

DISS. ETH NO 26431

**Unravelling molecular drivers of human embryonic stem cell
pluripotency and early cell fate specification**

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by
Panagiota Tsikrika
M.Sc., Karolinska Institute
born on 26.05.1987
citizen of Greece

accepted on the recommendation of

Prof. Anton Wutz
Prof. Ulrike Kutay
Prof. Laurent David
Dr. Matthias Gstaiger
Dr. Tobias Beyer

2019

1 Summary

Human pluripotent stem cells (hPSCs) have the ability to self-renew indefinitely in culture, as well as differentiate towards somatic cells and extraembryonic tissues in appropriate culture conditions. These properties make stem cells (SCs) central to emerging concepts in modern medicine, by providing an invaluable source of cells for basic research, disease modelling, drug testing and ultimately cell replacement therapies. The core transcription factors (TF) OCT4, SOX2 and NANOG maintain the hPSC state by occupying specific regulatory regions to control the expression of both pluripotency and pro-differentiation genes. This delicate balance is maintained by the extracellular signaling cues of the Fibroblast Growth Factor (FGF) and TGF- β families. Their spatiotemporal pattern of activity controls the re-localization of downstream TF complexes to specific regulatory elements and enhancers. Such enhancers are called “poised” as these are bookmarked in pluripotency with a unique chromatin signature. To gain deeper mechanistic insights in binding events on those enhancers during cell fate decisions, we took advantage of mass spectrometry (MS) - coupled to DNA pull down. For this, enhancer regions of genes important for pluripotency or Primitive streak (PS) and Definitive Endoderm (DE) formation were labelled by PCR amplification using biotinylated oligonucleotides and the resulting biotinylated fragments were incubated with cell lysates from pluripotent human embryonic stem cells (hESCs) or differentiated cells, followed by DNA pull down. The differentially bound proteins between the two states were analysed by MS and classified according to their enrichment over a control sequence in the different conditions. In this manner, we identified Transcriptional Adaptor 2-beta (TADA2B), as a novel protein binding to the enhancer of *NANOG*. We uncovered an essential role of TADA2B in maintenance of pluripotency. hESCs with deletion of *TADA2B* exhibit reduced expression of pluripotency factors *OCT4* and *SOX2*, upregulation of Mesendoderm genes *GATA6*, *GSC*, *T*, *EOMES* and phenotypically resemble endoderm-like cells. TADA2B is a member of the STAGA complex which is involved in several processes and especially in development and cancer. Therefore, the dissection of the mechanism of action of this complex and identification of its partners could shed light on cancer research and embryonic development. In conclusion, in this study, we captured binding events on active and poised enhancers in hESCs and discovered a protein that has an essential

Summary

role in self-renewal of hESCs and possibly in embryo development.

2 Zusammenfassung

Humane pluripotente Stammzellen definieren sich durch ihre Fähigkeit, sich in Kultur auf unbestimmte Zeit selbst zu erneuern und sich bei geeigneten Kulturbedingungen in somatische Zellen und in extraembryonales Gewebe zu differenzieren. Diese Eigenschaften machen Stammzellen zum Dreh- und Angelpunkt neuer Ansätze in der modernen Medizin. Dies basiert auf der Tatsache, dass sie eine unversiegbare Quelle für Zellen für Grundlagenforschung, Krankheitsmodellierung, Arzneimitteltests und letztendlich für Zellersatztherapien bieten. Auf der molekularen Ebene, erhält ein Komplex, bestehend aus den Grundtranskriptionsfaktoren OCT4, SOX2 und NANOG den hPSC-Zustand aufrecht, indem er spezifische regulatorische Regionen bindet um die Expression von Pluripotenz- und Prodifferenzierungsgenen zu regulieren. Dieses empfindliche Gleichgewicht wird durch die extrazellulären Signalwege der Fibroblast Growth Factor (FGF) - und TGF- β -Familien aufrechterhalten. Ein komplexes räumliches und zeitliches Aktivitätsmuster dieser Wachstumsfaktoren steuert die Bindung von TF-Komplexen an spezifischen regulatorischen Elementen, welche als Enhancer bezeichnet werden. Diese Enhancer werden als "poised" bezeichnet, da diese im pluripotenten Zustand eine sowohl aktive wie auch repressive Chromatin-Zusammensetzung aufweisen. Um tiefere mechanistische Erkenntnisse über Bindungsereignisse an diesen Enhancern während der Differenzierung von pluripotenten SCs zu gewinnen, setzten wir Massenspektrometrie gekoppelt an DNS-Pulldown ein. Hierzu wurden Enhancerregionen von Genen, die für die Pluripotenz oder Primitive streak (PS) - und Definitive Endoderm (DE) -Bildung wichtig sind, mittels PCR-Amplifikation mit biotinylierten Oligonukleotiden markiert und die resultierenden biotinylierten Fragmente mit Zelllysaten aus hPSCs oder differenzierten Zellen inkubiert. Anschliessend wurden assoziierte Proteine präzipitiert und mittels MS differentiell gebundene Faktoren identifiziert. Die unterschiedlich gebundenen Proteine zwischen den beiden Konditionen wurden mit MS analysiert und entsprechend nach der Anreicherung gegenüber eine Kontrollesequenz klassifiziert. Auf diese Weise identifizierten wir TADA2B als neues Protein, das an den Enhancer von NANOG bindet. Eine detaillierte Analyse der Funktion von TADA2B zeigte, dass TADA2B eine wesentliche Rolle bei der Aufrechterhaltung der Pluripotenz spielt. hPSCs mit einer Mutation im TADA2B Gens, zeigen eine verringerte Expression

Zusammenfassung

der Pluripotenzfaktoren OCT4 und SOX2, eine Induktion der Mesendoderm-Gene GATA6, GSC, T, EOMES und ähneln phänotypisch endodermartigen Zellen. TADA2B ist Mitglied des STAGA-Komplexes, der an mehreren Prozessen beteiligt ist, insbesondere an der Entwicklung und an Krebs. Die Aufschlüsselung des molekularen Wirkmechanismus dieses Komplexes und die Identifizierung seiner Interaktionspartner könnten daher Aufschluss nicht nur über die Steuerung der humanen embryonalen Entwicklung, sondern auch in die Krebsforschung geben. Zusammenfassend haben wir in dieser Studie die Bindung von TF an einzelnen DNS loci auf aktiven und „poised“ Enhancern in hPSCs systematisch erfasst und ein Protein entdeckt, das eine wesentliche Rolle bei der Selbsterneuerung von hESCs und möglicherweise bei der Embryonalentwicklung spielt.