REPROGRAMMING MOUSE EMBRYONIC FIBROBLASTS INTO FUNCTIONAL THYMIC EPITHELIAL CELLS AND ITS MECHANISM ANALYSIS

By

ZHONGYAO MA

(Under the Direction of Nancy R. Manley and Brian Condie)

ABSTRACT

The thymus is a primary lymphoid organ, and the major source of self-restricted, self-tolerant naïve T cells required for a robust adaptive immune system. However, through aging, the thymus is one of the earliest organs that starts losing its function. Thymus involution leads to decreased immune function, which significantly increases the risk of diseases. Thus, finding an method to rescue thymus function caused by thymus involution & thymus abnormalities is significant.

We previously demonstrated that by over-expressing transcription factor FOXN1 in mouse embryonic fibroblasts (MEFs), we could successfully reprogram MEFs into induced functional thymus epithelial cells (iTECs). The iTECs express many thymus related genes such as Dll4, Ccl25, and Krt5. Also, the reprogrammed iTECs successfully support T cell development both in vitro & in vivo. However, the detailed mechanisms of iTECs reprogramming and how similar it is compared to fetal TECs are still being uncovered.

This dissertation presents a detailed analysis of the reprogramming mechanisms during iTEC reprogramming, including a gene expression profile comparison to E14.5 fetal TECs. I utilized bulk RNA-sequencing to re-construct gene expression profiles through the different time

points of reprogramming. I analyzed these RNA-sequencing results and show that the iTECs reprogramming process acts in a step by step manner, in which the early stage reprogramming cell population and the late-stage reprogramming cell population are very distinct. I also identified and tested specific pathways that limit iTEC reprogramming, such as the Notch pathway and cell cycle-related genes. By modifying these pathways, we can make iTECs turn on the essential medullary transcription factor Aire, or increase the reprogramming efficiency and rescue the cell cycle arrest phenotype. Furthermore, I performed a gene expression profile comparison between different stages of iTEC reprogramming and E14.5 fetal TECs. This analysis suggests even the late-stage iTECs are still incompletely reprogrammed compared to fetal TECs. This analysis also identified several pathways that may be potential targets to increase the similarity between iTECs and fetal TEC gene expression. Finally, I discuss possible ways to generate iTECs using non-transgenic CRISPR-dCAS9 methods that would be potentially useful for future pre-clinical applications in both mouse and human cells.

INDEX WORDS: Thymus, Foxn1, iTECs, Reprogramming, Cell Cycle, Crispr-dCAS9, Thymus Involution.

REPROGRAMMING MOUSE EMBRYONIC FIBROBLAST INTO FUNCTIONAL THYMICEPITHELIAL CELLS AND ITS MECHANISM ANALYSIS

By

ZHONGYAO MA

B.S, Shanghai Jiao Tong University, China,2014

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2020

© 2020

ZHONGYAO MA

All Rights Reserved

REPROGRAMMING MOUSE EMBRYONIC FIBROBLAST INTO FUNCTIONAL

THYMICEPITHELIAL CELLS AND ITS MECHANISM ANALYSIS

by

ZHONGYAO MA

Major Professors: Nancy Manley, Brian Condie
Committee: James D Lauderdale
Doug Menke
Kimberly D Klonowski

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
December 2020

DEDICATION

I dedicate this dissertation to my wife, Kuangyan Chen, who has been supporting me with love through all these years in United states. You are an amazing mother to our son, Evan Ma. It has been a long journey for our family and I am very grateful to you for staying with me and supporting me. Thanks you so much for come to United states with me, where we get married and also have our pretty boy Evan. It's you that help me grow up and help me find the most important thing in my life. I also dedicate this to my little boy Evan Ma. You bring me so much joy even sometimes I was too busy to play with you. Even we were separated by COVID-19 pandemic, you are still accompany me via video and give me so much power to finish my Ph.D.

I would also dedicate this to my parents, Wei Ma and Jue Gao. I thank you for all the love and support these years and help me make up my mind to pursure my Ph.D career after my college.

ACKNOWLEDGEMENTS

I would like to thank my mentor and advisor Nancy Manley for providing me with her expertise and wisdom during my graduate research. Thanks you for supporting my research and providing me the opportunities to learn all the differernet techniques so that I can understand not only Genetics, but also cell bioloby and molecular biology. I have no idea of this project at beginning and your patience and guidance is the reason why I could develop my ability to independently design and execute research to finish it. Also, due to the pandemic, thanks so much for helping my defense in China and solve all those difficulties that I encountered. I would also like to thank my co-advisor, Brian condie, for all those brilliant ideas that help me solve so many issues in my research and inspire me of new method and technology to finish my project.

I would also like to thankd the rest of my committee members: Doug Menke, Jim Lauderdale, and Kimberly D Klonowski. Our conversation during committee meeting helped me so much for my professional development as well as my project design. All those advices that I got even out of committee meeting really helps me for accomplish my research goals. I would also like to thanks all current and previous members of the Manley/Condie lab, especially, Shiyun Xiao, Jie Li, Wen Zhang. These great scientist provided me so much help and training on lab techniques so that I can become a good scientist. Also, Julie Nelson in Flow cytometry core facility helps me so muchon flow cytomertry which was a crucial piece of my work and research. I must also thank Joelle Szendel for helping me so much on administrative lab staff.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS ....# **v**

LIST OF FIGURES ....# viii

CHAPTER

1 INTRODUCTION AND LITERATURE REVIEW ....# 1

Development and organization of the thymus ....# 1

T cell maturation in the thymus ....# 3

Hoxa3, the earliest player which regulates initial organ formation ....# 5

Foxn1, the key TEC-specific transcription factor ....# 6

The Notch pathway plays a critical role in medullary TEC development.......# 8

TEC differentiation and proliferation ....# 8

Thymus involution ....#10

Mechanisms to rescue thymus involution .... #11

Directly reprogramming of fibroblasts into thymic epithelial cells by Foxn1#12

Dissecting direct reprogramming using RNA-seq process............................#12

References ....#15

2 Dissecting direct reprogramming from fibroblast to induced thymic epithelial cell (iTECs) using RNA-seq ....#23

Abstract ....#24

Introduction ....#25

Results..........#27

Discussion..........#37

Reference..........#41

Figures and Figure Legends..........#42

3 Correlation Between Transgenic Foxn1-GFP Reporter and Endogenous Foxn1 Gene Expression differs in fetal & postnatal stages..........#51

Introdction..........#51

Results..........#53

Discussion..........#59

Figures and legends..........#62

4 Activating Endogenous Foxn1 using CRISPRa system and its application in non-transgenic iTECs reprogramming..........#70

Introduction..........#70

Results..........#73

Discussion..........#77

References..........#80

Figures and Figure Legends..........#82

5 DICUSSION AND FUTURE DIRECTION..........#91

APPENDICES

A CRISPR-based lenti virus transfection on MEFs..........#88

LIST OF FIGURES

Chapter 2:

Page

Figure 2.1: Generation and charateritics of R26CAG-Foxn1-IRES-GFP/+ MEFs. ....#45

Figure 2.2: Schematic diagram of RNA-seq experiment design and geneal transcriptome analysis ....#46

Figure 2.3: Detailed gene expression comparision analyisis of different stage of iTECs reprogramming..#47

Figure 2.4: Cell cycle analysis of iTECs and cell cycle rescue using iMYC transgen. ...#48

Figure 2.5: Notch pathway analyisis and DAPT treamtn experiment ....#49

Figure 2.6: Trancriptome and differential gene expression analyisis with E14.5 Fetal TECs ....#50

Chapter 3:

Figure 3.1: Foxn1-EGFP mice characterization ....#64

Figure 3.2: GFP reporter analyisis by immunohistochemistry in Foxn1-EGFP mice. ....#65

Figure 3.3: GFP reporter analyisis by immunohistochemistry in Foxn1-EGFP mice at late postnatal stages ....#66

Figure 3.4: Correlation between GFP signal and TECs identidy decrease as mice aging ....#67

Figure 3.5: T lymphocyte is not affected in Foxn1-EGFP transgene mice......................#68

Figure 3.6: Foxn1-egfp reporter analhyisis by immunohistochemistry in Foxn1-EGFP;Foxn1-Cre;CAG-floxp-tomato mice ....#69

Chapter 4:

Figure 4.1: CRISPR-activation system design diagram and endogenous Foxn1 activation ....#84

Figure 4.2: Track endogenous Foxn1 expression using Foxn1-EGFP reporter and endo-Foxn1 activated MEF morphology analysis ....#85

Figure 4.3: Epithelial cell marker expression and Foxn1 intron CRISPR-activation test ....#86

Figure 4.4: CRISPR-activation on human dermal fibroblast...#87

Appendix Figure:

Figure 6.1 Schematic diagram of lenti-virus activation of endogenous Foxn1 ....#90

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

My dissertation research was to designed to further investigate the mechanisms of reprogramming fibroblasts to an induced thymic epithelial cells (iTEC) fate. This analysis included a comparison of the iTEC gene expression profile to the expression profile in fetal TECs and methods to improve both reprogramming efficiency and similarity to functional fetal TECs. The following literature review will first cover thymus function and the thymus involution process to highlight the significance of developing iTEC reprogramming. I then present current information about the critical TEC-specific transcription factor Foxn1, which serves as a pioneer factor in iTEC reprogramming, as well as other important transcription factors including HOXA3. Finally, I discuss previous iTEC reprogramming research as well as other studies investigating reprogramming process mechanisms.

Function and development of the thymus

Development and organization of the thymus

The thymus is one of the primary lymphoid organs and the source of self-restricted, self-tolerant naïve T cells required for a robust adaptive immunity system. In mice, the third pair of pharyngeal pouches each forms a single epithelial organ primordium, which later develops into one thymus lobe and one parathyroid gland. (Gordon et al., 2004). A previous study concluded that endoderm and ectoderm both contribute to the thymic epithelium, the "dual-origin" model for thymus. (van Vliet, Jenkinson, Ton, Owen, & van Ewrjk, 1985). However, lineage-tracing studies

from the Manley lab showed that ectoderm did not contribute to the thymus and concluded that all TEC are of endodermal origin [Gordon, et al. 2004]. More recent studies find that cultured epithelial cells from embryonic day E12 mouse thymi (S. W. Rossi, Jenkinson, Anderson, & Jenkinson, 2006) or in vivo postnatal thymi (C. C. Bleul et al., 2006) have both cortical and medullary potential. Together with other findings, these data support a single endodermal origin for the thymic epithelium, the "endoderm-centric" model.

The thymus has two major distinct functional regions, the cortex and the medulla, each of which plays a unique and important role in T cell development and maturation. The cortex is the thymus's outer region, which supports the positive selection of T cell development, while the medulla is the inner region of the thymus, which supports negative selection. (Gordon & Manley, 2011). More than 99% of the cells in the young adult mouse thymus are T cells, and the remaining stromal component of the organ comprises thymus epithelial cells (TECs), dendritic cells, macrophages, fibroblasts, and endothelial cells. Surprisingly, all these stromal components provide an essential microenvironment for T cell development and interactions. However, these stromal elements themselves are dependent on the presence of lymphoid cells to differentiate and maintain their integrity, which suggests there is a 'cross-talk' between thymocytes and thymus stromal cells. (van Ewijk, Shores, & Singer, 1994)

Among these stromal cells, the thymic epithelial cells (TECs) are the major functional compartment that supports T cell development in both the cortex (cTECs) and medulla (mTECs). cTECs can be identified by the expression of beta5t, a proteasome component that plays a pivotal role in generating an immunocompetent repertoire of T cells (Ripen, Nitta, Murata, Tanaka, & Takahama, 2011). cTECs can also be recognized by their expression of intermediate filament (IF) proteins Keratin 8 and Keratin18, which are required to maintain the architectural

structure in thymus epithelium. (Odaka et al., 2013), and other markers such as CD205, a C-type lectin (Jenkinson, Nakamura, White, Jenkinson, & Anderson, 2013) and LY51(Gray, Chidgey, & Boyd, 2002). mTECs have distinct markers such as keratin 5 and keratin 14 (Lomada, Liu, Coghlan, Hu, & Richie, 2007) and bind the lectin Ulex Europe (UEA-1). mTEC progenitors can be identified by tight-junction components claudin-3 and claudin-4. (Hamazaki et al., 2007). A subset of mTECs also expresses the autoimmune regulator Aire, which functions as the transcriptome controller in mTECs for self-antigen expression(Speck-Hernandez et al., 2018). cTECs and mTECs are both required to develop a self-restricted, self-tolerant immune system by producing mature T cells.

T cell maturation in the thymus

T cell maturation in the thymus consists of several processes that require the dynamic migration of developing lymphocytes into, within, and out of the multiple microenvironments within the thymus (Takahama, 2006). Lymphoid progenitor cells first enter the thymus as early as embryonic day 11.5(E11.5) in mice (Owen & Ritter, 1969). After ~E14.5 and in the postnatal thymus, T cell progenitors enter the thymus at the cortical-medullary junction through the vasculature(Lind, Prockop, Porritt, & Petrie, 2001). This T cell progenitor entry process is regulated by the adhesive interaction between platelet (p)-selectin glycoprotein ligand 1 (PSGL1) expressed by T cell progenitor cells and P-selectin, which is expressed by the thymic endothelium. (F. M. V. Rossi et al., 2005) After entering the thymus, these T cell progenitors are broadly characterized by the expression of 4 different markers, CD4, CD8, CD44, and CD25. The first T cell population that enters the thymus is CD4-CD8-CD44+CD25- (DN1). They then migrate from the cortex-medullary junction to the thymic cortex's subcapsular region (Peschon, 1994). These migrating thymocytes become CD44+CD25+ (DN2) due to the interaction between the Notch1 receptor and TEC expressed Notch ligand, DLL4 (Hozumi et al., 2008). Several

chemokine receptors including CXCR4, CCR7, and CCR9 play crucial roles in this process, and deletion of these chemokine receptors leads to inefficient migration of DN thymocytes (Plotkin, Prockop, Lepique, & Petrie, 2003). The DN thymocytes begin to rearrange their TCRb locus. (Muegge, Vila, & Durum, 1993). T cells that succeeded in VDJ rearrangement will further develop into DP(CD4+CD8+) thymocytes and start migrating towards the medulla. During this process, the DP thymocytes interact via the TCR with peptide-MHC complexes expressed by cTECs. Only those DP thymocytes that have low-avidity interaction between TCR and the MHC complex will receive survival signals and further develop into single-positive (SP) thymocytes. This process, called positive selection, happens in the cortex and selects for T cells that could react to foreign antigens but will not react to self-MHC antigens. (Egerton, Scollay, & Shortman, 1990). Positively selected DP thymocytes are induced to differentiate into SP (either CD4+CD8- or CD4-CD8+) thymocytes, which migrate into the medulla and begin the negative selection process, which is required to establish central tolerance. During the negative selection process, one key gene is the autoimmune regulator, Aire, which regulates the expression of many proteins that are not typically expressed in thymic cells to identify and delete T cells with reactivity to self-peptides. (Zuklys et al., 2000)

To sum up, T cell progenitors derived from HSC enters the thymus from the cortical-medullary junction, experience a massive proliferation, further mature from DN thymocytes into DP thymocytes in the cortex, and later develop into SP positive thymocytes. By interaction with both cTECs and mTECs, these T cell progenitors receive all necessary survival and proliferation signals while they also undergo the positive selection in the cortex and negative selection in the medulla, and finally become function T cells that comprise our adaptive immune system.

Transcriptional Regulation of Thymus Organogenesis and Thymic Epithelial Cell Differentiation

Transcriptional regulatory networks and signaling pathways are the central regulatory mechanisms that control organ identity, patterning, and differentiation. During thymus organogenesis and fetal development, several transcription factors play crucial roles, including HOXA3, EYA1, PAX1, and PAX9 which are expressed in the 3rd pouch before thymus fate is established. These are followed by FOXN1, which plays multiple fundamental roles in TEC differentiation and proliferation. Additionally, several signaling pathways play necessary roles in different stages of thymus development and TEC differentiation. The Notch signaling pathway has been recently shown to control mTEC maturation and differentiation in the fetal thymus. At later stages, the RB signaling pathway controls postnatal TEC proliferation.

Hoxa3, the earliest player which regulates initial organ formation

HOXA3 is a transcription factor with a very early expression pattern, and is turned on at E9.5 at the third pharyngeal pouch endoderm and surrounding mesenchyme (N. R. Manley & Capecchi, 1995). *Hoxa3* mutants have multiple abnormalities in the 3rd and 4th pharyngeal arches, including deletion of the thymus and parathyroid (Prockop et al., 2002) (Kameda, Arai, Nishimaki, & Chisaka, 2004). The mesenchymal neural crest cell's amount and migration pattern surrounding the 3rd pouch where thymus forms are normal when *Hoxa3* is deleted (Nancy R. Manley & Capecchi, 1998) suggesting that the *Hoxa3* gene may affect the neural crest cells' capacity to induce surrounding tissues to differentiate appropriately. Other experiments show that *Hoxa3* appears to have primary functions in both endoderm and neural crest, as conditional deletion of *Hoxa3* in either tissue causes distinct organogenesis defects (Chojnowski et al., 2014). However, *Hoxa3* null mutants do initiate thymus patterning and *Foxn1* expression, although it is delayed, and the thymus then undergoes rapid apoptosis (Chojnowski et al., 2014). However, although Hoxa3 is also expressed in thymic epithelial cells after the initial development patterning,

the loss of Hoxa3 after the early stages does not result in any obvious defects in TEC differentiation, indicating that Hoxa3 is not critical during later TECs development. Together, these experiments show the *Hoxa3* gene is an early player in thymus development.

Foxn1, the key TEC-specific transcription factor

By about E11.5, all cells specified for TEC identity within the 3rd pouch express the forkhead transcription factor Foxn1. Null mutation of the *Foxn1* gene (the nude mutation) leads to disruption of both normal hair growth and thymus development, which causes nude mice and rats to be immune-deficient. The *Foxn1* gene was first discovered in 1994 (Nehls, Pfeifer, Schorpp, Hedrich, & Boehm, 1994). Several studies show that Foxn1 is required for multiple stages of thymic epithelial cell differentiation. The first study used wild-type/nude chimeric mice to show that even in the presence of wild-type cells, the nude-derived cells could not contribute to either cortical or medullary networks and arrest at thymic epithelial progenitor cell stages. (Blackburn et al., 1996). Another study generated a hypomorphic allele called Foxn1-delta, from which an N-terminal domain was deleted. In these mutant mice, the thymus shows incomplete development of the thymic structure, which leads to the disruption of normal thymus domain development and loss of cortical and medullary domains. However, unlike Foxn1-nude mice, the Foxn1-delta mutant mice can still support lymphocyte development, although significant defects are detected in the resulting T cell populations. This study shows that the N-terminal domain of Foxn1 is required for normal TEC development and is responsible for cross-talk dependent TEC differentiation (Su, Navarre, Oh, Condie, & Manley, 2003). A more recent study shows that *Foxn1* is activated in TEC progenitors, contributing to both cTECs and mTECs differentiation and compartment formation. This study used a conditional mutant allele of Foxn1 that utilized the Cre/Loxp system to revert a null allele of Foxn1 to wild-type function in single epithelial cells *in vivo*. After reversion of the

allele , small thymic lobules with clear cortical and medullary regions formed that could support development of mature T lymphocytes. (Conrad C. Bleul et al., 2006). Together these and other research findings provide evidence that *Foxn1* is both necessary and sufficient for TEC differentiation and maturation, and support the conclusion that FOXN1 is the key transcription factor required for TEC differentiation.

FOXN1 also plays a critical role in postnatal thymus function and maintenance. *Foxn1* is widely expressed in cTECs and mTECs during postnatal stages, but its level is significantly decreased through the aging process, which is considered the major cause of thymus involution. (Itoi, Tsukamoto, & Amagai, 2007). The first evidence of the relationship between *Foxn1* expression and thymus involution is that there is a significant *Foxn1* downregulation coincident with initial involution and the loss of thymus function. (Ortman, Dittmar, Witte, & Le, 2002). More direct evidence is from *Foxn1-lacZ* mutant mice. This *Foxn1* allele has normal expression during fetal stages but is downregulated to about 30% of normal expression levels at 1-week after birth. This reduced *Foxn1* expression in postnatal stages leads to an early thymus involution phenotype similar to accelerated normal thymus involution in aged mice. These findings show that *Foxn1* is not only playing an important role in the fetal thymus but also plays a key role in the postnatal thymus in a dosage-sensitive manner(Chen, Xiao, & Manley, 2009).

Foxn1 is also expressed in the skin, which is the origin of the 'nude' phenotype in null mutants. There are two identified promoters, located upstream of exon 1a and exon1b, that are differentially active in skin or thymus. The Exon 1b promoter can only rescue the skin and hair mutant phenotype in nude mice, while a construct containing both exon1a and 1b can rescue both thymus and hair phenotypes. These different promoters are important for developing strategies to activate Foxn1 expression properly for cellular reprogramming strategies.

The Notch pathway plays a critical role in medullary TEC development

We and others have recently shown that the Notch signaling pathway regulates both mTECs proliferation and differentiation(Li et al., 2020). In this study, a series of loss- and gain-of-function and lineage tracing experiments were used to investigate NOTCH1 signaling in fetal TEC development. *Notch1* expression is first detected in a few Foxn1+ TECs around E11.25， soon after the onset of *Foxn1* expression, and then expressed specifically in TEC progenitor cells by E16.5. Using a *Notch1-flox* conditional allele together with *Foxn1-Cre* mice, we performed a tissue-specific loss-of-function experiment. By knocking out Notch expression in *Foxn1+* cells, most mTEC progenitor cells are lost during early thymus development. Although cTECs are not significantly affected, mTECs frequency is reduced. Further analysis showed that mTEC specification absolutely depends on Notch signaling. Constitutive activation of *Notch1* expression on TECs resulted in expanded mTEC progenitor phenotypes with significantly fewer and smaller clusters of UEA1+ cells as well as very few AIRE+ cells. These results show that constitutive Notch1 expression blocks mTEC terminal differentiation. A similar study from the Blackburn lab deleting the Notch downstream target gene *Rbpj* also showed similar results [Blackburn paper ref]. Thus, Notch signaling is required to specify the mTEC progenitors and promote their proliferation, but must be shut off to allow terminal mTEC differentiation.

TEC differentiation and proliferation

The TEC compartment can be identified as two major populations, the medullary TECs (mTECs) and the cortical TECs (cTECs), which have different localization, molecular characteristics, and function in supporting T cell development. Identification of the putative Thymic epithelial progenitor cells (TEPCs) is one of the major challenges in the field. Several

studies have proposed the existence of bipotent TEPC, while others have argued that lineage-specific progenitors maintain mTECs and cTECs. The most well-defined surface marker for bipotent TECPs is PLET1, which is recognized by the monoclonal antibody MTS24. Evidence suggests that PLET1+ thymic epithelial cell progenitors are competent and sufficient to fully reconstitute the complex thymic epithelial microenvironment that supports normal T cell development (Bennett et al., 2002). A later study showed that single EpCam+CD45- cell isolated from the thymus at E12.5 can give rise to both cTEC and mTEC progeny after transplantation, suggesting that cortical and medullary epithelial cells share a common origin in bipotent precursors. Another study used *in vivo* lineage fate mapping to show that beta5t+ TEC give rise to both the cortical and medullary TEC compartments, and used single-cell resolution lineage tracing to show that the postnatal medulla is maintained by individual beta5t+ progenitors located at the cortico-medullary junction(Mayer et al., 2016). Recently a study using single-cell RNA sequencing (scRNA-seq) analysis, chromatin profiling, and gene targeting characterized the entire stromal compartment of the mouse thymus during ontogeny. They find that most embryonic TECs are characterized by a strong cTEC footprint (Bornstein et al., 2018).Taken together, these studies support the "cTEC first" model of TEC differentiation (Takahama, 2006), in which all TEC pass through a cTEC-like stage during differentiation into either lineage.

TECs differentiation and proliferation also require cross-talk between developing T cells (thymocytes) and TECs [Ritter and Boyd]. Different stages of thymocyte differentiation control the microarchitecture of thymic microenvironments, including TEC proliferation and differentiation. In 1994, researchers first visualized lympho-stromal interactions by scanning electron microscopy, and the interaction is particularly evident in the outer cortex region of the thymus. More experiments showed that the presence of lymphoid T cells is essential for TEC

differentiation and maintenance of epithelial morphology (van Ewijk et al., 1994). For example, the medullary microenvironment is under the direct control of mature T cells bearing TCRs. Depletion of single positive (CD4+ or CD8+) cells using mABs against TCR results in a reduced medullary region with few mTECs, while mutants with no single positive T cells also show no defined medullary region (Philpott et al., 1992).

To sum up, current evidence supports that the differentiation and proliferation of thymic epithelial cells begins with a bi-potent TEC progenitor and differentiate into either cTECs or mTECs under the regulatory influence of pathways such as Notch signaling. This process is not only controlled by the thymic epithelial cell gene expression profile, but also highly depends on cross-talk interactions with differentiating thymocytes.

Thymus involution and potential thymus function recovery methods

Thymus involution

Aging is a continuous and gradual process compromising the architecture and function of different organs and systems both in humans and animals. (Rose et al., 2012). Moreover, it is associated with a decline in the normal functioning of the immune system, termed "immunosenescence", resulting in reduced ability to respond to new infectious agents(Yan & Wei, 2011). A key feature of immunosenescence is the regression of the thymus both in size and function, termed thymic involution, which leads to diminished production of naïve T lymphocytes and a loss of T cell repertoire diversity(Britanova et al., 2014). The thymus is the first organ that declines during aging both in mice and humans. In mice, the first change in the thymus occurs within the first few months after birth (Hale, Boursalian, Turk, & Fink, 2006), while thymic involution can be observed as early as six months of age in humans(Contreiras et al., 2004). Also, the rate of thymic involution is not constant. In mice, 1-6-month mice experience a rapid decline

of thymus cellularity, while in 7-28-month-old mice the thymus cellularity decreases at a constant slow steady rate(Aw & Palmer, 2012). These early declines in thymus cellularity are caused by loss of TECs, leading to a decrease of naïve T cell export as early as 1-year of age in humans, and continues exponentially over time with a half-life of ~16 years (Koepp et al., 1998).

Mechanisms to rescue thymus involution

The decrease in thymus function and naïve T cell export due to thymic involution is believed to be a major reason why elderly individuals are immunodeficient, and at risk for novel infectious diseases. (Qi, Zhang, Weyand, & Goronzy, 2014). This decreased function of the immune system could also lead to a higher risk of cancer due to a decline in T cell-mediated immune surveillance (Wang, Thomas, Sizova, & Su, 2020). For this reason, researchers are trying different methods to reverse thymus involution. One method currently being tested in preclinical and clinical studies is based on cytokine or hormone stimulation, such as LFRH-A, IL-7, IL-22, KGF, and GH, or sex steroid inhibition. However, all these cytokine treatments rely on a pre-existing thymus tissue, and more importantly, the effect of thymus recovery is transient (Velardi, Dudakov, & van den Brink, 2013). Other studies show that manipulation of *Foxn1* levels could prevent or delay the decline of naïve T cell output due to thymus involution. For example, both Foxn1-transgenic mice (Zook et al., 2011) and TEC-specific increase of Foxn1 expression (Bredenkamp, Nowell, & Blackburn, 2014) lead to an increase of naïve T cell output as well as attenuation of age-associated thymic involution. However, all these methods are highly limited by their effective period as well as difficulties of application into human clinical treatment. Thus, generation function TECs directly from patient tissue or pluripotent stem cells is one of the most promising approaches for direct transplant and in vitro generated autologous T cells.

Directly reprogramming of fibroblasts into thymic epithelial cells by Foxn1

A recent study showed that the transcription factor Foxn1 is sufficient to reprogram mouse embryonic fibroblasts (MEFs) into functional thymic epithelial cells (TECs) (Bredenkamp et al., 2014). Using a transgenic mouse line they induced *Foxn1* expression directly in primary MEFs by CRE mediated excision of a transcriptional stop cassette. Within these Foxn1+ primary MEFs, a subset not only began to express FOXN1 downstream targets *Dll4, Ccl25*, and *Kit-l,* but also became positive for Keratin-8 and epithelial cell adhesion molecule (EpCAM), which are expressed by all TECs during early thymus development, and developed an epithelial-like morphology. Most importantly, these induced thymic epithelial cells (iTECs) can successfully support Early T cell progenitor (ETP) maturation from to the single positive (SP) T cell stage. Finally, by grafting iTECs with MEFs into the kidney capsule, the iTEC graft can successfully develop into a tissue that has a distinct cortex and medullary regions. iTEC-derived grafts can further support athymic nude mice to re-gain T cell generation, which indicates the potential therapeutic application of iTECs into rescue thymus involution. However, the process remains inefficient and has significant challenges to overcome before this potential is realized.

Dissecting direct reprogramming using RNA-seq process

Direct lineage reprogramming includes a dynamic remarkable conversion of cellular morphology and transcriptome states. Thus, using RNA-seq methods to analyze the process of direct reprogramming at different time points is necssary to understand the mchanisms underlying this reprogramming process. One recent study dissected the reprogramming pathway from fibroblasts into functional neurons using RNA-seq analysis at different time points of the

reprogramming process. They found that the direct reprogramming pathway is continuous. The overexpression of proneural pioneer factor ASCL1 initiates a process whereby the fibroblasts first exit the cell cycle and then express a series of distinct neuron-specific transcription factors. Although the first response to the pioneer factor ASCL1 is homogeneous among all the treated fibroblasts, the later emergence of the neural differentiation pathway competes with the myogenic program as well as other expression pathways, and is highly variable among the cells. Also, an intermediate state that is unique compared to both fibroblast cells and neurons was identified(Treutlein et al., 2016). Another study focused on the process in which somatic cells are reprogrammed into induced pluripotent stem cells (iPSCs). Using RNA-seq , they successfully divided the iPSC reprogramming process into several distinct steps. First, the cell will bifurcate into two categories, reprogramming potential (RP) or non-reprogramming (NR). Then, several other checkpoints were identified and important distinguishable marker genes such as Klf4 identified. This study shows the detailed mechanism of reprogramming and reveals the different stages as well as important marker genes during the iPSC reprogramming process. (Guo et al., 2019).

To sum up, due to the highly dynamic large scale gene expression profile changes during direct reprogramming process, as well as poor understanding of marker genes that can be used to distinguish the different stages of reprogramming, traditional ways of gene expression analysis are inadequate for understanding the mechanism of direct reprogramming. Using RNA-seq at different time points during the reprogramming process, we can better dissect the dynamics underneath the cell fate change, and discover potential genes, pathways, or markers that may be integral to the reprogramming process. In this thesis, I used this approach to analyze the mechanism by which mouse embryonic fibroblasts (MEFs) are reprogrammed into induced thymic epithelial cells

(iTECs). To get a detailed understanding of how this cell fate change happens, we choose five cell populations at time points spanning this process. To investigate how complete the reprogramming is, we also include E14.5 fetal TECs ato compare the gene expression profile difference and similarity between iTECs and fetal TECs. This study identified distinct stages of iTEC reprogramming, and identified and tested candidate pathways that limit the reprogramming process. Thus, this study provides new and critical information for developing iTECs as a useful experimental tool both for understanding TEC biology and for a variety of research and pre-clinical applications.

References

Aw, D., & Palmer, D. B. (2012). It's not all equal: a multiphasic theory of thymic involution. *Biogerontology, 13*(1), 77-81. doi:10.1007/s10522-011-9349-0

Bennett, A. R., Farley, A., Blair, N. F., Gordon, J., Sharp, L., & Blackburn, C. C. (2002). Identification and Characterization of Thymic Epithelial Progenitor Cells. *Immunity, 16*(6), 803-814. doi:https://doi.org/10.1016/S1074-7613(02)00321-7

Blackburn, C. C., Augustine, C. L., Li, R., Harvey, R. P., Malin, M. A., Boyd, R. L., . . . Morahan, G. (1996). The nu gene acts cell-autonomously and is required for differentiation of thymic epithelial progenitors. *Proceedings of the National Academy of Sciences, 93*(12), 5742. doi:10.1073/pnas.93.12.5742

Bleul, C. C., Corbeaux, T., Reuter, A., Fisch, P., Monting, J. S., & Boehm, T. (2006). Formation of a functional thymus initiated by a postnatal epithelial progenitor cell. *Nature, 441*(7096), 992-996. doi:10.1038/nature04850

Bleul, C. C., Corbeaux, T., Reuter, A., Fisch, P., Mönting, J. S., & Boehm, T. (2006). Formation of a functional thymus initiated by a postnatal epithelial progenitor cell. *Nature, 441*(7096), 992-996. doi:10.1038/nature04850

Bornstein, C., Nevo, S., Giladi, A., Kadouri, N., Pouzolles, M., Gerbe, F., . . . Amit, I. (2018). Single-cell mapping of the thymic stroma identifies IL-25-producing tuft epithelial cells. *Nature, 559*(7715), 622-626. doi:10.1038/s41586-018-0346-1

Bredenkamp, N., Nowell, C. S., & Blackburn, C. C. (2014). Regeneration of the aged thymus by a single transcription factor. *Development, 141*(8), 1627. doi:10.1242/dev.103614

Britanova, O. V., Putintseva, E. V., Shugay, M., Merzlyak, E. M., Turchaninova, M. A., Staroverov, D. B., . . . Chudakov, D. M. (2014). Age-Related Decrease in TCR Repertoire Diversity Measured with Deep and Normalized Sequence Profiling. *The Journal of Immunology, 192*(6), 2689. doi:10.4049/jimmunol.1302064

Chen, L., Xiao, S., & Manley, N. R. (2009). Foxn1 is required to maintain the postnatal thymic microenvironment in a dosage-sensitive manner. *Blood, 113*(3), 567-574. doi:10.1182/blood-2008-05-156265

Chojnowski, J. L., Masuda, K., Trau, H. A., Thomas, K., Capecchi, M., & Manley, N. R. (2014). Multiple roles for HOXA3 in regulating thymus and parathyroid differentiation and morphogenesis in mouse. *Development, 141*(19), 3697-3708. doi:10.1242/dev.110833

Contreiras, E. C., Lenzi, H. L., Meirelles, M. N. L., Caputo, L. F. G., Calado, T. J. C., Villa-Verde, D. M. S., & Savino, W. (2004). The equine thymus microenvironment: a morphological and immunohistochemical analysis. *Developmental & Comparative Immunology, 28*(3), 251-264. doi:https://doi.org/10.1016/S0145-305X(03)00134-4

Egerton, M., Scollay, R., & Shortman, K. (1990). Kinetics of mature T-cell development in the thymus. *Proceedings of the National Academy of Sciences, 87*(7), 2579.

Gordon, J., & Manley, N. R. (2011). Mechanisms of thymus organogenesis and morphogenesis. *Development, 138*(18), 3865-3878. doi:10.1242/dev.059998

Gordon, J., Wilson, V. A., Blair, N. F., Sheridan, J., Farley, A., Wilson, L., . . . Blackburn, C. C. (2004). Functional evidence for a single endodermal origin for the thymic epithelium. *Nat Immunol, 5*(5), 546-553. doi:10.1038/ni1064

Gray, D. H. D., Chidgey, A. P., & Boyd, R. L. (2002). Analysis of thymic stromal cell populations using flow cytometry. *Journal of Immunological Methods, 260*(1), 15-28. doi:https://doi.org/10.1016/S0022-1759(01)00493-8

Guo, L., Lin, L., Wang, X., Gao, M., Cao, S., Mai, Y., . . . Chen, J. (2019). Resolving Cell Fate Decisions during Somatic Cell Reprogramming by Single-Cell RNA-Seq. *Molecular Cell, 73*(4), 815-829.e817. doi:https://doi.org/10.1016/j.molcel.2019.01.042

Hale, J. S., Boursalian, T. E., Turk, G. L., & Fink, P. J. (2006). Thymic output in aged mice. *Proceedings of the National Academy of Sciences, 103*(22), 8447. doi:10.1073/pnas.0601040103

Hamazaki, Y., Fujita, H., Kobayashi, T., Choi, Y., Scott, H. S., Matsumoto, M., & Minato, N. (2007). Medullary thymic epithelial cells expressing Aire represent a unique lineage derived from cells expressing claudin. *Nature Immunology, 8*(3), 304-311. doi:10.1038/ni1438

Hozumi, K., Mailhos, C., Negishi, N., Hirano, K.-i., Yahata, T., Ando, K., . . . Habu, S. (2008). Delta-like 4 is indispensable in thymic environment specific for T cell development. *Journal of Experimental Medicine, 205*(11), 2507-2513. doi:10.1084/jem.20080134

Itoi, M., Tsukamoto, N., & Amagai, T. (2007). Expression of Dll4 and CCL25 in Foxn1-negative epithelial cells in the post-natal thymus. *International Immunology, 19*(2), 127-132. doi:10.1093/intimm/dxl129

Jenkinson, W. E., Nakamura, K., White, A. J., Jenkinson, E. J., & Anderson, G. (2013). Normal T Cell Selection Occurs in CD205-Deficient Thymic Microenvironments. *PLOS ONE, 7*(12), e53416. doi:10.1371/journal.pone.0053416

Kameda, Y., Arai, Y., Nishimaki, T., & Chisaka, O. (2004). The role of Hoxa3 gene in parathyroid gland organogenesis of the mouse. *J Histochem Cytochem, 52*(5), 641-651. doi:10.1177/002215540405200508

Koepp, M. J., Gunn, R. N., Lawrence, A. D., Cunningham, V. J., Dagher, A., Jones, T., . . . Grasby, P. M. (1998). Evidence for striatal dopamine release during a video game. *Nature, 393*(6682), 266-268. doi:10.1038/30498

Li, J., Gordon, J., Chen, E. L. Y., Xiao, S., Wu, L., Zúñiga-Pflücker, J. C., & Manley, N. R. (2020). NOTCH1 signaling establishes the medullary thymic epithelial cell progenitor pool during mouse fetal development. *Development, 147*(12), dev178988. doi:10.1242/dev.178988

Lind, E. F., Prockop, S. E., Porritt, H. E., & Petrie, H. T. (2001). Mapping Precursor Movement through the Postnatal Thymus Reveals Specific Microenvironments Supporting Defined Stages of Early Lymphoid Development. *Journal of Experimental Medicine, 194*(2), 127-134. doi:10.1084/jem.194.2.127

Lomada, D., Liu, B., Coghlan, L., Hu, Y., & Richie, E. R. (2007). Thymus Medulla Formation and Central Tolerance Are Restored in IKKα^{−/−} Mice That Express an IKKα Transgene in Keratin 5⁺ Thymic Epithelial Cells. *The Journal of Immunology, 178*(2), 829. doi:10.4049/jimmunol.178.2.829

Manley, N. R., & Capecchi, M. R. (1995). The role of Hoxa-3 in mouse thymus and thyroid development. *Development, 121*(7), 1989.

Manley, N. R., & Capecchi, M. R. (1998). HoxGroup 3 Paralogs Regulate the Development and Migration of the Thymus, Thyroid, and Parathyroid Glands. *Developmental Biology, 195*(1), 1-15. doi:https://doi.org/10.1006/dbio.1997.8827

Mayer, C. E., Žuklys, S., Zhanybekova, S., Ohigashi, I., Teh, H.-Y., Sansom, S. N., . . . Holländer, G. A. (2016). Dynamic spatio-temporal contribution of single β5t+ cortical epithelial precursors to the thymus medulla. *European Journal of Immunology, 46*(4), 846-856. doi:https://doi.org/10.1002/eji.201545995

Muegge, K., Vila, M. P., & Durum, S. K. (1993). Interleukin-7: a cofactor for V(D)J rearrangement of the T cell receptor beta gene. *Science, 261*(5117), 93. doi:10.1126/science.7686307

Nehls, M., Pfeifer, D., Schorpp, M., Hedrich, H., & Boehm, T. (1994). New member of the winged-helix protein family disrupted in mouse and rat nude mutations. *Nature, 372*(6501), 103-107. doi:10.1038/372103a0

Odaka, C., Loranger, A., Takizawa, K., Ouellet, M., Tremblay, M. J., Murata, S., . . . Marceau, N. (2013). Keratin 8 Is Required for the Maintenance of Architectural Structure in Thymus Epithelium. *PLOS ONE, 8*(9), e75101. doi:10.1371/journal.pone.0075101

Ortman, C. L., Dittmar, K. A., Witte, P. L., & Le, P. T. (2002). Molecular characterization of the mouse involuted thymus: aberrations in expression of transcription regulators in thymocyte and epithelial compartments. *International Immunology, 14*(7), 813-822. doi:10.1093/intimm/dxf042

Owen, J. J. T., & Ritter, M. A. (1969). TISSUE INTERACTION IN THE DEVELOPMENT OF THYMUS LYMPHOCYTES. *Journal of Experimental Medicine, 129*(2), 431-442. doi:10.1084/jem.129.2.431

Peschon, J. J. (1994). Early lymphocyte expanaion is severely impaired in interleukin 7 receptor-deficient mice. *J. Exp. Med., 180*, 1995.

Philpott, K. L., Viney, J. L., Kay, G., Rastan, S., Gardiner, E. M., Chae, S., . . . Owen, M. J. (1992). Lymphoid development in mice congenitally lacking T cell receptor alpha beta-expressing cells. *Science, 256*(5062), 1448-1452. doi:10.1126/science.1604321

Plotkin, J., Prockop, S. E., Lepique, A., & Petrie, H. T. (2003). Critical Role for CXCR4 Signaling in Progenitor Localization and T Cell Differentiation in the Postnatal Thymus. *The Journal of Immunology, 171*(9), 4521. doi:10.4049/jimmunol.171.9.4521

Prockop, S. E., Palencia, S., Ryan, C. M., Gordon, K., Gray, D., & Petrie, H. T. (2002). Stromal Cells Provide the Matrix for Migration of Early Lymphoid Progenitors Through the Thymic Cortex. *The Journal of Immunology, 169*(8), 4354. doi:10.4049/jimmunol.169.8.4354

Qi, Q., Zhang, D. W., Weyand, C. M., & Goronzy, J. J. (2014). Mechanisms shaping the naïve T cell repertoire in the elderly — Thymic involution or peripheral homeostatic proliferation? *Experimental Gerontology, 54*, 71-74. doi:https://doi.org/10.1016/j.exger.2014.01.005

Ripen, A. M., Nitta, T., Murata, S., Tanaka, K., & Takahama, Y. (2011). Ontogeny of thymic cortical epithelial cells expressing the thymoproteasome subunit beta5t. *Eur J Immunol, 41*(5), 1278-1287. doi:10.1002/eji.201041375

Rose, M., Flatt, T., Graves Jr, J., Greer, L. F., Martínez, D., Matos, M., . . . Shahrestani, P. (2012). What is Aging? *Frontiers in Genetics, 3*, 134.

Rossi, F. M. V., Corbel, S. Y., Merzaban, J. S., Carlow, D. A., Gossens, K., Duenas, J., . . . Ziltener, H. J. (2005). Recruitment of adult thymic progenitors is regulated by P-selectin and its ligand PSGL-1. *Nature Immunology, 6*(6), 626-634. doi:10.1038/ni1203

Rossi, S. W., Jenkinson, W. E., Anderson, G., & Jenkinson, E. J. (2006). Clonal analysis reveals a common progenitor for thymic cortical and medullary epithelium. *Nature, 441*(7096), 988-991. doi:10.1038/nature04813

Speck-Hernandez, C. A., Assis, A. F., Felicio, R. F., Cotrim-Sousa, L., Pezzi, N., Lopes, G. S., . . . Passos, G. A. (2018). Aire Disruption Influences the Medullary Thymic Epithelial Cell Transcriptome and Interaction With Thymocytes. *Frontiers in Immunology, 9*, 964.

Su, D.-m., Navarre, S., Oh, W.-j., Condie, B. G., & Manley, N. R. (2003). A domain of Foxn1 required for crosstalk-dependent thymic epithelial cell differentiation. *Nature Immunology, 4*(11), 1128-1135. doi:10.1038/ni983

Takahama, Y. (2006). Journey through the thymus: stromal guides for T-cell development and selection. *Nature Reviews Immunology, 6*(2), 127-135. doi:10.1038/nri1781

Treutlein, B., Lee, Q. Y., Camp, J. G., Mall, M., Koh, W., Shariati, S. A. M., . . . Quake, S. R. (2016). Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq. *Nature, 534*(7607), 391-395. doi:10.1038/nature18323

van Ewijk, W., Shores, E. W., & Singer, A. (1994). Crosstalk in the mouse thymus. *Immunol Today, 15*(5), 214-217. doi:10.1016/0167-5699(94)90246-1

van Vliet, E., Jenkinson, E. J., Ton, R. K., Owen, J. J. T., & van Ewrjk, W. (1985). Stromal cell types in the developing thymus of the normal and nude mouse embryo. *European Journal of Immunology, 15*(7), 675-681. doi:https://doi.org/10.1002/eji.1830150707

Velardi, E., Dudakov, J. A., & van den Brink, M. R. M. (2013). Clinical strategies to enhance thymic recovery after allogeneic hematopoietic stem cell transplantation. *Immunology Letters, 155*(1), 31-35. doi:https://doi.org/10.1016/j.imlet.2013.09.016

Wang, W., Thomas, R., Sizova, O., & Su, D.-M. (2020). Thymic Function Associated With Cancer Development, Relapse, and Antitumor Immunity – A Mini-Review. *Frontiers in Immunology, 11*, 773.

Yan, S.-x., & Wei, W. (2011). Castration reverses immunosenescence in aged mice. *Acta Pharmacologica Sinica, 32*(9), 1085-1086. doi:10.1038/aps.2011.110

Zook, E. C., Krishack, P. A., Zhang, S., Zeleznik-Le, N. J., Firulli, A. B., Witte, P. L., & Le, P. T. (2011). Overexpression of Foxn1 attenuates age-associated thymic involution and prevents the expansion of peripheral CD4 memory T cells. *Blood, 118*(22), 5723-5731. doi:10.1182/blood-2011-03-342097

Zuklys, S., Balciunaite, G., Agarwal, A., Fasler-Kan, E., Palmer, E., & Holländer, G. A. (2000). Normal Thymic Architecture and Negative Selection Are Associated with Aire Expression, the Gene Defective in the Autoimmune-Polyendocrinopathy-Candidiasis-Ectodermal Dystrophy (APECED). *The Journal of Immunology, 165*(4), 1976. doi:10.4049/jimmunol.165.4.1976

CHAPTER 2

Dissecting direct reprogramming from fibroblast to induced thymic epithelial cell (iTECs) using RNA-seq[1]

[1]Zhongyao, Ma and Nancy, Manley. To be submitted to Development

Abstract

The thymus a primary lymphoid organ required for a robust adaptive immunity system that loses its function through aging. Previously by using transgenic overexpression of the thymic epithelial cell transcription factor FOXN1, we successfully converted mouse embryonic fibroblasts (MEFs) into functional thymic epithelial cells (iTECs). However, the reprogramming process mechanisms, including a remarkable dynamic conversion of cellular and transcriptome states, are undefined. We used bulk RNA-seq at multiple points to further analyze the reprogramming process from mouse embryonic fibroblast (MEFs) to iTECs. We first find that the iTECs reprogramming process is a continuous step-by-step process with distinct intermediate cell stages. *Foxn1* overexpression results in immediate gene expression profile changes, which are progressive over time. The reprogramming process proceeds by step-wise activation of thymus-related gene expression, beginning with cTEC-associated genes such as keratin 8 (K8) and beta-5t, followed by mTEC-related genes, such as cldn4, K5, and K14. Furthermore, we identify several pathways that may limit complete iTEC reprogramming, including the cell cycle and Notch pathways. By comparing iTECs and fetal TECs, we find a significant transcriptome difference between iTECs and fetal TEC, but show that Epcam+ iTECs share significant similarity to fetal

TECs. To sum up, our data provide a more detailed understanding of the iTECs reprogramming process and offer a potential pathway that could further promote the iTECs to become more similar to fetal TECs.

Introduction

The thymus the primary lymphoid organ that is the major source of self-restricted, self-tolerant naïve T cells required for a robust adaptive immunity system. However, the thymus is ia also the first organ in our body to lose its function through aging, a process termed thymus involution, which leads to diminished production of naïve lymphocytes and a loss of diversity of T cell repertoire (Britanova et al., 2014). Thymus involution initiates as early as within the first few months of age in mice and causes an increased possibility of infection and disease. (Hale et al., 2006). Thus, finding a way to recover thymus function is significant. Many different studies have been focused on rescuing thymus function by variable methods. Some have been utilizing cytokine stimulation, while others tried to manipulate the gene expression in the thymus (Velardi et al., 2013). However, most of these methods are highly limited by their effective period and difficulties of further application into human clinical treatment. Previously, we showed that through the ectopic overexpression of thymus key transcription factor Foxn1, mouse embryonic fibroblast (MEFs) can be directly reprogrammed into functional induced thymic epithelial cells(iTECs) (Bredenkamp et al., 2014)These iTECs express many thymic specific genes and can support early T cell progenitor maturation into functional T cells. However, the detailed mechanisms of how MEFs convert into iTECs are still largely undefined, nor do we know how similar the iTECs are compared to fetal TECs at the whole gene transcriptome level.

It has recently been shown that RNA-seq analysis can be used to dissect the direct reprogramming process, revealing that direct lineage conversion often includes a remarkable dynamic cellular and gene expression change (Treutlein et al., 2016). Analyzing cells undergoing reprogramming at different time points and re-ordering by whole genome transcriptome similarity can further divide the reprogramming process into several distinct steps to generate a more complete picture of how the reprogramming process happens, as well as identifying candidates key players in each step of reprogramming(Guo et al., 2019). Here, we have generated iTECs using direct nucleofection, which provides a clean starting point for the direct reprogramming process. We then analyzed iTECs at different time points of reprogramming together with embryonic date 14.5 (E14.5) fetal TECs. These data show that iTECs reprogramming occurs in a step-by-step manner with distinct intermediate cell types. Also, by comparing and ordering iTECs samples from different time points and further functional experiments, we have found that cell cycle-related genes may play a key role in the iTEC reprogramming process as well as iTECs reprogramming efficiency. Finally, by comparing the iTECs with fetal TECs, several potential pathways and genes are potential candidates for facilitating iTECs reprogramming efficiency and competency. We also show that mTEC lineage specification does occur during reprogramming at a low level, but is limited by Notch signaling; by first allowing, then blocking Notch signaling, we can induce mTEC differentiation and *Aire* expression in iTEC cultures. Our study provides more detailed insights into the iTECs reprogramming mechanisms, provides a complete comparison of transcriptome similarity between iTECs and fetal TECs, and generates several potential pathways that can be used for future iTECs investigation.

Results

Generation of iTECs using nucleofection based methods

To generate samples representing different iTECs reprogramming time-points, we first utilized the nucleofection system based direct transfection method, instead of transgenic 4OHT treatment based iTEC generation methods. In this nucleofection-based system, only the transgene iFoxn1-IRES-GFP is needed in the MEFs with a stop-codon that can be removed by Cre recombinase (Fig2.1A). R26-iFoxn1-Stop-Foxn1-IRES-GFP/+ heterozygous MEFs from E13.5 embryos are used as starting cell material for all following experiments. After the MEFs reach optimal confluency for nucleofection, we transfect the MEFs with a PGK-Cre plasmid. After 48hrs of culture, a majority of the MEFs will remove the STOP cassette by Cre recombinase and activate the R26-iFoxn1-IRES-GFP locus, identified by GFP expression. We termed these Foxn1- and GFP-expressing cells iFoxn1 MEFs. (Fig2.1B). This method avoids the low level of “leaky” Cre expression, providing a clean background for RNA-seq. (Fig2.1C). By using this Cre-plasmid based transfection system, we can precisely control the onset of iTECs reprogramming.

As an initial characterization of the reprogramming process, we further analyzed marker expression of the iFoxn1 MEFs. Flow cytometry shows that more than 10% of the iFoxn1 MEFs express the epithelium-specific marker EpCam 13days after the initial activation of transgene iFoxn1; as in the previous report, MHC Class II is not induced. (Fig2.1D). Also consistent with the previous report, a significant proportion of the iFoxn1 MEFs are Keratin-8 positive 10days after the transfection. (Fig2.1E). No keratin-8 positive cells are observed in the mock transfected MEF controls.

We collected iFoxn1-MEFs at a different time points after the initial iFoxn1 activation and use qRT-PCR to analyze several thymic-related gene expression profiles. (Fig2.1F). The Foxn1

transgene is activated immediately after the transfection and maintained at a relatively high level compared to similar to 2 weeks old TECs. As expected, several FOXN1 target genes [PMID: 27548434] are also detected early after induction of *Foxn1* expression. *Dll4* and *Kit-L* are significantly increased as early as 24hr after transfection, and *Ccl25* is significantly up-regulated 2days after transfection. The increase in these FOXN1 target genes does not follow a linear pattern. Instead, all three genes undergo rapid expression increases 7days after transfection. This pattern suggests that reprogramming has two different stages an early stage with intitial low level expression of target genes, followed by a later stage in which FOXN1 target gene expression is dramatically accellerated. (Fig.1F). .

iTEC reprogramming has progressive gene expression changes with a distinct intermediate cell step

To further understand the dynamic process of iTECs reprogramming, we designed a time-point based bulk-RNA-seq experiment. (Fig2.2A). We used mock-transfected R26-iFoxn1/+ MEFs as our control, representing the initial MEFs transcriptome status (No Cre MEFs). We induced R26-iFoxn1/+ MEFs to express Foxn1 using Cre transfection, cultured for 2days, and sorted GFP+ cells using flow cytometry. These GFP+ iFoxn1 MEFs represent initial reprogramming, and are referred to as iTEC2d. We also cultured iTECs for a total of 10days after transfection to represent the late stage of reprogramming, termed iTEC10d. Additionally, we cultured iTECs for a total of 12 days and then used Epcam antibody staining to separate them into two distinct groups, designated EpCam+ & EpCam-. Each sample consisted of cells pooled from multiple independent iTEC cultures, with RNA-seq performed in duplicate (MEFs, iTEC2d, and

EpCam-) or triplicate (iTEC10d). The exception is the EpCam+ iTECs, for which only one pooled sample was analyzed in these initial experiments.

To get an initial understanding of the iTEC reprogramming progression, we first analyzed MEF and all iTEC whole transcriptome profiles using PCA analysis (Fig2.2B). We found that even 48hrs after the initial activation of ectopic Foxn1, the iTECs' gene expression has changed significantly compared to MEFs, indicated by the clear separation of iTEC2d and No Cre MEF samples. The iTEC10d sample is further separated from the iTEC 2d sample, but does not show a linear increase of variance from No Cre to iTEC10d through iTEC2d, indicating that the iTECs reprogramming process is not a continuous process at the gene expression profile level but includes distinct gene expression pattern changes at the early and late stages. Finally, the EpCam+ samples are more distinct than Epcam- sample compared to iTECs 10day. The whole transcrptome analysis shows a linear progression of differential gene expression, which indicating that Epcam+ cells represent a more advanced reprogramming state, and suggesting that 10d, EpCam-, and EpCam+ iTEC represent a linear progression of differentiation (Fig2.2B). The sample distance count matrix also shows this progression process from MEFs to EpCam+ iTECs, with a significant difference between each step of reprogramming (Fig2.2C). Furthermore, the reproducibility of the samples at each reprogramming stage is high, with the exception of the EpCam+ stage, for which only one sample was analyzed.

The gene count matrix results show that a subset of genes is transiently expressed in the iTEC2d stage (off in MEFs, on in iTEC2d, and off in iTECs10d and EpCam+). (Fig2.2D). We specifically assessed transcription factor expression changes during the reprogramming process, comparing MEFs, 2d, 10d, and EpCam+ (Fig2.2E). We find that many transcription factors are differentially expressed in the reprogramming process in both the early and late reprogramming

process. These data indicate that each stage of reprogramming is characterized by a unique gene expression profile, rather than a gradual accumulation of gene expression changes that build through reprogramming.

Together, our results show that there is a remarkable transcriptome change during the reprogramming process. The reprogramming process's intermediate step is significantly different from MEF controls and the late reprogramming stage.

cTECs & mTECs gene turned on at different stages of reprogramming & multiple thymic related gene pathways are involved

The initial iTEC report showed differentiation of iTECs into both cTEC and mTEC and formation of a well-organized organoid upon transplantation under the kidney capsule [bredemkamp]. However, it did not characterize cTEC and mTEC lineage markers in detail in iTECs in culture. We performed MA plot analysis to investigate TEC lineage specification and differentiation-related gene expression patterns during the reprogramming process (Fig2.3A-E). Overall, this analysis shows that shows more genes are up-regulated than down regulated as reprogramming progresses.

Comparison of MEFs to 2d (Fig2.3A) and 10d (Fig2.3B) stages shows progressive up-regulation of FOXN1 target genes *Dll4, Kit-L*, and *Ccl25*. After 2d, these genes remain relatively close to the median, with a slight up-regulation. Strikingly, cTEC markers appear very early, with Krt18 and to a lesser extent K8 already significantly expressed at iTEC2d (Fig2.3A), followed by strong up-regulation of Krt8, and Psmb11 (β5t) in iTEC10d samples (Fig2.3C). In contrast, mTEC markers (Krt5, Krt14, Cldn4) are not detected at early reprogramming stages (Fig2.

3A), indicating that cTEC markers are expressed before mTECs markers in iTEC reprogramming. mTEC markers first appear in iTEC10d stages, where they are highly differentially expressed compared to 2d samples, but still at a much lower expression levels than the cTEC markers (Fig2. 3B & Fig2.3C). Interestingly, while cTEC markers remain stable from the iTEC10d to EpCam+ iTEC, mTEC related genes are differentially upregulated at this latest stage of reprogramming. (Fig2.3E). This progressive up-regulation of TEC differentiation genes with delayed mTEC marker up-regulation is further evident in a heat map analysis (Fig2.3F).

Heatmap analysis further shows that thymus organogenesis-related transcription factors and signaling pathways have differential expression patterns through the reprogramming process, even if they are not direct FOXN1 targets. For example, *Hoxa3*, which is expressed prior to *Foxn1* during initial organ formation and is not a known *Foxn1* target, is turned on in iTEC2d samples and stays on during the whole reprogramming process similar to *Foxn1* itself. However, *Pax9, Pax1*, and *Eya1*, which have similar *in vivo* expression patterns during early organogenesis *in vivo*, are expressed in the early stages of reprogramming and the down regulated at later reprogramming stages. (Fig2.3G).

We also extracted the normalized transcript counts to better visualize the dynamic gene expression patterns during the reprogramming process. (Fig2.3H). For example, BMP4 signaling is thought to be an important regulator of *Foxn1* expression and TEC differentiation *in vivo*, although evidence is mixed. We find that BMP4 pathway-related genes have complex expression patterns across reprogramming, without a clear relationship to the reprogramming process, suggesting that BMP signaling may not be an important regulator during reprogramming. Consistent with its role as an antagonist of thymus organogenesis *in vivo* [Moore-Scott, Bain],

SHH pathway-related genes including Gli2, Ptch1, & Smo, are not significantly expressed during the reprogramming process, while the Gli3 repressor is up-regulated.

Finally, we separately analyzed up-regulated & down-regulated genes at different reprogramming stages by gene enrichment analysis (Fig2.3I). We find that in early reprogramming stages, iTECs are characterized by down-regulation of cell shape and fibroblast related genes, and up-regulation of cytokine signaling and other immune-related categories. In the late stage of reprogramming, the cell cycle and DNA replication are inhibited, consistent with the initial iTEC study showing that iTECs stop proliferating early in reprogramming [Bredenkamp]. Genes for epidermis development are highly enriched at this stage, indicating that the iTECs are transitioning into epithelial cells, which is not observed in the early stage of reprogramming.

To sum up, our data shows that the iTEC reprogramming process consists of discrete gene expression changes that differ in the early & late reprogramming stages (Fig2.3J). In the early reprogramming stages, iTECs are characterized by activation of FOXN1 downstream targets, cTECs markers, cell-matrix re-organization and cell-reshape, and inhibit cell cycle & cell proliferation. In the late reprogramming stages, iTECs up-regulate thymus epithelial markers, especially mTECs markers, and take on global characteristics of epithelial cells.

Cell cycle arrest is a barrier to the iTEC reprogramming process.

The cell cycle arrest observed early in iTEC reprogramming is similar to that seen in other reprogramming systems. Our RNA-seq gene expression data also showed that many cell cycle-promoting genes such as Myc, Mcm7, and p53 are inhibited in the late stage of reprogramming, while cell cycle inhibitors including Rb, Cdkn1a/p21, and Gadd45b are

upregulated beginning at the iTEC2d stage (Fig2.4A & B). However, *in vivo* FOXN1 has been shown to promote TEC proliferation. Thus, this cell cycle inhibition is consistent with reprogramming in general, but inconsistent with the known functions of FOXN1. Furthermore, this block results in an inability to expand iTECs in culture, limiting their utility as an experimental system.

If cell cycle arrest is necessary for iTEC reprogramming as it is for reprogramming other cell types, then if we rescue the cell cycle, the iTEC reprogramming could be inhibited. Conversely, if cell cycle arrest is not necessary, then the cell cycle's rescue would not affect iTEC reprogramming and could allow iTEC expansion in culture. To test these possibilities, we utilized an inducible Rosa26 transgene for the Myc gene, which can also be activated by Cre plasmid transfection, and also expresses the CD2 antigen so that the Myc expressing cell can be detected by staining for the CD2. We generated double transgenic MEFs (iFoxn1/+;iMYC/+) in which Cre expression during initial iTEC induction should activate both Foxn1 and Myc gene expression. (Fig2.4C). Cells that have activated both alleles are GFP+ and CD2+. Using PI staining and Cell trace analysis, we find that Myc overexpression can partially rescue the cell cycle arrest in iTECs. iFoxn1+iMyc+ iTEC cultures had increased S phase frequency (Fig2.4D) and a dramatic shift to the left in Cell Trace analysis, indicating dilution of the Cell Trace reagent consistent with most cells undergoing proliferation (Fig2.4E).

We then analyzed whether this rescue of cell cycle arrest would affect the iTEC reprogramming. Surprisingly, activation of the cell cycle does not block iTEC reprogramming, but instead further increases the reprogramming efficiency and the frequency of epithelial cell marker Epcam expression. (Fig2.4F). These results suggest that the cell cycle block may be a side-effect of the reprogramming process rather than a necessary step, and further indicates that rescuing the

cell cycle can increase reprogramming efficiency. Activation of Myc alone was unable to induce reprogramming or EpCam expression as detected by flow cytometry (Fig2.4G). We also tried using a Gadd45b inhibitor as an independent method to relieve the cell cycle block. However, Gadd45b inhibition did not increase EpCam+ cells (Fig2.4H), likely due to incomplete cell cycle rescue (Fig2.I). Further experiments need to be done to test this hypothesis.

Together, these data show that cell cycle arrest is likely to be a barrier for the iTEC reprogramming process, and suggest that rescuing the cell cycle arrest can significantly improve the efficiency of iTEC reprogramming. This provides potential avenues for future research to improve iTECs reprogramming.

Notch Pathway controls mTEC differentiation during iTEC reprogramming.

Our RNA-seq MAplot analysis found that cTEC marker expression and mTECs marker expression are turned on at different reprogramming stages, with cTEC markers turned on as early as 2days in reprogramming, and mTEC markers expressed at later reprogramming stages (Fig2.5A-D). Also, the mTECs progenitor marker Cldn4 is one of the highest mTEC markers expressed at these later stages. Thus, we hypothesized that mTEC differentiation is induced at the progenitor stage, but that further mTEC differentiation is inhibited or in-complete.

Recently we and others showed that the Notch pathway plays an essential role in mTEC proliferation and differentiation.Most critically, while Notch signaling is required to specify mTEC progentiors and their initial proliferation, too much Notch after initial mTECs development would lead to a block of mTEC terminal differentiation, in which most mTECs would stop at Cldn4+ mTEC progenitor states. These *in vivo* mechanisms are highly similar to what we observe in iTECs

reprogramming, and suggest the possibility that the Notch pathway may play the same role in mTECs differentiation during iTEC reprogramming.

To test whether Notch plays a crucial role in mTECs differentiation during iTEC reprogramming, we first analyzed gene expression of components of the Notch signaling pathway. (Fig2.5E). We find that many Notch ligands and downstream target genes are up-regulated during reprogramming and maintained in the late reprogramming process.

To test whether the maintained high level of Notch signaling is the reason why mTECs differentiation is blocked at a Cld4+ stage, we designed a temporally controlled Notch inhibition experiment. We allowed early reprogramming to progress, then at day5 after the initial transfection added the Notch signaling inhibitor DAPT during the late reprogramming process. We first tested whether inhibiting the Notch signaling pathway would impact the iTEC reprogramming (Fig2.5F). We find that inhibiting the Notch signaling pathway did not affect iTEC reprogramming efficiency based on the percentage of EpCam+ population. Also, DAPT treatment does not increase the MHCII expression, indicating that inhibiting Notch-related expression also does not promote this aspect of iTEC reprogramming.

Finally, we tested whether DAPT inhibition would promote mTECs differentiation and maturation. By qRT-PCR, we find that after DAPT inhibition, Cldn4 is strongly down-regulated, indicating the mTEC iTECs lose mTECs progenitor characteristics. (Fig2.5G). To assess mTEC differentiation, we used *Aire* expression, which is an critical marker for mTEC maturation and terminal differentiation (which is not expressed during normal iTEC induction) Significantly, we find that Aire is strongly up-regulated after the DAPT treatment compared to controls. This data shows that by inhibiting the Notch signaling pathway, we can successfully further promote the mTECs iTECs differentiation, as predicted by the *in vivo* Notch pathway mechanism.

Together, these data suggest that the Notch signaling pathway plays an important role in iTECs reprogramming. Inhibiting the Notch signaling pathway does not increase overall iTEC reprogramming efficiency, but does promote mTEC terminal differentiation in cultured iTECs, as shown by down-regulation of the mTECs progenitor marker Cldn4 expression and activation of *Aire*. These data indicate that iTEC reprogramming may follow the same mTEC differentiation mechanisms *in vivo*, and opens the possibility of establishing iTEC as a useful system for investigating mTEC differentiation and function.

iTECs are still distinct from fetal TECs, but EpCam+ sample shares more similarity in the gene expression profile

To test how similar iTECs are compared to real fetal TECs in transcriptome level, we isolated and sequenced total fetal TECs at E14.5 as well as a second sample of EpCam+ iTECs and analyzed these data in combination with out previously collected samples. PCA analysis clearly demonstrates that all iTECs samples are still quite different from fetal TECs, although the Epcam+ samples are the closest (Fig2.6A). The sample distance count matrix shows a similar result, with fetal TECs distinct from all iTECs samples, and Epcam+ samples as the closest to fetal TECs. (Fig2.6B). By using count matrix heatmap, we identified differentiatlly expressed genes that define these different stages. (Fig2.6C). This analysis also clearly shows that the two EpCam+ samples are not just different from each other, but that one is clearly more similar to fetal TECs, indicating that it represents a more advanced stage of reprogramming. Furthermore, four of the genes that fail to be up-regulated in any iTEC samples compared to fetal TEC (blue/white in all iTEC samples, red/organ in fetal TEC) are three MHCII genes (annotated as H2-A or H2-E) and

Tbata, which is known to regulate TEC proliferation *in vivo* [ref]. The gene that is the most down regulated in fetal TEC compared to all iTEC samples is Col6a3 (**), suggesting persistent retention of some fibroblast characteristics in iTECs. Several gene pathway analyses were also performed to detail analyze the difference between iTECs and fetal TECs (Fig2.6G). Not surprisingly, we find that the cell cycle is actively expressed in fetal TECs while highly inhibited in the iTECs. In contrast, several Notch related genes are down-regulated in fetal TECs while up-regulated in iTECs. These data are consistent with our results above, and suggest that these pathways may be potential candidates to promote iTECs further to be more similar to fetal TECs.

We also used an MAplot to analyze mTEC, cTEC, and Foxn1 target gene expression between EpCam+ iTECs and fetal TECs (Fig2.6E). We find that most Foxn1 targets and all cTEC and mTECs markers are differentially up-regulated in fetal TEC compated to EpCam+ iTEC, further underscoring the higher level of differentiation in fetal TEC.

To sum up, by analyzing the transcriptome of fetal TECs, we find that even the latest stage of iTEC analyzed, EpCam+ iTECs, are distinct in gene expression profiles from fetal TECs. However, Epcam+ samples share more similarities with fetal TECs than earlier stages. A critical difference is the expression of MHC Class II genes, which are still absent in iTECs, but are critical functional components of TECs.

Discussion

The results of this study provide a significant advance in our understanding of iTEC reprogramming and provide clear avenues for future improvement of this technique, although many questions need further study. We have shown that iTEC reprogramming is not linear, and defined characteristics of early and late stage reprogramming. Our data indicate that

reprogramming start with immediate up-regulation of FOXN1 direct target genes, but soon expands to include many additional changes not directly attributable to known FOXN1 *in vivo* functions. We further show that the cell cycle arrest that occurs early in iTEC reprorgramming is not required, and instead inhibits reprogramming efficiency. Further, we show that iTEC reprogramming follows a path consistent with the "cTEC first" model, with mTEC markers appearing later, and consistent with arrest at the mTEC progenitor stage. Finally, we show that this apparent block to mTEC differentiation can be relieved by blocking the Notch signaling pathway, suggesting that mTEC differentiation in iTECs follows similar pathways as those of endogenous fetal mTECs.

One persistent question is why iTEC differentiation is not uniform across iTEC cultures. In our current iTECs direct reprogramming, the EpCam positive population is always a subset of the cultures, even though all iTECs are expressing the same Foxn1 transgene for the same amount of time. One possibility is that the cell cycle may play a partial role in this process, since rescuing the cell cycle block did increase the percentage of Epcam+ iTECs (although it still was not 100%). MEFs are a heterogeneous population, and it is possible some are more susceptible to FOXN1-reprogramming than others [PMID: 30898844]. Another variable is endogenous *Foxn1* expression. We consistently detect activation at the endogenous locus at a low-level at the late stage of reprogramming, although it is unclear whether this occurs in all cells. Single-cell RNA-seq and other analyses of heterogeneity during reprogramming needs to be done to test these hypotheses.

Several previous studies show that cell cycle arrest is necessary for direct lineage reprogramming of other cell types, such as neuron direct reprograming. However, in iTECs reprogramming, our data show that while cell cycle arrest occurs, it is not necessary for iTEC direct linage reprogramming. In contrast, removing this cell cycle block by forced Myc expression

resulted in higher efficiency of EpCam+ iTEC generation. However, it is still unclear what exactly the cell cycle does in the iTEC reprogramming process, and to what extent these proliferating iTECs are similar to or different from cell cycle arrested iTECs. Further experiments will need to be done to analyze the impact of cell cycle arrest rescue and further understand the mechanisms underlying how the cell cycle regulates iTEC reprogramming.

Finally, we do detect both cTECs and mTECs marker expression during iTEC reprogramming. Our data are consistent with both the "cTEC first" model of TEC differentiation [Takahama review] and with successful induction of mTEC lineage specification, at least the Cld4+ mTEC progenitor stage. Our experiments also demonstrate that persistent Notch signaling may limit mTEC differentiation in iTEC cultures, and that blocking Notch signaling at later stages can result in significant mTEC differentiation, including expression of the critical marker *Aire.* However, it is still unknown how this process is occurring at the individual cell level. It is possible that all iTECs expresses cTEC markers first, and as they further differentiate during the reprogramming process, all or a subset start to express mTECs markers. Alternatively, it is possible that different subsets of cells initiate cTEC and mTEC differentiation directly. This question also is best approached using scRNA-seq methods.

Finally, none of of our iTEC cultures express MHC Class II, a crucial functional hallmark of TEC differentiation. In the original iTEC report, cross-talk with immature thymocytes either in culture or after transplant induced MHC Class II expression, allowing for the generation of single positive CD4+ T cells (Bredenkamp et al., 2014). We have also recapitulated this result (Appendix 4). Thus, future experiments to identify the crosstalk elements that activate MHC Class II expression will be essential to generation of a fully viable iTEC culture system.

In this chapter, we provide transcriptome analysis of different stages of reprogramming and fetal TECs and test several mechanisms that are important in regulating iTECs reprogramming. To further improve iTECs reprogramming in both efficiency and function, further studies can use these data as a platform to design further mechanism analysis of iTECs, at a single cell level. Such studies have the potential to improve the differentiation, proliferation, and maintenance of iTECs as a novel experimental platform for understanding TEC biology and and function.

Reference:

Bredenkamp, N., Nowell, C. S., & Blackburn, C. C. (2014). Regeneration of the aged thymus by a single transcription factor. *Development, 141*(8), 1627. doi:10.1242/dev.103614

Britanova, O. V., Putintseva, E. V., Shugay, M., Merzlyak, E. M., Turchaninova, M. A., Staroverov, D. B., . . . Chudakov, D. M. (2014). Age-Related Decrease in TCR Repertoire Diversity Measured with Deep and Normalized Sequence Profiling. *The Journal of Immunology, 192*(6), 2689. doi:10.4049/jimmunol.1302064

Guo, L., Lin, L., Wang, X., Gao, M., Cao, S., Mai, Y., . . . Chen, J. (2019). Resolving Cell Fate Decisions during Somatic Cell Reprogramming by Single-Cell RNA-Seq. *Molecular Cell, 73*(4), 815-829.e817. doi:https://doi.org/10.1016/j.molcel.2019.01.042

Hale, J. S., Boursalian, T. E., Turk, G. L., & Fink, P. J. (2006). Thymic output in aged mice. *Proceedings of the National Academy of Sciences, 103*(22), 8447. doi:10.1073/pnas.0601040103

Treutlein, B., Lee, Q. Y., Camp, J. G., Mall, M., Koh, W., Shariati, S. A. M., . . . Quake, S. R. (2016). Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq. *Nature, 534*(7607), 391-395. doi:10.1038/nature18323

Velardi, E., Dudakov, J. A., & van den Brink, M. R. M. (2013). Clinical strategies to enhance thymic recovery after allogeneic hematopoietic stem cell transplantation. *Immunology Letters, 155*(1), 31-35. doi:https://doi.org/10.1016/j.imlet.2013.09.016

Figures and Figure Legends

Figure 2.1: generation and charateritics of R26CAG-Foxn1-IRES-GFP/+ MEFs.

a) Gene stucuture of R26CAG-STOP-FOXN1-IRES-GFP transgene

b) Workflow of generation (iFoxn1)MEFS USING Cre-plasmid based nucleofection

C) IRES-GFP reporter signal analysis by flow cytometry for R26-ifoxn1/+ with cre plasmid and other three control groups

D) Epithelium marker EpCam/MHCII analyisis of R26-iFoxn1/+ with Cre+ plasmid 13days after transfection and other 2 control group.

E) epithelial Marker Keratin 8 staining for ITEC 10days and control MEFs 10days

F) gene expression analyisis using qRT-PCR at different time-point after Cre transfection.

Figure 2.2: schematic diagram of RNA-seq experiment design and geneal transcriptome analysis

A) Schematic diagram and sorting plot of generating iTECs samples for RNA-seq experiment

B) PCA analysis of all sample transcriptome except fetal TECs

C) Sample distance analyisis of all samples except fetal TECs

D) Top variable genes heatmap analyisis of all samples except felal TECs

E) Top differential expressed transcription factor heatmap analysis except fetal TECs

Figure 2.3: detailed gene expression comparision analyisis of different stage of iTECs reprogramming

A)To E) MAplot analyisis comparing different stages of iTEC reprogramming with important mTECs/cTEC markers

F) And G) gene expression heatmap for thymus related transcription factor and other important thymus genes

I) gene enrichment analyisis of up/down regulated genes in iTEC2d & iTEC10d

H) normalized counts plot of Bmp4 pahtway and Shh pathway

J) diagram of step-by-step iTECs repreogrmaming hypothesis

Figure 2.4: Cell cycle analysis of iTECs and cell cycle rescue using iMYC transgen.

A) to B) gene expression analyisis of cell-cycle related genes in iTECs reprogrammign

C) design of resuce cell cycle by overexpress iMYC and sorting gate

D) & F) PI staining & cell trace analyisis of cell cycle for iFoxn1/+ iTECs and iFoxn1/+; iMYC/+ iTECs

E) Epithelium marker EpCam staining of iFoxn1/+ iTECs and iFoxn1/+; iMYC/+ iTECs

G) Epithelium marker EpCam staining of iMYC/+ iTECs and iFoxn1/+; iMYC/+ iTECs

H) to I) GADD45B treatment experiment. H) EpCam expression analyisis I) Cell cycle resuce PI staining analysis.

Figure 2.5: Notch pathway analyisis and DAPT treamtn experiment.

A） diagram of cTEC/mTEC marker expression during iTECs reprogrmaming

B) C) D) MAplot of cTEC/mTEC marker expression at different stages B) No Cre MEFs compare to iTECs 2day C) iTECs 2day compare to iTEC 10day D) Epcam+ vs EpCam-

E) Notch pathway gene expression heatmap

F) Epcam/MHCII flow cytometry analysis of DAPT treated iTECs

G) qRT-PCR analyisis of DAPT treated iTECs

Figure 2.6: Trancriptome and differential gene expression analyisis with E14.5 Fetal TECs.

B) top differential expressed genes between iTECs 10day, Epcam+ and Fetal TECs

C) sample distance analysis of iTECs 10day, Epcam+ and Fetal TECs

D) mTECs, cTECs and other thymus related genes MAplot analysis of iTECs 2day vs iTEC10day & iTEC 10day vs fetal TECs.

E) PCA analyisis of different time point of iTECs sample and fetal TECs

H) mTECs, cTECs and other thymus related genes MAplot analysis of EpCampositive sample vs Fetal TECs

G) variable gene expression pathway heatmap analysis including RB pathway, Notch Pathway and cell cycle pathway.

Figure 2.1:

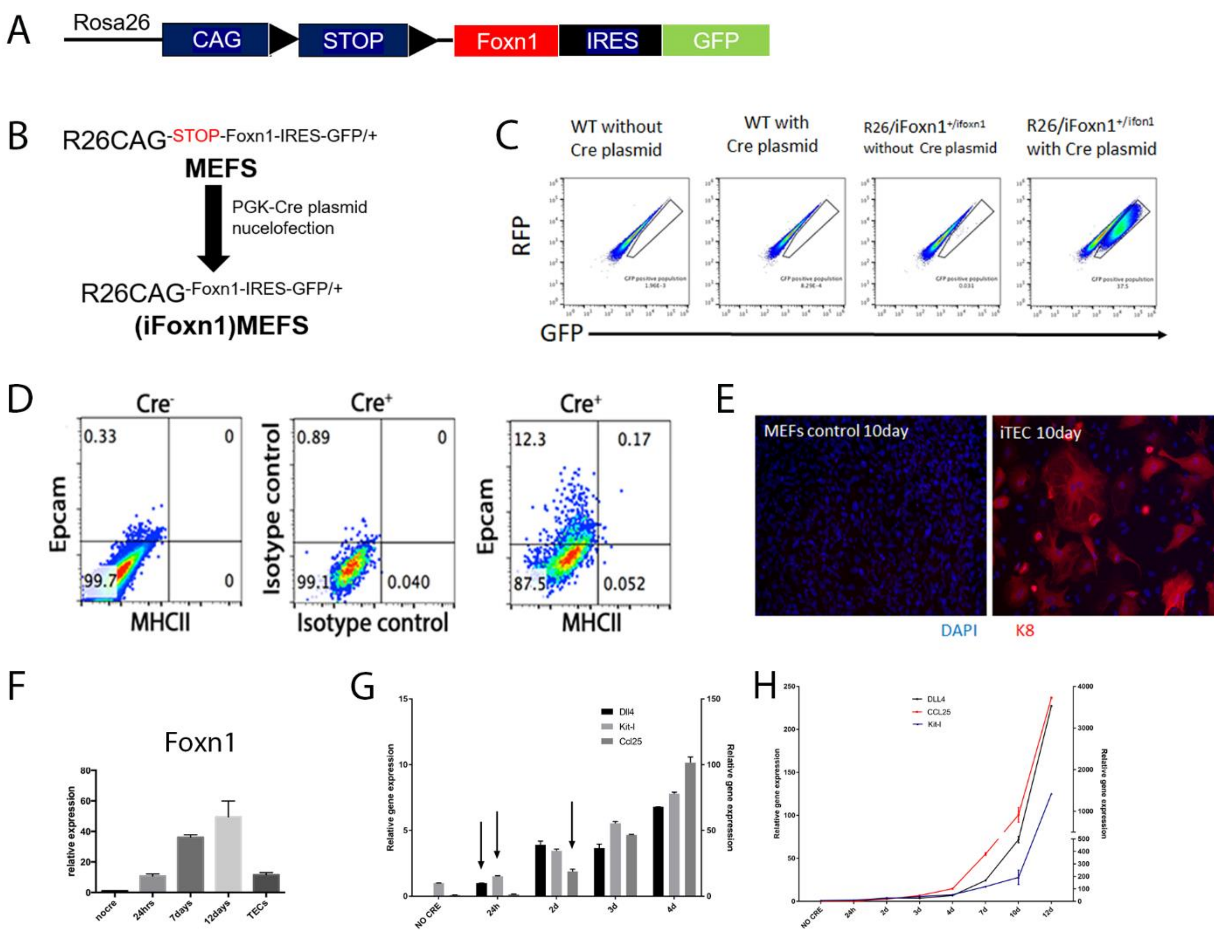


Figure 2.2:

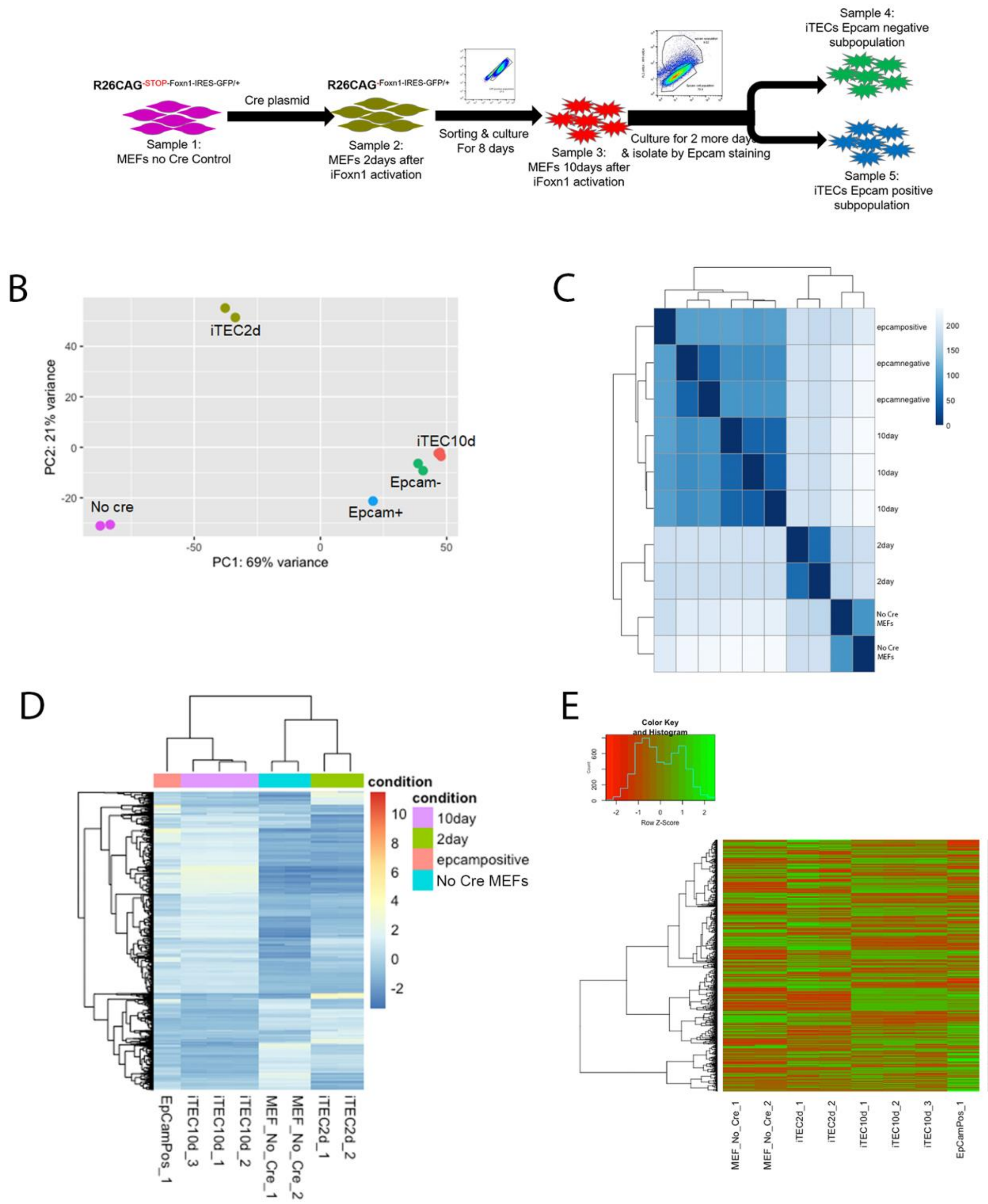


Figure 2.3:

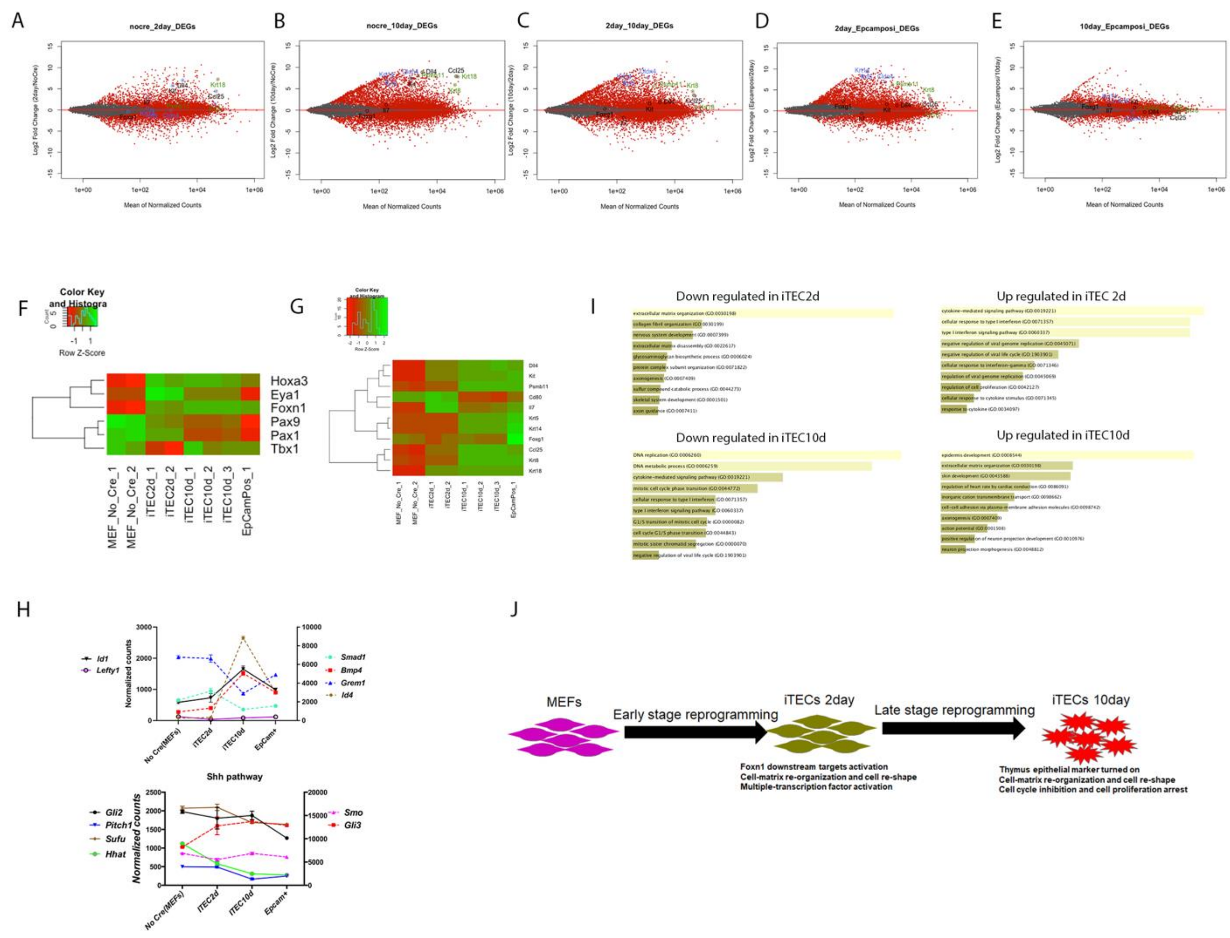


Figure 2.4:

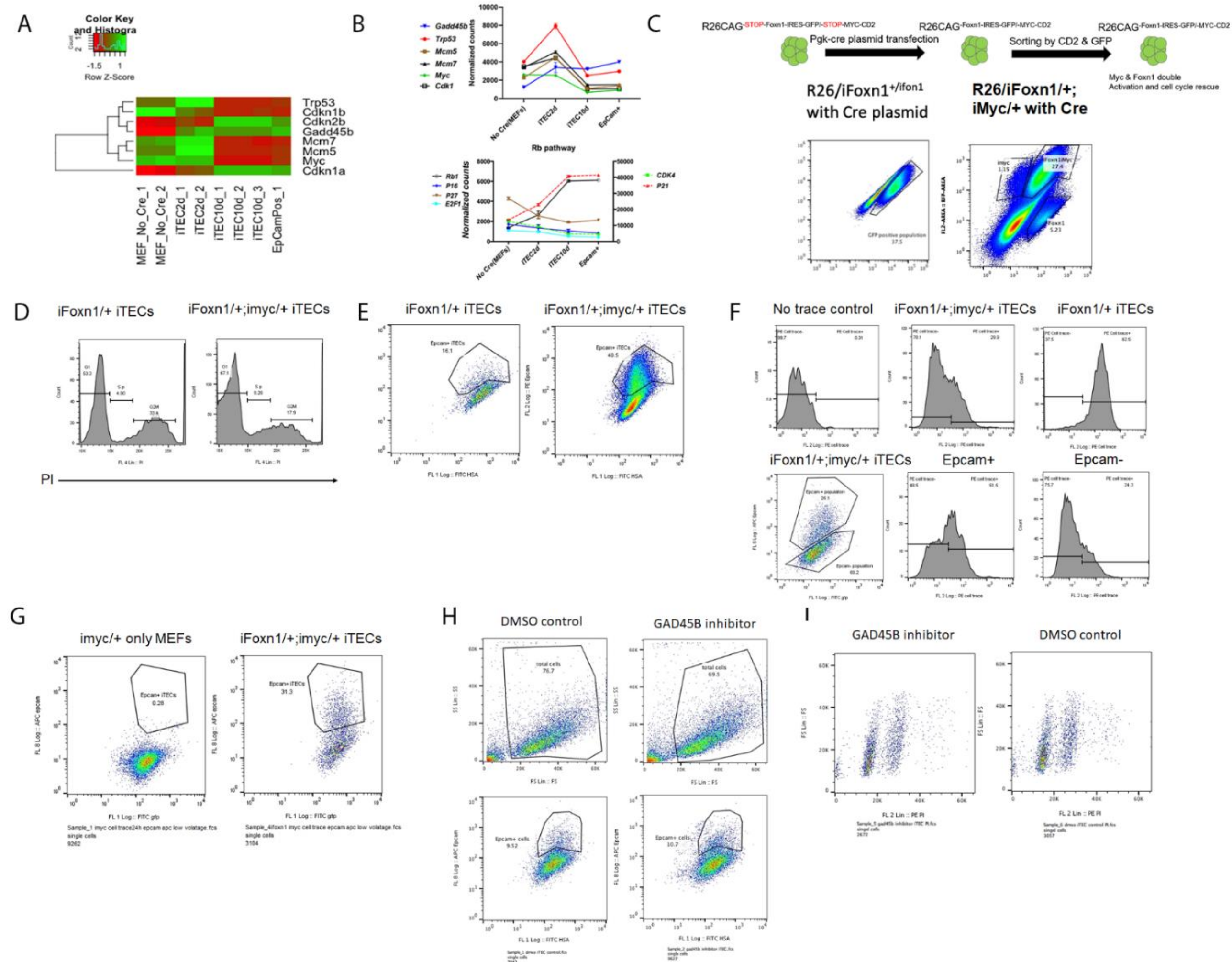


Figure 2.5:

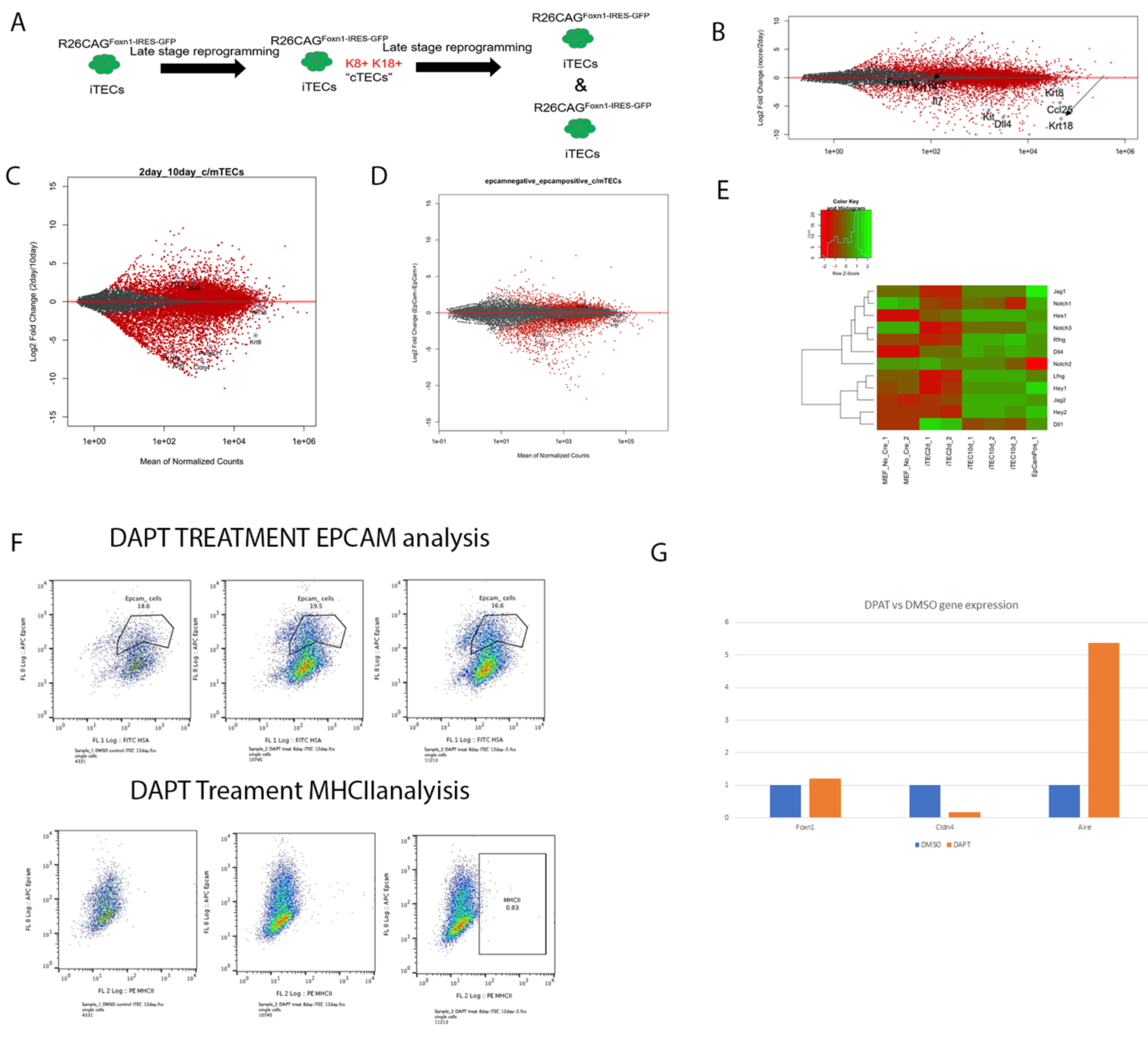


Figure 2.6:

Chapter 3:

Correlation Between Transgenic Foxn1-GFP Reporter and Endogenous Foxn1 Gene Expression differs in fetal & postnatal stages

INTRODUCTION

The thymus is a bilobed organ that is in the central compartment of the thoracic cavity and serve as the major primary lymphoid organ that produces self-restricted and self-tolerant T cells, which plays a very important role in our adaptive immune system.The T cell production in the thymus involve a lot of crosstalk between T cells and thymus stromal cells, especially thymus epithelial cells (TECs). The TECs can be roughly divided into two different cell populations, cortical thymus epithelial cells (cTECs) and medulla thymus epithelial cells (mTECs).

Several transcription factors play a significant role in both thymus development and function. One of the key genes that is involved in thymus development and TECs differentiation is Fork head Box N1 (Foxn1). Foxn1 nude mice loses their normal hair and also has disruption of thymus development. Researchers have identified Foxn1's crucial role in differentiation of TECs in the developing thymus, the maintenance of TECs in postnatal stages as well as TECs proliferation and thymus involution.

Several different mouse models have been developed to identify the spatial and temporal patterning of Foxn1 in thymus development as well as purify TECs by flow cytometry. One method to detect Foxn1 expression is the Foxn1-EGFP BAC transgene which has easily detectable, high-level GFP expression. This Foxn1-EGFP transgene mouse model has been widely used in both fetal and postnatal stages. However, the correlation between this transgene and endogenous Foxn1

gene expression and protein levels has not been verified in detail. Moreover, the Foxn1-EGFP BAC transgene mice shows significant hairless and smaller body size phenotype compare to wildtype mice. Hence, there remains a need to demonstrate the correlation between this EGFP reporter and endogenous Foxn1 gene. Also, validation of the Foxn1-EGFP|tg mice thymus phenotype would also generate enough reference information for applying this Foxn1-EGFP|tg mice into different stages of thymus study.

Here we demonstrate the correlation between the Foxn1-EGFP reporter and endogenous Foxn1 differs during various stages. In fetal stages and early postnatal stages (newborn to 4 weeks), this reporter has a strong correlation between endogenous Foxn1 and Foxn1-EGFP reporter both in location and levels. However, in aged 3-month-old mice, 6-month-old and 1 year old mice, the correspondence between GFP and FOXN1 protein is limited. Also, flow cytometry analysis shows that GFP-high cell population shows strong correlation to the EPCAM+/MHCII+ TECs population in early postnatal stages and can be used as marker for purification of TECs. In aged mice, The GFP-high cell population correlation to the EPCAM+/MHCII+ TECs population is limited. The GFP-LOW cell population correlation to EPCAM+/MHCII+ TECs population is always limited and should not be used for TECs purification. Finally, we demonstrate that the Foxn1-EGFP|tg mice's thymus is phenotypically normal and displays neither epithelial nor T cell cell defects in early postnatal stages.

Results

1) **Foxn1-EGFP transgene mice has slow growth and a hairless phenotype without affecting thymus/bodyweight ratio.**

Foxn1-EGFP mice has a clear body size and hair phenotype when compare to BL6 WT mice postnatally. (Fig3.1A to E). At two weeks stages, Foxn1-EGFP heterozygous mice is at least 25% less than normal BL6 WT control mice in body size (Fig3.1E). Also, they can be easily distinguished by the sparse hair phenotype as well as less length in general (Fig3.1A). The thymus in the 2weeks Foxn1-EGFP heterozygous mice is also significantly smaller than BL6 WT control mice. However, since the Body size of the Foxn1-EGFP heterozygous mice is smaller too, the overall body weight to thymus weight ratio is not significantly different (Fig3.1A). When the mice reach 5weeks, the sparse hair phenotype is still observable, but the body size difference is decreased to about 5% to 10% in general (Fig3.1E). Also, the thymus size and weight difference is also smaller at 5 weeks stages and the thymus to body weight ratio is not significantly different at this stage too (Fig3.1B and Fig3.1E). Another morphology phenotype of Foxn1-EGFP transgene mice is that the hair phenotype and body size phenotype is transgene-allele copy dependent, which means homozygous has a more severe phenotype. Foxn1-EGFP homozygous mice has no hair since newborn stages and grow slower than both heterozygous Foxn1-EGFP mice and BL6 WT mice. At 6-month age, Foxn1-EGFP homozygous mice has a very clear no-hair phenotype while the body size reaches the same level of the BL6 WT mice. (Fig3.1D). Thus, GFP heterozygous mice has a growth slowdown phenotype as well as hairless phenotype from newborn stages to at least

3 months. Also, Foxn1-EGFP transgene mice can provide clear GFP reporter signal starting from embryo stages. (Fig3.1C). To sum up, since these Foxn1-EGFP reporter transgene mice is widely used in analyzing thymus biology but have some suspicious phenotype such as slow-down in growth and hair-less, we decide to further analyze the correlation between real Foxn1 in thymus to the GFP reporter expression in different stages as well as the thymocytes phenotype.

2) Foxn1-EGFP reporter has good correlation to thymus Foxn1 gene expression in early postnatal stages and fetal stages

We first analyze the correlation between EGFP reporter and Foxn1 gene expression in early postnatal stages and fetal stages using IHC antibody staining. In Newborn stages, EGFP reporter is mainly observed in nuclei and has a very nice overlap with our Foxn1 antibody (more than 90% of the GFP reporter positive cells are also overlapping with Foxn1 antibody staining). These indicates EGFP transgene reporter can correctly indicating the Foxn1 expression cells, but not the expression level of Foxn1. (Fig3.2 A to C) At 2 weeks stages, we start to observe GFP reporter show up in cytoplasm region, but we can still observe a clear nuclei expression. When we compare the expression correlation between EGFP reporter and Foxn1 antibody, we can see the correlation is lower compare to newborn stages but still more than 80% of the EGFP positive cells are also Foxn1 positive. So, at 2 weeks stages, we can still use EGFP reporter as our thymus Foxn1 gene expression indicator. (Fig 3.2D to F). At fetal stages, we analyzed E14.5 embryos using Foxn1 antibody and GFP reporter. Although GFP reporter shows different signal strength in different region of feta thymus compared to Foxn1 antibody (Fig3.2G to I), the overall overlapping is still high at

fetal stages. Almost every GFP positive cell is also staining positive for Foxn1 antibody. To sum up, from fetal stages to early postnatal stages such as 2 weeks, the Foxn1-GFP reporter is expressed mainly in nuclei and highly overlapping with Foxn1-antibody staining in cell-to-cell co-localization, not in gene expression level. So we can make the conclusion that Foxn1-EGFP transgene reporter can be used as reliable indicator of endogenous Foxn1 expression in fetal stages as well as early postnatal stages.

3) Foxn1-EGFP reporter has poor correlation to thymus Foxn1 gene expression after early postnatal stages.

We also investigate postnatal stages after 2 weeks to check whether the correlation of GFP reporter gene and endogenous Foxn1 remain good. At 4 weeks postnatal stages, even only 2 weeks older than 2 weeks stages, we start to see significant decrease of the correlation between Foxn1-EGFP reporter and endogenous Foxn1. We start to see certain cells and regions that are GFP reporter positive but negative for Foxn1-antibdy staining (Fig3.3 A to C). Also, GFP expression is no longer limited in cell nuclei and has been extended into cytoplasm. Overlapping analysis shows that less than 75% GFP positive cells is still Foxn1 antibody staining positive. To further analyze the effect of age to the EGFP reporter gene, we also test the correlation at 6-month stages, in which thymus has been highly involuted. At 6 month stages, large proportion of the GFP reporter detected in the cytoplasm which makes it very hard to distinguish single cell at this stages. Also, we see the a significant increase of loss overlapping between Fonx1-antibody positive cells and GFP reporter positive cells at this stage compare to 4 weeks stages (Fig3.3 D to F). Only less than 60 % of the total GFP positive cells is now still staining positive for foxn1 antibody. These results indicate that though the aging process, the Foxn1-EGFP reporter is keep losing correlation to the endogenous Foxn1 expression. After 4 weeks, the

correlation is already poor and when it reaches 6 months, the GFP reporter can no longer be used as a reliable indicator of the endogenous Foxn1 expression cells. (Fig3.3G) the Foxn1-EGFP transgene mice has a clear hair-less phenotype which is highly similar to hypo Foxn1 phenotype such as Foxn1-nude mice. Foxn1 gene is highly exressed in both thymus and skin, so we also using qRT-PCR to analyze the Foxn1 gene expression in thymic epithelial cell and total skin cells（Fig3.3H）. We found that the Foxn1 gene expression is not significantly affected in Foxn1-EGFP transgene mice TECs, but significantly down-regulated in Foxn1-EGFP mice skin. This down-regulation of Foxn1 expression in skin may explain the hairless phenotype but further experiments is still needed to fully understand the hairless phenotype.

4) Foxn1-EGFP cell population's correlation to mature TECs depends on GFP expression level and stages postnatally.

Foxn1-EGFP reporter mice is also widely used as a tool for sorting Thymic epithelial cells based on the TECs Foxn1 expression. However, detailed analysis of the thymic epithelial cells in flow cytometry is unclear in Foxn1-EGFP mice. Here we isolate the thymic epithelial cells from Foxn1-EGFP transgene mice at different stages and test whether they are GFP positive too. In our thymic epithelial cells population from newborn stages to 4 weeks stages(CD45-/Epcam+/MHCII+), more than 90% of these cell population is also GFP signal positive(Fig3.4 A-F). This indicates at early postnatal stages, GFP reporter can be used as for sorting Thymic epithelial cells. However, at 3month 15% of the thymic epithelial cell is now GFP low(Fig3.4E). And at 6 month stages, almost 30% of the thymic epithelial cell population is either GFP low or GFP negative (Fig3.4F). Then we ask the reverse question that how many GFP positive cells is really thymic epithelial cells. Within CD45- cell population that isolated

from thymus, we first finds that the CD45- cell population can be divided into 3 different cell population: GFP negative, GFP low and GFP high cell population (Fig3.4 G to K).

The GFP low cell population has detectable GFP signal but only 15% or less are real Epcam/MHCII double positive thymic epithelial cells from all stages.(Fig3.4M). However in GFP high cell population, almost all GFP high cell population is also Epcam/MHCII double positive thymic epithelial cells from newborn stages to 3 month stages (Fig3.4L). At 6 month stages, the GFP high cell population is no longer contains high percentage of double positive thymic epithelial cells, only 42 % of the GFP high cell population.(Fig3.4L) This may due to the lost of thymic epithelial cell identity during the thymus involution process but the GFP protein still remains inside of the cells.

These results together shows that at early postnatal stages (from newborn stages to as late as 3 month postnatally), most thymic epithelial cells (more than 90%) is GFP signal positive. Also, in early postnatal stages the GFP high cell population can always be used as a reporter for the Epcam+/MHCII+ thymic epithelial cells. To make a conclusion, The GFP high signal can be used for sorting thymic epithelial cells at early postnatal stages but the GFP low signal can't be used for sorting at any stages.

5) Foxn1-EGFP transgene mice does not effect T cell phenotype.

Since Foxn1-EGFP reporter mice has altered growth speed and skin phenotype, we further analyze the T cell phenotype within this reporter mice. At 4 weeks stages, the CD4/CD8 DN, CD4 SP, CD8 SP & CD4/CD8 DP T cell proportion is not changed at all in both hemizybous and homozybous GFP reporter mice compared to WT (Fig3.5A). This suggests that the Foxn1-EGFP mice has no effect on the T lymphocytes development and maturation at early postnatal stages. At 6 months(Fig3.5B), the CD4/CD8 DN, CD4 SP, CD8 SP & CD4/CD8 DP T cell proportion

are not significantly changed within Foxn1-EGFP hemi mice and Wt mice, indicating that even in aging mice the Foxn1-EGFP transgene is not affecting the T cell development and production in thymus.

Together these results shows that although the Foxn1-EGFP reporter mice has very similar phenotype to Foxn1-nude mice, the thymus can still support T cells for normal development at both early postnatal stages and late postnatal stages. This results indicating that the thymus's primary lymphoid organ function is not changed by the Foxn1-EGFP transgene and no significant evidence suggesting these mice is different compared to WT mice at immune system.

6) Foxn1-EGFP reporter expression represent Foxn1-lineage cells

Since previously we showed that some Foxn1-EGFP low expression cells are no longer stained positive for thymic epithelial markers such as EpCam & MHCII, thus detailed analzye the relationship between EGFP reporter expression and Foxn1 lineage expression is significant. This is to ensure EGFP expression is only representing those Foxn1-lineage cells in thymus, not randomly leaky expressed on non-Foxn1 lineage cells.

To accomplish this lineage tracing experiment, I crossed Foxn1-EGFP mice into Foxn1-cre; Cag-loxp-Tomato mice and generated Foxn1-EGFP;Foxn1-Cre;Cag-Tomato transgene mice. By using these mice, I was able to trace Foxn1 lineage because any cell that ever expressed Foxn1 would express the Foxn1-Cre and thus activate the Tomato and becomes tomato+. Then I could further analzye the overlapping and co-locoliaztion of GFP signal and tomato signal of the cells within thymus.

I first analzye whether all GFP postivie is also tomato positive by imaging 4weeks mice at 10x & 20x. I find that almost all GFP positive cells is also tomato positive (Fig3.6A to Q), which

indicating that Foxn1-EGFP reporter suceesfully represent Foxn1-lineage cells and no leaky GFP signal is observed in non-Foxn1 lineage cells. Further flowcytometry analysis also shows that compare to WT mice, in the lineage tracing mice, all GFP+ cells are also gated Tomato positive (Fig3.6S). Then I treid reverse analiyisis to check whether all Tomato positive cell are also GFP positive (Fig3.6M to R). I found that in contrast to our previous analiyisis, not all Tomato positive cells are also GFP positive, since several tomato+ cells are clearly not positive for GFP reporter. (white arrow in Fig3.6M to R). This results is also verified by flow cytometry analysis that we can observe several cells that are Tomato+ GFP- and some other cells are a stronger signal of tomato compare to GFP. This missing GFP signal in tomato+ cell population is possibly due to some TECs that previously expressing Foxn1 is gradually losing Foxn1 expressing because of aging.

Together these lineage tracing results shows that Foxn1-EGFP reporter is reliably representing the Foxn1 lineage and thus GFP positive cells can be identified as Foxn1-expressing cells. Also, the lineage tracing experimetns hows some TECs lose Foxn1 expression, which makes them becomes Tomato+ but GFP reporter negative. This data is consistant with our previous finding that more TECs are identified GFP signal low or even GFP signal undetected in flow cytometry in aging mice.

Discussion

The result of this study provide a detailed analyisis of the correlation between Foxn1-EGFP signals and endogenous Foxn1 expression at different fetal & postnatal stages. I first performed general charateriaziton of the Foxn1-EGFP transgene mice and shows that the Foxn1-EGFP mice is not only expressing GFP reporter, but have a significant phenotype in hair. The Foxn1-EGFP mice has significant less hair and interesintly, more copy of Foxn1-EGFP would leads to less

hair.Furthermore, I found that Foxn1-GFP transgene reporter mice is significantly grows slower than WT mice at early postnataly stages, leads to a smaller body weight, which the casue is still unknown and need further investigation. Secondly, I collected thymus sample and performed Foxn1 antibody staining at various fetal & postnatal stages. Thus, I was able to compare endogenous Foxn1 expression to Foxn1-EGFP reporter expression. Our resuls shows that GFP signal has good overlapping with endogenous Foxn1 expression at fetal stages as well as early postnatal stages, but decreased rapidly after 4 weeks postnataly. Flowcytomery analyisis further prove this correlation decrease through aging and further showed that GFP high cell population always representing TECs until 6 month while GFP low cell population is never reliable indicator for TECs. Finally, we analzye the T cell within Foxn1-EGFP mice and didn't detect significant lymphocyte change between Foxn1-EGFP mice and WT mice, indicating the Foxn1-EGFP mice's T cell is not impacted by the Transgene. Together, our data provide detailed analyisis of Foxn1-egfp transgene mice and shows the correlation between GFP reporter and endogenous Foxn1 at different age. These data would be helpful for further study that plan to untilize the GFP reporter to either purify TECs by sorting or track Foxn1 expression both in-vivo & vitro.

Several further studies is still necessary for for characterize this Foxn1-EGFP transgene reporter mice. Firstly, we still have no idea why the Foxn1-EGFP has hair-defect phenotype which is highly related to Foxn1-GFP transgene copy numbers.Also, although thymus is not highly impacted, the growth rate is significantly slowed down in Foxn1-EGFP transgene mice. The reason why these mice shows obvious hair phenotype and growth rate phenotype is still undefined and needs further investigation. One possibility is that the transgene is inseted into some function genes that regulate hair development and body growth. To investigate this potential cause, DNA choromatin sequencing need to be performed to identify all insertion site of the Foxn1-EGFP

transgene. Another possibility is that since the Foxn1-EGFP transgene shares the same promoter region as endogenous Foxn1, it's possible that the transgene promoter region compete with the endogenous Foxn1 and thus result in Foxn1 regulatory difference.

Thid chapter provide detailed analyisis of Foxn1-EGFP reporter mice which would help further study that utilize this GFP reporter in thymus and Foxn1-related research. Despite the aging related correlation decrease, Foxn1-GFP mice can be reliable used in fetal and early postnatal stages. However, further analyisis is still necessary to further undertand the hair and growth rate phenotype.

Figures and legends:

Figure 3.1: Foxn1-EGFP mice characterization. (A ,B & D) morphology and thymus comaprsion of GFP hets and WT mice in A) 2 weeks, B) 5 weeks & D) 6 months.

C) fluorescence microscope imaging of fetal Foxn1-EGFP embyors.

E) body weight comparision between Foxn1-EGFP hets & WT

F) 5 weeks Foxn1-EGFP hets, homozygous and WT thysmus weight/body weight comparision

Figure 3.2: GFP reporter analyisis by immunohistochemistry in Foxn1-EGFP mice. (A to I) staining of thymus with Foxn1 (red) and GFP(Green) in (a to c) newborn stages, (d to f) 2 weeks stages and (G to I) E14.5 stages.

Figure 3.3: GFP reporter analyisis by immunohistochemistry in Foxn1-EGFP mice at late postnatal stages. (A to F) staining of thymus with Foxn1 (red) and GFP(Green) in (A to C) 4 weeks stages and (D to F) 6 months stages.

G) analyisis of Foxn1+ vs GFP+ cell population overlap percentage

H) qRT-PCR analsyis of Foxn1 expression in Foxn1-EGFP mice thymus and skin.

Figure 3.4: Correlation between GFP signal and TECs identidy decrease as mice aging.

(A to F) flow cytometry staining of Epcam/MHCII for Foxn1-EGFP mice in different stages.

(G to K) flow cytometry gating example of GFP-low cell population and GFP-high cell population at different stages.

L) the percentage of Epcam/MHCII doulbe positive TECs analyisis in GFP high cell population

M) the percentage of Epcam/MHCII doulbe positive TECs analyisis in GFP low cell population.

Figure 3.5: T lymphocyte is not affected in Foxn1-EGFP transgene mice.

Foxn1-EGFP mice and WT mice T cell analyisis by flow cytometry staining of CD4 and CD8 in A) 4 weeks and B) 6 months.

Figure 3.6: Foxn1-egfp reporter analhyisis by immunohistochemistry in Foxn1-EGFP;Foxn1-Cre;CAG-floxp-tomato mice.

(A to F) 4x staining of thymus with Foxn1 (pink) , Tomato (Red) & GFP (green).

(G to L) 10x staining of thymus with Foxn1 (pink) , Tomato (Red) & GFP (green).

(M to R) 20x staining of thymus with Foxn1 (pink) , Tomato (Red) & GFP (green).

S) flow cytometry analysis of Tomato signal vs EGFP signal in 4 weeks Foxn1-EGFP;Foxn1-Cre;CAG-floxp-tomato mice.

Figure 3.1:

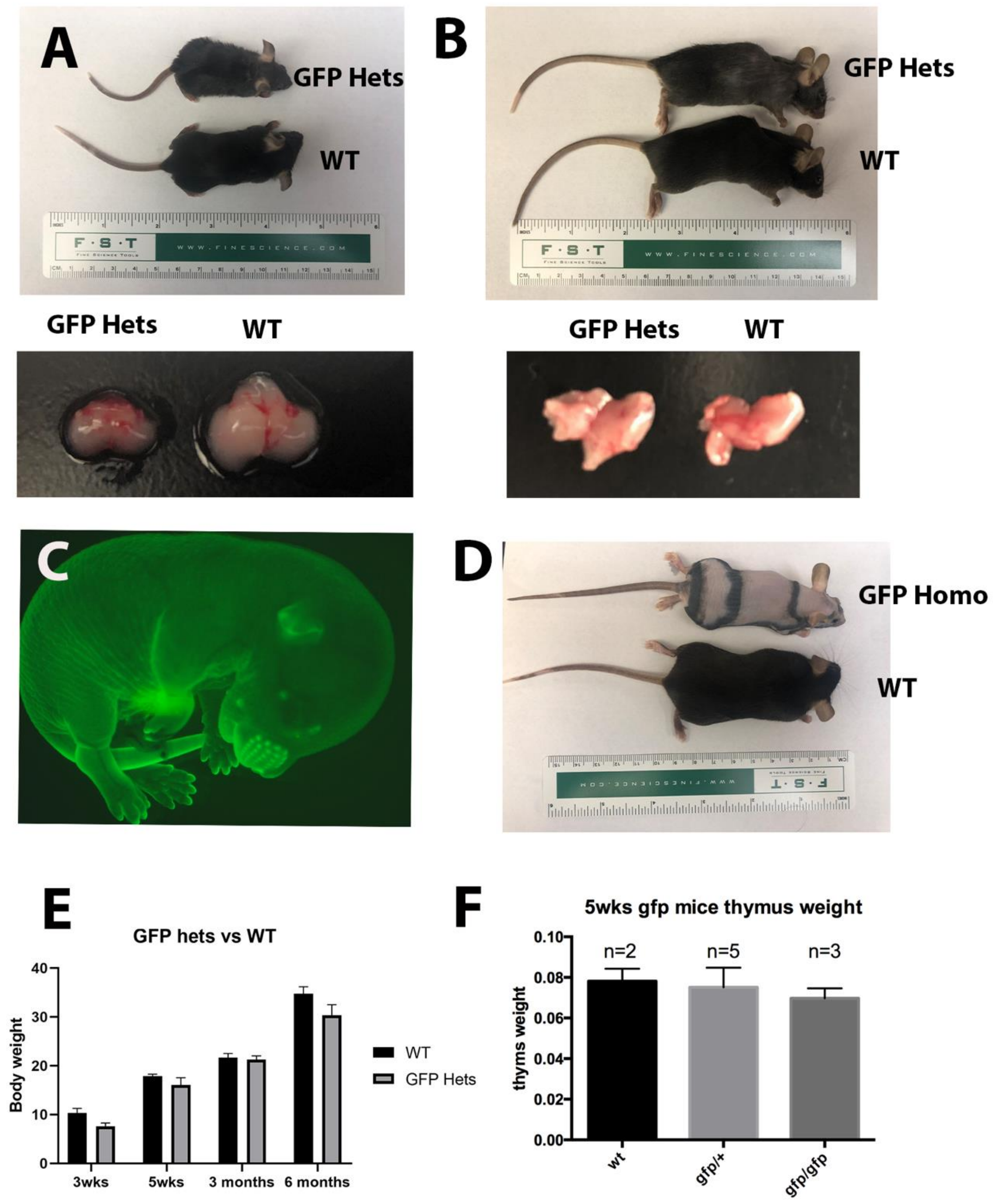


Figure 3.2:

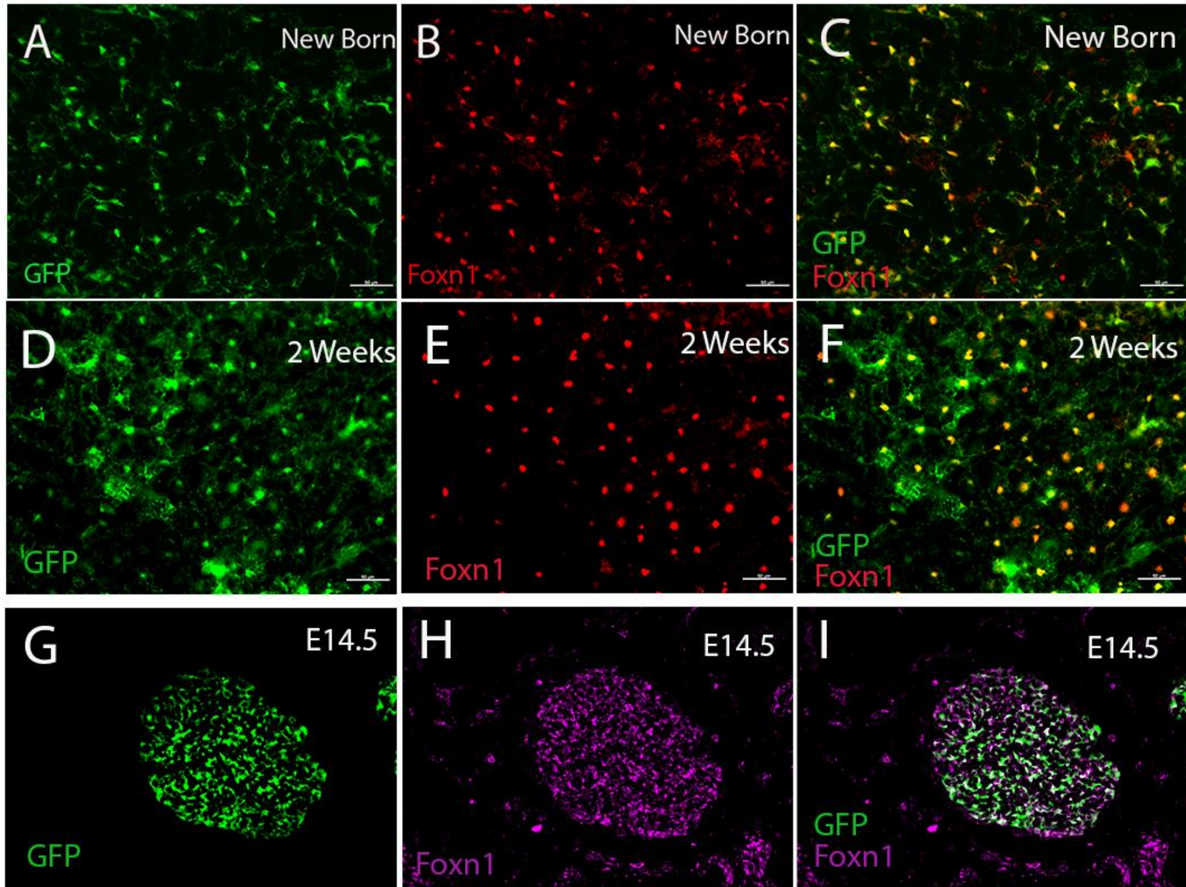


Figure 3.3:

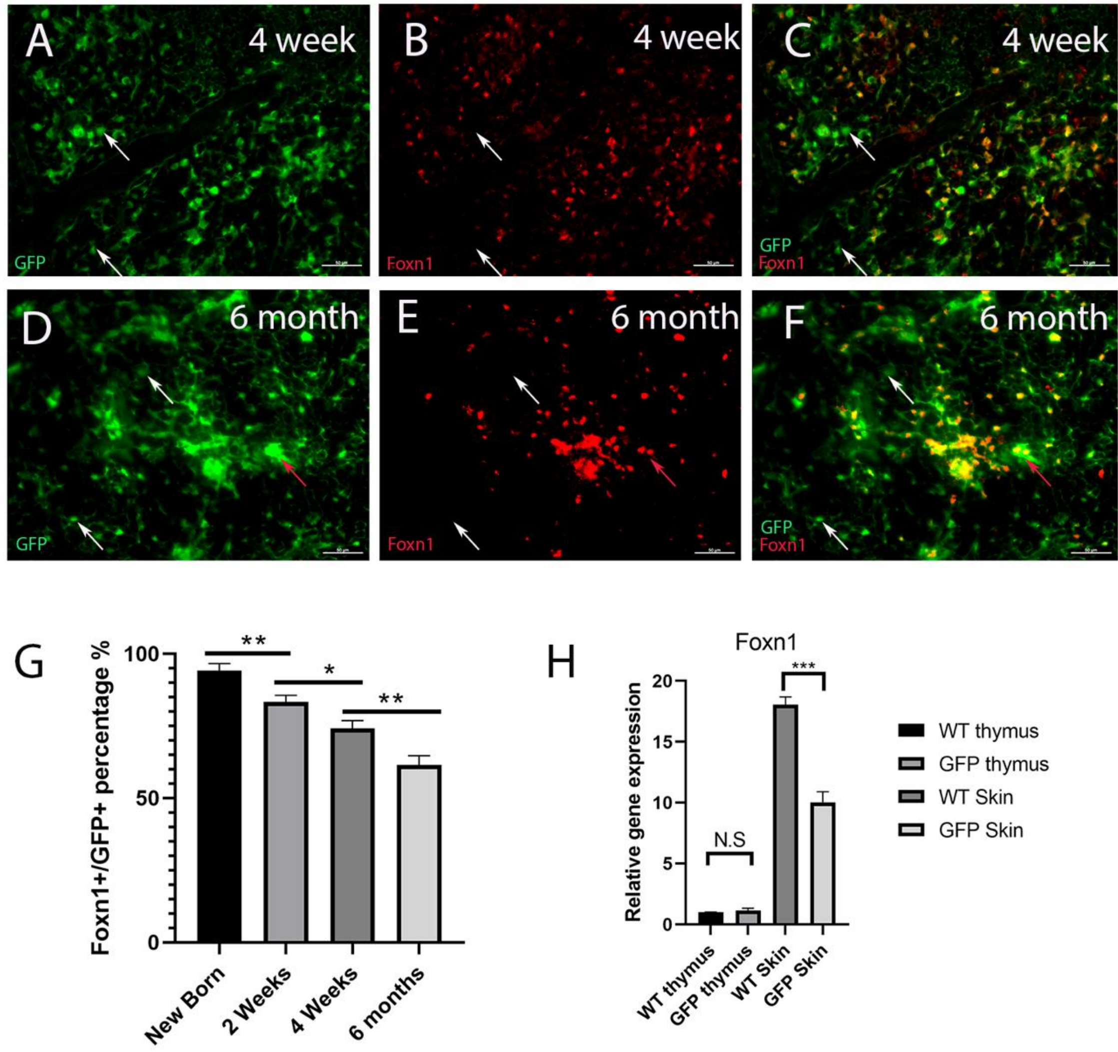


Figure 3.4:

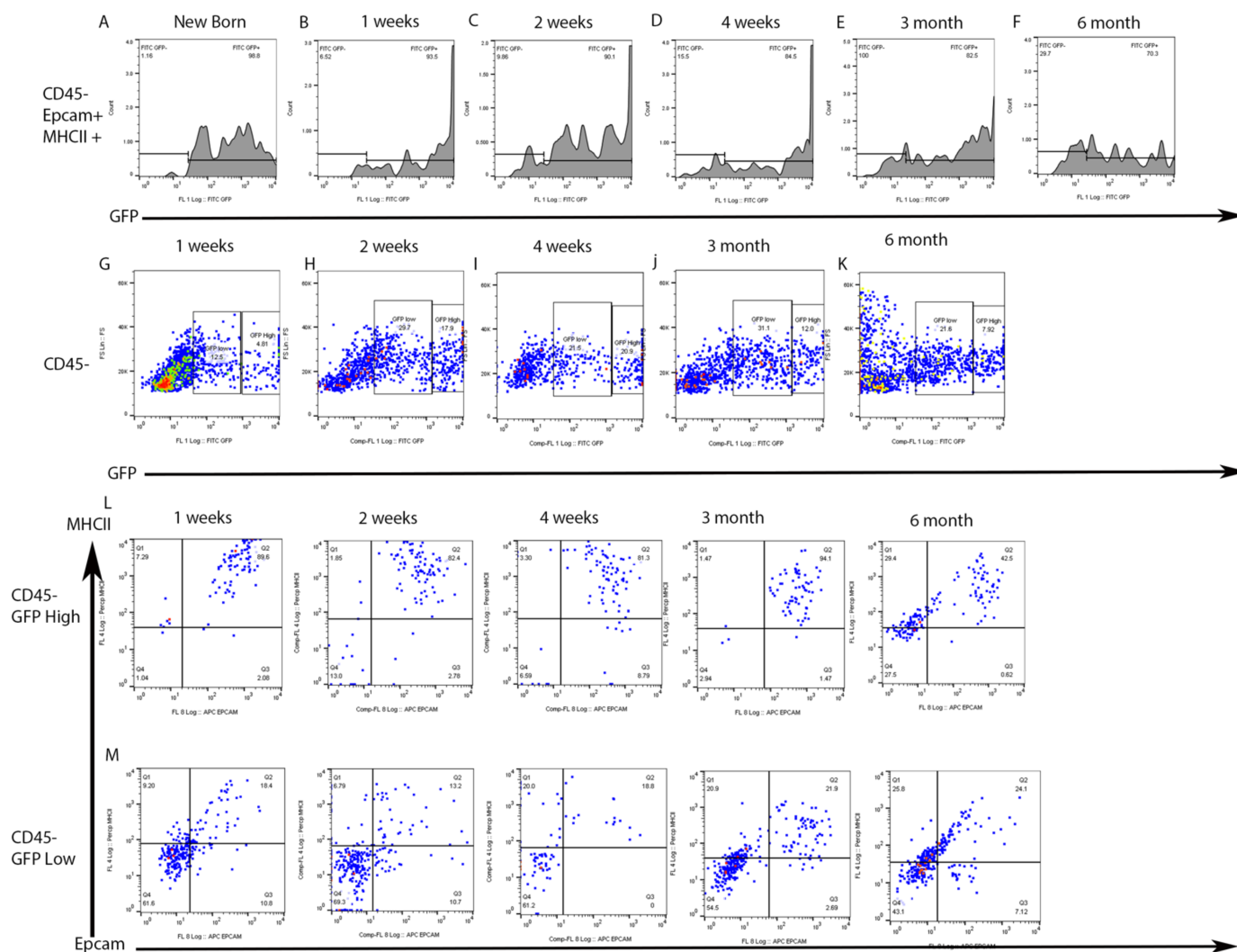


Figure 3.5:

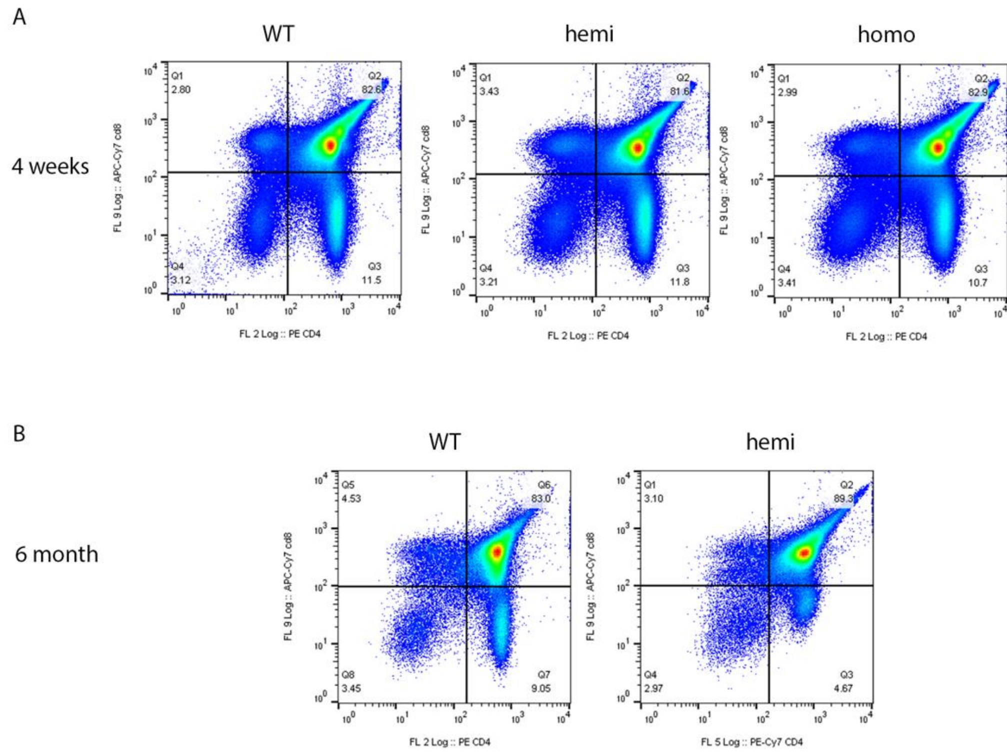
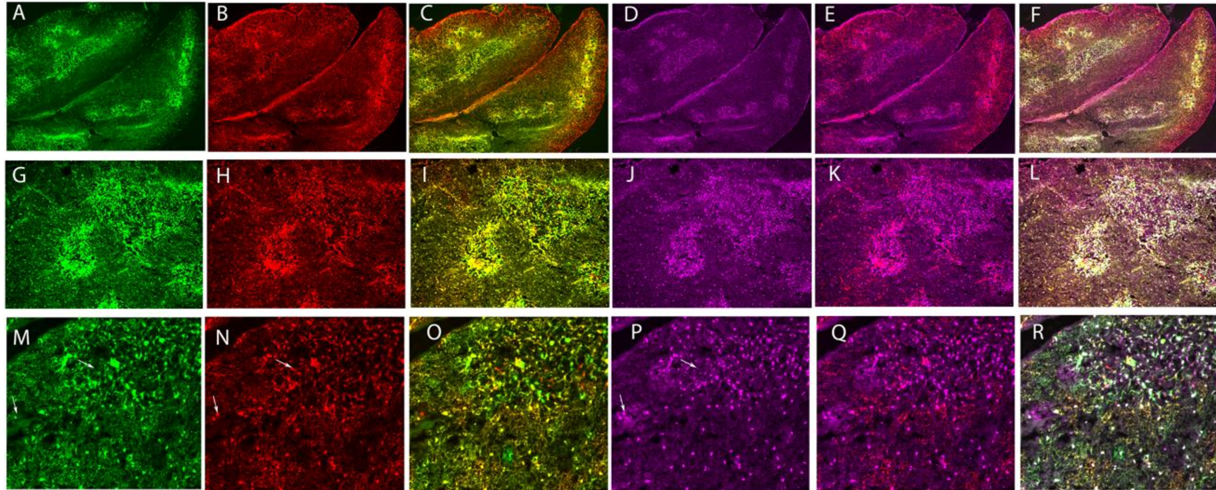
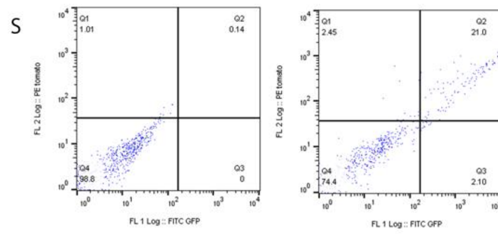


Figure 3.6:





Chapter 4: Activating Endogenous Foxn1 using CRISPRa system and its application in non-transgenic iTECs reprogramming

Introduction:

For decades, scientists have been focusing on discovering and utilizing gene-editing methods such as ZFNs and TALENs so that we can correct a variety of human diseases and defects by gene editing. The newly discovered Clustered Regularly-interspaced Short Palindromic Repeats (CRISPR) method has been a great potential gene-editing tool for gene therapy applications, especially after its rapid improvement and follow-up investigations(Lino, Harper, Carney, & Timlin, 2018). CRISPR was first discovered in the early 1990s as two genomic sequences from the halophilic archaeon *haloferax meditterranei*. It has also been considered an important candidate for gene regulation as a response to environmental conditions(Mojica, Juez, & Rodriguez-Valera, 1993). Scientists have focused on CRISPR's function, which is composed of 25-50 bp repeats separated by unique sequence spacers of similar length found in many bacterial species. CRISPR has been strongly associated with extrachromosomal elements' invasions and may provide adaptive cell immunity against phage infection and foreign DNA expression(Bolotin, Quinquis, Sorokin, & Ehrlich, 2005). However, the application of CRISPR as a gene-editing tool was not accomplished until the early 2010s. A family of endonucleases that uses dual-RNAs for site-specific DNA double-strand break was discovered. This CRISPR-associated (Cas) system was then deployed in RNA-programmable genome editing(Jinek et al., 2012). It was then shown that using the CRISPR-Cas system, which comprises complexes of the Cas9 protein and artificial

chimeric RNAs, it is possible to cleave genomic sites in human cells and introduce indels into the site of the double strand break(Cho, Kim, Kim, & Kim, 2013).

The thymus, the primary lymphoid organ, continuously loses its function and architecture through aging, which leads to an increased possibility of getting a variety of diseases. Thymus transplantation can reconstitute DiGeorge disease's immune function, but the thymus tissue availability for transplantation is strongly limited(Rice et al.) One way to solve this problem is by utilizing regenerative medicine, which can potentially restore diseased, injured, and aging tissues and even whole organs. For example, the Carticel, which is FDA approved regenerative medicine treatment, uses ex vivo expanded autologous chondrocytes to treat focal articular cartilage defects(Mao & Mooney, 2015). In the thymus, the key function cells are TECs, which support T cell maturation and development by crosstalk. Previously It has been shown that by overexpressing transcription factor Foxn1, it's possible to direct reprogram MEFs into functional induced thymic epithelial cells (iTECs) and also restore the immune system in immune-deficient mice. (Bredenkamp et al 2014). Thus, iTECs has the potential to be a regenerative medicine method for further clinical application.

However, several obstacles must be solved for further clinical application of iTECs. One major challenge is that current iTECs generation methods strongly rely on high-level overexpression of an exogenous Foxn1 transgene, which is not favored for clinical usage. Thus, it is necessary to find a non-transgenic way to reprogram cells into functional iTECs. One potential method is directly activating endogenous Foxn1 gene expression to facilitate iTEC reprogramming. To achieve this goal, we decide to utilize the CRISPR-activation system. The gene expression/amplification CRISPR (CRISP-activation, or CRISPRa), was first developed in 2014 by creating a fusion of a catalytically dead Cas9 enzyme and a repeating peptide array or activation

domains from transcription factors(Tanenbaum, Gilbert, Qi, Weissman, & Vale, 2014). Later, using the Cas9-VP64 activation domain, strong target-specific endogenous gene transcriptional activation could be achieved(Mali et al., 2013), which provides a potential solution for endogenous Foxn1 activation.

We first designed several gRNAs targeting the putative Foxn1 promoter region. By transfection the gRNAs and the dCas9-VPR activation domain, we show that the CRISPR-activation system can successfully activate endogenous Foxn1 to a robust level. Furthermore, by mixing several gRNAs targeting different regions of the Foxn1 promoter, we show that the synthetic effect of multiple gRNAs can significantly increase the expression level of endogenous Foxn1. Also, by using a Foxn1-EGFP reporter, we show that the activation of endogenous Foxn1 in MEFs can change the MEFs morphology, cell size and activate the Foxn1 downstream target Dll4 as well as the epithelial cell marker EpCam. Finally, we tested the same CRISPR-activation system in human dermal fibroblasts and induced robust expression of endogenous Foxn1. Together these results show that the CRISPR-activation system can successfully activate endogenous Foxn1 in MEFs and human dermal fibroblasts. This system has great potential to be further used as a reprogramming method to convert human fibroblast into functional iTECs for future clinical applications.

Results:

Design of gRNA targeting Foxn1 promoter region and endogenous Foxn1 activation by CRISPR-activation system

Most CRISPR-a systems are designed to target the gene promoter region, although some studies also show it is possible to activate genes by targeting the regulatory region. Since the Foxn1 gene has two promoters (Fig4.1a), 1a and 1b, we designed our gRNAs upstream of promoter 1a which is active in the thymus. Since different locations may have different activation effects, we decide to design as many gRNAs as possible to cover most of the 400bp upstream region of Foxn1 promoter 1a (Fig4.1b). We designed seven gRNAs targeting both sense and anti-sense strands of Foxn1 promoter with proper PAM sequences and low off-target scores.

In our design, we used a strong dCas9 activation domain dCas9-Vpr to activate the endogenous Foxn1. We co-transfected each gRNA individually with the dCas9-Vpr into MEFs to test whether we can successfully activate the Foxn1 gene. (Fig4.1C). Most of the gRNAs successfully activated endogenous Foxn1 compared to gRNA only and dCas9-vpr only control. (Fig4.1D). gRNA1, the gRNA closest to Foxn1 promoter 1a, had the most effective activation, which suggests this region may have an important role in Foxn1 promoter regulation. We further tested whether we can improve the endogenous Foxn1 activation level by combining several gRNAs targeting different regions. (Fig4.1E). Interestingly, some combinations of the gRNAs would lower the endogenous Foxn1 expression level, such as g1+g7, while other combinations increased the endogenous Foxn1 expression level (Fig4.1F). Moreover, by using several gRNAs, we obtained a stable, robust expression of endogenous Foxn1 compared to single or dual gRNA combinations.

To sum up, we designed several gRNAs targeting the upstream region of the Foxn1 promoter. Several gRNAs can successfully activate endogenous Foxn1 expression and gRNA1, which is closest to the Foxn1 promoter, has the strongest activation level. By combining several gRNA targeting different regions, we successfully further increased the Foxn1 activation level, which provides the possibility to utilize this method in iTECs reprogramming.

Tracking Endogenous Foxn1 expression and cell morphology change after CRISPR-activation using the Foxn1-EGFP reporter

While combining several gRNAs would result in the highest expression of endogenous Foxn1, this approach also requires transfection of multiple plasmids, and it is difficult to determine which cells receive all gRNAs as well as the dCAS9-vpr, much less correlate with relative endogenous Foxn1 expression level. Moreover, based on our RNA-seq data, the iTEC reprogramming process is a step-by-step process that requires sufficient time for the transcriptome change. Considering the endogenous Foxn1 expression from the CRISPR-a system is lower than the iFoxn1 transgene expression level, it is possible that it requires a longer time to reprogram MEFs into iTECs. Thus, developing a system to track endogenous Foxn1 expression quantitively in each cell is necessary during these technology development stages.

To successfully track endogenous Foxn1 expression, we generated MEFs from a mouse line containing a Foxn1-EGFP transgene (see Chapter 2). The Foxn1-EGFP reporter transgene consists of 27970bps of Foxn1 promoter fragment and a 4729bp EGFP reporter (Fig4.2A). Since this transgene includes the Foxn1 promoter as well as the upstream region (Fig4.2B), we proposed that the CRISPR-activation complex would target not only the endogenous Foxn1 gene but also the Foxn1-EGFP transgene. Thus, after the CRISPR-activation, we would not only activate the

endogenous Foxn1 but also the Foxn1-EGFP transgene. In this way, we could tract endogenous Foxn1 expression by tracking EGFP.

After the co-transfection of gRNA and dCAS9-vpr complex, we successfully activated Foxn1-EGFP expression (Fig4.2C). We did not detect GFP signals in our gRNA only controls in flow cytometry analysis, but did detect a distinct GFP+ cell population in our 4gRNA+dCAS9-vpr mix group. Moreover, the GFP signals covered a broad range of signal intensity, consistent with our assumption that CRISPR-based activation level would vary. Thus, we further divided our GFP+ cell population into GFP high and GFP low populations, and further analyzed the endogenous Foxn1 expression level (Fig4.2D). We find that the GFP high cell population has a higher endogenous Foxn1 expression level compared to GFP low cell population. The GFP high cell population endogenous Foxn1 expression was about 25% of the iFoxn1 transgene expression. These data suggest that we can successfully and accurately track endogenous Foxn1 expression by using Foxn1-EGFP reporter. Also, the expression level of endogenous Foxn1 is relatively high, suggesting it is sufficient to reprogram MEFs into functional iTECs.

Since iTEC reprogramming causes significant cell morphology changes, we analyzed cell morphology after CRISPR-activation. (Fig4.2E). After 12 days of CRISPR-a complex transfection, we surprisingly find that there are still several cells that have strong GFP signals, indicating that they are still expressing endogenous Foxn1 at a high level. Moreover, we find that these GFP positive cells have distinct cell morphology that is big and flat compared to normal MEFs. Multiple nuclei, which also happens in iTEC reprogramming due to the cell cycle arrest, were also observed in some GFP positive cells. Flow cytometry analysis also showed that the CRISPR transfected GFP+ cells are significantly larger than control MEFs. (Fig4.2F). At the early time point (day 2 after transfection), there is no significant cell size difference between GFPhigh,

GFPlow and GFP negative cell populations. However, at the late time point (15 days after transfection), the GFP high cell population is significantly larger than the GFP negative cell population. These data suggest that the activation of endogenous Foxn1 does result in cell morphology change consistent with iTEC reprogramming.

CRISPR-activation of endogenous Foxn1 leads to iTEC marker expression, and CRISPR-activation's effect in the intron.

Since we successfully activated endogenous Foxn1 by CRISPR-activation complex, and the cells changed morphology, we further analyzed other gene and marker expression. We first assessed the epithelial marker EpCam (Fig4.3A). Flow cytometry analysis identified a cell population with higher EpCam levels compared to the GFP negative cell population & MEF controls. Also, we detected significant Dll4 expression four days after CRISPR-activation transfection. Dll4 expression increased to a relatively high level 6 days after CRISPR-activation transfection. (Fig4.3B).

We also tested whether other regions could be used as a potential CRISPR-activation targeting regions for endogenous Foxn1 activation. To test this, we designed two gRNAs targeting the intron between Foxn1 exon two and exon 3. (Fig4.3C). By co-transfecting gRNAs with the dCAS9-VPR, we find that endogenous Foxn1 is successfully activated by gRNAs targeting the intron (Fig4.3D). Although the activation level is not as high compared to the expression from targeting the promoter region, these results suggest that it is possible to activate the endogenous Foxn1 gene by targeting intronic enhancers.

Activating endogenous Foxn1 in human dermal fibroblasts using CRISPR-activation

To further develop the iTEC method for clinical applications, we tested whether it is possible to activate endogenous Foxn1 and reprogram human fibroblasts into iTECs. We first analyzed the DNA sequence between human Foxn1 promoter region and the mouse Foxn1 promoter region (Fig4.4A). We find that the promoter region sequence is highly conserved between human Foxn1 and mouse Foxn1. This indicates that we can target the same upstream promoter region to activate endogenous Foxn1 in human dermal fibroblast. We designed four gRNAs targeting this potential human Foxn1 promoter region (Fig4.4B). We first used nucleofection to transfect a GFP control plasmid into primary human dermal fibroblasts. We find that nucleofection can achieve great transfection efficiency in these primary cells. (Fig4.4C). After co-transfecting human gRNAs with the dCAS9-vpr complex, we successfully activated endogenous Foxn1 in human dermal fibroblast. (Fig4.4D). However, only gRNA1 successfully activated endogenous human Foxn1, and the expression was lost by 6 days after transfection. Moreover, the expression level was relatively low compared to the level of endogenous Foxn1 after similar activation in MEFs. Together these results suggest that it is possible to use this method tod activate endogenous Foxn1 in human dermal fibroblasts. However, further optimization is required to increase the Foxn1 expression level to that needed for iTEc reprogramming.

Discussion:

The results of this preliminary study show the great potential of using CRISPR-activation system to reprogram MEFs and human dermal fibroblast into functional iTECs by endogenous Foxn1 expression. Also, these data reveal several difficulties that need to be solved in

future studies so that we can further apply this technology into clinical application. We have shown that by targeting Foxn1 promoter 1a upstream region, we can successfully activate endogenous Foxn1 to a relatively high level. Also, by using several gRNAs that targeting different region near the promoter, we can further increase the endogenous Foxn1 expression to a level approaching that of the iFoxn1 transgene. More importantly, by activating endogenous Foxn1 using CRISPR-activation system, we not only change the cell morphology of MEFs, but also successfully turn on other critical TEC marker genes such as EpCam and Dll4. Finally, we show that this system could be potentially be applied in human fibroblasts. Together these data show that CRISPR-activation system has great potential to be utilized as the system to reprogram iTECs by directly activating endogenous Foxn1.

One persistent question is how to maintain the CRISPR-activation expression long enough in the MEFs to obtain stable and high level expression of endogenous Foxn1. In our current experiment, due to the limitation of transient transfections, only a few cells still showed endo-Foxn1 expression after more than one week of culture. This low number of endo-Foxn1 positive cells makes it extremely difficult to analyze the cell characteristics, gene expression profiles, and immune function of these cells in detail, even when using the Foxn1-EGFP reporter gene to facilitate purification. Thus, developing a method to stably and efficiently express the CRISPR-activation complex is crucial. One possibility is by using Lentivirus and AAV vectors. However, our initial results suggest that Lentivirus vectors do not provide high efficiency of Endogenous Foxn1 activation (Appendix 1). Furthermore, AAV virus is limited by the maximum packaging size of the plasmid, which is less than the size of the dCAS9-Vpr coding sequence. Further optimization of the Lentivirus and AAV based CRISPR-activation system is required.

Another problem that needs to be resolved is that the currently obtainable endogenous Foxn1 expression may be still not high enough, especially in human fibroblasts. Although we tried using different gRNAs targeting the promoter region as well as design gRNAs that targeting potential regulatory intron region of Foxn1, the endogenous Foxn1 expression level is still less than half that of the iFoxn1 transgene. Since the CRISPR-activation domain is developing rapidly, from the VP64 into current VPR, one solution to further increase the endogenous Foxn1 expression level would be further optimizing the dCAS9-activation domain.

In this chapter, we provide CRISPR-activation system design and gene expression analysis of directly activating the endogenous Foxn1 gene. To further improve this method and efficiently reprogram MEFs and human fibroblast into functional iTECs, we need to develop new methods to stably express the CRISPR-activation domain as well as further increase the endogenous Foxn1 expression to a higher level. Further experiments are also necessary to study the immune function as well as transcriptome of CRISPR-activation generated endo-Foxn1 expression iTECs.

References

Bolotin, A., Quinquis, B., Sorokin, A., & Ehrlich, S. D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology, 151*(8), 2551-2561. doi:https://doi.org/10.1099/mic.0.28048-0

Cho, S. W., Kim, S., Kim, J. M., & Kim, J.-S. (2013). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nature Biotechnology, 31*(3), 230-232. doi:10.1038/nbt.2507

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science, 337*(6096), 816. doi:10.1126/science.1225829

Lino, C. A., Harper, J. C., Carney, J. P., & Timlin, J. A. (2018). Delivering CRISPR: a review of the challenges and approaches. *Drug Delivery, 25*(1), 1234-1257. doi:10.1080/10717544.2018.1474964

Mali, P., Aach, J., Stranges, P. B., Esvelt, K. M., Moosburner, M., Kosuri, S., . . . Church, G. M. (2013). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature Biotechnology, 31*(9), 833-838. doi:10.1038/nbt.2675

Mao, A. S., & Mooney, D. J. (2015). Regenerative medicine: Current therapies and future directions. *Proceedings of the National Academy of Sciences, 112*(47), 14452. doi:10.1073/pnas.1508520112

Mojica, F. J. M., Juez, G., & Rodriguez-Valera, F. (1993). Transcription at different salinities of Haloferax mediterranei sequences adjacent to partially modified PstI sites. *Molecular Microbiology, 9*(3), 613-621. doi:https://doi.org/10.1111/j.1365-2958.1993.tb01721.x

Rice, H. E., Hoehner, J. C., Gupton, I. K. C., Hale, L. P., Kepler, T. B., He, M., . . . Skinner, M. A. Review of 54 patients with complete DiGeorge anomaly enrolled in.

Tanenbaum, Marvin E., Gilbert, Luke A., Qi, Lei S., Weissman, Jonathan S., & Vale, Ronald D. (2014). A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging. *Cell, 159*(3), 635-646. doi:https://doi.org/10.1016/j.cell.2014.09.039

Figures and Figure Legends

Figure 4.1: CRISPR-activation system design diagram and endogenous Foxn1 activation

a) Gene structure of endogenous Foxn1 & the region for gRNAs design

b) The 7 gRNAs relative location in Foxn1 promoter 400bp upstream region

c) The schematic diagram of how dCAS9-VPR & gRNAs activating endogenous Foxn1

d) The different endogenous Foxn1 expression level by different gRNA

e) The schematic diagram of how dCAS9-VPR & multiple gRNAs combination activating endogenous Foxn1

f) The synthetic effect of different gRNAs combination on endogenous Foxn1 activation.

Figure 4.2: Track endogenous Foxn1 expression using Foxn1-EGFP reporter and endo-Foxn1 activated MEF morphology analysis

a) Gene structure of Foxn1-EGFP reporter and its relative location on Foxn1 allele

b) The schematic diagram of how CRISPR-activation complex could activate both endogenous Foxn1 & Foxn1-reporter gene

c) Flow cytometry analysis of GFP reporter signal proves that gRNA plus dCAS9-vpr complex could turns on Foxn1-EGFP transgene

d) Endogenous Foxn1 expression in CRISPR-activation experiment vs iFOXN1 transgene.

e) Fluorescence image of MEFs 12days after CRISPR-activation complex transfection

f) Cell size comparison between control MEFs and CRISPR-transfected GFP+ cells

g) Cell size analysis of GFP high population, GFP low population & GFP negative population at 2days & 15days after transfection

Figure 4.3: Epithelial cell marker expression and Foxn1 intron CRISPR-activation test

a) Flow cytometry analysis on GFP+, GFP- and MEF control for EpCam expression

b) Foxn1 and Foxn1 direct downstream target DLL4 expression after CRISPR-activation

c) Diagram of gRNA design location in Foxn1 intron region

d) Endogenous Foxn1 expression after gRNA targeting intron region and dCAS9-vpr activation complex transfection

Figure 4.4: CRISPR-activation on human dermal fibroblast

a) Human Foxn1 & mouse Foxn1 promoter region sequence conservation analysis

b) gRNAs relative design location on Human Foxn1 gene

c) GFP control plasmid nucleofection test on Primary human dermal plasmid

d) Foxn1 expression analysis after CRISPR-activation transfection on human dermal fibroblast

Figure 4.1:

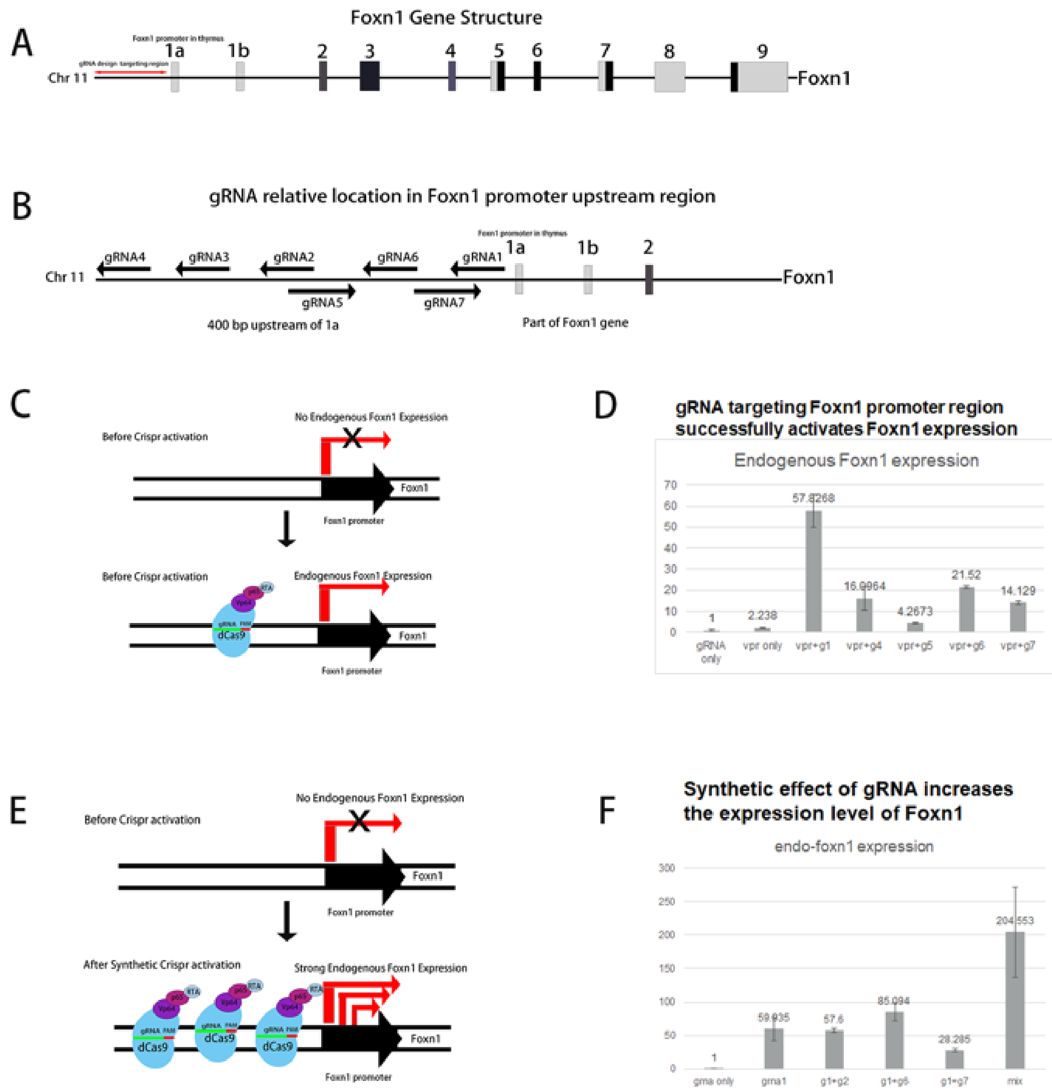


Figure 4.2:

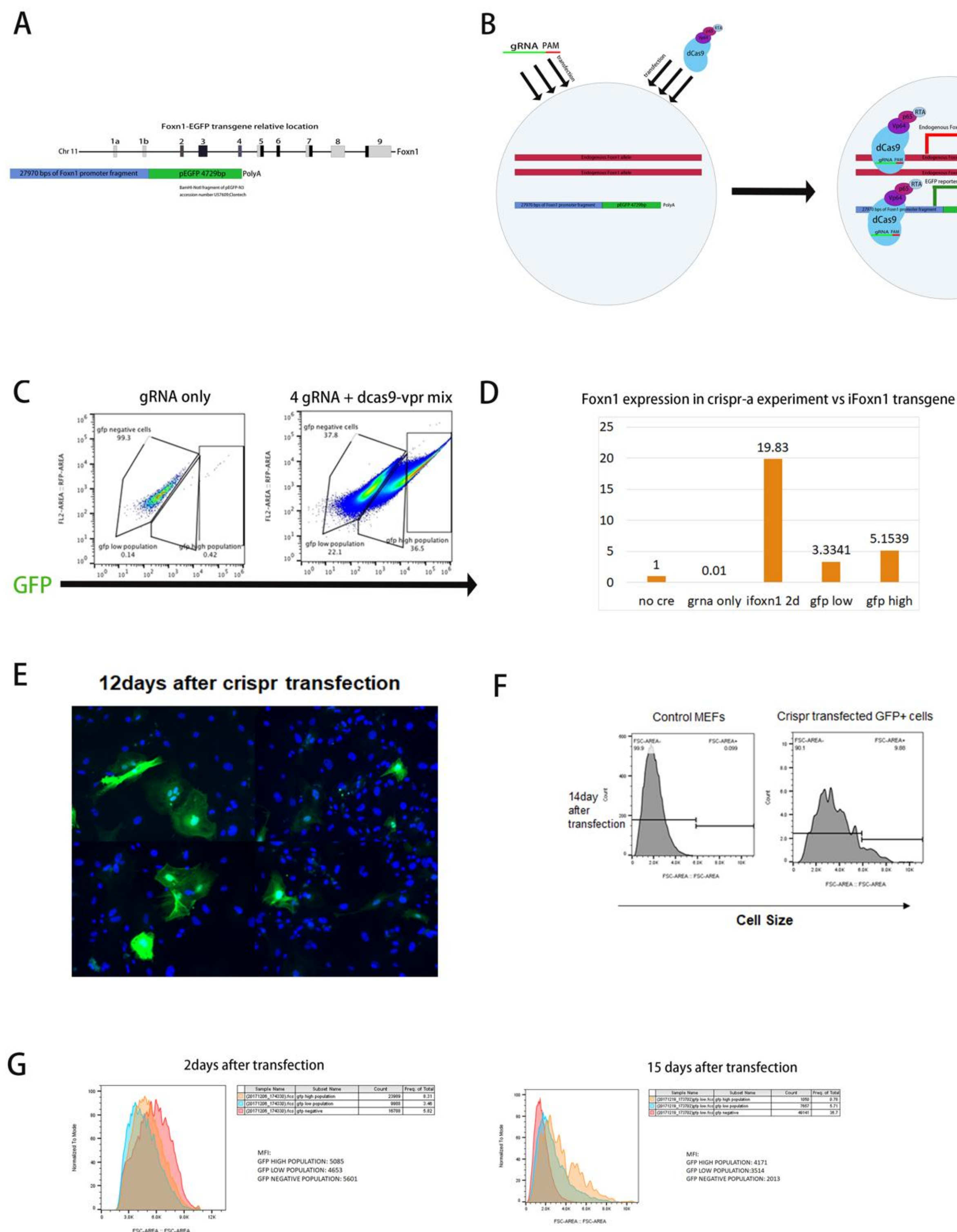


Figure 4.3:

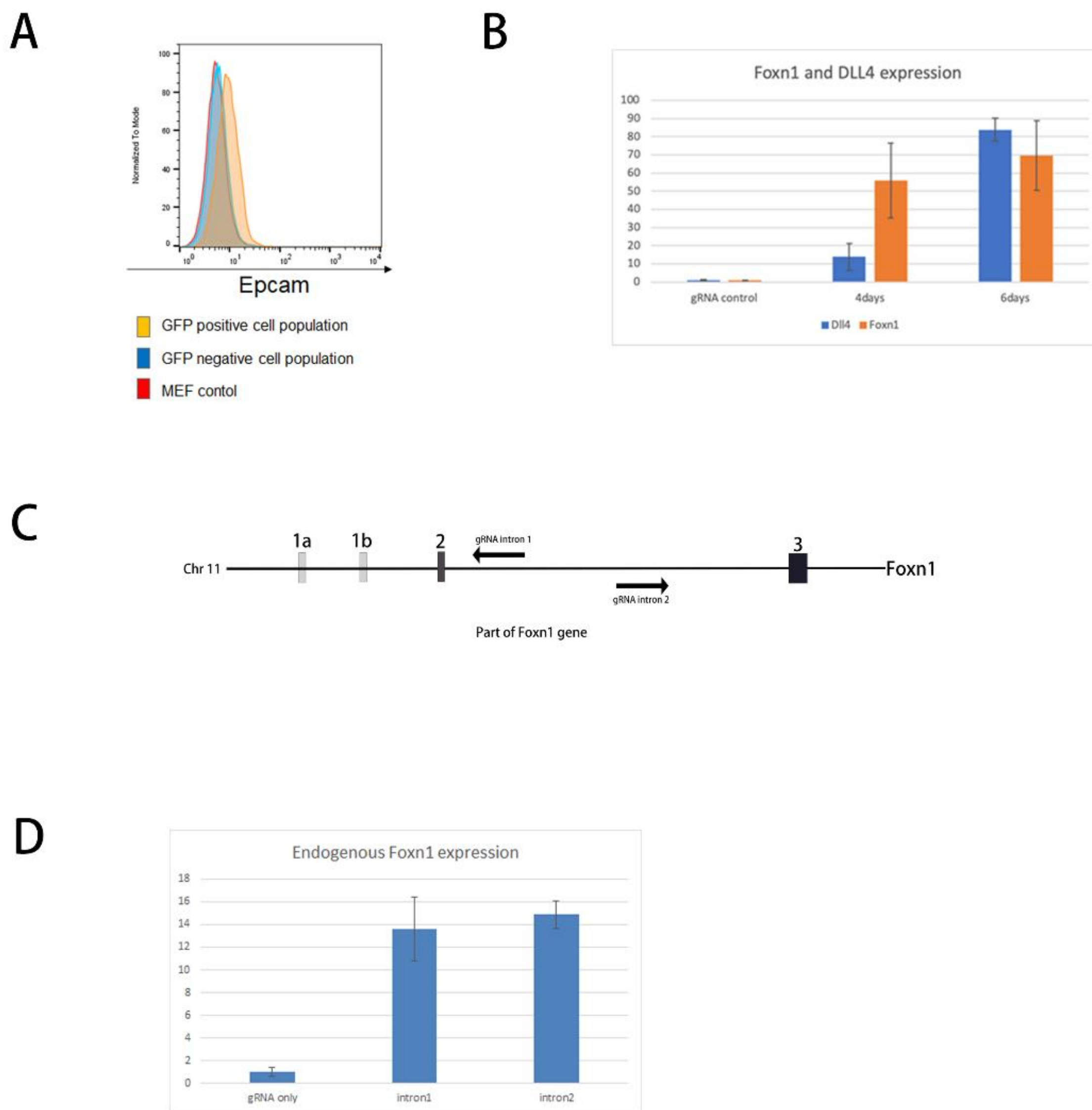
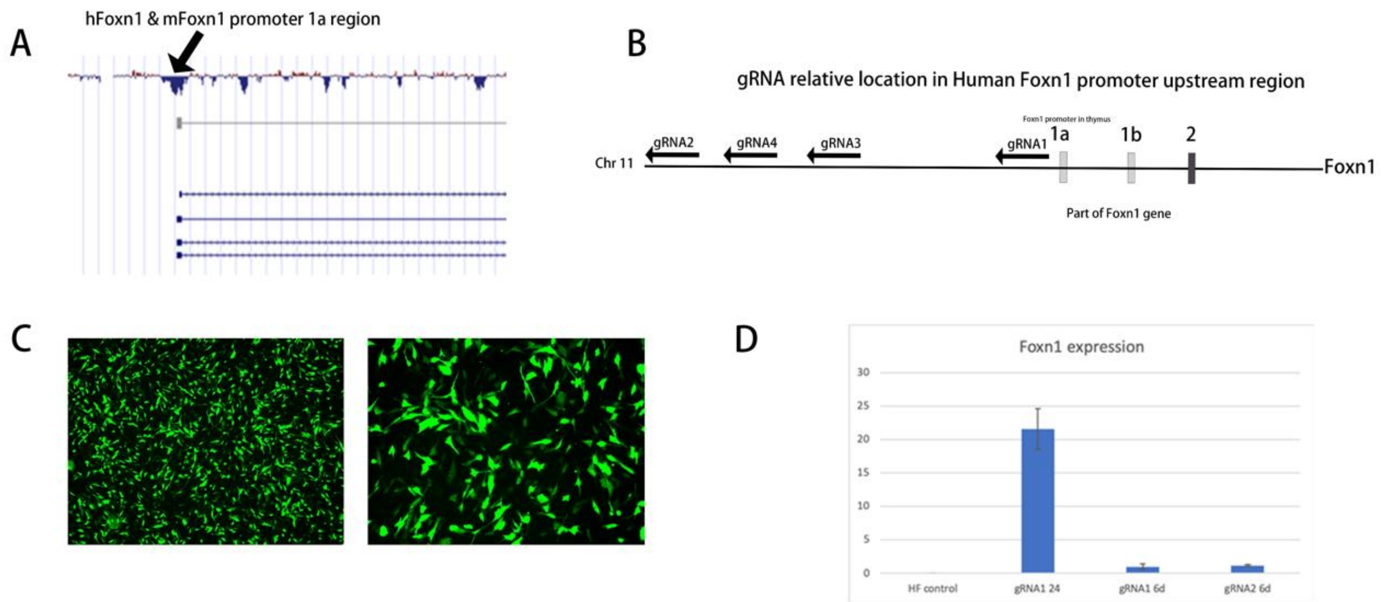


Figure 4.4:



Chapter 4 Appendix

CRISPR-based lenti virus transfection on MEFs

As we previously discussed in Chapter 4, in order to get stable and high expression of endogenous Foxn1 from CRISPR-activation, one solution is to utilize lentivirus/AAV based technology. To test whether these methods would facilitate our non-transgenic reprogramming methods, we developed a lenti-EF1a-dCas9-VPR-Puro/TOPO-gRNA system and tested its ability to activate endogenous Foxn1.

We first use lenti-EF1a-dCas9-VPR-Puro from Addgene^{TM} to activate endogenous Foxn1 (Fig6.1A). We also generated a lentivirus plasmid that contains the gRNA targeting the Foxn1 promoter region. Then, we co-transfected gRNA lentivirus and lenti-dCas9-VPR-puro into MEFs, (Fig.1B), so that dCAS9-VPR DNA and gRNA DNA would be integrated into the MEFs chromosome and stably express the dCAS9-VPR/gRNA activation complex.

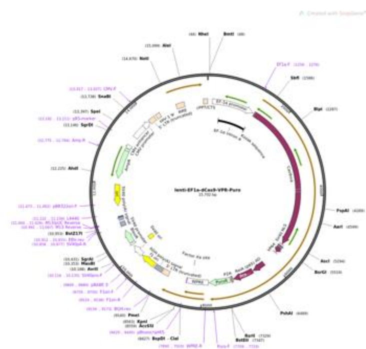
To test how lentivirus transduction works on primary MEFs, we first test the lenti-GFP. (Fig6.1C). We find that the lenti-GFP could generate good transduction results on MEFs, with more than 80% of MEFs being GFP+ after the lenti-GFP transduction. Then we further applied the lenti-EF1a-dCas9/lenti-gRNA into Foxn1-EGFP MEFs that we previously discussed in Chapter 4. After transduction, we found that only 8% of the primary MEFs were GFP+, indicating that they received the lenti-gRNA (Fig6.1D). Then we mixed the lenti-gRNA and lenti-dCas9 and transfected the primary MEFs. However, in contrast to the lenti-GFP experiment and the lenti-gRNA experiment, only 0.11% of the primary MEFs were GFP positive, indicating they turned on endogenous Foxn1 after the lenti-virus transduction. (Fig6.1E). Then we further

increased the lentivirus dose to test whether this would stably activate endogenous Foxn1 in more cells. (Fig6.1F). In this dosage gradient test, we found that a higher dose of lentivirus leads to more GFP+ cells that have activated endogenous Foxn1 expression. Around 0.5% of the GFP+ cells were detected in the flow cytometry indicating they are stably expressing the CRISPR-activation complex and thus turning on the endogenous Foxn1.

In this study, we successfully transduced lenti-dCAS9-vpr and lenti-gRNA into primary MEFs and used co-transduction of the lenti-dCAS9-vpr and lenti-gRNA to successfully activate the endogenous Foxn1 gene. Moreover, higher dosage of the lentivirus led to higher percentage of MEFs expressing endogenous Foxn1. However, this percentage is still well below that needed for further reprogramming and functional analyses. More study needs to be done to further improve this system.

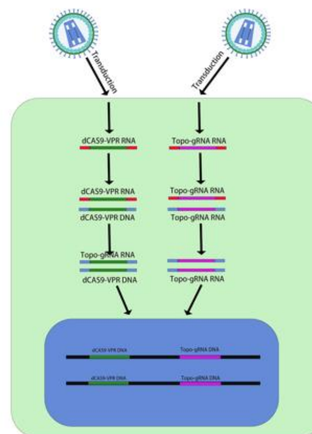
Figure 6.1

A

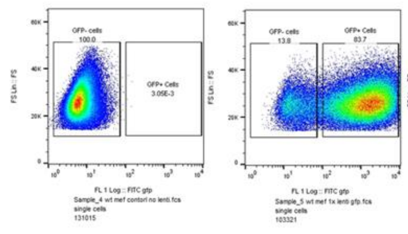


lenti-EF1a-dCas9-VPR-Puro was a gift from Kristen Brennand
(Addgene plasmid # 99373 ; http://n2t.net/addgene:99373 ; RRID:Addgene_99373)

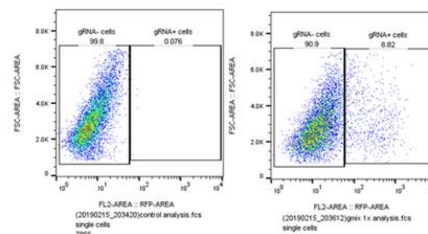
B



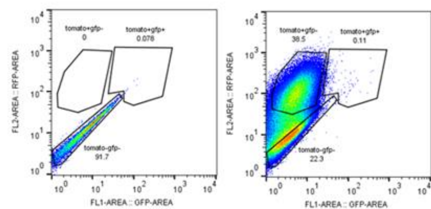
C



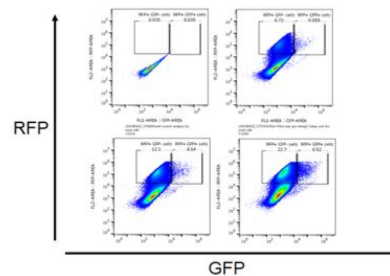
D



E



F



Chapter 5

DICUSSION AND FUTURE DIRECTION

Understanding the detailed mechanisms of thymus development and how thymic epithelial cells (TECs) support T cell development and maintain its function is required to devise both experimental and therapeutic approaches to prevent TEC from losing function due to thymus involution and maintain generation of naïve T cells. Furthermore, we do not currently have in vitro systems that recapitulate most aspects TEC function or allow detailed mechanistic or biochemical analysis of TEC differentiation or TEC-thymocyte interactions. Our previous collaborative work showing that overexpressing the key TEC transcription factor FOXN1 could directly reprogram mouse embryonic fibroblasts (MEFs) into functional induced thymic epithelial cells (iTECs) provided an avenue to bridge this gap. That original study showed that these iTECs not only express TEC markers but also can successfully support T cell development and rescue immune-deficient mice after kidney capsule transplant. However, the mechanisms underlying the direct reprogramming process from MEFs to iTECs is mostly undefined. Thus, it is unclear how many different mechanisms regulate iTEC reprogramming or control its efficiency and differentiation. Also, while it is clear that iTECs have significant differences compared to functional fetal TECs, more information is needed to design hypothesis-driven strategies to improve iTEC reprogramming before applying iTECs in clinical applications human cells. For these reasons, I undertook the current studies, which provide a foundation for future studies to

analyze the iTEC reprogramming process in detail, preferably at the single-cell level, and test impacts on their functionality.

The purpose of my Ph.D. research was first to understand the mechanism of how MEFs are directly reprogrammed into iTECs and to identify candidate gene pathways that may play essential roles in iTECs reprogramming efficiency and differentiation. To accomplish this goal, I collected iTECs from different reprogramming time points as well as fetal TECs at E14.5, utilizing a flow-cytometry based sorting strategy, and performed bulk-RNA-seq to analyze the transcriptome of iTECs & fetal TECs. I then analyzed the sequencing data using DEseq2 and related R-studio packages to identify the transcriptome differences between different time points of iTECs reprogramming and fetal TEC. The resulting differential gene expression analysis shows the reprogramming process has distinct intermediate cell types which suggests the iTECs reprogramming may follow a step-by-step process that is non-linear. I defined an early stage of reprogramming in which FOXN1 direct targets are activated quickly, but at a low level. This rapid activation of targets is consistent with FOXN1 acting as a pioneer factor. Pioneer transcription factor play a primary role in establishing competence for gene expression and start the cell fate direct change and cell reprogramming. (doi.org/10.1002/wsbm.1427). Thus, FOXN1's role in iTECs reprogramming is similar to other forkhead proteins, such as FoxA, which can bind to their target sites on compacted nucleosome DNA and open a local domain of compacted chromatin in vitro. (doi.org/10.1016/S1097-2765(02)00459-8)

Future studies to investigate this function at the chromatin level would provide direct evidence for the mechanism of FOXN1 function.

My approach was verified to provide candidate mechanisms that may mediate some of the main shortcomings of the initial iTEC study. Two critical issues of the iTEC protocol are 1) a rapid cell cycle arrest after Foxn1 expression; and 2) an apparent absence of mTEC lineage specification. Using pathway-specific gene expression profile analysis, I discovered several potential molecular signatures that suggested candidate approaches to improve iTEC differentiation and utility. For the cell cycle, I identified multiple cell cycle inhibitors that are activated during reprogramming, including the RB pathway, as well as reduced Myc and p53 expression. The response of iTECs to overexpressing Myc showed that cell cycle arrest is not required for reprogramming to occur, and indeed that the frequency of EpCam+ Tec was increased in these iTEC cultures. These very promising initial results should be followed by further experiments which would be discussed in later this chapter.

The identification of low levels of mTEC markers, including the mTEC progenitor marker Cld4, was a surprise, and suggested that mTEC specification was actually occurring at some level. The further identification of persistent Notch signaling across iTEC differentiation also provided a possible mechanism to enhance mTEC differentiation. Our results using DAPT inhibition and the resulting Aire expression are exciting developments and open a door to a variety of future experiments such as Notch-related gene regulation experiment and cell-based Notch related gene IHC staining which would be detailed discussed separately. By comparing the transcriptome between iTECs and fetal TECs, I showed the significant gene expression profile difference, suggesting further modification is necessary for improving iTECs reprogramming. Among the main gene expression differences, the absence of MHCII expression was a primary characteristic of fetal TECs that is absent even in our most mature EpCam+ iTEC sample. Thus, further experiments that focus on understanding

MHCII expression regulation mechanisms in iTECs and improve MHCII expression through different strategy, such as T cell co-culture.

However, currently, we still have several difficulties for further detailed analyze the thymic-related mechanism. One of them is it's hard to track thymus key transcription factor Foxn1. Foxn1 is an important thymus transcription factor, and Fonx1-nude mice would lead to athymic phenotype and completely immune deficiency. Foxn1 is a nucleus-located transcription factor, and for this reason, assessing Foxn1 in TECs by flow cytometry is difficult. This issue makes it hard to track Foxn1-expressing cells quantitively and measure Foxn1-expressing levels within the TECs at different stages. Furthermore, lacking a strategy for tracking Foxn1 expression also makes it hard to sort out Foxn1-expressing cells, in most cases, TECs directly, and perform transcriptome analysis of TECs at different development stages. Therefore, developing a Foxn1 tracking strategy is necessary further to understand thymus development and TECs gene expression mechanisms.

My second Ph.D. goal is to further examine the Foxn1-EGFP reporter expression in Foxn1-EGFP transgene mice in both fetal thymus and different postnatal thymus stages and determine whether Foxn1-EGFP reporter can be used as a reliable marker to trace Foxn1 transcription factor expression and localization. To accomplish this goal, I generated Foxn1-EGFP reporter mice samples from different stages, from embryonic and postnatal stages from E15.5 through 6 months of age. The resulting samples were further stained by Foxn1 antibody, together with some other antibodies. I examined the correlation between endogenous Foxn1 and the Foxn1-EGFP reporter expression and show that the correlation was good until the early postnatal stages but decrease quickly after 4 weeks. This result suggests that the Foxn1-EGFP transgene reporter mice can be used for reliable tract real

endogenous Foxn1 expression in fetal stages and early postnatal stages, but can't be used after around 4 weeks postnatally because of the loss of correlation.

Cre-plasmid based iTECs reprogramming methods & potential other iTEC reprogramming strategy

As mentioned above, by overexpressing Foxn1 gene, it is sufficient to reprogram embryonic fibroblast into TEC which can successfully support T cell development as well as rescue immune deficient mice. (Bredenkamp et al.2014). However, the previous system including two transgene system, in which one of them is CreER that can be activated by 4OHT injection. However, this CreER transgene-based system has possibility to cause leaking issue that activate inducible Foxn1 in MEFs before the 4OHT treatment. To solve this problem, we remove the CreER transgene and instead used the Cre-plasmid based nucleofection system to activate the inducible transgene Foxn1. By utilizing this system, we can successfully activate the Foxn1 at the time we want and avoid the un-wanted to express of inducible Foxn1. Using this Cre-plasmid based system, I was able to generate iTECs from different time points with reliable Foxn1 activation time and thus perform the RNA-seq transcriptome analysis.

However, this Cre-plasmid based system still have some issues that needs to be solved for further improving iTECs reprogramming and transfer it into potential clinical application. First, this system still needs a previously generated transgene iFxon1 allele so that the iTECs reprogramming could happens. the problem of including a transgene is that transgene is generally not allowed in clinical applications and thus make the iTECs reprogramming hard to transfer into real therapeutic application. Secondly, due to the cell number limitation of nucleofection, the number of iTECs that could be generated by each nucleofection reaction is highly limited. This

limitation of the number iTECs generated leads to several difficulties including hard to generating enough samples for different kinds of vivo & vitro experiments. To solve these issues in order to further improve iTECs reprogramming, I did some preliminary experiments and study to discover the possibility of reprograming MEFs into functional iTECs using non-transgenic methods.

One potential method that we can use is CRISPR-Cas, which is an RNA-mediated adaptive immune system found in bacteria and archaea, in which it protects host cells from invasion by foreign DNA elements. (Barrangou, R. et al 2007). The discovery of a small transactivating CRISPR RNA (trancrRNA), which directs the post-transcriptional processing and maturation of the CRISPR RNA (crRNA) through sequence complementarity. (Deltcheva, E et al 2011). Through the binding of Cas9 protein can bind to a tracrRNA-crRNA complex or to a designed, chimeric sgRNA to generate a double-strand break (DSB) at a specific site of the target DNA in vitro. (Gasiunase, G et al 2012). Further study of CRISPR technology shows that we could use deactivated Cas9 fused with certain protein to recruit transcription activator to mediate gene activation. (Gilbert, L et al 2013).

Since the CRISPR technology can be used to mediate gene activation, I decide to utilize the CRIPSR activation (CRISPRa) to improve iTECs reprogramming. I designed several guide-RNA targeting potential Foxn1 promoter region to check whether it’s possible to activate endogenous Foxn1. By design several guide RNA, I successfully activate endogenous Foxn1 expression in the MEFs. Different guide RNA has different effect on activating endogenous Foxn1, which the one that is ~100bp upstream to the endogenous Foxn1 promoter 1a has the most efficient effect on activaint Foxn1. By measuring the expression level of endogenous Foxn1, we find that endogenous Foxn1 can be activated to about one-fourth compare to iTECs Foxn1 transgene

expression. Further analysis shows that by combining several guide-RNA, the activation level of endogenous Foxn1 can be further increased.

After successfully activate endogenous Foxn1 to a relatively high level, I was able to further analyze whether it's possible to direct reprogram MEFs into functional iTECs by CRISPRa system. We first check whether activating endogenous Foxn1 can also activate Foxn1 direct downstream targets. After activating endogenous Foxn1 for 48hrs, we do detect the expression of DLL4, CCL25 & Kit-L. Also, the activation of endogenous Foxn1 leads to a proportion of MEFs change morphology and also express low level of EpCam.

Together these results show that it's possible to direct reprogramming MEFs into function iTECs by activating endogenous Foxn1 by CRISPRa system. However, the expression level of endogenous-Foxn1 still needs to be further increased as well as maintained in the cells in order to successfully reprogram MEFs into iTECs. One possible solution is using AAV virus to stable express CRISPRa system gene compartment in the targeting cells. Another potential method would be use more strong activators fused to dCAS9 since now the VPR activator and VP64 can still be further improved.

Potential application of iTECs reprogramming into human cells system

To further apply iTECs into clinical application, we also need to investigate whether it's possible to direct reprogram iTECs in human cells using over expression of Foxn1 too. To test this hypothesis, I tried to activate endogenous Foxn1 in human dermal fibroblast using CRISPRa system. By transfecting the dCAS9 compartment complex as well as guide RNA that targeting

upstream of human Foxn1 promoter region, I was able to activate endogenous Foxn1 in human dermal fibroblast. However, I didn't detect most of the Foxn1 downstream targets after activating Foxn1 and also no cell morphology change is observed. One potential reason is the endogenous Foxn1 expression in human dermal fibroblast is not high enough to drive the direct cell linage switch to happen; another possibility is that reprogram human dermal fibroblast needs additional transcription factor beside thymus transcription factor Foxn1. This is possible because in some other recent study, researchers show that human cells may need some different combination of transcription factor to make the same reprogramming happen, such as the different transcription factor requirements to reprogram MEFs into functional neurons and reprogram human fibroblast into function human neurons. For future experiments, a cocktail-transcription factor candidates screening is highly necessary to investigate the potential combination of transcription factor for human iTECs reprogramming. In order to do so, we need to add different transcription factor that has been identified to be important in thymus development together with the Foxn1 overexpression. This can help us identify potential candidates that can further help human iTECs reprogramming in addition to Foxn1.

Understanding the role of cell cycle arrest in iTECs reprogramming and potential effect of cell cycle arrest rescue on iTECs reprogramming utility

Previously, we always observe that iTECs slows down proliferation, if not completely stopped, after reprogramming. For this reason, we have to seed the iTECs at very high density, which is around 170K to 200K cells per 24 wells to maintain the iTECs' close cell-cell interaction. We have no idea whether this cell cycle arrest is necessary for iTECs reprogramming or it's a unnecessary side-effect that developed by the iTEC reprogramming.

By using RNA-seq analysis, I was able to detail analyze the cell cycle related gene

expression change during the iTECs reprogramming. I identified several cell cycle inhibitors that are activated during reprogramming, as well as decreased Myc and p53 expression. In-vitro rescue of cell cycle experiments shows that by over-expressing cell cycle gene Myc, we can not only successfully make the iTECs reprogramming happens, but also further improve the reprogramming differentiation in which the EpCam+ iTECs population has been increased.

For further analyze the impact of cell cycle arrest in iTECs reprogramming, several additional experiments need to be performed. Firstly, although we successfully rescue the cell cycle by overexpression Myc, the Myc expressing iTECs is still largely undefined. We need to further perform IHC staining for TECs markers as well as detailed transcriptome expression profile analysis for these Myc expression iTECs using RNA-seq. Secondly, we need to further characterize the Myc expression iTECs both in morphology and functionality. I noticed that Myc expressing iTECs are significantly different compared to normal iTECs, especially in size that Myc expressing iTECs are smaller. Detailed morphology analysis including continuous microscope imaging are necessary for characterize these two different kinds of iTECs and understand how the morphology changes through time. Also, these Myc expression iTECs need to be test for their functionality such as support T cell development in intro and in vivo to make sure cell cycle rescue would not affect the iTECs function. Finally, several other cell cycle related genes needs to tested such as the RB pathway to see whether we could repeat the Cell cycle rescue as well as improve iTECs reprogramming efficiency using a completely different cell cycle mechanisms.

Identification the role of Notch pathway in mTEC marker expression in iTECs and potential methods to further improve mTECs differentiation in iTECs

Previously mTECs markers are rarely detected in iTECs reprogramming. In IHC staining, only cTEC markers K8 is stained positive and in qPCR analysis, most mTEC genes are not detected. However, by using RNA-seq analysis which have a higher sensitivity on gene expression profile, low level of mTEC markers, including mTEC progenitor marker Cld4, are detected in late reprogramming stages. Also, we successfully turned on Aire in iTECs by using DAPT inhibition, which indicating Notch pathway may plays a crucial role in mTEC marker expression during iTEC reprogramming. However, to further understanding the mTEC marker expression and iTECs differentiation, a variety of future experiments need to be done.

Firstly, we need to further analyze the DAPT treated iTECs. Although the DAPT inhibition successfully activate AIRE, we still have little understanding about those DAPT inhibited iTECs both in gene expression level and functionality. Thus, we need to further analyze the iTECs after DAPT inhibition to see whether there are any other mTEC differentiation markers changed other than Aire and Cld4. Also, functionality test such as T cell co-culture is also necessary to check whether these mTEC marker expressing iTECs can also support T cell development as normal, or even better. Secondly, whole transcriptome analysis is necessary to check what gene pathway is highly changed after we modify Notch pathway in iTECs and thus, we can possibly further improve the mTEC differentiation in iTECs. Also, we still have no clue that to what expression level and which exact Notch gene is the key that regulating mTECs differentiation in iTECs reprogramming. thus, further potential Notch pathway candidate genes and gradient dosage test is necessary to complete the understanding of the relationship between Notch pathway and mTEC differentiation in iTECs. Finally, we have no idea whether it’s the same cell population that express

both cTECs markers and mTECs markers or it's distinct cell population that express one of them. This is important because it would help us better understand the path of iTEC reprogramming, that whether it can be bi-potent, or it all follows the same pathway of gene expression pattern. To test this hypothesis, single cell RNA-seq should be performed on the later stage of iTEC reprogramming so that we can compare each individual iTECs gene expression based on the cTECs marker and mTECs marker.

Foxn1-EGFP reporter and its potential application based on its correlation to endo-Foxn1

Previously we find that Foxn1-EGFP reporter has good correlation to endogenous Foxn1 at fetal stages as well as early postnatal stages. However, after as early as 4 weeks postnatally, the correlation between the Foxn1-EGFP reporter and the endogenous Foxn1 has been significantly decreased both in relative expression level and co-localization. We also find that certain cell population that is GFP positive but not stained positive for epithelial marker Epcam/MHCii, especially for those GFP low cell population.

These results indicating that we could use the Foxn1-EGFP reporter mice to track Foxn1 expression at early postnatal stages and fetal stages. We can also sort our thymic epithelial cells at early postnatal stages and fetal stages based on gfp signal. But at any stages that later than 4 weeks postnatally, only GFP high signal cell population can be used marker for thymic epithelial cells.

Furthermore, in addition to the usage of Foxn1-EGFP reporter in thymus development study and sorting thymic epithelial cells, several other potential applications is possible. One of them is using Foxn1-EGFP reporter to purify endo-foxn1 expression MEFs that activated by CRISPRa system which we previously mentioned. This is because the Foxn1-EGFP reporter shares a same promoter upstream sequence as the endogenous Foxn1 gene. And the Foxn1-EGFP reporter

signal is strong so that we can directly sort these cells out by flow cytometry. Our preliminary data shows that the same design of gRNAs can also activating Foxn1-EGFP signal when activating the endogenous Foxn1. The signal is maintained at a high level and can be sort out by flow cytometry. Also, the GFP signal can be used to visualize the cell morphology which would be super helpful to track the iTECs reprogramming process in real time.

Supplementary Appendix list

1) Lenti-virus based cirspr-activation system in MEFs and potential application on clinical