, 22-33.
Hodges, S.J., Yoo, J.J., Mishra, N., and Atala, A. (2010). The effect of epigenetic therapy on congenital neurogenic bladders--a pilot study. Urology *75*, 868-872.
Holt, P.G., Strickland, D.H., Wikström, M.E., and Jahnsen, F.L. (2008). Regulation of immunological homeostasis in the respiratory tract. Nature Reviews Immunology *8*, 142-152.
Hon, G.C., Song, C.X., Du, T., Jin, F., Selvaraj, S., Lee, A.Y., Yen, C.A., Ye, Z., Mao, S.Q., Wang, B.A.*, et al.* (2014). 5mC oxidation by Tet2 modulates enhancer activity and timing of transcriptome reprogramming during differentiation. Mol Cell *56*, 286-297.
Hsin, J.P., and Manley, J.L. (2012). The RNA polymerase II CTD coordinates transcription and RNA processing. Genes Dev *26*, 2119-2137.
Hu, L., Lu, J., Cheng, J., Rao, Q., Li, Z., Hou, H., Lou, Z., Zhang, L., Li, W., Gong, W.*, et al.* (2015). Structural insight into substrate preference for TET-mediated oxidation. Nature *527*, 118-122.
Huang, Y., Chavez, L., Chang, X., Wang, X., Pastor, W.A., Kang, J., Zepeda-Martinez, J.A., Pape, U.J., Jacobsen, S.E., Peters, B.*, et al.* (2014). Distinct roles of the methylcytosine oxidases Tet1 and Tet2 in mouse embryonic stem cells. Proc Natl Acad Sci U S A *111*, 1361-1366.
Inoue, A., Shen, L., Matoba, S., and Zhang, Y. (2015). Haploinsufficiency, but not defective paternal 5mC oxidation, accounts for the developmental defects of maternal Tet3 knockouts. Cell Rep *10*, 463-470.
Iurlaro, M., Ficz, G., Oxley, D., Raiber, E.A., Bachman, M., Booth, M.J., Andrews, S., Balasubramanian, S., and Reik, W. (2013). A screen for hydroxymethylcytosine and formylcytosine binding proteins suggests functions in transcription and chromatin regulation. Genome Biol *14*, R119.
Jambhekar, A., Dhall, A., and Shi, Y. (2019). Roles and regulation of histone methylation in animal development. Nat Rev Mol Cell Biol *20*, 625-641.
James, A.L., Elliot, J.G., Jones, R.L., Carroll, M.L., Mauad, T., Bai, T.R., Abramson, M.J., McKay, K.O., and Green, F.H. (2012). Airway smooth muscle hypertrophy and hyperplasia in asthma. Am J Respir Crit Care Med *185*, 1058-1064.
Jensen, S., and Thomsen, A.R. (2012). Sensing of RNA viruses: a review of innate immune receptors involved in recognizing RNA virus invasion. J Virol *86*, 2900-2910.
Jin, B., Li, Y., and Robertson, K.D. (2011). DNA methylation: superior or subordinate in the epigenetic hierarchy? Genes Cancer *2*, 607-617.
Jin, S.G., Zhang, Z.M., Dunwell, T.L., Harter, M.R., Wu, X., Johnson, J., Li, Z., Liu, J., Szabo, P.E., Lu, Q.*, et al.* (2016). Tet3 Reads 5-Carboxylcytosine through Its CXXC Domain and Is a Potential Guardian against Neurodegeneration. Cell Rep *14*, 493-505.
Jones, P.A. (1999). The DNA methylation paradox. Trends Genet *15*, 34-37.
Jung, M., and Pfeifer, G.P. (2015). Aging and DNA methylation. BMC Biol *13*, 7.
Kariko, K., Buckstein, M., Ni, H., and Weissman, D. (2005). Suppression of RNA recognition by Toll-like receptors: the impact of nucleoside modification and the evolutionary origin of RNA. Immunity *23*, 165-175.

Kawai, T., and Akira, S. (2010). The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. Nat Immunol *11*, 373-384.
Khan, M.A. (2013). Inflammation signals airway smooth muscle cell proliferation in asthma pathogenesis. Multidiscip Respir Med *8*, 11.
Khare, T., Pai, S., Koncevicius, K., Pal, M., Kriukiene, E., Liutkeviciute, Z., Irimia, M., Jia, P., Ptak, C., Xia, M.*, et al.* (2012). 5-hmC in the brain is abundant in synaptic genes and shows differences at the exon-intron boundary. Nat Struct Mol Biol *19*, 1037-1043.
Kiyokawa, H., and Morimoto, M. (2020). Notch signaling in the mammalian respiratory system, specifically the trachea and lungs, in development, homeostasis, regeneration, and disease. Dev Growth Differ *62*, 67-79.
Knoop, K.A., and Newberry, R.D. (2018). Goblet cells: multifaceted players in immunity at mucosal surfaces. Mucosal Immunol *11*, 1551-1557.
Knudsen, L., and Ochs, M. (2018). The micromechanics of lung alveoli: structure and function of surfactant and tissue components. Histochem Cell Biol *150*, 661-676.
Ko, M., An, J., Pastor, W.A., Koralov, S.B., Rajewsky, K., and Rao, A. (2015). TET proteins and 5-methylcytosine oxidation in hematological cancers. Immunol Rev *263*, 6-21.
Kodzius, R., Kojima, M., Nishiyori, H., Nakamura, M., Fukuda, S., Tagami, M., Sasaki, D., Imamura, K., Kai, C., Harbers, M.*, et al.* (2006). CAGE: cap analysis of gene expression. Nat Methods *3*, 211-222.
Kohli, R.M., and Zhang, Y. (2013). TET enzymes, TDG and the dynamics of DNA demethylation. Nature *502*, 472-479.
Kotton, D.N., and Morrisey, E.E. (2014). Lung regeneration: mechanisms, applications and emerging stem cell populations. Nat Med *20*, 822-832.
Kovach, M.A., and Standiford, T.J. (2011). Toll like receptors in diseases of the lung. Int Immunopharmacol *11*, 1399-1406.
Kudo, M., Melton, A.C., Chen, C., Engler, M.B., Huang, K.E., Ren, X., Wang, Y., Bernstein, X., Li, J.T., Atabai, K.*, et al.* (2012). IL-17A produced by alphabeta T cells drives airway hyper-responsiveness in mice and enhances mouse and human airway smooth muscle contraction. Nat Med *18*, 547-554.
Kumar, H., Kawai, T., and Akira, S. (2009). Pathogen recognition in the innate immune response. Biochem J *420*, 1-16.
Lambrecht, B.N., and Hammad, H. (2012). The airway epithelium in asthma. Nat Med *18*, 684-692.
Lavelle, E.C., Murphy, C., O'Neill, L.A., and Creagh, E.M. (2010). The role of TLRs, NLRs, and RLRs in mucosal innate immunity and homeostasis. Mucosal Immunol *3*, 17-28.
LeBleu, V.S., Taduri, G., O'Connell, J., Teng, Y., Cooke, V.G., Woda, C., Sugimoto, H., and Kalluri, R. (2013). Origin and function of myofibroblasts in kidney fibrosis. Nat Med *19*, 1047-1053.
Li, G., and Reinberg, D. (2011). Chromatin higher-order structures and gene regulation. Curr Opin Genet Dev *21*, 175-186.
Li, T., Yang, D., Li, J., Tang, Y., Yang, J., and Le, W. (2015). Critical role of Tet3 in neural progenitor cell maintenance and terminal differentiation. Mol Neurobiol *51*, 142-154.
Li, X., Yao, B., Chen, L., Kang, Y., Li, Y., Cheng, Y., Li, L., Lin, L., Wang, Z., Wang, M.*, et al.* (2017). Ten-eleven translocation 2 interacts with forkhead box O3 and regulates adult neurogenesis. Nat Commun *8*, 15903.
Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923-930.
Lio, C.J., and Rao, A. (2019). TET Enzymes and 5hmC in Adaptive and Innate Immune Systems. Front Immunol *10*, 210.
Lio, C.W., Zhang, J., Gonzalez-Avalos, E., Hogan, P.G., Chang, X., and Rao, A. (2016). Tet2 and Tet3 cooperate with B-lineage transcription factors to regulate DNA modification and chromatin accessibility. Elife *5*.

Liu, R., Jin, Y., Tang, W.H., Qin, L., Zhang, X., Tellides, G., Hwa, J., Yu, J., and Martin, K.A. (2013). Ten-eleven translocation-2 (TET2) is a master regulator of smooth muscle cell plasticity. Circulation *128*, 2047-2057.
Liu, R., Leslie, K.L., and Martin, K.A. (2015). Epigenetic regulation of smooth muscle cell plasticity. Biochim Biophys Acta *1849*, 448-453.
Luo, Z., Ge, M., Chen, J., Geng, Q., Tian, M., Qiao, Z., Bai, L., Zhang, Q., Zhu, C., Xiong, Y.*, et al.* (2017). HRS plays an important role for TLR7 signaling to orchestrate inflammation and innate immunity upon EV71 infection. PLoS Pathog *13*, e1006585.
Mailleux, A.A., Kelly, R., Veltmaat, J.M., De Langhe, S.P., Zaffran, S., Thiery, J.P., and Bellusci, S. (2005). Fgf10 expression identifies parabronchial smooth muscle cell progenitors and is required for their entry into the smooth muscle cell lineage. Development *132*, 2157-2166.
Marina, R.J., Sturgill, D., Bailly, M.A., Thenoz, M., Varma, G., Prigge, M.F., Nanan, K.K., Shukla, S., Haque, N., and Oberdoerffer, S. (2016). TET-catalyzed oxidation of intragenic 5-methylcytosine regulates CTCF-dependent alternative splicing. EMBO J *35*, 335-355.
McDaniel, S.L., and Strahl, B.D. (2017). Shaping the cellular landscape with Set2/SETD2 methylation. Cell Mol Life Sci *74*, 3317-3334.
Melamed, P., Yosefzon, Y., David, C., Tsukerman, A., and Pnueli, L. (2018). Tet Enzymes, Variants, and Differential Effects on Function. Front Cell Dev Biol *6*, 22.
Milewicz, D.M., Kwartler, C.S., Papke, C.L., Regalado, E.S., Cao, J., and Reid, A.J. (2010). Genetic variants promoting smooth muscle cell proliferation can result in diffuse and diverse vascular diseases: evidence for a hyperplastic vasculomyopathy. Genet Med *12*, 196-203.
Moiseenko, A., Kheirollahi, V., Chao, C.M., Ahmadvand, N., Quantius, J., Wilhelm, J., Herold, S., Ahlbrecht, K., Morty, R.E., Rizvanov, A.A.*, et al.* (2017). Origin and characterization of alpha smooth muscle actin-positive cells during murine lung development. Stem Cells *35*, 1566-1578.
Morselli, M., Pastor, W.A., Montanini, B., Nee, K., Ferrari, R., Fu, K., Bonora, G., Rubbi, L., Clark, A.T., Ottonello, S.*, et al.* (2017). Correction: In vivo targeting of de novo DNA methylation by histone modifications in yeast and mouse. Elife *6*.
Nelson, J.D., Denisenko, O., Sova, P., and Bomsztyk, K. (2006). Fast chromatin immunoprecipitation assay. Nucleic Acids Res *34*, e2.
Neri, F., Incarnato, D., Krepelova, A., Rapelli, S., Anselmi, F., Parlato, C., Medana, C., Dal Bello, F., and Oliviero, S. (2015). Single-Base Resolution Analysis of 5-Formyl and 5-Carboxyl Cytosine Reveals Promoter DNA Methylation Dynamics. Cell Rep *10*, 674-683.
Neri, F., Rapelli, S., Krepelova, A., Incarnato, D., Parlato, C., Basile, G., Maldotti, M., Anselmi, F., and Oliviero, S. (2017). Intragenic DNA methylation prevents spurious transcription initiation. Nature *543*, 72-77.
O'Dwyer, D.N., Dickson, R.P., and Moore, B.B. (2016). The Lung Microbiome, Immunity, and the Pathogenesis of Chronic Lung Disease. J Immunol *196*, 4839-4847.
Pan, Y., Liu, L., Li, S., Wang, K., Ke, R., Shi, W., Wang, J., Yan, X., Zhang, Q., Wang, Q.*, et al.* (2018). Activation of AMPK inhibits TGF-beta1-induced airway smooth muscle cells proliferation and its potential mechanisms. Sci Rep *8*, 3624.
Papaioannou, A.I., Spathis, A., Kostikas, K., Karakitsos, P., Papiris, S., and Rossios, C. (2017). The role of endosomal toll-like receptors in asthma. Eur J Pharmacol *808*, 14-20.
Pastor, W.A., Aravind, L., and Rao, A. (2013). TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. Nat Rev Mol Cell Biol *14*, 341-356.
Pelaia, G., Renda, T., Gallelli, L., Vatrella, A., Busceti, M.T., Agati, S., Caputi, M., Cazzola, M., Maselli, R., and Marsico, S.A. (2008). Molecular mechanisms underlying airway smooth muscle contraction and proliferation: implications for asthma. Respir Med *102*, 1173-1181.
Perera, A., Eisen, D., Wagner, M., Laube, S.K., Kunzel, A.F., Koch, S., Steinbacher, J., Schulze, E., Splith, V., Mittermeier, N.*, et al.* (2015). TET3 is recruited by REST for context-specific hydroxymethylation and induction of gene expression. Cell Rep *11*, 283-294.
Portela, A., and Esteller, M. (2010). Epigenetic modifications and human disease. Nat Biotechnol *28*, 1057-1068.

Que, J., Wilm, B., Hasegawa, H., Wang, F., Bader, D., and Hogan, B.L. (2008). Mesothelium contributes to vascular smooth muscle and mesenchyme during lung development. Proc Natl Acad Sci U S A *105*, 16626-16630.
Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842.
Ramirez, F., Dundar, F., Diehl, S., Gruning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res *42*, W187-191.
Ramos-Barbon, D., Fraga-Iriso, R., Brienza, N.S., Montero-Martinez, C., Verea-Hernando, H., Olivenstein, R., Lemiere, C., Ernst, P., Hamid, Q.A., and Martin, J.G. (2010). T Cells localize with proliferating smooth muscle alpha-actin+ cell compartments in asthma. Am J Respir Crit Care Med *182*, 317-324.
Rampal, R., Alkalin, A., Madzo, J., Vasanthakumar, A., Pronier, E., Patel, J., Li, Y., Ahn, J., Abdel-Wahab, O., Shih, A.*, et al.* (2014). DNA hydroxymethylation profiling reveals that WT1 mutations result in loss of TET2 function in acute myeloid leukemia. Cell Rep *9*, 1841-1855.
Rasmussen, K.D., and Helin, K. (2016). Role of TET enzymes in DNA methylation, development, and cancer. Genes Dev *30*, 733-750.
Rawlins, E.L., Okubo, T., Xue, Y., Brass, D.M., Auten, R.L., Hasegawa, H., Wang, F., and Hogan, B.L. (2009). The role of Scgb1a1+ Clara cells in the long-term maintenance and repair of lung airway, but not alveolar, epithelium. Cell stem cell *4*, 525-534.
Rawlins, E.L., Ostrowski, L.E., Randell, S.H., and Hogan, B.L. (2007). Lung development and repair: contribution of the ciliated lineage. Proc Natl Acad Sci U S A *104*, 410-417.
Rensen, S.S., Doevendans, P.A., and van Eys, G.J. (2007). Regulation and characteristics of vascular smooth muscle cell phenotypic diversity. Neth Heart J *15*, 100-108.
Risse, P.A., Jo, T., Suarez, F., Hirota, N., Tolloczko, B., Ferraro, P., Grutter, P., and Martin, J.G. (2011). Interleukin-13 inhibits proliferation and enhances contractility of human airway smooth muscle cells without change in contractile phenotype. Am J Physiol Lung Cell Mol Physiol *300*, L958-966.
Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol *29*, 24-26.
Schlee, M., and Hartmann, G. (2016). Discriminating self from non-self in nucleic acid sensing. Nat Rev Immunol *16*, 566-580.
Shah, A.S., Farmen, S.L., Moninger, T.O., Businga, T.R., Andrews, M.P., Bugge, K., Searby, C.C., Nishimura, D., Brogden, K.A., Kline, J.N.*, et al.* (2008). Loss of Bardet-Biedl syndrome proteins alters the morphology and function of motile cilia in airway epithelia. Proceedings of the National Academy of Sciences *105*, 3380-3385.
Sharma, S., Fitzgerald, K.A., Cancro, M.P., and Marshak-Rothstein, A. (2015). Nucleic Acid-Sensing Receptors: Rheostats of Autoimmunity and Autoinflammation. J Immunol *195*, 3507-3512.
Shen, L., Song, C.X., He, C., and Zhang, Y. (2014). Mechanism and function of oxidative reversal of DNA and RNA methylation. Annu Rev Biochem *83*, 585-614.
Shiraki, T., Kondo, S., Katayama, S., Waki, K., Kasukawa, T., Kawaji, H., Kodzius, R., Watahiki, A., Nakamura, M., Arakawa, T.*, et al.* (2003). Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. Proc Natl Acad Sci U S A *100*, 15776-15781.
Shukla, S., Kavak, E., Gregory, M., Imashimizu, M., Shutinoski, B., Kashlev, M., Oberdoerffer, P., Sandberg, R., and Oberdoerffer, S. (2011). CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. Nature *479*, 74-79.
Smith, E., Lin, C., and Shilatifard, A. (2011). The super elongation complex (SEC) and MLL in development and disease. Genes & Development *25*, 661-672.
Smith, Z.D., and Meissner, A. (2013). DNA methylation: roles in mammalian development. Nat Rev Genet *14*, 204-220.
Solary, E., Bernard, O.A., Tefferi, A., Fuks, F., and Vainchenker, W. (2014). The Ten-Eleven Translocation-2 (TET2) gene in hematopoiesis and hematopoietic diseases. Leukemia *28*, 485-496.

Sood, A., Qualls, C., Schuyler, M., Arynchyn, A., Alvarado, J.H., Smith, L.J., and Jacobs, D.R., Jr. (2013). Adult-onset asthma becomes the dominant phenotype among women by age 40 years. the longitudinal CARDIA study. Ann Am Thorac Soc *10*, 188-197.
Spallotta, F., Cencioni, C., Atlante, S., Garella, D., Cocco, M., Mori, M., Mastrocola, R., Kuenne, C., Guenther, S., Nanni, S.*, et al.* (2018). Stable Oxidative Cytosine Modifications Accumulate in Cardiac Mesenchymal Cells From Type2 Diabetes Patients: Rescue by alpha-Ketoglutarate and TET-TDG Functional Reactivation. Circ Res *122*, 31-46.
Spin, J.M., Maegdefessel, L., and Tsao, P.S. (2012). Vascular smooth muscle cell phenotypic plasticity: focus on chromatin remodelling. Cardiovasc Res *95*, 147-155.
Spruijt, C.G., Gnerlich, F., Smits, A.H., Pfaffeneder, T., Jansen, P.W., Bauer, C., Munzel, M., Wagner, M., Muller, M., Khan, F.*, et al.* (2013). Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. Cell *152*, 1146-1159.
Stamatiou, R., Paraskeva, E., Gourgoulianis, K., Molyvdas, P.A., and Hatziefthimiou, A. (2012). Cytokines and growth factors promote airway smooth muscle cell proliferation. ISRN Inflamm *2012*, 731472.
Takahashi, H., Lassmann, T., Murata, M., and Carninci, P. (2012). 5' end-centered expression profiling using cap-analysis gene expression and next-generation sequencing. Nat Protoc *7*, 542-561.
Tan, L., Xiong, L., Xu, W., Wu, F., Huang, N., Xu, Y., Kong, L., Zheng, L., Schwartz, L., Shi, Y.*, et al.* (2013). Genome-wide comparison of DNA hydroxymethylation in mouse embryonic stem cells and neural progenitor cells by a new comparative hMeDIP-seq method. Nucleic Acids Res *41*, e84.
Tata, P.R., and Rajagopal, J. (2017). Plasticity in the lung: making and breaking cell identity. Development *144*, 755-766.
Teissandier, A., and Bourc'his, D. (2017). Gene body DNA methylation conspires with H3K36me3 to preclude aberrant transcription. EMBO J *36*, 1471-1473.
Tipping, P.G. (2006). Toll-like receptors: the interface between innate and adaptive immunity. J Am Soc Nephrol *17*, 1769-1771.
Travillian, R.S., Gennari, J.H., and Shapiro, L.G. (2005). Of mice and men: design of a comparative anatomy information system. AMIA Annu Symp Proc, 734-738.
Tsagaratou, A., Aijo, T., Lio, C.W., Yue, X., Huang, Y., Jacobsen, S.E., Lahdesmaki, H., and Rao, A. (2014). Dissecting the dynamic changes of 5-hydroxymethylcytosine in T-cell development and differentiation. Proc Natl Acad Sci U S A *111*, E3306-3315.
Tsagaratou, A., Gonzalez-Avalos, E., Rautio, S., Scott-Browne, J.P., Togher, S., Pastor, W.A., Rothenberg, E.V., Chavez, L., Lahdesmaki, H., and Rao, A. (2017). TET proteins regulate the lineage specification and TCR-mediated expansion of iNKT cells. Nat Immunol *18*, 45-53.
Tsagaratou, A., and Rao, A. (2013). TET proteins and 5-methylcytosine oxidation in the immune system. Cold Spring Harb Symp Quant Biol *78*, 1-10.
Turvey, S.E., and Broide, D.H. (2010). Innate immunity. J Allergy Clin Immunol *125*, S24-32.
Veloso, A., Kirkconnell, K.S., Magnuson, B., Biewen, B., Paulsen, M.T., Wilson, T.E., and Ljungman, M. (2014). Rate of elongation by RNA polymerase II is associated with specific gene features and epigenetic modifications. Genome Res *24*, 896-905.
Venkatesh, S., and Workman, J.L. (2013). Set2 mediated H3 lysine 36 methylation: regulation of transcription elongation and implications in organismal development. Wiley Interdiscip Rev Dev Biol *2*, 685-700.
Volckaert, T., and De Langhe, S.P. (2015). Wnt and FGF mediated epithelial-mesenchymal crosstalk during lung development. Dev Dyn *244*, 342-366.
Wagner, E.J., and Carpenter, P.B. (2012). Understanding the language of Lys36 methylation at histone H3. Nat Rev Mol Cell Biol *13*, 115-126.
Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., and Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell *153*, 910-918.

Wang, L., Zhou, Y., Xu, L., Xiao, R., Lu, X., Chen, L., Chong, J., Li, H., He, C., Fu, X.D., *et al.* (2015). Molecular basis for 5-carboxycytosine recognition by RNA polymerase II elongation complex. Nature *523*, 621-625.
Weber, A.R., Krawczyk, C., Robertson, A.B., Kusnierczyk, A., Vagbo, C.B., Schuermann, D., Klungland, A., and Schar, P. (2016). Biochemical reconstitution of TET1-TDG-BER-dependent active DNA demethylation reveals a highly coordinated mechanism. Nat Commun *7*, 10806.
Wen, L., Li, X., Yan, L., Tan, Y., Li, R., Zhao, Y., Wang, Y., Xie, J., Zhang, Y., Song, C., *et al.* (2014). Whole-genome analysis of 5-hydroxymethylcytosine and 5-methylcytosine at base resolution in the human brain. Genome Biol *15*, R49.
Wierda, R.J., Geutskens, S.B., Jukema, J.W., Quax, P.H., and van den Elsen, P.J. (2010). Epigenetics in atherosclerosis and inflammation. J Cell Mol Med *14*, 1225-1240.
Williams, K., Christensen, J., Pedersen, M.T., Johansen, J.V., Cloos, P.A., Rappsilber, J., and Helin, K. (2011). TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature *473*, 343-348.
Wossidlo, M., Nakamura, T., Lepikhov, K., Marques, C.J., Zakhartchenko, V., Boiani, M., Arand, J., Nakano, T., Reik, W., and Walter, J. (2011). 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. Nat Commun *2*, 241.
Wu, H., D'Alessio, A.C., Ito, S., Xia, K., Wang, Z., Cui, K., Zhao, K., Sun, Y.E., and Zhang, Y. (2011). Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. Nature *473*, 389-393.
Wu, H., Wu, X., Shen, L., and Zhang, Y. (2014). Single-base resolution analysis of active DNA demethylation using methylase-assisted bisulfite sequencing. Nat Biotechnol *32*, 1231-1240.
Wu, X., and Zhang, Y. (2017). TET-mediated active DNA demethylation: mechanism, function and beyond. Nat Rev Genet *18*, 517-534.
Xie, L., Pelz, C., Wang, W., Bashar, A., Varlamova, O., Shadle, S., and Impey, S. (2011). KDM5B regulates embryonic stem cell self-renewal and represses cryptic intragenic transcription. EMBO J *30*, 1473-1484.
Xu, Y., Wu, F., Tan, L., Kong, L., Xiong, L., Deng, J., Barbera, A.J., Zheng, L., Zhang, H., Huang, S., *et al.* (2011). Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. Mol Cell *42*, 451-464.
Xue, S., Liu, C., Sun, X., Li, W., Zhang, C., Zhou, X., Lu, Y., Xiao, J., Li, C., Xu, X., *et al.* (2016). TET3 Inhibits Type I IFN Production Independent of DNA Demethylation. Cell Reports *16*, 1096-1105.
Yao, B., Christian, K.M., He, C., Jin, P., Ming, G.L., and Song, H. (2016). Epigenetic mechanisms in neurogenesis. Nat Rev Neurosci *17*, 537-549.
Yu, H., Su, Y., Shin, J., Zhong, C., Guo, J.U., Weng, Y.L., Gao, F., Geschwind, D.H., Coppola, G., Ming, G.L., *et al.* (2015). Tet3 regulates synaptic transmission and homeostatic plasticity via DNA oxidation and repair. Nat Neurosci *18*, 836-843.
Yu, M., Hon, G.C., Szulwach, K.E., Song, C.X., Zhang, L., Kim, A., Li, X., Dai, Q., Shen, Y., Park, B., *et al.* (2012). Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. Cell *149*, 1368-1380.
Zakeri, A., Hansen, E.P., Andersen, S.D., Williams, A.R., and Nejsum, P. (2018). Immunomodulation by Helminths: Intracellular Pathways and Extracellular Vesicles. Front Immunol *9*, 2349.
Zentner, G.E., and Henikoff, S. (2013). Regulation of nucleosome dynamics by histone modifications. Nat Struct Mol Biol *20*, 259-266.
Zepp, J.A., and Morrisey, E.E. (2019). Cellular crosstalk in the development and regeneration of the respiratory system. Nat Rev Mol Cell Biol *20*, 551-566.
Zerbino, D.R., Johnson, N., Juettemann, T., Wilder, S.P., and Flicek, P. (2014). WiggleTools: parallel processing of large collections of genome-wide datasets for visualization and statistical analysis. Bioinformatics *30*, 1008-1009.

Zhang, Q., Zhao, K., Shen, Q., Han, Y., Gu, Y., Li, X., Zhao, D., Liu, Y., Wang, C., Zhang, X.*, et al.* (2015). Tet2 is required to resolve inflammation by recruiting Hdac2 to specifically repress IL-6. Nature *525*, 389-393.
Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.
Zhang, Y.W., Wang, Z., Xie, W., Cai, Y., Xia, L., Easwaran, H., Luo, J., Yen, R.C., Li, Y., and Baylin, S.B. (2017). Acetylation Enhances TET2 Function in Protecting against Abnormal DNA Methylation during Oxidative Stress. Mol Cell *65*, 323-335.
Zhong, J., Li, X., Cai, W., Wang, Y., Dong, S., Yang, J., Zhang, J., Wu, N., Li, Y., Mao, F.*, et al.* (2017). TET1 modulates H4K16 acetylation by controlling auto-acetylation of hMOF to affect gene regulation and DNA repair function. Nucleic Acids Res *45*, 672-684.
Zhu, T., Brown, A.P., and Ji, H. (2020). The Emerging Role of Ten-Eleven Translocation 1 in Epigenetic Responses to Environmental Exposures. Epigenet Insights *13*, 2516865720910155.
Zhu, X., Girardo, D., Govek, E.-E., John, K., Mellén, M., Tamayo, P., Mesirov, Jill P., and Hatten, Mary E. (2016). Role of Tet1/3 Genes and Chromatin Remodeling Genes in Cerebellar Circuit Formation. Neuron *89*, 100-112.

APPENDIX

Abbreviations

%	Percent
&	And
°C	Degree Celsius
5caC	5-carboxycytosine
5fC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
Acetyl-CoA	Acetyl Coenzyme-A
ACTA2	smooth muscle α-actin
ADP	Adenosine triphosphate
AP-1	activator protein-1
APS	Ammonium persulfate
Arhgap18	Rho GTPase Activating Protein 18
Asf1	histone chaperone ASF1
ASM	airway smooth muscle
ASMC	airway smooth muscle cell
ATP	Adenosine triphosphate
BALF	bronchoalveolar lavage fluid
BER	base excision repair
bp	Base pairs
BSA	Bovine serum albumin
CAMs	cell adhesion molecules
CCL21	CC-chemokine ligand 21
CCSP	Club cell secrete protein
CD4	cluster of differentiation 4
CD8	cluster of differentiation 8
CF	Cystic fibrosis
CFTR	CF transmembrane conductance regulator
ChIP	Chromatin immunoprecipitation qPCR
ChIP	Chromatin immunoprecipitation
cm	Centimeter
CMV	Cytomegalovirus
Co-IP	Co-immunopricipitation
COPD	chronic obstructive pulmonary disease
CpG	Cytosine-guanine dinucleotide
CTCF	CCCTC-binding factor
CXCL10	interferon-g-inducible protein-10
DAB	3,3'-DIAMINOBENZIDIN TETRAHYDROCHLORID

DAMPs	damage-associated molecular patterns
DAPI	4',6-diamidino-2-phenylindole
DAPI	4',6-diamidino-2-phenylindole
DC	dendritic cells
DKO	double knockout
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
DRB	5,6-dichloro-1-β-D-ribofuranosyl-1H-benzimidazole
dsRNA	double-stranded RNA
DTT	1,4-Dithiothreit
Eaf3	chromodomain protein mortality factor 4-like protein 1
ECM	extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EEA1	early endosome antigen 1
EGTA	Ethylene glycol-bis (2-aminoethylether)-N N N'N'-tetraacetic acid
ESC	embryonic stem cell
FACS	Fluorescence-activated cell sorting
FCS	Fetal calf serum
FGF10	fibroblast growth factor 10
FoxJ1	Forkhead box protein J1
g	Gram
G	Guanine
g	Gravitational acceleration
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
h	Hour
H&E	Hematoxylin and eosin
H3K18ac	H3 lysine 18 acetylation
H3K27me3	Histone H3 trimethyl Lysine 27
H3K36	Histone H3 lysine 36
H3K36me3	Histone H3 trimethyl Lysine 36
H3K36me3	Histone H3 trimethyl Lysine 36
H3K79	Histone H3 Lysine 79
H3K9me2	Histone H3 dimethyl Lysine 9
H3K9me3	Histone H3 trimethyl Lysine 9
H3R8me2s	Histone H3 symetrical dimethyl Arginine 8
H4K20me1	Histone H4 monomethyl Lysine 20
H4K20me2	Histone H4 dimethyl Lysine 20
H4K20me3	Histone H4 trimethyl Lysine 20
H4K4me3	Histone H4 trimethyl Lysine 4
HA	Hemagglutinin
HATs	histone acetyltransferases

hCASMCs	human coronary artery SMCs
HDACs	histone deacetylases
HDM	house dust mites
hMeDIP	hydroxymethylcytosine immunoprecipitation qPCR
hMOF	histone H4 lysine K16 specific acetyltransferase
HSC	hematopoietic stem cells
IAV	influenza A virus
ICF	facial anomalies syndrome
IF	Immunofluorescence
IFN beta	Interferon beta
IFNγ	Interferon γ
IL-17A	Interleukin 17A
IL-1b	interleukin 1b
IL-4	Interleukin 4
IL-5	Interleukin 5
IL-6	Interleukin 6
IL-8	Interleukin 8
IL-9	Interleukin 9
IL13	Interleukin 13
IRFs	IFN (interferon) regulatory factors
kb	kilo-base pairs
KD	knock-down
kDa	kilodalton
KDM5B	lysine specific demethylase 5B
kg	kilogram
KLF4	Kruppel-like factor 4
KO	Knock-out
KRT5	cytokeratin 5
M	molar
MAPKs	mitogen-activated protein kinases
MBD1	methyl-CpG binding domain protein 1
MCC	mucociliary clearance
MECP2	methyl CpG binding protein 2
MeDIP	methylcytosine immunoprecipitation qPCR
min	Minute
miR-145	microRNA-145
ml	milliliter
MLL	mixed lineage leukemia
mm	millimeter
mM	millimolar
MRG15	transcription factor like protein MRG15
mRNA	messenger RNA
MUC5AC	Mucin 5AC
MUC5B	Mucin 5B

MYD88	myeloid differentiation primary response 88
Myh10	myosin heavy chain 10
MYH11	myosin heavy chain 11
Myh7	myosin heavy chain 7
MYOCD	myocardin
ncRNAs	noncoding RNAs
NF-κB	nuclear factor κB
NGFR	nerve growth factor receptor
NK	natural killer
NP-40	Tergitol-type NP-40
NPCs	neuronal progenitor cells
NSD	nuclear receptor-binding SET domain
NTP	nucleoside triphosphate
O/N	over night
O-GlcNAc	O-linked β -D-N-acetylglucosamine
OGT	O-linked β-D-N-acetylglucosamine (O-GlcNAc) transferase
P300	E1A binding protein p300
PAMPs	pathogen-associated molecular patterns
PAS	Periodic Acid-Schiff
PCR	Polymerase chain reaction
PDGF	platelet-derived growth factor
Pdlim3	PDZ and LIM domain 3
PDPN	podoplanin
PFA	Paraformaldehyde
PGCs	primordial germ cells
PH	pulmonary hypertension
PLA	proximity ligation assay
PMSF	Phenyl methyl sulfonyl fluoride
Pol II	RNA Polymerase II
PRC2	Polycomb repressive complex 2
PRRs	pattern recognition receptors
qPCR	quantitative PCR
Rbp1	retinol binding protein 1
rDNA	Ribosomal DNA
REST	RE1-silencing transcription factor
RSV	respiratory syncytial virus
RT	Room temperature
RT-qPCR	Quantitative reverse transcription polymerase chain
s	Second
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SetD2	SET Domain Containing 2
SETD2	SET domain containing 2
SIN3A	SIN3 transcription regulator family member A

SM22α	TAGLN, Transgelin
SMC	smooth muscle cell
Spt6	transcription elongation factor
SRF	serum response factor
SRF	serum response factor
ssRNA	single-stranded RNA
Susd2	sushi domain containing 2
Synpo2	Synaptopodin 2
TAM	tamoxifen
TBS	Tris-Buffered Saline
TDG	thymine DNA glycosylase
TEMED	Tetramethylethylenediamine
TET	Ten to eleven translocation dioxygenase
TET3	Ten to Eleven Translocation dioxygenase 3
TGF-b1	transforming growth factor b1
TH17	T helper 17 cells
TH2	T help type 2 cell
TIR	Toll-interleukin-1 receptor
TLR7	Toll-like receptor 7
TLR8	Toll-like receptor 8
TLRs	toll-like receptors
TNF-α	tumor necrosis factor alpha
Tpm4	tropomyosin 4
TRIF	TIR-containing adaptor-inducing IFN
TSSs	transcription starts sites
Tween 20	Polyoxyethylene-sorbitan-monolaurate
Vim	vimentin
VSMC	vascular smooth muscle cell
WB	Western blot
YY1	Ying-yang1
µg	micro gram
µl	micro liter
µM	micro molar
µm	micro meter

List of figures

Acknowledgments

First, I would like to express my gratitude to my Ph.D supervisor Prof. Dr. Dr. Thomas Braun for providing me this invaluable opportunity to pursue my Ph.D study in such an excellent scientific research center. He encouraged me at the first day of my Ph.D study by a sentence "You will have a lot of data then", and he supported and guided throughout my time for motivating me to think and develop ideas to my project. His excellent knowledge impressed me deeply and inspired me to develop a scientific perspective to cellular and molecular biology. I am grateful to my former supervisor Dr. Yonggang Zhou for encouraging me to find the value and potential of myself towards scientific work. He had taught me how to conduct scientific research by carefully and cautiously thinking about every detail, express ideas by confidently and logically presenting the project and answering questions, and unveil the behind of negative results by developing inspiring ideas to complex questions. His spirits towards science motivated me to passionately overcome many critical situations. I would like to express my gratitude to my Ph.D supervisor Dr. Xuejun Yuan for her constant support and advice at various stages of this work, and for thorough and constructive criticism. She is always a friend and a guide, providing me unconditional guidance, encouragement, and advice whenever I needed any help during my study. This work was built and facilitated by every discussion with her, which inspired and encouraged me to accomplish it by engraving every single angle to its perfection. Her patience, diligence and toughness taught me how to bravely confront and deal with every difficulty of this study. I express my appreciation to Prof. Dr. Reinhard Dammann for his willingness to be my second corrector and for his valuable advice as my thesis committees. Thanks to all AG Braun members for a very supporting and inspiring atmosphere. Many thanks to my group members/friends, Dr. Ting Zhang, Dr. Hui Qi, Dr. Angelina Georgieva, Hang Liu, Xinyue Guo, Dr. Dong Ding for the friendly working atmosphere especially Ting for my first training for molecular cloning and methodological training. Thanks to my dear boyfriend, Xiang Li, for his helpful discussion, cuteness and love alleviating all the bitter during this study. I would also like to express my appreciation to Dr. Mario Looso for the bioinformatic analysis.

Finally, I would like to express my gratitude to my parent and my grandparent, for the understanding, support and agreement with sending their only daughter and granddaughter abroad to get the education she wanted. I dedicate this dissertation to them.

Curriculum Vitae

FAN WU

Dept. I - Cardiac Development and Remodelling	Phone: +49(0) 6032 7051190
Max Planck Institute for Heart and Lung Research	Mobile: +49(0) 176 80758301
Ludwigstrasse 43	E-mail: Fan.Wu@mpi-bn.mpg.de
D-61231 Bad Nauheim	

EDUCATION

2015-present	Graduate student, in Dept. I, Max-Planck Institute for heart and lung research Bad Nauheim, Germany.
Project:	Unraveling epigenetic mechanisms underlying lung pathogenesis
2012-2014	**M.S.,** major in Physiology, Department of Physiology, LKS Faculty of Medicine, The University of Hong Kong SAR, China.
Project:	Exercise ameliorates high fat diet induced cardiac hypertrophy
2006-2011	**B.S.,** major in Clinical Pharmacy, School of Pharmacy, Harbin Medical University, China.

PUBLICATIONS

Xuejun Yuan, Hui Qi, Xiang Li, Fan Wu, Jian Fang, Eva Bober, Gergana Dobreva, Yonggang Zhou, Thomas Braun: Disruption of spatiotemporal hypoxia signaling causes congenital heart disease in mice, *J Clin Invest. 2017*

MEMBERSHIPS

2015-present	International Collaborative Research Centre - SFB TRR81 (Transregional Collaborative Research Centre 81) Chromatin Changes in Differentiation and Malignancies
2015-2017	The International Max Planck Research School for Heart and Lung Research (IMPRS-HLR)

CONFERENCES AND MEETING

International Symposium 2017 "Chromatin Changes in Differentiation and Malignancies" September 18- 20, 2017 Egmond aan Zee, The Netherlands –poster

Summer School on Chromatin and Metabolism, August 25-31, 2018, Spetses Island, Greece – poster

11th TRR81 Mini-symposium on "Epigenetics & Metabolism", November 28th, 2018, Bad Nauheim, Germany –organizer

EMBL Symposia "Metabolism Meets Epigenetics", November 20-23, 2019, Heidelberg,

Germany –attended

EMBL Conference "Chromatin and Epigenetics", May 17-20, 2021, Virtual - poster