Functional Characterization of LEF1 Isoforms in Mouse Embryonic Stem Cells

by

Mohammad Shokouhian

A Thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
in partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

Department of Biochemistry and Medical Genetics
University of Manitoba
Winnipeg

ABSTRACT

The expression of Wnt/β-catenin signaling target genes is regulated by nuclear translocation of β-catenin, mainly in complex with T-cell factor (TCF)/lymphoid enhancer binding factor (LEF) factors. LEF1, a member of the TCF/LEF family of transcription factors, holds significant importance in both normal developmental processes and various pathological conditions such as cancer. Our objective was to characterize the function of two LEF1 isoforms, namely full-length LEF1 and ΔN-LEF1, via detecting the proteins they interact with by using TurboID-based proximity labeling. Additionally, we aimed to map their binding profiles throughout the genome by using the CUT&RUN technique. This investigation was conducted in the context of differentiating mouse embryonic stem cells under two conditions: the presence of a GSK-3 inhibitor known as CHIR99021 (CHIR), which serves to mimic Wnt pathway activation, and the absence of Wnt pathway activation. In our TurboID data, we detected both well-established LEF1 interactors such as Wnt enhanceosome components and new isoform-specific interactors, such as NuRD complex proteins and the transcription factor, ZSCAN10, in the presence and absence of CHIR, respectively. Notably, the enrichment of RNA-binding proteins, chromatin modifiers, and numerous novel transient interactors such as ubiquitin ligases, deubiquitinases, kinases and phosphatases with FL-LEF1 in the presence of CHIR underscores the impact of Wnt/β-catenin signaling on modulating LEF1 function in an isoform-specific manner. The CUT&RUN data discovered the binding of LEF1 isoforms to the *Lef1* gene and intriguingly revealed an unexpected interaction with *Zscan10*. Our findings reveal new insights into isoform-specific interactions and the underlying mechanisms through which LEF1 isoforms function in regulating gene expression and warrant further studies characterizing the function of other TCF/LEF isoforms in a similar fashion.

ACKNOWLEDGEMENTS

I am indebted to my supervisor, Dr. Doble, for his unwavering support and guidance throughout my research journey, especially during the pandemic led challenges I had with my visa application and transition to Canada. Despite my numerous mistakes, he has always responded with patience and positivity. Dr. Doble's open-door policy and welcoming smile have made it easy for me to approach him with any personal or work-related concerns. Working under his supervision has been a privilege, as his creativity and research strategy made the learning process and research enjoyable. I consider myself incredibly fortunate to have had the opportunity to master cutting-edge research techniques under his mentorship.

I would also like to extend my sincere appreciation to Professor Tamra Werbowetski-Ogilvie and Dr. Susan Logue, who served on my advisory committee. Thank you both for your support. Your encouragement has been vital in my growth during these two years. Dr. Ogilvie, your invaluable lessons in time-management and discipline have been instrumental in shaping my approach to scientific work. Thank you, Dr. Logue, for unwavering positivity and challenging me to strive for excellence even when faced with limited results or imperfect presentations. I am deeply grateful to the Ogilvie lab members Ludivine Coudière-Morrison and Dr. Jamie Zagozewski for their assistance in providing necessary reagents and material whenever we needed them. Furthermore, I would like to thank Leo McKay for his mentorship over the past two years. I am obliged to acknowledge and thank Dr. Jim Davie's lab for generously allowing us to use their Bioanalyzer and Dr. Alexandre Blais from the University of Ottawa for his expertise in analyzing CUT&RUN data.

This research would have not been possible without the generous funding from the National Sciences and Engineering Research Council of Canada (NSERC).

Lastly, I want to acknowledge the support and encouragement of my family and friends here and back home in Iran. Their belief in my abilities and their constant support have always been the main source of motivation for me to move forward and pursue my education.

DEDICATION

To my family.

My father passed away 12 years ago, but he has always been my role model for kindness, generosity, perseverance, commitment, and contribution. My beloved mom and stepmom who raised me. My dear wife, for being supportive, always being there either I fail or succeed. You are beautiful, smart, talented, and kind. My only sister, Fatemeh; my brothers, Sajad, Javad, and Alireza. My dear nieces and nephews: Hossein, Zahra, Maryam, and Diana. I miss you!

Love you all!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

~	Approximately
°C	Degrees Celsius
<	Less than
≥	Greater than or equal to
>	Greater than
µg	Microgram(s) (weight)
µL	Microliter(s) (volume)
µM	Micromolar (concentration)
%	Percent
Å	Angstrom
aa	Amino Acid(s)
AMP	Adenosine Monophosphate
ATP	Adenosine triphosphate
AVE	Anterior Visceral Endoderm
BAF	BRG1-Asscociated Factor
BFDR	Bayesian False Discovery Rate
bp	Base Pair(s)
BSA	Bovine Serum Albumin
Cas9	CRISPR Associated Protein 9
CHIR	CHIR99021
CK1	Casein Kinase 1
CLL	Chronic Lymphocytic Leukemia
cm	Centimeter
CO2	Carbon Dioxide
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CUT&RUN	Cleavage Under Target and Release Using Nuclease
DC	Detergent Compatible
DE	Definitive Endoderm
DMEM	Dulbecco's Modified Eagle Medium
DVL	Dishevelled
E3	Ubiquitin Ligase
EB	Embryoid Body
EDTA	Ethylenediaminetetraacetic Acid
EMT	Epithelial Mesenchymal Transition
ESC	Embryonic Stem Cell
FBS	Fetal Bovine Serum
FL	Full Length

FZD	Frizzled
GAPDH	Glyceraldehyde-3-phosphate Dehydrogenase
GATA	GATA Binding Protein
GSK-3	Glycogen Synthase Kinase-3
h	Hour(s)
H3	Histon 3
HDAC	Histon Deacetylase
HMG	High Mobility Group
HRP	Horseradish Peroxidase
ICM	Inner Cell Mass
Indels	Insertions or Deletions
JAK	Janus Kinase
kb	Kilobase(s)
KCl	Potassium Chloride
kDa	Kilodalton(s)
KLF4	Kruppel-Like Factor 4
KO	Knockout
LB	Lysogeny Broth
LDS	lithium Dodecyl Sulfate
LEF1	Lymphocyte enhancer-binding factor
LF3K	Lipofectamine 3000
LIF	Leukemia Inhibitory Factor
LRP5/6	Low Density Lipoprotein Receptor-Related Protein 5/6
MBD	Methyl-CpG Binding Domain
mEB	Mouse embryoid body medium
mESC	Mouse embryonic stem cell
min	Minute(s)
mL	Milli Liter
mm	Millimeter
mM	Millimolar
MTA2	Metastasis-Associated Protein 2
ng	nanogram (weight)
NGS	Next Generation Sequencing
NLS	Nuclear Localization Signal
nm	Nanometer(s) (length)
nM	Nanomolar (concentration)
Nt	Nucleotide(s)
NuRD	Nucleosome Remodeling and Deacetylase
O/N	Overnight
O2	Oxygen

OCT	Octamer-Binding Transcription Factor 4
P	Passage
P-value	Probability Value
PAM	Protospacer Adjacent Motif
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PL	Proximity Labeling
PPI	Protein-Protein Interaction
PRC2	Polycomb Repressive Complex 2
PSC	Pluripotent Stem Cell
PTC	Premature Termination Codon
QC	Quality Control
RIPA	Radio-Immunoprecipitation Assay
RPM	Revolution per Minute
RT	Room Temperature
s	Second(s)
SAINT	Significance Analysis of Interactome
SDS	Sodium Dodecyl Sulfate
SEACR	Sparse Enrichment Analysis for CUT&RUN
sgRNA	Synthetic Guide RNA
SHH	Sonic Hedgehog
STAT	Signal Transducer and Activator of Transcription
SWI/SNF	Switch/Sucrose Non-Fermentable
TAE	Tris-acetate Ethylenediaminetetraacetic Acid
TBS	Tris-Buffered Saline
TBST	Tris-Buffered Saline wit 0.1% Tween 20 detergent
TCEP	Tris(2-Carboxyethyl) Phosphate
TCF	T-Cell Factor
TLE	Transducin-like enhancer of Split
UTR	Untranslated Region
V	Volts
vs.	Versus
w/v	Weight per volume
WB	Western Blot
WRE	Wnt Response Element
YY1	Yin Yang 2
×g	Acceleration

CHAPTER 1. INTRODUCTION

1.1. THE WNT/β-CATENIN SIGNALING PATHWAY

Wnt signaling is conserved across evolution. Most mammalian genomes, including those of mice and humans, express at least 18 distinct Wnt genes, while the cnidarian Hydra has only one Wnt gene and Drosophila has seven.[1] Wnt signaling plays an instrumental role in the origin of multicellular organisms. It regulates several processes essential for development, including tissue patterning, cell fate determination, cell polarity, and cellular migration.[2,3] Wnt proteins are secreted hydrophobic cysteine-rich glycoproteins that act as autocrine and paracrine signaling proteins. Wnts can bind a combination of transmembrane receptors and coreceptors and elicit different signaling pathways in a tissue-specific, cell-type-dependent, and cell-state-dependent manner. The Wnt signaling network is activated when Wnt ligands bind the cysteine-rich domain (CRD) of Frizzled (Fzd). The β-catenin-independent/planar cell polarity (PCP) pathway (non-canonical) and the β-catenin-dependent pathway (canonical) are the two Wnt signalling pathways that are best understood. However, it is known that Wnt ligands can also activate several other downstream responses.[4]

In the absence of Wnt ligands, β-catenin is phosphorylated sequentially by casein kinase 1 (CK1) at serine 45 and glycogen synthase kinase-3 (GSK-3) at threonine 41, serine 37, and serine 33 in a multi-protein destruction complex. The destruction complex, or Axin degradasome, comprises the scaffold protein Axin 1/2, Adenomatous Polyposis Coli (APC), CK1, and GSK-3. Phosphorylated β-catenin is then ubiquitinated by the E3 ubiquitin ligase, β-transducing repeat-containing protein (β-TrCP), targeting β-catenin for subsequent proteasomal degradation.[5,6] In the absence of available β-catenin, TCF/LEFs serve to repress gene transcription primarily through

interaction with Groucho/Transducin-like enhancer of Split (Gro/TLE) repressor family members (***Figure 1-1A***).[7]

In the canonical signaling pathway, Wnt proteins (e.g., Wnt1, Wnt3a) bind to a heterodimeric receptor complex, also known as the Wnt signalosome, consisting of a seven-pass transmembrane receptor of the Frizzled family (FZD_{1-10}) and one of the low-density lipoprotein receptor-related proteins 5/6 (LRP5/6) single-pass transmembrane receptors.[6,8] Wnt binding to its receptors and coreceptors induces a conformational change of the Fz receptors (Fz or Fzd) followed by the recruitment of cytoplasmic Dishevelled (DVL) via interaction with the cytoplasmic part of the Fz receptor. This then triggers and facilitates interaction between LRP5/6 and Axin, thereby promoting phosphorylation of the LRP5/6 cytoplasmic tail by two kinases, GSK-3 and CK1.[9-11] Enhanced Axin recruitment to the receptors following LRP6 phosphorylation inhibits the Axin-mediated β-catenin phosphorylation and thereby leads to β-catenin stabilization.[11,12] β-catenin cytoplasmic accumulation leads to its translocation to the nucleus, where it interacts with the leading nuclear effectors of the Wnt/β-catenin signaling cascade, TCF/LEF family members, and regulates the expression of Wnt target genes (***Figure 1-1B***).

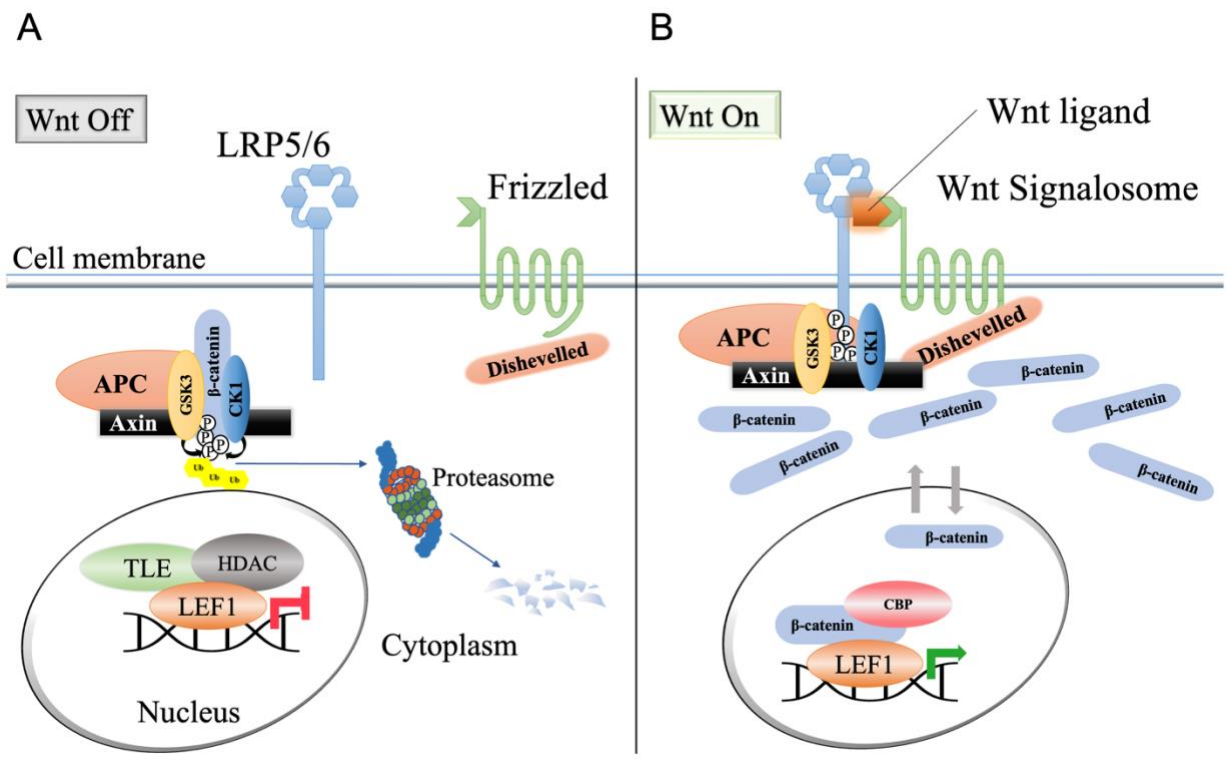

Figure 1-1. Overview of the Wnt/β-Catenin Signaling Pathway.
(A) In the absence of Wnt ligand (Wnt//β-catenin "Off"), β-catenin is retained in a destruction complex with Axin, GSK-3, and CK1, which results in β-catenin phosphorylation and subsequent ubiquitination, triggering its proteosomal degradation. In the nucleus, TCF/LEFs repress Wnt/β-catenin target genes primarily through interaction with TLE co-repressors. ***(B)*** In the presence of Wnt (Wnt//β-catenin "On"), Wnt receptor complex conformational change leads to Axin recruitment together with the other components of the destruction complex. β-catenin is devoid of phosphorylation and thus is stabilized. β-catenin cytoplasmic accumulation leads to its translocation to the nucleus, where it interacts with Wnt/β-catenin nuclear effectors (primarily TCF/LEFs) and recruits co-activators.

1.2. T-CELL FACTORS AND LYMPHOID ENHANCER-BINDING FACTOR

While almost all invertebrates possess only one *TCF/LEF* gene, humans and other vertebrates have expanded the *TCF/LEF* family to encompass four members: *Tcf7* (previously known as *Tcf1*), *Tcf7l1* (previously known as *Tcf3*), *Tcf7l2* (previously known as *Tcf4*), and *Lef1*, which combined with isoform complexity, result in functional diversification and specification in the Wnt signaling

pathway.[13,14] Other mechanisms affecting mRNA production, transport, and stability, plus post-translational modifications (PTMs), add more complexity to their functions.

TCF/LEFs are multidomain proteins, and their full-length isoforms display common structural features in the following order (see **Figure 1-2**): a β-catenin-binding domain at the amino terminus; a context-dependent regulatory domain (CRD) that participates in both transcriptional repression by recruiting TLE (in mammals; in *Drosophila*, it is known as Groucho) and transcriptional activation; a highly conserved monomeric high mobility group box DNA-binding domain (HMG-DBD) near the carboxyl terminus that recognizes the consensus sequence (5′-CTTTGAC(T or A)-3); and alternatively spliced C-terminal tails containing a stretch of basic residues proximal to the HMG domain.[7,15] Each TCF/LEF transcription factor has various isoforms generated either via alternative splicing and/or alternative promoters. Only *Tcf7* and *Lef1* have alternative promoters. Two other differences are the presence of a C-clamp in the C-termini of TCF7 and TCF7L2 and

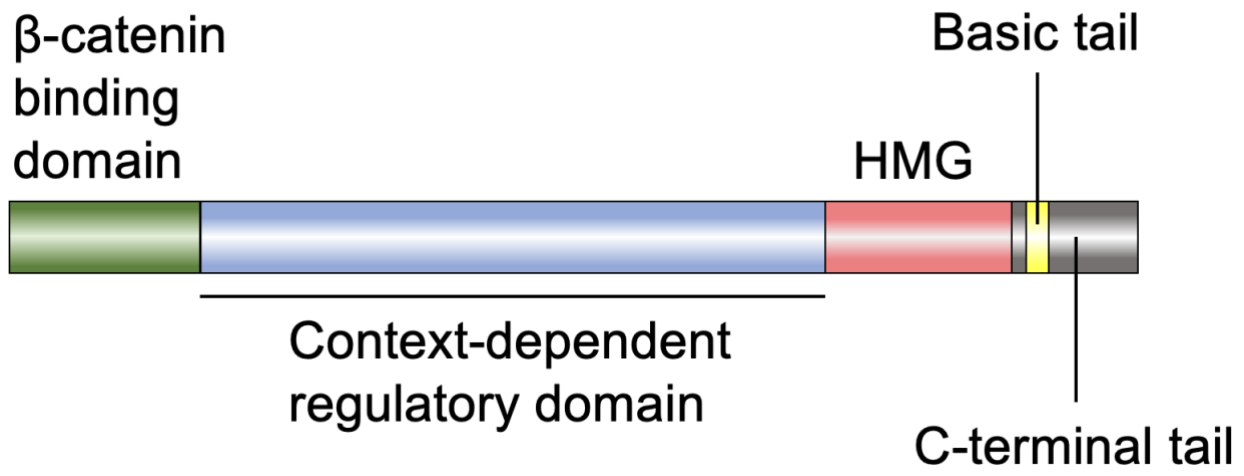

Figure 1-2. The Conserved Domains of T-Cell Factor/Lymphoid Enhancer-binding Factor (TCF) Family of Transcription Factors.
Schematic of common structural features of a generic TCF, showing the β-catenin binding domain (Green), the context-dependent regulatory domain (CRD) (Blue), the high mobility group (HMG) domain (Red), the basic tail (Yellow), and the C-terminal tail (Black).

its absence in the LEF1 and TCF7L1, and the presence of a C-terminal binding protein (CtBP) domain in TCF7L1 and TCF7L2 (***Figure 1-3***).

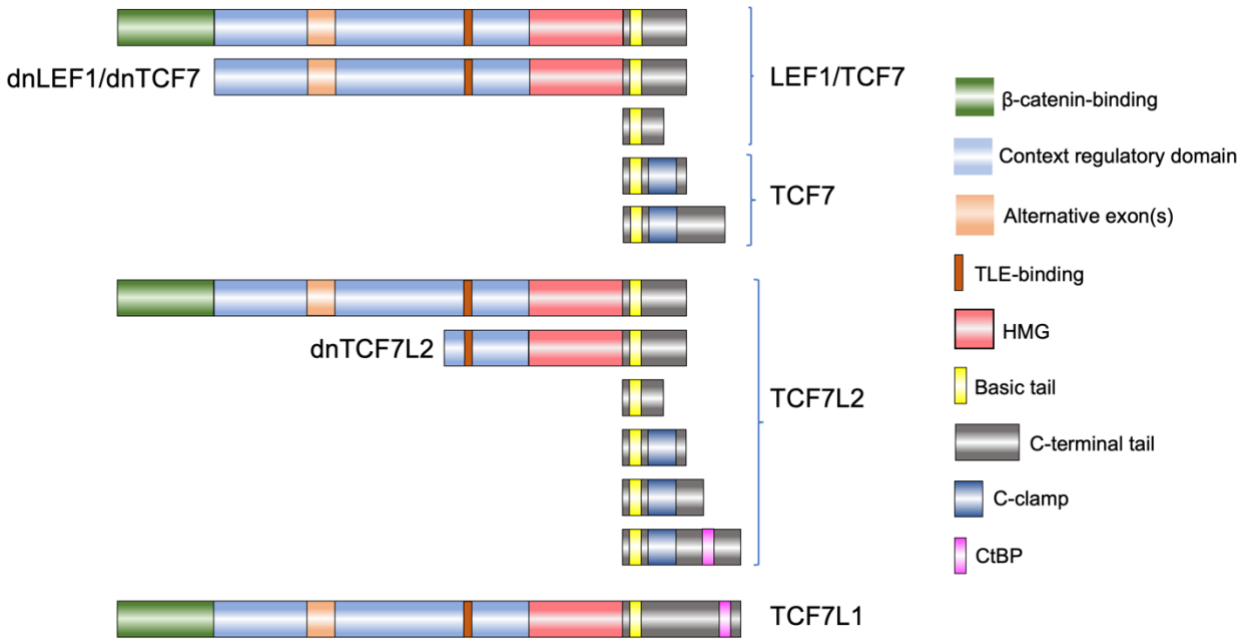

Figure 1-3. Schematic Representation of a Variety of TCF-LEF Isoforms Generated via Alternative Promoters and/or Alternative Splicing.
Short forms of TCF7 and LEF1 are generated via alternative promoters and lack the N-terminal domain (β-catenin binding domain (Green)). In addition, a TCF7L2 dominant negative isoform lacks the β-catenin binding domain. The context -dependent regulatory domain (CRD) (Light Blue) has the least conservation compared to other regions of TCF/LEF proteins and contains alternatively spliced exons (Orange) and a TLE binding domain (Brown). The highly conserved domain, high mobility group (HMG) domain (Red), and a basic tail that serves as the nuclear localization signal (NLS) (Yellow) are present in all TCF/LEF proteins. An alternatively spliced C-terminal tail is a highly variable domain, with some TCF7 and TCF7L2 isoforms containing an additional DNA-binding domain called the cysteine-clamp or C-clamp (Dark Blue). A motif found only in the C-terminus of TCF7L2 (some isoforms) and TCF7L1 is C-terminal binding protein (CtBP) which interacts with a co-repressor protein called CtBP.

Although TCF/LEF members intrinsically lack transcriptional activity, they play a crucial role in gene regulation by virtue of their ability to bind and flex DNA, serving as platforms for the recruitment of β-catenin and coregulators such as chromatin remodeling complexes.[16]

Furthermore, they can interact with transcriptional corepressors, such as Gro/TLE, through their CRD domain.[17]

In TCF/LEF quadruple knock-out/rescue experiments it has been shown that TCF7L1 can mediate both transcriptional activation and repression, although it typically acts as a repressor.[18] The activity of TCF7L1 as a transcriptional repressor is counteracted by its association with β-catenin, which displaces TCF7L1 from chromatin and promotes its destruction.[18,19] TCF7 and LEF1 behave mainly as β-catenin-dependent transcriptional activators. TCF7L2, as the most versatile member of the family, appears to suppress or activate gene expression in a context-dependent manner.[13,16,20]

TCF/LEF factors have different levels in various settings, and their expression is controlled at multiple levels, from transcription and translation to post-translationally regulated protein stability. Post-translational modifications (PTMs) of TCF/LEF proteins, including phosphorylation, ubiquitination, acetylation, and sumoylation, not only affect their half-life, and thus protein abundance, but also regulate their DNA-binding and transcriptional activity and their interactions with transcriptional co-factors such as β-catenin.[21-24] TCF/LEFs can autoregulate their own expression and form positive or negative feedback by regulating other Tcf/Lef genes and/or their target genes.[25-27]

1.3. WNT/β-CATENIN ROLE IN PLURIPOTENCY AND GASTRULATION

Wnts are defined as morphogens, as they have a crucial role in pattern formation and symmetry breakage and instructing cell fate during development.[10] In the embryo, Wnts promote stemness and, in collaboration with other morphogens in a BMP, WNT, and NODAL signaling network,

initiate the patterning of pluripotent epiblast cells into the three germ layers: endoderm, mesoderm, and ectoderm.[28,29]

Pluripotent stem cells (PSCs) possess the capacity to differentiate into all three germ layers. PSCs have various sources: gonadal primordial germ cells, pre-implantation gastrula (embryonic stem cells or ESCs), and post-implantation epiblast cells (EpiSCs).[30] When mouse ESCs (mESCs) are introduced into blastocysts, they can integrate into the inner cell mass and develop into various cell types within the embryo.[31] These mESCs most closely resemble epiblast cells at embryonic day 4.5 (E4.5) and exhibit characteristics of naïve pluripotency, which differs from 'primed' pluripotency in which cells acquire a predisposition towards specific germ layers.[32] At E4.75-5.0, following implantation into the maternal uterus and stimulation by Wnt/β-catenin and two other factors (Egf2/Extracellular Signal-Regulated Kinase and Activin), amorphous epiblast cells undergo a transformation into polarized epithelial cells.[33] Then at ~E5.5 EpiSCs generate the anterior visceral endoderm (AVE). At ~E6.0, mouse EpiSCs, corresponding to ESCs in a state of primed pluripotency, exit pluripotency, and differentiate to form the primitive streak, a temporary structure located in the posterior region, which serves as the hallmark of gastrulation occurring at E6.5-7.5.[34]

There is reciprocal communication between the embryonic and extra-embryonic tissues, where two proteins called Furin and Pace are secreted from the extra-embryonic ectoderm.[35] These proteins play a crucial role in promoting the maturation and release of Nodal protein from the epiblast.[36] Upon release, Nodal stimulates the expression of BMP4 in the extra-embryonic ectoderm, which subsequently activates Wnt3 in the epiblast, further amplifying Nodal signaling.[37] The coordinated signaling network leads to the formation of AVE cells. These AVE cells secrete

antagonists of Wnt and Nodal, which drive the specification of the primitive streak and initiate gastrulation. At this early gastrulation stage, the embryo consists of only a few thousand cells.[38]

Gastrulation is a pivotal developmental process during which PSCs undergo an irreversible loss of pluripotency and acquire specific cell fates along with spatial coordination; in other words, pluripotent cells of the epiblast spatially organize into the three embryonic (or definitive) germ layers; definitive endoderm, mesoderm and (neuro-) ectoderm.[33,39,40] In the region of the posterior epiblast, where the primitive streak emerges, PSCs undergo a significant event called epithelial-to-mesenchymal (EMT) transition. This transition is considered a critical element of the gastrulation process.[41,42]

EMT is evolutionary conserved and is essential for gastrulation and embryonic morphogenesis. It is activated in posterior epiblast cells upon their induction by high levels of Wnt/β-catenin, transforming growth factor-β, and fibroblast growth factor signaling.[43] Induced epiblast cells lose their epithelial morphology and adopt a mesenchymal fate. These cells then ingress and migrate into the primitive streak. EMT also has crucial roles in tissue remodeling, organ development, and wound healing.[44,45] Abnormalities in EMT regulation can lead to many malignant tumors and cancer metastasis.[46,47] Therefore, understanding gastrulation is very helpful in studying the molecular basis of EMT and provides a framework for understanding mechanisms of cancer metastasis, stem cell differentiation, and congenital diseases.[48,49]

1.3.1. mESC Cell Culture

Our understanding of human post-implantation development events (such as gastrulation and organogenesis) and their cellular and molecular mechanisms are limited, as when these cells are implanted in the uterus, they are no longer accessible for visualization and examination.[50] Ethical concerns arise regarding the established legal restriction of a 14-day limit for the *in vitro*

cultivation of fertilized human eggs. It has been suggested that the break of the radial symmetry of the epiblast following the appearance of the primitive streak symbolizes the emergence of human individuality.[40] Therefore, researchers have successfully developed simplified *in vitro* embryo models. These models enable the study of post-implantation development *in vitro*.[41,51]

A common method of sustaining the self-renewal of mouse embryonic stem cells *in vitro*, to model the mouse pre-implantation embryo, involves culturing them in a conventional fetal bovine serum/leukemia inhibitory factor (LIF) medium. This medium is also known as mES medium or SL medium (for Serum + LIF). In this condition, Wnt signaling promotes self-renewal, whereas when LIF is removed, mESCs will convert into a primed state of pluripotency in which they are no longer mESCs, but rather mEpiSCs, where Wnt signaling promotes their differentiation in a process corresponding to post-implantation development and early gastrulation (***Figure 1-4A***). In

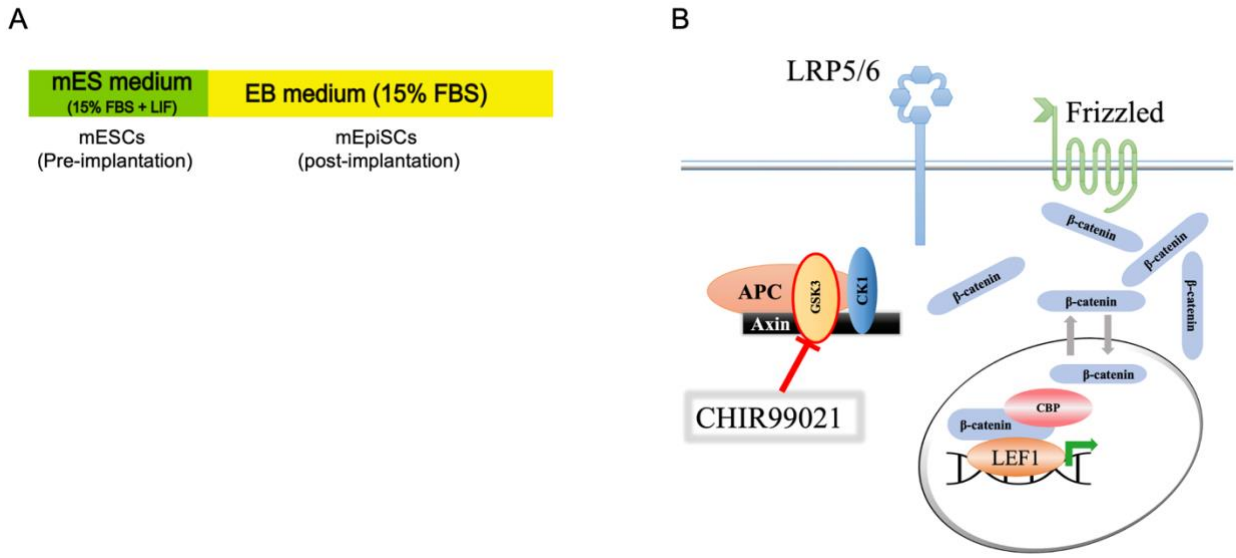

Figure 1-4. An Overview of the *in vitro* Model for Maintaining and Differentiating mESCs. ***(A)*** mESCs cultured in mouse ES medium (mES), which contains 15% Fetal Bovine Serum (FBS) and LIF, model pre-implantation epiblast cells. mESCs transition out of pluripotency and will convert into mEpiSCs after LIF withdrawal, modeling post-implantation and early gastrulation stages of mouse development. ***(B)*** A Wnt-activated condition is mimicked by using GSK-3 inhibitors such as CHIR99021 (CHIR).

these *in vitro* models, Wnt/β-catenin signaling pathway activation is mimicked by using GSK-3 inhibitors such as CHIR99021 (hereafter referred to as CHIR) (***Figure 1-4B***).[52,53]

It should be noted that GSK-3 not only has Wnt-associated activities such as phosphorylation of β-catenin, Axin, APC, and TCF,[54] it also possesses non-Wnt related activities including insulin signaling, NFAT phosphorylation and hedgehog signaling.[55-57] Nonetheless, CHIR has been widely adopted as a Wnt-β-catenin-inducing agent that is routinely used in differentiation protocols for pluripotent stem cells from many species.[58]

GSK-3 inhibition by CHIR99021 promotes mouse ES cell self-renewal and propagation of ES cells cultured in mES medium by regulating multiple pathways. GSK-3 has roles in phosphoinositide 3-kinase, ERK and Wnt/β-catenin pathways, all of which are key regulators of pluripotent stem cell self-renewal and differentiation.[59,60] By contrast, adding GSK-3 inhibitor to mouse embryoid body medium (LIF removed) induces mEpiSC differentiation.

Wild type mESCs will undergo differentiation in the absence of LIF; however, they are unable to differentiate properly when *Gsk-3* (*Gsk-3*α/β double-knockout) is genetically deleted.[61] Recombinant LIF has been used for over 20 years in cell culture media as the substitute for inactivated embryonic fibroblasts for maintaining propagation of mouse ESCs *in vitro*.[62] LIF is a 20 kDa glycoprotein cytokine and a member of the interleukin-6 family of cytokines with highly pleiotropic biological activity. LIF signals through a heterodimer receptor, consisting of cytokine receptor gp130 and LIF receptor β (LIFRβ), and activates the Janus kinase-signal transducer and activator of transcription 3 (JAK-STAT3), the Phosphoinositide 3-kinase, and Mitogen-activated protein kinase pathways, with JAK/STAT being considered the dominant pathway.[63,64]

The maintenance of mESC self-renewal by LIF involves increased expression of *Myc* via STAT3. MYC also functions to suppress expression of differentiation programs, thereby

preventing mESC differentiation into mesoderm and endoderm lineages.[65] It has been demonstrated that LIF withdrawal induces differentiation in mESCs by promoting cell-cycle inhibitory mechanisms and upregulating the expression of D cyclins and cyclin-dependent kinase inhibitors *p21* and *p27*.[66-68]

1.3.2. TCF/LEF Factors Regulate mESC Self-Renewal and Differentiation

The canonical Wnt signalling pathway plays a key role in maintaining the equilibrium between self-renewal, pluripotency, and differentiation of mESCs. β-Catenin, a key effector of this pathway, exerts its nuclear function by interacting with TCF/LEF family members.[69] The expression pattern of TCF/LEF proteins (TCF7, TCF7L1, TCF7L2, LEF1) varies in undifferentiated (self-renewing) and differentiating mESCs, suggesting unique functions for the different TCF/LEF proteins where a balance between their expression is important for the regulation of self-renewal and differentiation in mES cells.[70] *Tcf7l1* is the predominant TCF/LEF family member expressed in mESCs. Somewhat counter-intuitively, it is known to act as a pro-differentiation factor that limits the self-renewal of mESCs by transcriptionally repressing pluripotency genes such as *Nanog*, *Essrb*, *Tcfp2l1* and *klf2*.[71,72] Recently, it has been reported that TCF7L1 can also repress *Lef1*, *Otx2*, and *Dnmt3b*, driving mESCs toward extraembryonic primitive endoderm.[73] Furthermore, it has been shown that TCF7L1 is recruited to the promoters of Wnt targets (*Axin2* and *Lef1*) and pluripotency transcription factors (*Oct4*, *Tbx3*, *Nanog*).[74] *Tcf7l1* deletion (*Tcf7l1*$^{-/-}$) displays the most severe embryonic defects of all *Tcf/Lef* single-gene knockout mice and is lethal, with delayed mesoderm specification, duplicated axial mesodermal structures, and impaired lateral mesoderm formation.[75]

Tcf7 is the second most abundantly expressed *Tcf/Lef* factors in mESCs after *Tcf7l1*. TCF7 is recognized for its role in promoting mESC differentiation, but a study has also revealed that ectopic

expression of *Tcf7* supports mESC self-renewal and inhibits differentiation of mESCs in the absence of LIF. This effect is achieved by enhancing Nanog expression.[70,76] In addition, it has been demonstrated that TCF7 could negatively affect mESC proliferation via regulating key cell cycle repressor genes such as $p15^{Ink4b}$, $p16^{Inka}$ and $p19.^{Arf}$.[74,77] $Tcf7^{-/-}$ mice are viable, but they show severe defects in thymocyte differentiation and reduced numbers of these cells.[78,79]

The expression of *Tcf7l2* is low in undifferentiated mESCs, but its level gradually increases during mESC differentiation (removing LIF), suggesting that it functions in late differentiation. It has been shown that it supresses TCF reporter activity.[70,80] Mice lacking *Tcf7l2* ($Tcf7l2^{-/-}$) display postnatal lethality and show an absence of crypt stem cells of the small intestine.[81]

Lef1 exhibits limited expression in undifferentiated mESCs, and its levels transiently increase upon LIF removal (mEB medium). *Lef1* expression reaches its highest level 24-48 hours after LIF withdrawal and subsequently declines in the following days. This pattern suggests that LEF1 may be associated with the initial stages of trilineage differentiation. [70] In response to Wnt signalling, LEF1 helps mESCs maintain their pluripotency through modulation of *Oct4* and *Nanog*.[82] Mice lacking *Lef1* are undersized, lack hair and teeth, and die shortly after birth.[83]

1.4. LEF1

In the mouse genome, *Lef1* is positioned on chromosome 3. It consists of 12 exons, and various splicing events yield distinct mRNA variants, resulting in the translation of diverse protein isoforms. *Lef1* is transcribed from two closely spaced promoters [***Figure 1-5; adapted form Tony et al. (2006)***].[84] The first promoter generates a 3.6 kb mRNA containing a long (1.2 kb) GC-rich 5′-untranslated region (UTR), and is translated via a cap-independent mechanism to produce the full-length activating form of LEF1 protein with a β-catenin binding domain at its N-terminus. The second promoter, found within the second intron, contains a TATA box positioned 50 nucleotides

upstream of the third exon. It produces a truncated dominant-negative protein isoform (or ΔExon-1 RNA isoform) with 283 amino acids and 38-kD molecular weight, which lacks a β-catenin binding domain.[85] Truncated LEF1 is generated through conventional cap-dependent ribosome scanning, as truncated *Lef1* mRNA (2.2 kb) contains a short 60-nucleotide 5′-UTR.[85,86] The reason for calling it a "dominant-negative" isoform (DN) is because its overexpression represses luciferase-based *Tcf/Lef* reporter activity driven by co-overexpressed β-catenin. While both LEF1 isoforms can bind to Wnt-responsive elements (WRE), the main difference between them is the ability of full-length LEF1 to recruit β-catenin and, thus, activate its target genes. By contrast, DN-LEF1, lacking the ability to interact with β-catenin, acts as a suppressor of Wnt target genes, opposing the actions of FL-LEF1.[85]

The use of *Lef1* alternative promoters or otherwise modulated changes in LEF1 protein levels is associated with cellular responses to received signals, which can alter cell fate decisions.[87] Therefore, it is important to understand the patterns of alternative promoter usage and the function LEF1 isoforms in various cellular contexts such as the switch and/or change in the expression of the full-length variant that occurs during mESC differentiation.[88,89] Differential LEF1 isoform usage is not only observed in mESCs. The second LEF1 promoter is highly active in lymphocytes, while the first promoter of *LEF1* is silent.[90] However, the opposite occurs in colon cancer.[91] During osteoblast differentiation, full-length *LEF1* expression decreases while truncated *LEF1* expression is induced.[92]

Both *Lef1* promoters are directly regulated and can be activated by TCF/LEF-β-catenin complexes; however, it has been shown that other repressor and activator elements can control their expression. For instance, transforming growth factor-β3 upregulates *Lef1* expression (full-length) via SMAD2,4/LEF1 interaction in the absence of β-catenin during mouse palate

development.[93] In colon cancer cells, the MYC transcription factor functions as a transcriptional activator of *LEF1* expression. Kruppel-like factor 4 (KLF4) exhibits the capacity to bind the *Lef1* promoter and to stimulate its expression in mouse type II alveolar epithelial cells (AECs). OCT4 upregulates *LEF1* expression in the esophageal squamous cell carcinoma (ESCC) cell line Eca109 and hepatocellular carcinoma (HCC) cells and induces EMT in the HCC cells.[94-96] Caspase 8 associated protein 2 (CASP8AP2) binds the *LEF1* promoter and inhibits its expression via interaction with C-terminal binding protein 2 (CtBP2) and zinc-finger E-box binding homeobox 2 (ZEB2).[97] The second promoter, which contains at least three WREs, is repressed by Yin Yang 1 (YY1) transcription factor in colon cancer cells, where *LEF1* P2 is poorly acetylated. It has been suggested that YY1 represses *LEF1* P2 by interfering with TCF-β-catenin actions.[91]

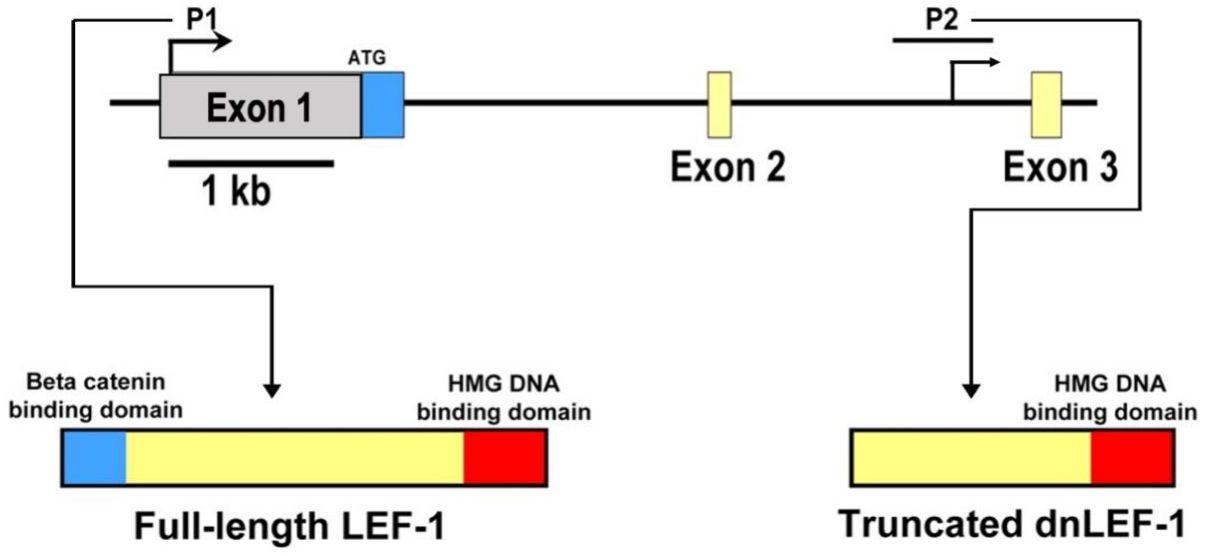

Figure 1-5. Alternative *Lef1* Promoter Usage. Adapted from Tony et al. (2006).
The first promoter (P1) is responsible for the production of FL-LEF1, while the second promoter (P2), located within the second intron, generates the dominant-negative LEF1 (DN-LEF1/ΔN-LEF1). The primary distinction FL-LEF1 and ΔN-LEF1 lies in the presence or absence of β-catenin binding domain. Specifically, FL-LEF1 contain the β-catenin binding domain, whereas ΔN-LEF1 lacks it.

LEF1 has a high-mobility group DNA-binding domain known as HMG-1, which is highly conserved. HMG DNA-binding domains have a ~85 amino-acid region of homology called an HMG box (HMGB). The HMGB DNA-binding domain (HMGB-DBD) in LEF1 is a 200-amino-acid region near its carboxyl-terminus that binds specific DNA sequences, which is not a feature shared by all HMG box proteins. Cell type-specific transcription factors such as LEF-1 and the other TCF proteins have only one HMG DNA-binding domain. HMG-1 proteins such as HMG-1 and upstream-binding factor, which are expressed in all cell types, have two or four HMG boxes and are non-sequence-specific DNA-interactors that generally act as chromatin remodeling proteins.[98]

1.4.1. LEF1 and Cancer

Abnormal expression and dysregulated function of LEF1 have been implicated in the development of various cancer types, including colon cancer, leukemia (acute lymphoblastic leukemia, acute myeloblastic leukemia, and chronic lymphocytic leukemia), esophageal squamous cell carcinoma, gastric cancer, breast cancer.[99-102] Several studies have indicated that LEF1 plays an oncogenic role in acute lymphoblastic leukemia by regulating gene expression, thereby promoting cell proliferation.[103] LEF1 has been used as a prognostic biomarker in chronic lymphocytic leukemia, acute lymphoblastic leukemia, liver metastasis, and colorectal cancer.[104-107] Studies have provided evidence of the significant involvement of LEF1 in the growth of prostate cancer through the regulation of androgen receptor expression.[108] Additionally, LEF1 has been proposed as a potential therapeutic target in prostate cancer, hematological malignancies such as chronic lymphocytic leukemia, breast cancer, colorectal cancer, and hepatocellular carcinoma.[109-113]

1.5. OVERVIEW OF GENE EXPRESSION

Gene expression is a complex process whereby the information encoded in genes is converted into functional protein or non-coding RNA products. It is controlled by a series of regulatory mechanisms at various levels, including transcriptional, post-transcriptional, translational, and post-translational levels.

The control of gene transcription is orchestrated in a dynamic and regulated fashion by an intricate network of proteins. This complex interplay of proteins consists of DNA-binding proteins, such as histones and transcription factors (TFs), as well as co-factors that do not directly bind to DNA but are recruited by TFs and/or influenced by histone modifications.[114] TFs, as the primary regulators of gene transcription, interact with non-DNA-binding proteins, such as chromatin remodeling proteins, basal transcriptional machinery, enzymes involved in modulating TF activity (e.g., phosphatases and kinases), dimerization partners, and inhibitory proteins. These interactions collectively form a complex network that governs gene expression.[115] There are approximately 1639 proteins (14,000-19,000 proteins corresponding to 6-9% of proteins) in humans identified as DNA transcription factors. Therefore, it is important to understand TF protein-protein and protein-DNA interactions and how these interactions affect transcriptional regulation.[116,117]

1.5.1. Chromatin Remodellers

Nucleosomes, as core structural elements of chromatin, regulate chromatin accessibility to TFs and other chromatin-binding factors, and nucleosome positioning across the genome is determined by chromatin remodellers. Furthermore, TFs have a key role in regulating chromatin accessibility via interacting with the chromatin remodellers.[118,119] Chromatin remodellers are characterized by their ATP-dependent nature, which enables them to use the energy from ATP hydrolysis for various process such as nucleosome movement, destabilization, ejection, or restructuring.[120]

Several chromatin remodeling complexes, including polycomb repressive complex 2 (PRC2), BAF (canonical switching (SWI) /sucrose nonfermenting (SNF) complex), and nucleosome remodelling and deacetylase (NuRD), share a common ATPase subunit.[121] However, they associate with additional protein subunits to carry out distinct functions. These accessory protein partners have the ability to interact with DNA, histones, and histone PMTs, facilitating the targeted modulation of chromatin remodelling at specific genomic regions.[118] They mainly regulate the localization of nucleosomes, which is essential for genome packaging and transcription regulation.[122] Of these chromatin remodeller complexes, two have been shown to have key roles in regulating mESC self-renewal and differentiation: SWI/SNF and NuRD.[123-125]

The SWI/SNF family of chromatin-remodeling complexes includes the BRG-1 associated factor (BAF) complex and the polycomb-associated BRG1-associated factor (PBAF) complex, both of which share common and unique subunits.[126] These complexes are highly conserved across evolution and typically consist of approximately 12-15 protein subunits.[127,128] Core members of these complexes include SMARCA4 (also known as BRG1) and SMARCC1/SMARCC2 (also known as BAF155 and BAF170).[129,130] The PBAF complex, consisting of 12 subunits, shares eight subunits with the BAF complex and possesses an additional four unique subunits.[131] These unique subunits include the DNA-binding subunit ARID2, as well as the histone tail-binding subunits PBRM1, BRD7 and PHF10.[132] Among the eight shared subunits, two are transcription factor-binding subunits SMARCE1 and SMARCD1, while one is the scaffold subunit SMARCC2.[132]

The NuRD complex is a multiprotein complex composed of two subcomplexes linked via Methyl-CpG Binding Domains (MBD) 2 and 3.[133] The nucleosome remodeling subcomplex is composed of GATAD2 and CHD4 and the deacetylase subcomplex is composed of MTA, HDAC, and RBBP proteins.[134] NuRD-deficient mouse ES cells struggle to differentiate and tend to self-

renew indefinitely.[122] Knockdown of different NuRD components has different results, suggesting that they may function independently of the NuRD complex. For instance, RNAi-mediated knockdown of chromodomain helicase DNA-binding protein 4 (CHD4) and MBD3 suppressed neural differentiation.[135]

1.5.2. The Wnt Enhanceosome

The Wnt enhanceosome, a transcriptional complex, is formed by proteins involved in regulating Wnt target genes [***Figure 1-6; adapted from Hongyang et al. (2023)***].[136] The current model depicts the Wnt enhanceosome as a complex comprising 18 proteins, including β-Catenin and TCF/LEF proteins, CBP/p300, and the BAF complex.[6,137] Within this structure, there exists a stable subcomplex known as the structural core, consisting of BCL9, LDB1, PYGO, and the ChiLS complex (composed of LDB/Chip and SSDP proteins). Regardless of the activation status of Wnt, this structural core is situated on Wnt-responsive enhancers.[136-139] The connection between the structural core and TCF/LEF is facilitated through Gro/TLE. In the inactive or "Off" state of Wnt signaling, TCF/LEFs remain bound to chromatin and establish complex associations with transcriptional corepressor factors Gro/TLE and histone deacetylase HDAC.[6,140] Upon Wnt/β-catenin activation, Gro/TLE undergo either displacement by β-catenin or inactivation through ubiquitylation mediated by UBR5. Nevertheless, Gro/TLE continue to maintain their association with the Wnt enhanceosome.[6,17,141]

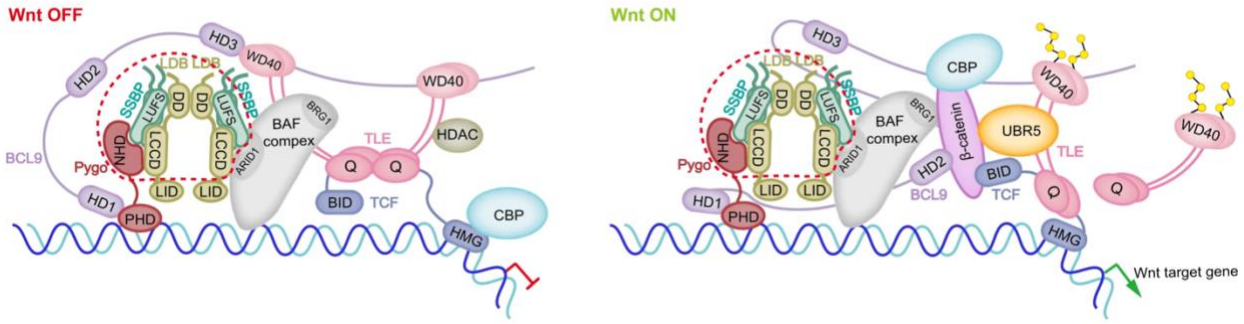

Figure 1-6. Schematic Overview of the Current Wnt Enhanceosome Model. Adapted from Hongyang et al. (2023).
The Wnt enhanceosome is composed of a stable subcomplex (including BCL9, Pygo, and ChiLS), BAF complex, TLE, HDAC, CBP, and TCF/LEFs in both Wnt "Off" and "On" state. In Wnt "On" state, there is a conformational change in the Wnt enhanceosome caused by β-Catenin recruitment to Wnt target genes.

1.6. PROTEIN FUNCTIONAL CHARACTERIZATION

Proteins exert their biological functions by forming transient interactions and/or stable complexes with molecules such as proteins and nucleic acids. To gain insights into the function of proteins, various methods have been developed to systematically analyse protein binding partners, including proteins (protein-protein interactions (PPI)) and nucleic acids (protein-nucleic acid (DNA/RNA) interactions). Transcription factors are proteins with the ability to bind DNA and are among the fastest evolving genes in humans.[142] Understanding their molecular interactions, which affect fundamental processes such as transcription, translation, splicing, and replication, would help us in elucidating cellular processes like intracellular signal transduction pathways and cell fate determination both in physiological and pathological contexts.[116,117,143]

PPIs are mainly mediated by either domain-domain interactions, with a large contact interface (~2000 Å; equals to 200 nm) or domain-motif interactions, with a small contact interface (~300-500 Å; equals to 30-50 nm).[144,145] Currently, there are two main types of methods for characterizing PPIs: 1) experimental 2) *in silico*. Experimental methods are based on biophysical, biochemical,

or genetic principles, employing a range of techniques both *in vitro* and *in vivo*. These techniques include affinity purification, Y2H (yeast 2-hybrid) systems, and TAP (tandem affinity purification).[146] *In-silico* methods rely on computational approaches and can be categorized into sequence-based approaches, structure-based approaches, chromosome proximity analysis, gene fusion analysis, *in silico* 2-hybrid analysis, phylogenetic tree analysis, phylogenetic profile analysis, and gene expression-based approaches.[147,148]

The method we used in our experiments is proximity labeling (PL), which is based on biochemical principles. Two primary types of enzyme-based PL methods exist: peroxidase-related methods and proximity dependent biotinylation. The latter is used in our study in which a biotin ligase, TurboID (35 kD, is used to covalently and promiscuously biotinylate proximal proteins. The first versions of biotin ligase PL method were BioID, derived from the *Escherichia coli* biotin ligase (BirA), which was mutated to ablate its substrate specificity yielding a 'promiscuous' biotin ligase called BirA*. This enzyme has very low activity (≥18 h of labeling time required). TurboID was developed by using a yeast display-based assay which used to screen for mutants of BioID with enhanced activity resulting from directed evolution of BirA*.[149] In both TurboID and BioID, the biotin ligase uses biotin and ATP as substrates to generate biotin-adenosine monophosphate (biotin-5´-AMP), which then labels lysine side chains of nearby proteins in a reactive radius of ~10 nm.[149] A thorough evaluation of the labelled potential interactors is required to be performed by computational pipelines to eliminate false-positive identifications.[150] The reason for choosing TurboID over BioID in our experiments is that the biotin ligase utilized in BioID (BirA*) has low activity and requires more than 18 h for labeling while the TurboID needs only 15 minutes (min).[151] Considering that Lef1 in mESCs has the highest protein levels at 48 hours after culture in mEB medium supplemented with the GSK-3 inhibitor, CHIR, a narrow temporal window of proximity

labeling is optimal and more informative; therefore, TurboID is the best choice for our mESC model.

To map the distribution of Lef1 on the genome we applied the cleavage under target and release using nuclease (CUT&RUN) technique. This is an *in-situ* antibody-targeted genome-wide chromatin profiling method where cells are immobilized on lectin-coated magnetic beads and then permeabilized and incubated with transcription factor-specific antibody. Then, a protein A/protein G-MNase fusion enzyme, which binds with high affinity to the antibody, is induced to cleave DNA adjacent to the antibody through the addition of calcium. TF-DNA complexes bound to the antibody are released from the nucleus, collected, purified, and sequenced. This method is an alternative to ChIP-seq, in which total chromatin is fragmented first and then the protein of interest bound to the DNA is immunoprecipitated, and finally the antibody-bound chromatin is recovered for DNA extraction and sequencing.[152]

CHAPTER 2. RATIONALE, HYPOTHESIS AND RESEARCH AIMS

2.1. RATIONALE

LEF1 plays a crucial role in gastrulation and other developmental processes and is a driving force in the development and metastasis of several high-grade cancers. One of the primary determinants of the biological functions of TFs, including LEF1, are interactions with multi-protein complexes. It is unclear whether the two main isoforms of LEF1 occupy the same Wnt-response elements in various contexts. Therefore, identifying proteins interacting with the two main isoforms of LEF1 and identifying LEF1 isoform target genes would help us understand the specificity of their chromatin interactions and variations in their biological function in our mESC model. Given the ubiquitous nature of TCF TFs and Wnt signaling during development and tissue maintenance and their crucial role in pluripotent and somatic stem cells, the knowledge gained could have important implications across a diverse range of developmental and disease contexts, such as developmental neoplasms and cancers (e.g., leukemia), regenerative medicine and cell therapy.

2.2. HYPOTHESIS AND RESEARCH AIMS

Hypothesis: The primary hypothesis is that ΔN-LEF1 regulates self-renewal and helps sustain ES cell pluripotency independent of Wnt/β-catenin signaling, whereas FL-LEF1 regulates lineage-specific genes by recruiting β-catenin and gene expression coregulators in differentiating mESCs.

Aim 1: To detect endogenous interaction partners of full-length LEF1 and ΔN-LEF1 in differentiating mESCs in the presence and absence of Wnt pathway activation.

Aim 2: To determine the genomic occupancy of LEF-1 and ΔN-LEF1 in mESCs differentiating *in-vitro* by using the Cleavage Under Target and Release Using Nuclease (CUT&RUN) method.

CHAPTER 3. MATERIALS AND METHODS

3.1. BIOINFORMATIC APPROACHES

3.1.1. Protein-Protein Interaction (PPI) Analysis

TurboID samples were processed by the Manitoba Centre for Proteomics and Systems Biology and data was provided as a separate excel spreadsheet (xlsx file) for each sample. A Python script, written by Dr. Doble, was used to extract data from the xlsx files that was required for downstream analyses. The script generated three data files: i) bait.dat, ii) prey.dat, and iii) inter.dat. The bait file is a 3-column list of unique sample names, bait names, and an indicator of whether the sample is a test or control (T or C). The prey file lists all identified putative preys identified in sample in a 3-column list of protein IDs, number of amino acids, and gene IDs. The interaction file is a 4-column list of sample names, bait names, associated prey names, number of peptide counts detected for the associated prey (***Figure 3-1***).

3.1.2. SAINT Analysis and Data Visualization

Processed data files were uploaded to the SAINTexpress tool in the publicly available APOSTL Galaxy Server (http://apostl.moffitt.org/) to score proximity interactions. The SAINTexpress tool reads the three .dat files and analyzes them by using the latest version of SAINTexpress. SAINT scores, ranging from 0 to 1, were used for predicting biologically significant PPIs by employing a cutoff value of 0.8.[153] SAINTexpress uses a comparison between the spectral counts of preys identified by a bait and a negative control to calculate the probability of a genuine interaction for each prey. Bait runs (three biological replicates) were compared against three negative controls (TurboID-only samples). Preys with a Bayesian false discovery rate (BFDR) of ≤ 0.01 were considered high-confidence proximity interactors. The SAINT report file was uploaded to ProHits-

Viz (a web tool for analyzing and visualizing protein-protein interaction data; https://prohits-viz.org).

Data submitted to ProHits-viz proceed through several steps to yield a desired visualization, plus a list of significant readouts. Data is transformed in five ways before visualization. First, a control value is subtracted from the abundance column, wherein SAINT input files have a column allocated to control values (CtrlCount). Then, as the length of the potential proximate proteins is presented in the prey file, the abundance value is normalized to the length of the readout (prey). By default, the data is normalized between conditions in the subsequent step. In fourth step, a log transformation is implemented to reduce or remove the skewness of the data. Finally, the highest abundance amongst the significant preys for a bait are used to simulate the multimerization of the bait with itself. After these transformation steps, data are filtered based on the primary filter value (0.01 to define significant readouts) and the minimum abundance (default: 0). Displayed readouts on the resultant dot plot image are considered significant readouts with minimum abundance in at least one condition. The intensity edges of the dot plot readouts are defined by the secondary filter with BFDR cutoff ≤ 0.01 indicated by the color of the edge.

The ProHits-viz results were used for gene ontology (GO) enrichment analysis and data visualization in PANTHER (pantherdb.org) and STRING (https://string-db.org/), Cytoscape (3.10.0), and Enrichr (https://maayanlab.cloud/Enrichr/). Venn diagrams were generated at www.BioVENN.nl and bioinfogp.cnb.csic.es (Venny 2.1.0).

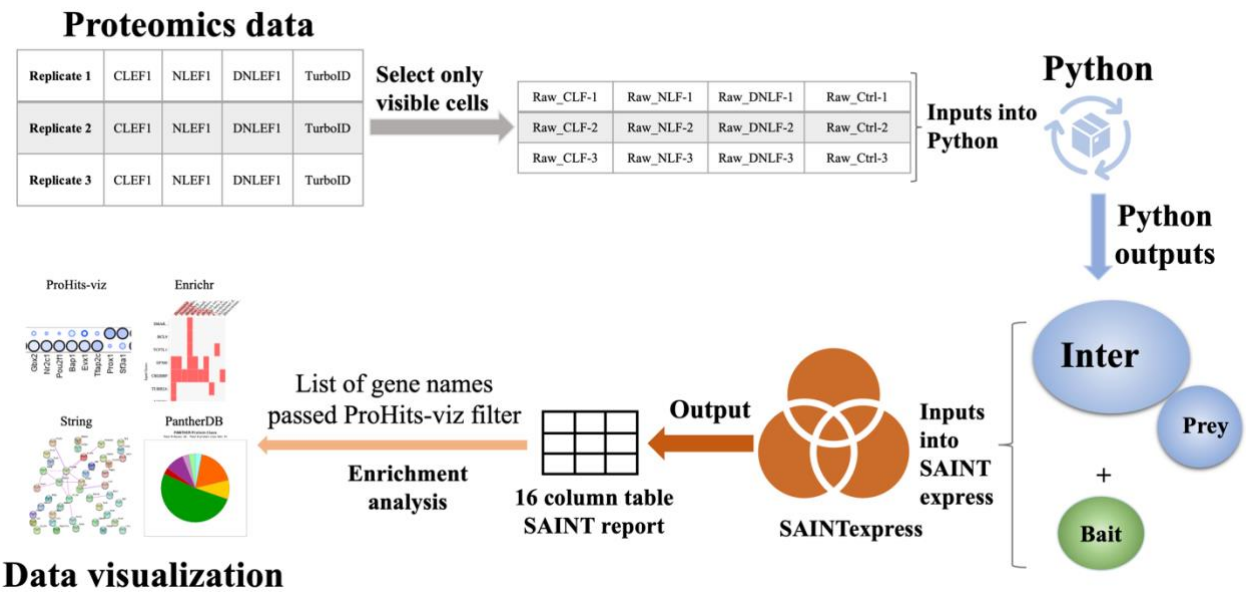

Figure 3-1. Scheme of Overall Data Analysis of TurboID Data (Proteomics Analysis Pipeline).

After preprocessed raw data (Excel) was received from the Proteomics facility it was passed through a customized Python script to generate interaction, prey and bait data files. The resulting three .dat files were uploaded to the SAINTexpress tool in the APOSTL Galaxy Server. The SAINTexpress tool was used to read the three .dat files and analyze them using the latest version of SAINTexpress. A SAINT report file was uploaded to ProHits-Viz. The data was then transformed and genes that passed the filter (significant readouts with BFDR cutoff < 0.01) were used for enrichment analysis and data visualization in Panther, STRING, Cytoscape (3.10.0), and Enrichr.

3.1.3. CUT&RUN Data Analysis

The analysis of raw sequencing data from CUT&RUN experiments followed a similar protocol as described for ATAC-seq in the study by Madden et al.,[154] with the exception of omitting ATAC-shifting of reads. The alignment process utilized the mouse mm10 genome version. Read coverage density plots were generated by employing deepTools2 bamCoverage, extending reads to the complete span of the fragment defined by pairs.[155] Average coverage tracks, as well as input-subtracted coverage tracks, were generated using wiggletools.[156] The bigWig files were uploaded to the Integrative Genomic Viewer tool (IGV_2.16.1) to visualize the CUT&RUN data (https://software.broadinstitute.org/software/igv/).

3.2. REAGENTS

A comprehensive list of the reagents and solutions used in this study is listed in Appendix A. In general, these materials were purchased from the following suppliers: Sigma-Aldrich, Gibco (Life Technologies), New England Biolabs, Cell Signaling Technology, Thermo Scientific, and EpiCypher.

3.3. CELL CULTURE

Table 3-1 provides a summary of the characteristic features of the mESCs used in this study. The mouse ES cell line E14TG2a (American Type Tissue Collection CRL-1821™) was maintained on plates pre-coated with 0.1% gelatin (Sigma, G9391) with mouse ES standard medium: DMEM (sigma, D5796), 15% fetal bovine serum (Sigma, F1051), 1000 U/milliliter (mL) mouse LIF (AMS-263-100, 10 ng/mL), 55 µM 2-mercaptoethanol (Gibco, 2198023), 1 mM GlutaMax (Gibco, 35050061), 1 mM non-essential amino acids (Sigma, M7145), sodium pyruvate (Sigma, S8636) in standard tissue culture incubator at 37 degrees Celsius (°C) with 5% CO2 (Appendix A).[157] mESCs were thawed onto 60 mm 0.1% gelatin-coated plates in standard medium and the standard media was changed after 24 hours. Another cell line used in our experiment was quadruple TCF/LEF1-Knockout (qKO) mESCs generated in the Doble lab, which lack all full-length TCF/LEFs (referred to as BD-qKO cells hereafter. This qKO cell line was used to confirm the specificity of the anti-LEF1 antibody (sc-374522) as being able to detect FL-LEF1 and ΔN-LEF1 isoforms, which was a requirement for our subsequent CUT&RUN experiments.

Table 3-1. Properties of Mouse Embryonic Stem Cells Employed in This Study.

E14TG2a	
Sex	Male
Culture Properties	Adherent
Doubling Time	~ 12-14 hours
Growth Medium	Standard mES (with LIF) + 15% FBS
Source	American Type Tissue Collection (ATTC) CRL-1821™

3.3.1. Cell Passaging

To sustain sub-confluent and actively proliferating cultures, cell passaging was performed every 48 hours within a biological safety cabinet. When the cells reached 70-80% confluency, they were split at a ratio of 1:5 to 1:6 onto new culture dishes coated with 0.1% gelatin. To liberate single cells from the tightly packed colonies, the adherent mESCs were rinsed with Dulbecco's phosphate buffer saline (DPBS, D8537) and were dissociated by adding accutase and incubating them at room temperature (RT) for 5 min. Once detached, cells were pipetted up and down to achieve a single-cell suspension and were then collected from the plate, transferred to a 15 mL or 50 mL conical centrifuge tube [FroggaBio; TB15-500 (220313); TB50-500 (220315)] and pelleted by centrifugation at 1200 relative centrifugal force (×g). After removing the supernatant, the cell pellet was suspended in 8-10 mL of standard mESC medium. Subsequently, cells were counted by using an EVE™ automatic cell counter (Montreal Biotech Inc.), and the appropriate quantity was used according to the desired seeding density for each size of culture plates (***Table 3-2***).

3.3.2. Cell Counting and Seeding

To measure cell count and viability, 50 microliters (µL) of cell the suspension was added to 50 µL of 0.4% trypan blue stain and mixed gently by pipetting. Subsequently, 10 µL of the resulting mixture was dispensed onto an EVE™ cell counting slide, in duplicate. The determination of cell count, representing the viable cell number in the cell suspension, measured in cells/mL, was obtained with the EVE™ counter. This device differentiates between live and dead cells by employing trypan blue dye exclusion as an indicator of viability. The average count of viable cells was used to determine the required cell dilution for all experiments (***Table 3-2***).

Table 3-2. Cell Seeding Density and Pipetting Volumes Employed for Various Vessels.

Experiment	Vessel	Cell Seeding Density (cell/well)	Growth Medium (mL)	Accutase (mL)	DNA per well (ng)
Western Blots	60mm	0.6×10^6	3	1	-
	100mm	2×10^6	10	5	-
TurboID	100mm	2×10^6	10	5	5000
***Lef1* KO**	6-well	0.5×10^6	2	0.5	2500
	12-well	0.1 × 106	1	0.2	-
	24-well	0.05 × 106	0.5	0.1	-
CUT&RUN	6-well	0.25×10^6	2	0.5	-

3.4. TURBOID PROCEDURE

To induce maximal LEF1 expression, based on previously conducted time-course experiments, 2 million cells were cultured on 100 mm plates pre-coated with sterile 0.1% gelatin in

differentiation medium (embryoid body (EB) medium), which has the same composition as standard medium but does not contain LIF, for 48 hours. Cells were transfected with TurboID plasmid constructs at the time of plating in mEB medium by using Lipofectamine 3000 (ThermoFisher) as per the manufacturer's guidelines on 100 mm gelatin-coated plates and incubated at 37 °C for 48 hours (***Table 3-3***). Apart from the volumes specified in ***Table 3-3***, an LF3K dilution was prepared for each reaction. This involved mixing 485 µl of OPTI-MEM medium with 15 µl of LF3K in separate tubes, resulting in a total volume of 500 µL. Subsequently, the DNA dilution of each plasmid was mixed with one LF3K tube and incubated at RT for 15 min. Finally, 500 µL of this mixture was added dropwise to the corresponding 100mm dish; each tube contained a 1000 µL mixture, which was evenly distributed between two 100mm dishes, with 500 µL added to each dish.

Each plasmid (5 µg/plate) was used to transfect cells grown on two 100 mm 0.1% pre-coated gelatin plates with 5 µM CHIR99021. The same transfection conditions were used without adding CHIR99021 for each plasmid. Then (48 hours after transfection) cells were pulsed with 500 µM biotin (ThermoFisher, B1595) in 10 mL mEB medium and incubated for 15 min at 37 °C (Appendix A).

To prepare whole-cell lysates, cells were washed twice in D-PBS and then detached with 5 mL Accutase (Sigma, A6964). Each sample duplicate was then taken from its 100 mm dish, and duplicates were pooled before proceeding to the next step. Once pooled, cells were washed again in PBS and centrifuged at 1200 ×g for 3 min. Following the removal of the supernatant, the pellet was suspended in another D-PBS wash (10 mL) and subsequently centrifuged at 1200 ×g for 3 min. Finally, the supernatant was discarded, and the pellet was rapidly frozen (snap-frozen with

liquid nitrogen) until three experimental replicates were ready for Western blotting and affinity purification.

Table 3-3. Pipetting Volumes Employed for TurboID Transfection.

	LEF1-TueboID (1210.9 ng/µL) (µl)	TurboID-LEF1 (1198 ng/µL) (µl)	TurboID-ΔNLEF1 (411 ng/µL) (µL)	TurboID-NLS Ctrl (250 ng/µL) (µl)
DNA	8.3	8.4	24.4	40
P3000	20	20	20	20
OPTI-MEM	471.7	471.6	455.6	440
Total volume	500	500	500	500

A total of three independent replicates were employed for both control and experimental samples. When all replicates were ready, samples were thawed on ice and resuspended in 2 mL/pellet ice cold Radio-immunoprecipitation Assay (RIPA) buffer (50 mM Tris pH 8.0, 150 mM NaCl, 1% NP-40, 0.5% deoxycholate, 0.1% sodium dodecyl sulfate (SDS), 1 mM EDTA) containing 1X protease and phosphatase inhibitors (ThermoFisher, 78440) for 10 min (Appendix A). To fragment DNA, the samples were subjected to sonication using two 10-second pulses at 30% amplitude (Fisherbrand™ Model 50 Sonic Dismembrator). Subsequently, the lysed samples were centrifuged at 13,200 ×g for 10 min at 4 °C. The protein amount of the supernatant of each sample was quantified with technical duplicates by using a Detergent Compatible (DC) assay (Bio-Rad, 5000111; microplate assay protocol) in accordance with the instructions provided by the manufacturer, using a standard curve created from a series of seven bovine serum albumin (BSA) protein standards. Samples and standards were read at 700 nm on a SpectraMax M3 plate reader (Molecular Devices). After the quantification process, 50 µL aliquots of streptavidin-Sepharose bead slurry (GE17-5113-01) were washed once in 1 mL of 1X RIPA per sample. The washed bead aliquots were subsequently resuspended in 1 mL of 1X RIPA containing 1 mg of quantified protein

supernatant and subjected to overnight rotation at 4°C. Following this, the samples underwent centrifugation at 1000 ×g for 5 min. The resulting bead pellets were then washed twice with 1 mL of 2% SDS solution (in ddH2O), followed by three washes with 1 mL of Wash Buffer (50 mM Tris-HCl pH 7.4, 8 M Urea). Each wash was carried out for 8 min at room temperature while rotating. Samples were then resuspended in Storage Buffer (285 µL of ammonium bicarbonate (50 mM) and 15 µL of 1 mM biotin) to saturate biotin binding sites on streptavidin and prevent peptide recapture during on-bead digestion (Appendix A). Samples were maintained at low temperature by placing them on ice and were delivered to the Manitoba Centre for Proteomics and Systems Biology for subsequent mass spectrometric analyses (MS) [***Figure 3-2; reproduced from Henning and Stephan (2020)***].[158]

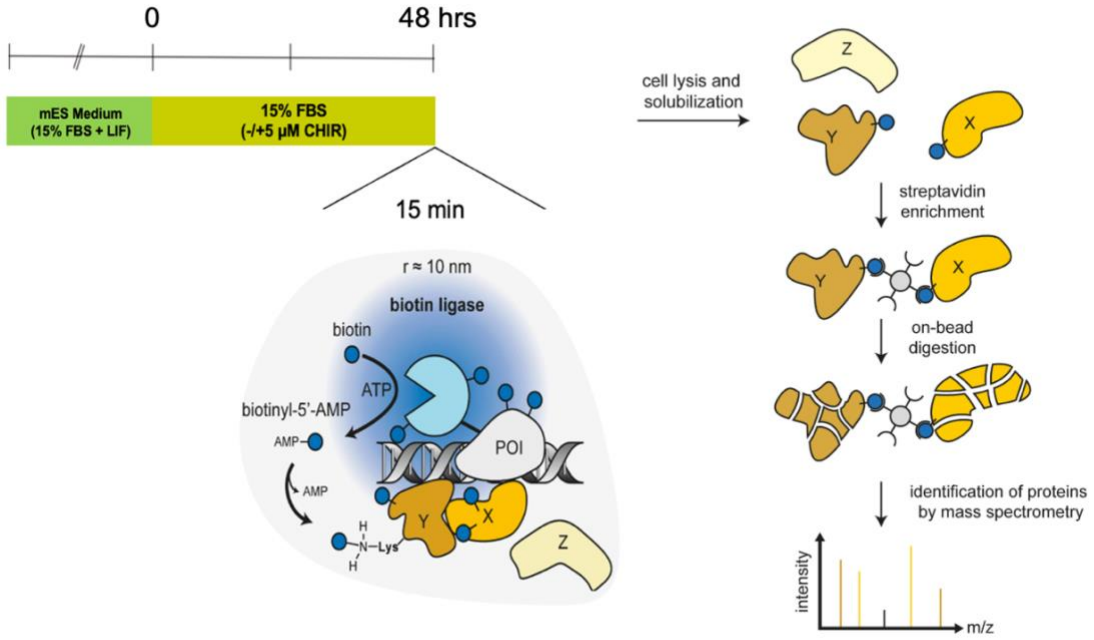

Figure 3-2. Overall Planned Scheme of the Experimental Strategy to Conduct TurboID Experiments. Reproduced from Henning and Stephan (2020).
Cells were expanded in standard LIF-supplemented medium, which is embryonic stem (ES) medium (15% FBS + LIF (10 ng/mL)). To identify FL-LEF1 and ΔN-LEF1 interactors, transient expression vectors were generated with CAG promoter-driven TurboID fusions with FL-LEF (C terminal fusion: TurboID FL-LEF1, and N terminal fusion: FL-LEF1 TurboID) or ΔN-LEF1 (N terminal fusion: TurboID ΔN-LEF1). To serve as an experimental control for the identification of non-specific TurboID interactors, control vectors (TurboID) were generated. These vectors contained a CAG-driven H2B-mAG-P2A-3xHA-TurboID cassette. E14TG2a mESCs cultured in EB medium in two conditions, with and without CHIR. Cells were transfected with each vector individually. Following a 48-hour incubation period in medium lacking LIF, the cells were exposed to a 500 µM Biotin pulse for 15 min before sample collection. Then cells were lysed, and the protein was solubilized and quantified. The biotinylated proteins were then enriched via Streptavidin-Sepharose beads and prepared for mass spectrometry analysis via on-bead digestion.

3.5. WESTERN BLOT ANALYSIS

Western blot analysis was used to quantify endogenous protein levels of FL-LEF1 and ΔN-LEF1, evaluate TurboID fusion to FL-LEF1 and ΔN-LEF1, and to screen and then verify CRISPR/Cas9-mediated *Lef1* knockout by identifying clones lacking detectable FL-LEF1 bands.

3.5.1. Whole Cell Protein Extraction

Protein extraction was conducted from cells cultivated in either 60 mm or 100 mm dishes. Cell culture medium was removed, and cells were washed twice with 2 mL or 5 mL cold PBS (4 °C). Following the rinsing steps, Radio-immunoprecipitation Assay (RIPA) buffer containing 1X protease and phosphatase inhibitors was added. The volume of RIPA buffer used was 300 µL for 60 mm dishes and 1 mL for 100 mm dishes. Cell culture plates were cooled on ice and allowed to incubate for 10 min, with gentle tapping every 5 min. Subsequently, the cells were collected using cell scrapers and transferred to chilled 1.5 mL microcentrifuge tubes kept on ice. To facilitate cell lysis, the lysate was subjected to sonication with two pulses of 10 seconds (s) each at 30% amplitude. Following sonication, the samples were centrifuged at 13,200 ×g for 10 min at 4 °C. The resulting supernatant was carefully removed and transferred to new, pre-chilled 1.5 microcentrifuge tubes for subsequent quantification.

3.5.2. Protein Quantification via Detergent Compatible (DC) Assay

Samples were quantified by using the Bio-Rad DC protein assay following the microplate assay protocol provided by the manufacturer. To prepare the working reagent A', a total of 20 µL of reagent S was added to each mL of reagent A, which contained alkaline copper tartrate. Then 5 µL of standards (a set of 7 BSA protein standards mentioned in section 3.3.3), and test protein samples were dispensed in duplicate wells of the clean, dry microtiter plate. 5 µL of 1X RIPA (containing

1X protease and phosphatase inhibitors) was used as blank. Then 25 µL of the prepared working reagent A' was added per well, followed by the addition of 200 µL of reagent B (a diluted Folin reagent) to each well. The plate was gently agitated to mix the reagents. After 15 min, absorbance was acquired for each well at 700 nm on SpectraMax M3 plate reader (Molecular Devices)

3.5.3. Gel Electrophoresis and Western Blot

After quantifying the proteins, samples were prepared by adding sample buffer (4X) containing lithium dodecyl sulfate (LDS) (ThermoFisher, B0007) along with Bond-Breaker TCEP solution (ThermoFisher, 77720) (Appendix A). The addition aimed to achieve equal protein quantities for each condition (10 µg/well). Samples were then heated at 70 °C in a mini dry bath (FroggaBio) for 5 min to denature proteins. The chamber of the electrophoresis tank (Mini Gel Tank; A25977) was filled with 400 mL of 1X running buffer (MES SDS Running Buffer (20X)) and 2 mL of sample reducing agent (200X) (Appendix A). Gel cassettes were placed into the chamber of the electrophoresis tank. Samples were loaded to each well (10 µg/well). To assess the relative position of the protein of interest, 10 µL of PiNK Plus Prestained Protein Ladder (GeneDireX, Cat. PM005-0500) was added into a single well.

Samples were run through 4-12% gradient Bis-Tris Plus mini protein gel (ThermoFisher, NW04122BOX) at 150 volts (V, constant voltage) for 50 min using a FisherBrand (Fisher Scientific) electrophoresis power supply. 10 min before the end of electrophoresis, pre-cut membranes (nitrocellulose, mini size) and filters were equilibrated in 1X Power Blotter 1-Step Transfer Buffer (Appendix A) for 10 min before proceeding with the transfer step. After the electrophoresis run, the cassette was rinsed with distilled water and then opened. The gel was dislodged from the cassette into distilled water and incubated at room temperature for 3 min to ensure uniform wetting, facilitate appropriate gel positioning, and enhance contact between the gel

contact and the membrane. Water was replaced with 20% ethanol in distilled water and incubated for 2 min.

The gel and membrane were assembled using the Power Blotter System (ThermoFisher). The power blotter was set to pre-programmed blotting methods depending on the number of the gels and the protein weight range. For our samples all transfers were set at Mixed Range Molecular Weight: 25-150 kDa, and electrical current 1.3 amps for 1 mini-sized gel and 2.5 amps for 2 mini-sized gels.

The membranes were subjected to blocking with 5% milk/Tris-buffered saline (TBS) for 1 hr at room temperature (RT), and then incubated overnight at 4 °C with primary antibodies diluted in 3% milk/TBST (tris-buffered saline tween 20) at the concentrations specified in ***Table 3-4*** (Appendix A). Following three washes (5 min each) in 3% milk/TBST, the blots were treated with secondary antibodies conjugated to horseradish peroxidase (HRP), which were diluted in 3% milk/TBST. The blots were then incubated at RT for 1 hour. After five washes with 1x TBST, membranes were incubated at RT for 5 min with SuperSignal West Pico PLUS Chemiluminescent Substrate (ThermoFisher). Subsequently, blots were imaged using a ChemiDoc MP Imaging System (Bio-Rad). Membranes were probed with antibodies against LEF1 (1:1000, Santa Cruz Biotechnology, sc-374522), hemagglutinin/HA tag (1:1,000, Cell Signalling Technology, C29F4), GAPDH (1:10,000, Sigma, G8795), goat anti-mouse IgG (H+L)-HRP conjugate (1:5,000, Bio-Rad 172-1011), goat anti-rabbit IgG (H+L)-HRP conjugate (1:10,000, Bio-Rad, 170-6515).

Table 3-4. List of Antibodies Employed for Western Blot Analysis.

Primary Antibodies				
Antibody	Source	Catalogue Number	Species	Dilution
LEF1	Santa Cruz Biotechnology	sc-374522	Mouse	1:500 and 1:1000
HA-Tag	Cell Signalling Technology	C29F4	Rabbit	1:1000
GAPDH	Sigma	G8795	Mouse	1:10,000
Secondary Antibodies				
Goat α Mouse HRP[A]	Bio-Rad	172-1011	Goat	1:10,000
Goat α Rabbit HRP[A]	Bio-Rad	170-6515	Goat	1:10,000

[A]HRP (Horseradish peroxidase)

3.6. PLASMID PREPARATION FOR TURBOID CONSTRUCTS AND CRISPR/CAS9-MEDIATED FL-LEF1 KNOCKOUT

3.6.1. TurboID Plasmid Design

TurboID fused to the N terminus of LEF1 was generated by previous Ph.D. student in the Doble lab, Dr. Victor Gordon (***Figure 3-3A***). To fuse TurboID to the C terminus of LEF1 with a long linker, primers were designed for the amplification of LEF1 (1188 bp, excluding the stop codon) and 3XHA TurboID (1050 bp) (Appendix B). A linker (oligonucleotide) was purchased from IDT

as a ssDNA oligo that could be inserted into a linearized N-vector buy using NEBuilder methodology (New England Biolabs). CAG-H2B-mAG-P2A-3xHA-TurboID-FL-Lef1 plasmid was linearized using KpnI (NEB #R3142) and ClaI (NEB #R0197s), and the ~6.3 kb band was extracted as the linearized vector. Finally, CAG-H2B-mAG-P2A-LEF1FL-3XGGGS-3XHA-TurboID (8597 bp) was generated via NEBuilder HiFi DNA Assembly Cloning Kit (NEB #M5520A) (***Figure 3-3B)***. P2A is a 66 bp sequence that is translated into 22 amino acids, facilitating the cleavage of polypeptides through ribosome skipping.[159] In our expression constructs, P2A is positioned between fluorescent marker gene H2B-mAG (histone 2B fused to the green fluorescent protein monomeric Azami Green[160]) and TurboID fusion proteins. This arrangement allows for the translation of both the fluorescent protein and the fusion products into separate cleaved proteins.[161]

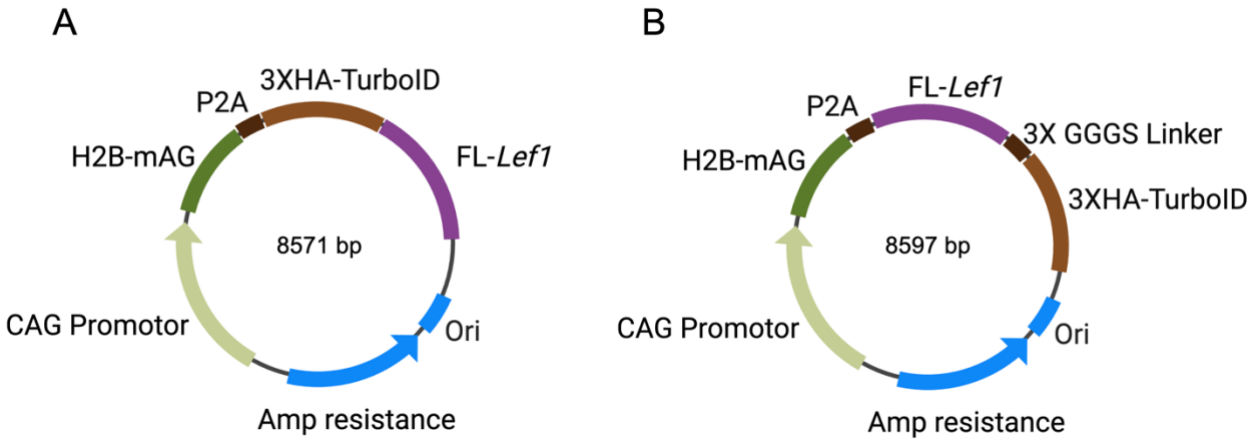

Figure 3-3. Schematic Circular Map of TurboID Fusion to the C- and N-terminus of FL-LEF1.

To drive high levels of *Lef1* expression in mammalian expression vectors, a CAG promoter composed of a cytomegalovirus (CMV) major immediate-early enhancer fused to the chicken beta-actin (CBA) promoter is used. The H2B-mAG is used as fluorescent chromatin marker, which then is followed by 2A "self-cleaving" peptide used for cloning of multiple genes in a single vector. Then, whether the protein of interest (POI) required to be fused N- or C-terminally with TurboID, the TurboID is located before or after the POI gene, respectively. The HA (hemagglutinin) tag, derived from the human influenza virus HA protein, is used extensively as a general antibody epitope tag. ***(A)*** TurboID-LEF1 plasmid (8571 bp) where TurboID gene is fused to the N-terminus of *Lef1 gene*. ***(B)*** FL-LEF1-longlinker-TurboID plasmid (8597 bp) generated by fusing TurboID gene to the C-terminus of Lef1gene via a long linker oligonucleotide.

To generate truncated LEF1 (ΔN-LEF1) TurboID plasmid, CAG-H2B-mAG-P2A-3xHA-TurboID (Control) vector was linearized using ClaI, and the resulting linearized vector (~7.3 kb) was used as the recipient for PCR-amplified ΔN-LEF1 (852 bp). ΔN-LEF1 was PCR amplified (from CAG-H2B-mAG-P2A-3xHA-TurboID-FL-LEF1) using primers designed to exclude the first 339 bp nucleotides and the reverse primer had 18 bp homology with the backbone of the linearized recipient vector (Appendix B). The assembly reaction resulted in the generation of the CAG_H2B_mAG_P2A_3xHA_TurboID_Short-LEF1 plasmid (8228 bp) (***Figure 3-4A***).

To generate TurboID-NLS vector, the TurboID-LEF1 plasmid was linearized via ClaI and KpnI-HF and the ~6.3 kb linearized vector was extracted. The TurboID gene for this plasmid was PCR amplified and the ~1.2 kb amplified fragment was extracted (Appendix B). Following the assembly reaction, the pCAG_H2BmAG_3xHA-Turbo-NLS (7455 bp) was generated (***Figure 3-4B).***

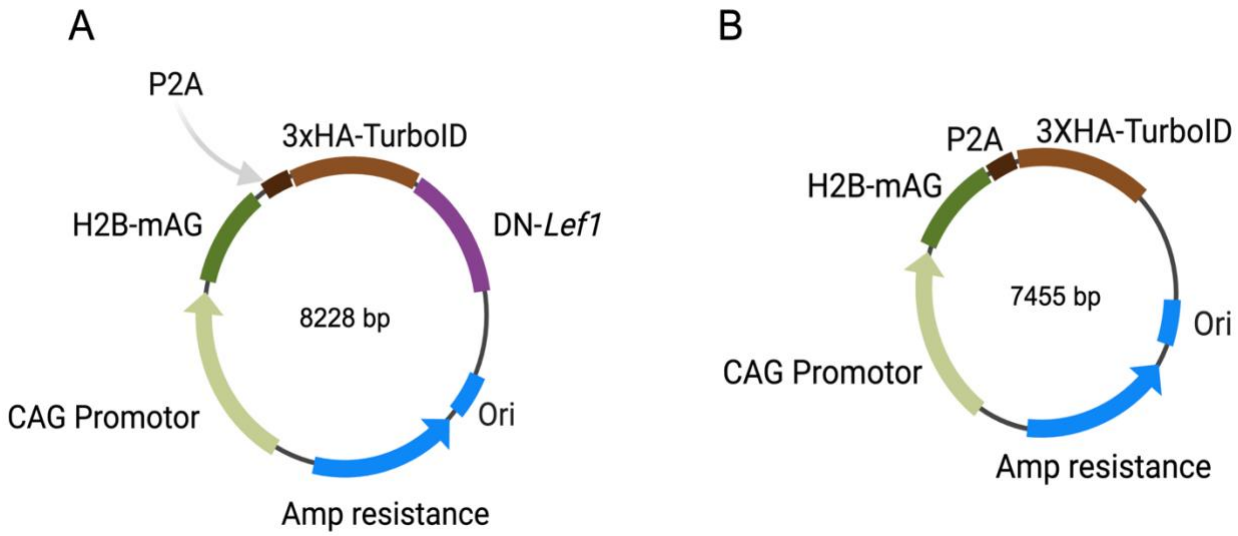

Figure 3-4. Schematic Circular Map of TurboID-DN-LEF1 and TurboID-NLS Control Plasmids.
(A) TurboID fused to the N-terminus of short *Lef1* gene (8228 bp) lacking the first 339 bp. ***(B)*** The TurboID-only plasmid (7455 bp) contains the TurboID), expressing biotin ligase TurboID without being fused to any other protein.

3.6.2. *Lef1* Knockout Plasmid Design

Single-guide RNAs (sgRNA) were designed using the Invitrogen TrueDesign Genome Editor. A target sequence of 20 nt (5'-GATCAGTCATCCCGAAGAGG-3'), with the protospacer adjacent motif (PAM; 5′-AGG-3′) in exon 1 of the mouse *Lef1* gene (target locus: 131111584) was selected based on the predicted efficiency (99.90%), reduced off-target effects (one: GATCACTCATCCCCAAAAGG; PAM: AGG; mismatches: 3(6,14,17); Genomic Locations: chr14[24597608]) (Appendix B). To generate pSpCas(*Lef1*-sgRNA) (9.1 kb), the synthesized *Lef1* sgRNA (IDT) was subjected to annealing and subsequently cloned into pSpCas9(BB)-2A-Puro (PX459) V2.0 (Addgene plasmid # 62988) where the sgRNA is driven by a U6 promoter according to the protocol developed by Dr. Zhang's laboratory (targeting sequence cloning protocol).[162] Guide RNAs were validated in mESCs by transfecting cells using Lipofectamine 3000. The targeted region for CRISPR was amplified using PrimeSTARMax polymerase (Takara; R045), and each resulting amplicon was subjected to digestion with T7 Endonuclease I, which cleaves at sites with DNA mismatches such as the insertions/deletions (indels) created by localized Cas9 endonuclease activity (NEB; M0302S). The activity of the sgRNA/Cas9 in creating indels was quantified by analysing the digested amplicons on agarose gels. The primer sequences for amplifying the *Lef1* gRNA target site for digestion with T7E1 are described in Appendix B.

3.6.3. *Escherichia Coli* Transformation and Plasmid Preparation

Chemically competent *Escherichia coli* cells were transformed with plasmids to allow them to be prepared in large amounts (e.g., plasmid midi-preps). Exogenous expression constructs, including TurboID plasmids and pSpCas9 (PX459) plasmids were prepared according to the manufacturer's protocol for transformation (High Efficiency Transformation Protocol; New England Biolabs). Briefly, 2 µL (containing 1 pg-100 ng) of the plasmid DNA (ligation mixture)

to the cell mixture and the mixture was placed on ice for 30 min. After being incubation on ice, bacteria underwent heat shock in 42 °C water bath for 30 s, followed by a return to ice for a duration of 5 min. Subsequently, a volume of 950 µL of Super Optimal broth with Catabolite repression (SOC) outgrowth medium (NEB) at room temperature was added to the mixture. The resulting mixture was then placed in an incubator shaker (New Brunswick Scientific) set to 37 °C and 250 revolutions per minute (rpm) for 60 min. For purpose of positive selection of cells containing desired plasmids, aliquots of 50 µL and 200 µL of transformed cells in SOC medium were spread onto petri dishes containing Luria-Bertani (LB) agar supplemented with 100 µg/mL carbenicillin (refer to Appendix A). The plated dishes were subsequently incubated overnight (O/N) at 37 °C.

The next day, a volume of 3 mL of LB broth with 100 µg/mL carbenicillin was used to inoculate single bacterial colonies (screening ~10 colonies per desired expression construct) (Appendix A). The inoculated colonies were then incubated O/N in a shaker incubator at 37 °C and 250 rpm. After 15-18 hours, the plasmid DNA was extracted using the Presto™ Mini Plasmid Kit (Geneaid; PDH300). The concentration and purity of the DNA were assessed using a Nanodrop spectrophotometer (Thermo Scientific). Out of 10 selected colonies, those with the expected bands, after cutting them into specific sized pieces and analyzing the resulting fragments by gel electrophoresis (diagnostic digestion using restriction enzyme(s)), were further amplified by inoculating 50 mL LB broth containing 100 µg/mL carbenicillin and incubated O/N at 37 °C at 250 rpm. The next day, plasmid DNA was extracted using Presto™ Midi Plasmid Kit (Geneaid; PIF025) and again diagnostic digestion was performed to confirm the amplification of the desired expression construct. All plasmid constructs were confirmed by Sanger Sequencing (Eurofins Genomics).

3.6.4. Generation and Isolation of Cas9 Clones

FL-LEF1 knockout cell lines were generated by using CRISPR-Cas9. The *E. coli* strain Stbl3 was used for the transformation of plasmids, and they were grown on Luria-Bertani (LB) agar plates with 100 µg/mL carbenicillin and incubated overnight. The plasmid DNA was extracted, and successful sgRNA integration was verified by double digestion with BbsI and EcoRI restriction enzyme according to the NEB restriction digest protocol. The sgRNA sequence in the final construct was validated by Sanger Sequencing (Eurofins Genomics).

E14GT2a cells (~3.0×10^5 E14GT2a) in ES medium were seeded on one well of 0.1% gelatin-coated 6-well plate. Two hours later, cells were transfected with 2.5 µg of CRISPR/Cas9 plasmid using Lipofectamine 3000 (ThermoFisher) per manufacturer's guidelines and were incubated for 48 hours at 37 °C (***Table 3-5***). Apart from the volumes specified in ***Table 3-5***, an LF3K dilution was prepared for each reaction. This involved mixing 121.25 µl of OPTI-MEM medium with 3.75 µl of LF3K in separate tubes, resulting in a total volume of 125 µL. Subsequently, the DNA dilution of each plasmid was mixed with one LF3K tube and incubated at RT for 15 min. Finally, 250 µL of this mixture was added dropwise to the corresponding well of 6-well plate.

Table 3-5. Pipetting Volumes Employed for CRISPR/Cas9 Transfection.

	Lef1*/exon 1 **(517.7 ng/µL) (µl)**	**PX459 ** (440 ng/µL) (µl)**	**TurboID-NLS Ctrl*** (221 ng/µL) (µl)**
DNA	4.83	5.68	11.3
P-3000	5	5	5
OPTI-MEM	115.2	114.3	108.7
Total volume	125	125	125

*Lef1/exon1 is the PX459 containing the *Lef1* sgRNA.
**PX459 (V.2.0) serving as the control negative as it does not have the *Lef1* sgRNA.
***TurboID -NLS was used to assess the success of transfection.

Cells were washed once with 1 mL PBS 48 hours post transfection, dissociated using 500 µL Accutase, and incubated for 5 min. The single cell suspension was transferred to a 50 mL tube containing 8 mL ES medium and centrifuged at 1200 ×g for 3 min. After centrifugation, the pellet was resuspended in 50 mL ES medium and then 10 µL Puromycin (stock: 10 mg/mL) was added and mixed to yield a final concentration of 2 µg/mL. 10 mL of this diluted cell suspension was plated per each of 5, 0.1% gelatin-coated 100 mm plates, and incubated for 36 hours. After 36 hours, cells were washed once with 5 mL PBS and the medium was changed (ES medium without puromycin). By diluting the cells and plating them in selective medium as described above, clonal colonies arising from single cells were generated.

3.6.5. Clone Expansion and Screening for full length *Lef1* Gene Knock out

To isolate individual CRISPR-edited clones, after 5 to 7 days in mES medium, with medium changed every 48 hours, single clones were visible to the naked eye and ready for picking. Cells were washed once with 5 mL PBS, followed by the addition of 10 mL of PBS to each plate. Using a 25 µL pipet, single clones were picked under a microscope, and one clone was transferred to one well of a U-bottom 96-well plate containing 50 µL Accutase. Following a 5-min incubation, a single-cell suspension was prepared by pipetting. Single-cell suspensions of single clones were seeded into single wells of 0.1% gelatin-coated 24-well plates containing 1 mL ES medium and were incubated for 48 hours (some clones needed more time to reach optimal confluency).

After 48 hours, each clone (one well of a 24-well plate containing 1 mL ES medium) was split into two wells of a 12-well plate. Cells were washed once with 500 µL PBS, then 150 µL Accutase was added to the one well (24-well plate) and incubated for 5 min at room temperature (RT). Then single-cell suspension was made by pipetting up and down and 75 µL of single-cell suspension was taken and added to one well of a gelatin-coated 12-well plate already containing 2 mL ES

medium. Cells were cultured for a duration of 48 hours at a temperature of 37 °C, although some clones took more time to reach the optimal confluency. The cells from one of each pair of 12-well replicates were frozen down after 48 hours and the cells in the remaining well were used to screen for successful CRISPR-Cas9-mediated FL-Lef1 knockout after expanding them to a single-well of a 6-well plate.

After 48 hours, cells were washed once with 1 mL PBS, then 200 µL Accutase was added per well of the 12-well plate and incubated for 5 min. Between 3.0-5.0 × 10^5 cells (50 to 100 µL of single cell suspension depending on confluency) were added to one well of 0.1% gelatin-coated 6-well plates containing 2.5 mL ES medium and incubated for 48 hours at 37 °C.

For freezing cells, cells on a 12-well plate were washed twice with 1 mL PBS. Each well was treated with 200 µL Accutase and incubated at RT for 5 min. Then 100 µL of single cell suspension was transferred to a cryovial containing 500 µL freezing medium (25% FBS + 10% DMSO). Immediately cryovials were put in cell freezing boxes (e.g., Mr. Frosty Freezing Container) and transferred to -80 °C. After 24 hours, the cryovials were moved to liquid nitrogen for long-term storage.

3.6.6. Genomic DNA Extraction

Genomic DNA was isolated from wild type mESCs, transfected mESCs (for T7E1 experiment), and from putative full length *Lef1* knockout clones by using Monarch DNA Purification Kit as per manufacturer's instructions (NEB #T3010). In brief, cells harvested from 60 mm dishes (0.1% gelatin-coated) that reached 80% confluency were treated with Accutase, collected and centrifuged at 1000 ×g for 1 min. The supernatant was discarded, and the cell pellet was resuspended in 100 µL cold PBS through gentle pipetting. The 100 µL cell suspension was subsequently treated following the manufacturer protocol (Monarch Genomic DNA Purification Kit). The DNA was

eluted in a preheated (60 °C) gDNA elution buffer, with elution volume ranging from 35 to 100 µL. Subsequently, the concentration and purity of the DNA were evaluated using a Nanodrop spectrophotometer (Thermo Scientific). The DNA samples were kept at -20 °C until further use for polymerase chain reaction (PCR) amplification and DNA sequencing.

3.6.7. Polymerase Chain Reaction

PCR was used for the amplification of targeted fragments used in construction plasmids, to verify the *Lef1* sgRNA-Cas9 target site after T7E1 experiment, and to amplify the edited region (Exon 1) of *Lef1* to send for sequencing. The PrimerQuest Tool (https://www.idtdna.com/PrimerQuest) was employed to design PCR primer pairs. Primer sequences used in all experiments in this thesis are mentioned in Appendix B. High-Fidelity DNA polymerase, PrimeStar Max Premix (Takara), was used to ensure robust high-fidelity amplification. Reactions were conducted following the manufacturer's instructions, using the specified reagents and volumes outlined in ***Table 3-6***. The PCR mixture was thoroughly mixed, followed by PCR amplification with a MiniAmp PCR thermocycler (Themo Fisher Scientific), according to the parameters described in ***Table 3-7***. Following PCR amplification, the resulting products underwent either gel electrophoresis for visualization and analysis or were stored at 4 °C before subsequent DNA sequencing.

Table 3-6. Reagents and Volumes Employed in PCR Reactions.

Reagent	Volume/50 µL Reaction	Final Concentration
Template DNA	< 200 ng (genomic DNA) *	< 5 ng/µL
Primer 1	1 µL (10 µM)	0.2 µM
Primer 2	1 µL (10 µM)	0.2 µM
PrimStar Max Premix	25 µL	1X
Nuclease-Free Water	To reaction volume of 50 µL	-

*Recommended quantities of template DNA (50 µL reaction) are 5-200 ng for genomic DNA and 10 pg – 1 ng plasmid DNA.

Table 3-7. Thermocycler Conditions for PCR.

Step	Temperature	Time	Number of Cycles
Initial Denaturation	98 °C	3 min	1
Denaturation	98 °C	10 s	35
Annealing	58 °C	15 s	
Extension	72 °C	15 s	
Final Extension	72 °C	3 min	1
Hold	4 °C	-	-

3.6.8. Agarose Gel Electrophoresis

PCR products and restriction enzyme digested products were visualized, analyzed, and/or extracted using agarose gel electrophoresis. To prepare a 1% agarose gel, 1 g of agarose (Invitrogen) was dissolved in 100 mL of 1X Tris-acetate-EDTA (TAE) buffer in an Erlenmeyer

flask (refer to Appendix A). The mixture was microwaved until the complete dissolution of the agarose. For the visualization of DNA electrophoresis, 10 µL of EZ-Vision Bluelight DNA Dye 10000X (VWR #10791-798) was added to the agarose prior to its solidification (above gelling temperature, 45 °C). The gel-dye mixture was poured into a casting set. Once the gel solidified, it was transferred to the electrophoresis tank filled with 1X TAE buffer. To load the samples, DNA (PCR product or RE digested product) was mixed with 10 µL Gel Loading Dye Purple 6X (NEB #B7024S), and 10 µL of the resulting mixture was loaded per well. The first well of the gel was loaded with 10 µL 1 kb DNA Ladder (FroggaBio #DM010-R500). Electrophoresis was performed for 45 min at 100 V, and the gel was visualized using a blue light transilluminator (VWR).

3.6.9. DNA Sequencing and Sequence Analysis

The identification, characterization, and verification of recombinant plasmids, including the DNA sequences of inserts, determination of insert orientation, and examination of the junctions between the plasmid and insert DNA, were accomplished through Sanger DNA sequencing (Eurofins Genomics). Moreover, DNA sequencing was employed to validate and analyze the gene edits mediated by CRISPR/Cas9 in putative *Lef1*$^{-/-}$ clones. We used two primers (forward and reverse) to perform bi-directional DNA sequencing using primers listed in Appendix B. DNA sequencing results were analyzed using the online sequence alignments tool of Benchling (https://benchling.com).

3.7. CUT&RUN

mESC cell lines (wild type, and three biological FL-*Lef1*$^{-/-}$ KO clones) were cultured in one well of a 6-well plate (0.1% gelatin coated; 2.5 × 10^5 cells per well) in EB (15% FBS) media in the presence and absence of CHIR (5 µM) for 48 hours at 37 °C. The CUT&RUN protocol was performed using a CUT&RUN Assay Kit (Cell Signalling Technology #86652S) supplemented

with DNA Purification and Spin Columns (Cell Signalling Technology #14209S) [***Figure 3-5; adapted from Peter et. al. (2018)***].[152] We assessed 4 CUT&RUN conditions and conducted three experimental/biological (knockout lines) replicates. One input sample was prepared for each cell line type and condition. In total, we had 12 CUT&RUN samples and 4 Input samples (***Table 3-8***).

Table 3-8. CUT&RUN Samples and Input Samples Used in CUT&RUN Experiment

Replicate	**With CHIR**		**Without CHIR**	
Replicate 1	WT	A5 (*Lef1*$^{-/-}$)	WT	A5 (*Lef1*$^{-/-}$)
Replicate 2	WT	D5 (*Lef1*$^{-/-}$)	WT	D5 (*Lef1*$^{-/-}$)
Replicate 3	WT	F1 (*Lef1*$^{-/-}$)	WT	F1 (*Lef1*$^{-/-}$)
Inputs	WT	D5 (*Lef1*$^{-/-}$)	WT	D5 (*Lef1*$^{-/-}$)

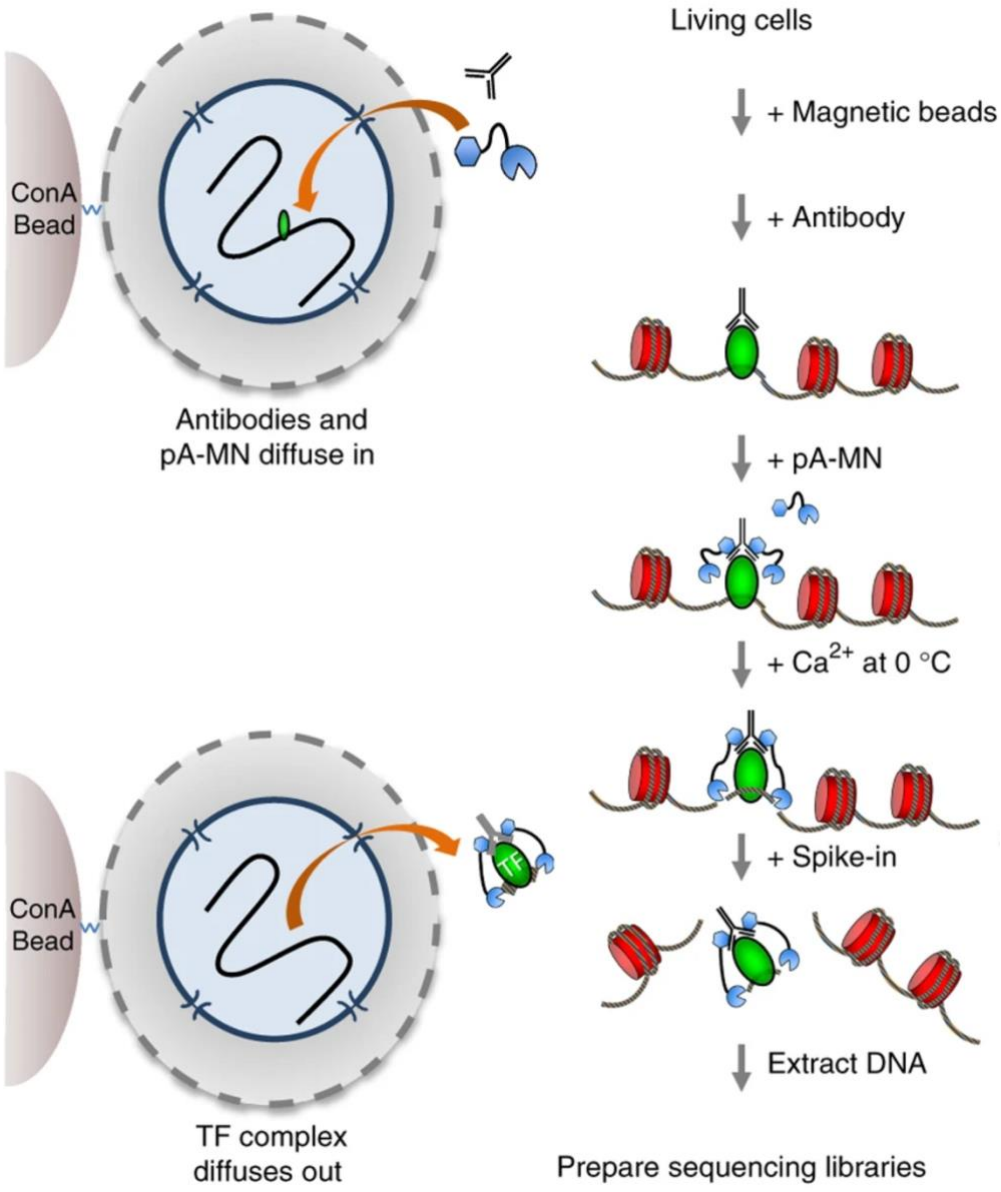

Figure 3-5. Schematic Overview of CUT&RUN Technique. Adapted from Peter et. al. (2018).

First, cells are harvested and immobilized on magnetic beads coated with ConA. Subsequently, the immobilized cells are concurrently permeabilized and incubated with the antibody against POI. Following an overnight incubation, the protein A/G-MNase is added to the reaction to be attached to the antibody already bound to POI. Then the MNase nuclease activity is activated by adding Ca^{2+}, leading to the cutting and the release of DNA sites following enzymatic digestion. The cut DNA fragments are purified and used for library preparation, and the samples are subsequently sent for sequencing.

3.7.1. Preparation of Live Cells

Microcentrifuge tubes (1.5 mL) were labeled according to the sample IDs. Accutase was added to detach cells (500 µL/well of 6-well plates) and incubated for 5 min, and cells were then pipetted up and down to make single cell suspensions. Cells were counted, and 2×10^5 cells (in < 1 mL media) were transferred to the labeled 1.5 mL microcentrifuge tubes and used for subsequent experiments. After centrifugation at 600 × g for 3 min at room temperature, the supernatant was removed (it is acceptable to retain a volume of cell medium per reaction in the tube that is equal to or less than 40 µL). The pellet was resuspended in 1 mL 1X Wash Buffer at RT by pipetting. The cell suspension was centrifuged at 600 × g for 3 min at RT, and the supernatant was removed. The pellet resuspended in 100 µL 1X Wash Buffer by gently pipetting (100 µL per tube containing 200,000 cells). One sample per condition and cell line (four inputs) were stored at 4˚C to serve as input samples.

3.7.2. Binding of Concanavalin A Beads and Primary Antibody

10 µL of the activated Concanavalin A (ConA) bead suspension was added per reaction to the washed cell suspension, mixed well by pipetting up and down, and incubated for 5 min at RT. To activate beads, ConA beads were resuspended (stored at 4 °C) gently and carefully by pipetting up and down. 45 µL of ConA was transferred to a 1.5 mL microcentrifuge tube. 450 µL ConA bead activation buffer (stored at 4 °C and kept on ice while using) was added to the ConA beads and beads were gently mixed by pipetting up and down. Tubes were placed on a magnetic rack for around 30 sec to 2 min till solution becamde clear. The liquid was removed using a pipette (not a vacuum aspirator). Then 45 µL ConA activation buffer was added to the ConA beads and resuspended gently. Activated beads were stored on ice until use (activated beads should be used within one day).

10 µL of the activated ConA bead suspension was added per reaction to the washed cell suspension, mixed well by pipetting up and down, and incubated for 5 min at RT. The suspension was placed on a magnetic rack until the solution turned clear (30 sec to 2 min). Then the supernatant was removed. Tubes were removed from the stand and 100 µL of Antibody Binding Buffer (+ spermidine + PIC + 0.025% digitonin) was added per reaction. 1 µg Ab [5 µL of anti-LEF1 per reaction (Santa Cruz B-6; 200 µg/mL); 4.2 µL of anti-H3K43me per reaction (241 µg/mL); 10 µL of Rabbit mAb IgG per reaction (100 µg/mL)] was added to the corresponding reactions and mixed gently by pipetting up and down. Tubes were incubated overnight at 4 °C.

3.7.3. Binding of pAG-MNase Enzyme

The next day, bead-bound cells were centrifuged at 100 ×g for 2 sec, then tubes were placed on a magnetic rack (30 sec to 2 min) and the liquid was removed. Tubes were removed from the rack and 1 mL of Digitonin Buffer (2.15 mL per reaction: 215 µL 10X Wash Buffer + 21.5 µL 100X Spermidine + 10.75 µL 200X Protease Inhibitor Cocktail + 53.75 µL Digitonin Solution + 1.849 mL Nuclease-free Water) was added to each tube. Beads were resuspended in 1 mL Digitonin Buffer by gently pipetting up and down (making sure no beads were stuck to the tube wall). Tubes were placed on the magnetic rack (30 sec to 2 min) and then the liquid was removed. Tubes were removed from the rack, and then 50 µL of pAG-MNase pre-mix (50 µL Digitonin Buffer + 1.5 µL pAG-MNase Enzyme per reaction) were added to each tube and gently mixed by pipetting up and down. Samples were incubated at 4 °C for 1 hour.

3.7.4. DNA Digestion and Diffusion

Samples were briefly centrifuged at 100 ×g for 2 sec. Samples were washed twice with Digitonin buffer. Each wash was as follows: 1 mL of Digitonin Buffer (per sample) was added to

magnetized samples and resuspended by gently pipetting up and down. Tubes were placed back on the magnetic rack for 30 sec to 2 min, and liquid was removed. After the second wash, samples were resuspended in 150 µL Digitonin Buffer and mixed gently by pipetting up and down. Tubes were placed on ice for 5 min to cool before digestion. MNase was activated by adding 3 µL cold calcium chloride to each tube and mixed by pipetting up and down. Samples were incubated at 4 °C on a cooling block (not on ice) for 30 min.

Following the incubation, 1X Stop Buffer (150 µL per sample) was added, consisting of 37.5 µL 4X Stop Buffer, 3.75 µL Digitonin Solution, 0.75 µL RNAse A + 108 µL Nuclease-free Water. Additionally, 10 µL of Spike-in DNA (200 pg; 1 ng/µL stock Spike-In DNA was diluted 1:50 in ddH2O) was added to each reaction to adjust the normalization reads to around 0.5% of total reads. The mixture was thoroughly mixed by pipetting up and down. Subsequently, the samples were incubated at 37 °C for 10 min, then centrifuged at 4 °C for 2 min at 16,000 ×g and were placed on a magnetic rack (30 sec to 2 min). The supernatant, containing the MNase-liberated chromatin samples, was then transferred to a new 2 mL centrifuge tube.

3.7.5. Preparation of the Input Sample

200 µL of DNA Extraction Buffer (2 µL Proteinase K + 0.5 µL RNAse A + 197.5 µL DNA Extraction Buffer) was added to each 100 µL input sample (cell preparation step). The input sample was placed at 55 °C for 1 hr with shaking (ThermoMixer ~750 rpm).[163] Then it was placed on ice for 5 min to completely cool the sample. The input sample was sonicated (eight cycles of 30 s on and off at 4 °C) to lyse the cells and fragment the chromatin. Samples were incubated on ice for 30 s between pulses. Lysates were clarified by centrifugation at 18,500 × g for 10 min at 4 °C. The supernatant was transferred to a new 2 mL microcentrifuge tube.

3.7.6. DNA Purification

The DNA from the CUT&RUN and input samples was purified using Spin Columns. The eluates (purified DNA in 50 µL elution buffer) were quantified by using a Qubit™ Fluorometer (Thermo Fisher Scientific) and Qubit 1X dsDNA high sensitivity assay kit (Thermo Fisher Scientific # Q33230). Purified CUT&RUN DNA samples were stored at -20 °C till all replicates were ready for library preparation. Quantification is not necessary for the CUT&RUN samples at this stage, as the DNA concentration is too low (0.5-10 ng). However, the Qubit assay is required for input samples as whole-genome chromatin DNA is used in input samples (higher DNA concentration; >10 ng), which has to be diluted appropriately for library preparation (5 ng starting material).

3.7.7. Next Generation Sequencing Library Preparation

The DNA samples from three CUT&RUN replicates were quantified via Qubit and 25 µL of the purified DNA was used for library preparation using CUTANA™ CUT&RUN Library Prep Kit Primer Set 1 (EpiCypher. #14-1001). The library was prepared based on the manufacturer's protocol (CUT&RUN Library Prep Kit Manual). Briefly, 25 µL of each CUT&RUN sample and 5 ng of each Input (the volume of Input was adjusted to 25 µL with 0.1X TE Buffer) was transferred to separate wells of an 8-strip PCR Tube. Then 5 µL of End Repair Master Mix (composed of End Prep Buffer and End Prep Enzyme) was added to each sample. Samples were pipetted up and down, gently vortexed and briefly centrifuged. The reactions were placed in a thermocycler with the heated lid set to 75 °C and conditions described in ***Table 3-9***.

Table 3-9. Thermocycler Conditions for End Repair Reaction

Step	Temperature	Time	Cycles
Reaction temperature	20 °C	20 min	1
Enzyme activation	65 °C	30 min	1
Hold temperature	4 °C	-	-

Following the end repair reaction, samples were put on ice. Then 1.25 µL Adaptor for Illumina and 15.5 µL Ligation Master Mix (containing Ligation Mix and Ligation Enhancer) was added to each reaction on ice. Samples were incubated in a thermocycler for 15 min at 20 °C without a heated lid. Then tubes were placed in a RT rack. U-Excision Enzyme (1 µL) was added to each reaction. Samples were gently vortexed, centrifuged briefly and incubated in a thermocycler at 37 °C for 15 min with a heated lid set to 47 °C. Samples were removed from the thermocycler and centrifuged briefly.

To perform DNA cleanup, 47.75 µL (1X reaction volume) SPRIselect reagent (i.e., NGS cleanup magnetic beads) was added to each reaction. Samples were gently vortexed, centrifuged briefly, and incubated at RT for 5 min. 8-strip tubes were placed on a magnet (Invitrogen™ DynaMag™- 96 Side Magnet #12331D) for 2 min. The supernatant was discarded, and while tubes were on the magnet, 180 µL of Fresh 85% Ethanol (EtOH) was directly added onto beads. The supernatant was discarded. This wash step was repeated for a total of two washes. After the final wash, samples were centrifuged briefly and replaced on the magnet. The supernatant was removed, and beads were air-dried for 2 min. The target DNA was eluted by adding 12 µL 0.1X TE Buffer to each reaction (vortex followed by a quick spin), then incubated at RT for 2 min. Then samples

were placed on the magnet and incubated for 2 min. The eluted DNA was transferred to new 8-strip PCR Tubes.

Index primers were added to the eluted DNA via an indexing PCR reaction (***Table 3-10***). 1 µL of each assigned i7 and i5 index primer was added to each sample and then 12.5 µL of High Fidelity 2X PCR Master Mix was added. Samples were vortexed and centrifuged briefly. The PCR reaction was performed based on the conditions mentioned in ***Table 3-11***.

Table 3-10. Indexing Primer Organization

	i5.1	**i5.2**	**i5.3**	**i5.4**	**i5.5**	**i5.6**	**i5.7**	**i5.8**
i7.1	WT Input	WT+CHIR Input	KO Input	KO+CHI Input	WT1	WT1+CHIR	*Lef1*$^{-/-}$ A	*Lef1*$^{-/-}$ A +CHIR
i7.2	WT2	WT2+CHIR	*Lef1*$^{-/-}$ C	*Lef1*$^{-/-}$ C +CHIR	WT3	WT3+CHIR	*Lef1*$^{-/-}$ B	*Lef1*$^{-/-}$ B +CHIR

Table 3-11. Thermocycler Conditions for Indexing PCR

Step	**Temperature**	**Time**	**Number of Cycles**
Hot Start Activation	98 °C	45 s	1
DNA melting	98 °C	15 s	14
Hybrid annealing/extension	60 °C	10 s	
Final extension	72 °C	60 s	1
Hold temperature	4 °C	-	-

Following indexing PCR, 25 µL of SPRIselect reagent was added per reaction to perform the DNA cleanup as mentioned above. After this cleanup, we did an additional cleanup (adding 10.5 µL SPRIselect reagent) to remove adaptor dimers (identified as an additional ~147 bp peak in Bioanalyzer results) as recommended by the Genome Québec.

The CUT&RUN libraries were quantified by using 1 µL of each library via Qubit fluorometer. Based on the Qubit results, the libraries were prepared (10 ng/µL) for Library QC via Bioanalyzer DNA Analysis (Agilent 2100 Bioanalyzer System). For the Bioanalyzer assay, High Sensitivity DNA Reagents and High Sensitivity DNA Chips (Agilent 5067-4626) were used. Briefly, the chip was put on a priming station and 9 µL of Gel-DNA dye mix (High Sensitivity DNA Gel Matrix and DNA Dye Concentrate) was loaded at the bottom of the well was marked G (microfluidic chip). The gel-dye mix was dispensed via the priming station syringe and held for 1 min. Three other wells assigned for gel-dye mix were loaded with 9 µL gel-dye mix. Then, 5 µL High Sensitivity DNA marker was added to the wells assigned for samples (11 wells) and one well marked as the ladder. 1 µL of High Sensitivity Ladder was added to the well which was marked as ladder. 1 µL of each library was loaded to other sample wells. The chip was vortexed in an IKA vortex mixer for 60 s at 2400 rpm. The chip was inserted into the Agilent 2100 Bioanalyzer and run using the 2100 Expert software, which was used for analyzing the Bioanalyzer data as well.

The Qubit results of the libraries (concentrations in ng/µL) and the average library size (measured via Bioanalyzer) were used in the online Library Concentration Conversion Calculator (Integrated DNA Technologies (IDT); https://www.idtdna.com/Calc/library-concentration-conversion) to give us library concentrations in nanomolar (nM), which is required to pool the libraries (all 16 samples were pooled together). The results calculated by the IDT convertor were then used as inputs to calculate the amount of each library required to be used to make 30 µL of 5

nM Pooled Library using the online illumina Pooling Calculator. The volume and concentration of pooled libraries are sequence type dependent. In our case, the required volume and concentration for illumina NextSeq PE50 (100 million reads) were ≥ 25 µL and ≥ 3 nM, respectively.

Once the pooled library was prepared, its quantity and quality were measured via Qubit and Bioanalyzer, respectively. The pooled library was pipetted into one well of 96-well fully skirted PCR microplate (Axygen PCR-96-FS-C), the plate was then sealed with aluminum adhesive film (adhesive PCR Plate Foils AB0626), and samples were sent for sequencing to Genome Québec. The quantity and the quality of the pooled library was verified by the Genome Québec sequencing facility using qPCR and LabChip (more sensitive than Bioanalyzer), respectively. The samples entered the sequencing run when the pooled library's quality control data was proven to be within the usual range of the sequencing facility.

CHAPTER 4. RESULTS

4.1. Determining the Endogenous Protein Levels of FL-LEF1 and ΔN-LEF1 in Undifferentiated and Differentiating mESCs

To identify the levels of FL-LEF1 and ΔN-LEF1 in self-renewing and differentiating wild-type (WT) E14TG2a mouse ESCs, cells were cultured in standard medium containing 15% FBS plus LIF for 48 hours and then were cultured in mEB medium lacking LIF (15% FBS) in the presence and absence of CHIR for 24, 48, 60, and 72 hours. FL-LEF1 protein levels began to increase 24 hours after growing cells in the absence of LIF and presence of CHIR and reached its maximal level at 48 hours, after which its level began to decrease at 60 hours and reaching to the same level at 72 hours as it was at 24-hour timepoint. ΔN-LEF1 expression was consistently detected in undifferentiated and differentiating cells, and it was more highly expressed than FL-LEF1 in all conditions tested. Treatment with CHIR increased protein levels of FL-LEF1 while it had no significant effect on ΔN-LEF1 protein levels (***Figure 4-1***).

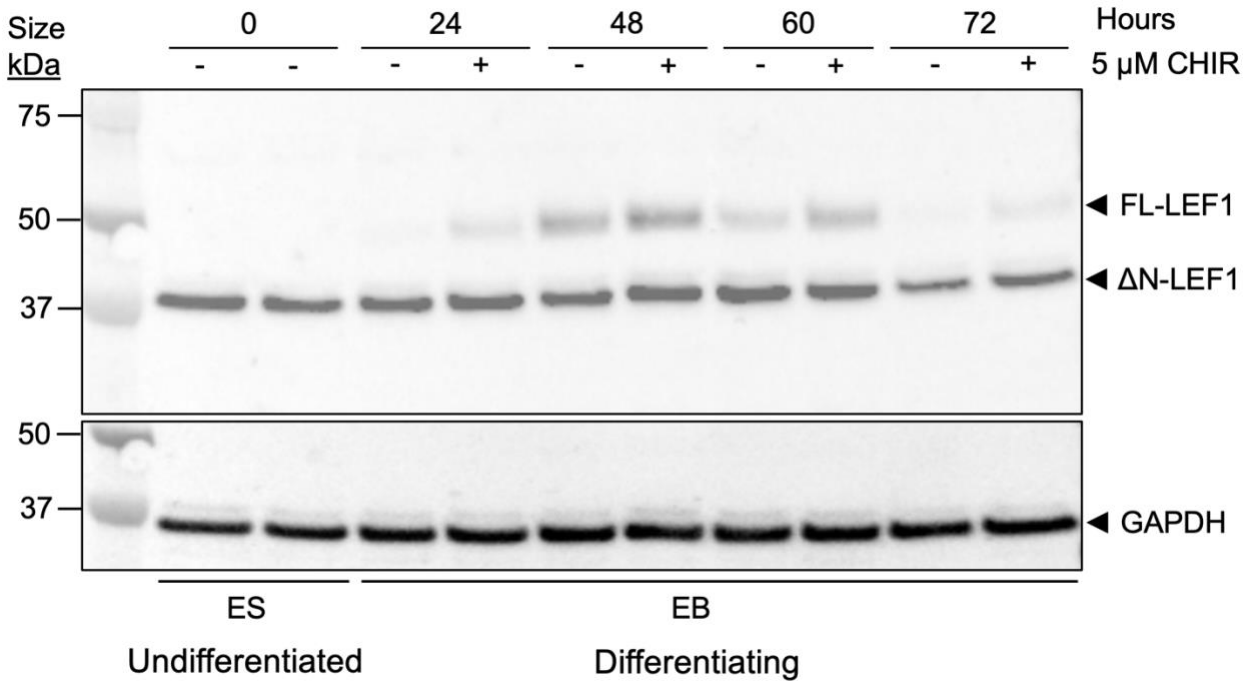

Figure 4-1. Western Blot Analysis of Endogenous Protein Levels of FL-LEF1 and ΔN-LEF1 in Wild Type (WT) mESCs in the Presence and Absence of CHIR.
Cells were expanded in standard LIF-supplemented medium for 48 hours (here is labeled as zero time) and were then cultured in EB medium in two groups: 1) treated with 5 μM CHIR). 2) without CHIR (-CHIR), across four time points [0 (Undifferentiated cells), 24, 48, 60, and 72 hours]. Lysates were probed with antibodies against LEF1 and GAPDH, as indicated.

4.2. Aim 1: Detecting Interaction Partners of Full-Length LEF1 and ΔN-LEF1 in Differentiating mESCs

4.2.1. Validation of the TurboID-LEF1 Fusion Protein Expression Constructs

Wild type mESCs were transfected separately with the following tandem HA-tagged (3×HA) overexpression plasmids: 1) TurboID fused to the C terminus of FL-LEF1; 2) TurboID fused to the N terminus of FL-LEF1; 3) TurboID fused to the N terminus of ΔN-LEF1; 4) TurboID-alone plasmid (negative control). Protein expression was driven by a CAG (Chicken-beta-actin) promoter. After 48 hours, cells were assessed by immunofluorescence (IF) and cell lysates by western blotting (WB) after the addition of 500 μM biotin (final concentration) for 15 min. Fluorescence was observed under an epifluorescence microscope, which indicated that the cells

had been successfully transfected with the corresponding expression constructs, which also drive the expression of a green-fluorescent protein-tagged histone 2B (mAG-H2B). ***Figure 4-2*** shows mESCs observed under a fluorescence microscope 48 hours after plasmid transfection. Western blot analysis revealed that exogenous constructs were successfully expressed, and their expression was analysed by using anti-LEF1 and anti-HA (hemagglutinin) antibodies (***Figure 4-3***).

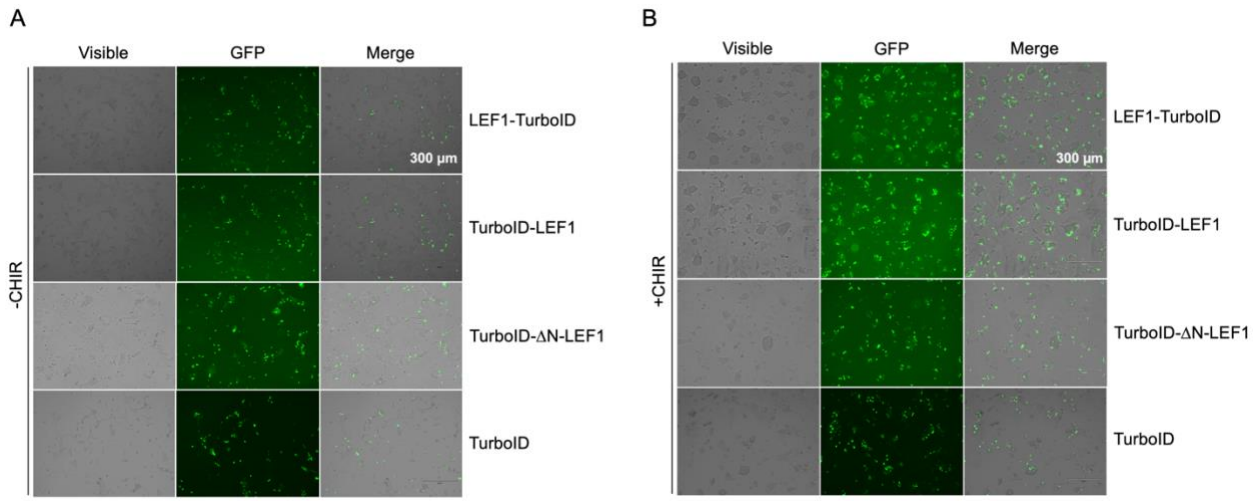

Figure 4-2. Immunofluorescence images of exogenous TurboID expression constructs.
(A and B) Images taken 48 hours following the transfection of WT mESCs with exogenous expression constructs (TurboID-LEF1, LEF1-TurboID, TurboID-ΔN-LEF1, and TurboID) in 15% FBS (EB) medium in the presence ***(A)*** and absence ***(B)*** of 5 µM CHIR. Scale bar, 300 µm.

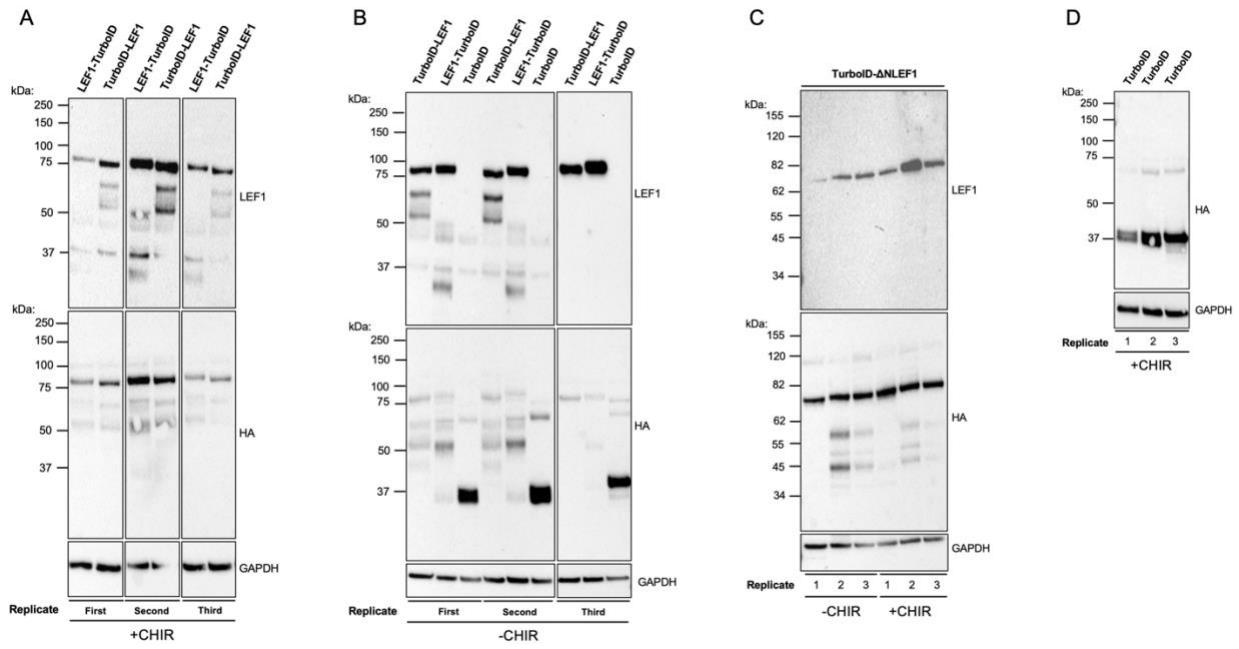

Figure 4-3. Western Blot Analysis of TurboID-LEF1 Fusion Proteins in mESCs Cultured in EB Medium in the Presence and Absence of CHIR.

(A) Three replicates of mESCs transfected with C&N terminal fusion of TurboID to FL-LEF1 in mEB medium in the presence of 5 μM CHIR for 48 hours. ***(B)*** Three replicates of mESCs transfected with C&N terminal fusions of TurboID to FL-LEF1 in EB medium in the absence of CHIR for 48 hours. ***(C)*** Three replicates of mESCs transfected with N-terminal fusion of TurboID to ΔN-LEF1 in mEB medium in the presence and absence of 5μM CHIR for 48 hours. ***(D)*** Three replicates of mESCs transfected with TurboID-NLS plasmid in the presence of CHIR to serve as negative control detecting non-specific interactors. Lower faint bands are protein degradation products resulting from freeze-thawing samples. These lower faint bands are not present in the third replicate of -CHIR condition in subfigure ***B***, which were fresh samples that had not undergone freezing and thawing.

4.2.2. Higher Number of LEF1 Hits in the Presence of CHIR versus in the Absence of CHIR

Our TurboID results revealed that there is a dramatic increase in the number of significant LEF1 interactors (BFDR ≤ 0.01) in the presence of CHIR (933 hits) compared to in the absence of CHIR (17 interactors). Some of the detected interactors are the same proteins with different identifiers. After deduplication, the vast majority (921 hits) are unique significant interactors in the presence of CHIR while only 5 are unique interactors in the absence of CHIR. There are 12 common significant interactors present both in the presence and absence of CHIR (***Figure 4-4A*** ;Venn diagram generated via https://www.biovenn.nl/). [164]

Five unique LEF1 interactors in the absence of CHIR are PTBP1, TLE3, RPS18, ZSCAN10, and GNL2. Of these five interactors, two proteins, ZSCAN10 and RPS18, are abundantly detected only with ΔN-LEF1 as the bait, while three proteins, PTBP1, GNL2, and TLE3 are preferentially detected with FL-LEF1 as the bait. PTBP1 and GNL2 are significant interactors of CLEF1 (TurboID fused to the C-terminus of FL-LEF1), whereas TLE3 is detected as being proximal to both C&N terminal FL-LEF1 TurboID fusions (***Figure 4-4B***).

12 common interactors are MYBBP1A, KMT2D, LEF1, NOP58, TBX3, ARID2, ELAVL1, TFAP2C, JPT2, RPL34, RPS13, and KMT2C. Of these 12 interactors, only LEF1 is found associating with both ΔN-LEF1 and C&N terminal FL-LEF1 TurboID fusions in the presence and absence of CHIR. LEF1, KMT2D, and TBX3 associate with both FL-LEF1 (C&N LEF1) and ΔN-LEF1 in the presence of CHIR. KMT2D interacts with both ΔN-LEF1 and C&N terminal FL-LEF1 TurboID fusions in the presence of CHIR, while in the absence of CHIR it associates with C&N terminal FL-LEF1 TurboID fusions but not ΔN-LEF1. MYBBP1A, NOP58, TBX3, ARID2, ELAVL1, JPT2, and RPL34 favor associating with C&N terminal FL-LEF1 TurboID fusion

proteins in the presence of CHIR. However, in the absence of CHIR, these interactors, except RPL34, favor associating with different baits: MYBBP1A and ELAVL1 favor NLEF1, NOP58 and TBX3 favor CLEF1, and ARID2, JPT2 and RPS13 favor ΔN-LEF1. TFAP2C has a significant interaction with Clef in the presence of CHIR, while in the absence of CHIR it interacts with ΔN-LEF1.

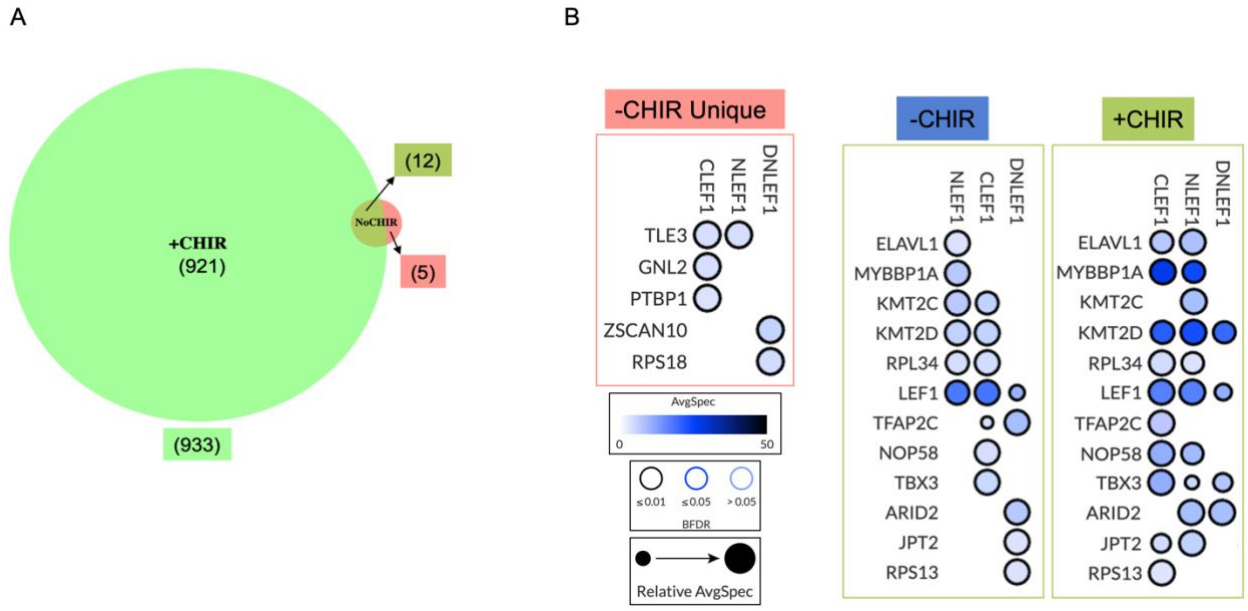

Figure 4-4. Common and Unique Interactors of ΔN-LEF1 and FL-LEF1.
(A) Venn diagram showing the higher number of interactors in the presence of CHIR (Wnt/*β*-catenin "On"). ***(B)*** Dot plot of relative average spectral count of LEF1(either ΔN-LEF1 or FL-LEF1) common and unique interactors in the presence and absence of CHIR (-/+CHIR). *Left (light red)*: 5 LEF1 unique hits in the absence of CHIR interacting with either ΔN-LEF1 (ZSCAN10 and RPS18) or FL-Lef1 (TLE3, GNL2, and PTBP1). *Right (green)*: 12 common LEF1 hits in -/+CHIR favoring different LEF1 isoforms (ΔN-LEF1 and FL-LEF1).

4.2.3. FL-LEF1 and ΔN-LEF1 Have Isoform-Specific and Isoform-Independent Interactomes in the Presence and Absence of CHIR

Significant interactors of individual C&N terminal FL-LEF1 TurboID fusion proteins and TurboID-ΔN-LEF1 in the presence and absence of CHIR are shown in ***Figure 4-5A and B*** (Venn diagram generated via https://bioinfogp.cnb.csic.es/tools/venny/).[165] We reported LEF1 interactors in the absence of CHIR in the previous section (Section ***4.2.2***). Here, we focus on the interactors in the presence of CHIR. Of the 921 hits interacting with LEF1 (both ΔN-LEF1 and FL-LEF1) in the presence of CHIR, the vast majority (889 hits) are FL-LEF1 unique (FL-LEF1-specific) interactors. There are 19 interactors common between FL and ΔN-LEF1 in the presence of CHIR, of which four are Wnt enhanceosome components (EP300, ARID1B, SMARCD1, and LDB1), three transcription factors (TCF7L1, ARNT, and ARID3A), three transcriptional co-activators (NCOA3, NCOA2, AND ARID5B), two PTM regulators (TRIM33 AND ASXL2), one transcriptional repressor (NCOR2), the RNA exome nuclease complex subunit (EXOSC10), cohesion-loading factor NIPBL, the GSK-3-dependent Wnt signaling regulator ALDOA, and two other proteins TMPO and HSPA5 (***Figure 4-5B;*** *brown*).

13 high confidence hits uniquely interact with ΔN-LEF1, of which six are transcription factors (GTF2IRD1, KLF4, ZFHX3, ZFP609, GATA-6, and GATAD2A), two epigenetic regulators (BCOR and KMTB2), two ribosome-associated proteins (RPL18A and ESF1), and three other proteins (MIDEAS, MRE11A, and LASP1) (***Figure 4-5B;*** *light green*).

Regarding the differences in the number of hits between TurboID fusions to the C-terminus and N-terminus FL-LEF1 (CLEF1 vs NLEF1), of the 920 high confidence FL-LEF1 (LEF1-TurboID and TurboID-LEF1) interactors, most hits (624 hits) are common C&N-terminal TurboID FL-LEF1 interactors while 130 hits only favor LEF1-TurboID, and 148 hits only favor TurboID-LEF1

(***Figure 4-6A***). ***Figure 4-6B*** represents GO Molecular Function enrichment analysis (maayanlab. cloud/Enrichr; integrative gene-list enrichment analysis) of the top 10 enriched terms ranked by enrichment score (P-value; top ten ≤ 0.001) overlapping for the three gene-lists: LEF1-TurboID, C&N-TurboID LEF1 common hits, and TurboID-LEF1.

As for LEF1-TurboID fusion protein interactors, the top 10 gene sets ranked by P-value are RNA-binding, keratin filament-binding (such as VIM, KRT14, and SIRT1), thioesterase binding (RAC1, ARF6, and HAUS7), cadherin binding (SEPTIN9, RAB1ATJP1/2, and PPP1CA), histone deacetylase activity (HDAC2/3 and SIRT1), snoRNA-binding (SNU13, IGF2BP3, and TSR1), U3 snoRNA-binding (SNU143 and TSR1), GTPase activity (RSB5A, RAB18, PAC1, ARF6, SEPTIN9, RAB1A, and TSR1), N-acetyltransferase activity (NMT2, ESCO2, and SIRT1), and ribonucleoside triphosphate phosphatase activity (RAB5A, RAB18, PAC1, ARF6, SEPTIN9, RAB1A, and TSR1).

Regarding TurboID-LEF1 fusion protein interactors, the top 10 gene sets ranked by P-value are RNA-binding (such as PUF60, PFN1, PRL23A), DNA-binding (such as RPS27, CNBP, CREBBP, ARID3A/B, CDK9, PATZ1, SOX6, TCF7L1, ZNF219, ZNF451, TAF4/6, H2AZ1, and PROX1), cadherin binding (such as PUF60, CTNND1, and NUDC), nucleosome binding (PARP1, SMARCC2, ARID1B, SMARCCD2, and H2AZ1), RNA polymerase II-specific DNA transcription factor binding (JUN, PARP1, CREBBP, TAF4/6, H2AZ1, PROX1, and MED1), damaged DNA-binding (SDE2, HMGB2, PARP1, and CREBBP), supercoiled DNA-binding (PSIP1 and HMGB2), LBD domain binding (PROX1 and MED1), aryl hydrocarbon receptor binding (TAF4/6), and nuclear receptor binding (PROX1, MED1, NCOR1, PHB2, and MED14).

As for common C&N terminal FL-LEF1 TurboID fusion proteins, the top 10 gene sets ranked by P-value are RNA-binding (such as THARP3, KDM1A, RTF1, HMGB1, RPL5/6/7,

MYBBP1A, POU5F1, HNRNPA0, HNRNPLL, HNRNPK, and ILF3), mRNA-binding (such as RNF20/40, CCT5, and SHMT1/2), cadherin binding (such as HDLBP, PRDX1, HSPA5/8, MYH9, CTNNB1, PARK7, JUP, and TMPO), DNA-binding (such as EP300, KDM6A, MBD3, ATRX, KMT2D, MCM4/5/6/10, POLA1, LEF1, and TARDBP), mRNA 3'-UTR binding (such as HNRNPU, RSL1D1, RPL5, ILF3, PABPC1/4, ELAVL1, and HNRANPA0), purine ribonuclease triphosphate binding (such as CAPRIN1, EIF4G1, HSPA8, ATP1A1, ARF3, CAD, DRG1, UBA2, NOL9, MSH6, CHD8, MCM6, and HRNPU), chromatin DNA-binding (such as SUZ12, KDM3B, KDM6A, MBD3, ATRX, and ACTL6A), nuclear receptor binding (such as DDX5, YWHAH,

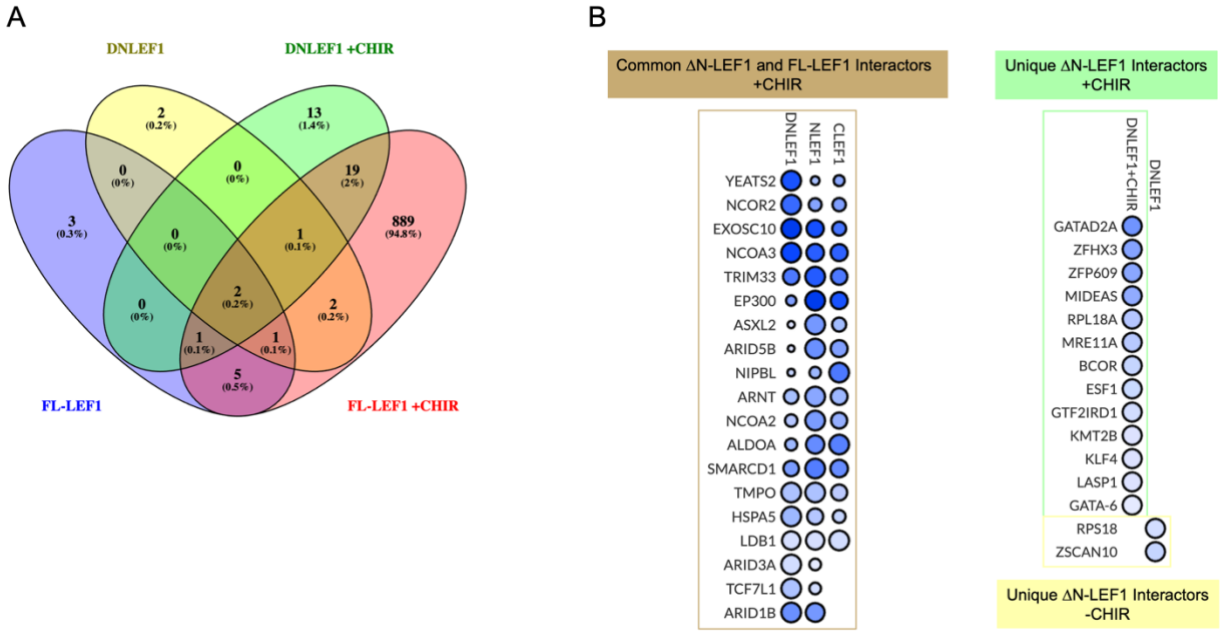

Figure 4-5. Comparing Isoform-Independent (common) and Isoform-Specific (unique) LEF1 Protein Interactions in the Presence and Absence of CHIR.
(A) Venn diagram showing the common and unique interactors of FL- and ΔN-LEF1 with and without CHIR. ***(B)*** Dot blot of 19 common FL- and ΔN-LEF1 hits (brown) and 13 unique ΔN-LEF1 hits (light green).
TAF7, TRIP12, and PARK7), mRNA 5'-UTR binding (PRL5, RSL1D1, SYNCRIP, RPS3A, NCL, DDX3X, SHMT1/2, and IGF2BP1), and single-stranded DNA-binding (such as POLA1, RPA1, TDP2, RTF1, HSPD1, and FUBP1/2/3).

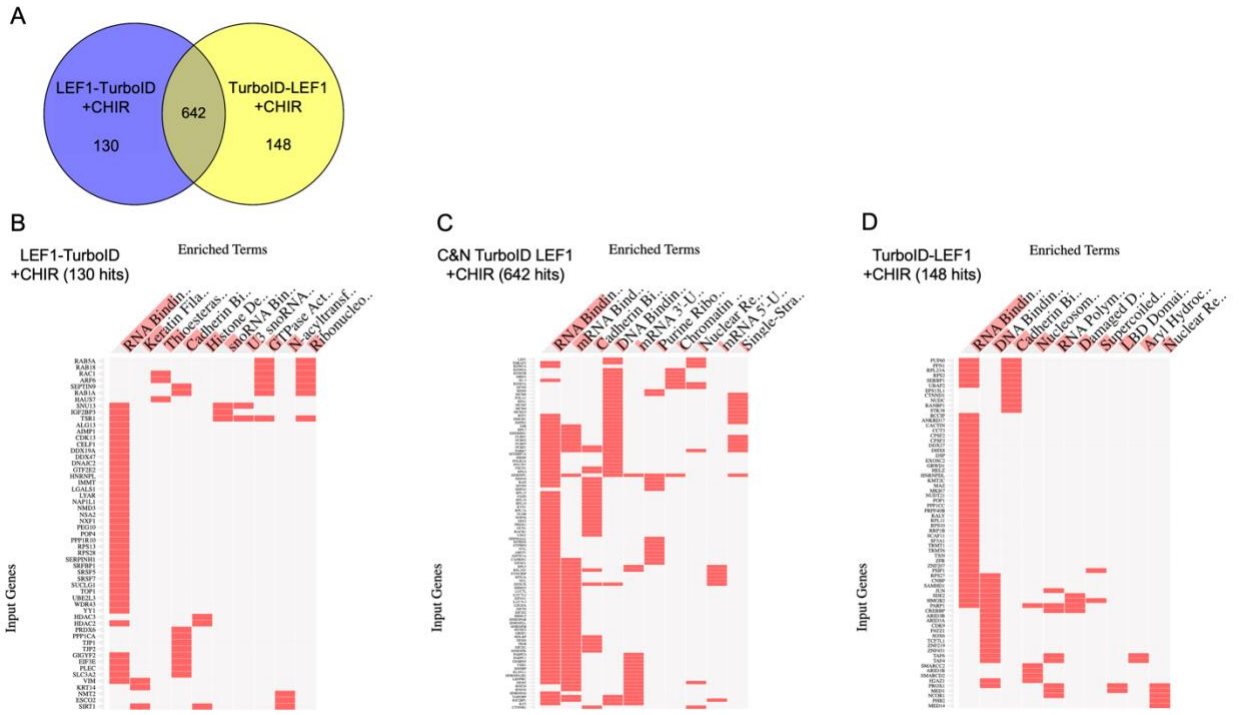

Figure 4-6. Differences in the Interactome of TurboID Fusion to C-terminus versus N-Terminus of FL-LEF1 in the Presence of CHIR.
(A) Venn diagram of interactors of C and N terminal FL-LEF1 TurboID fusion proteins in the presence of CHIR. Blue circle represents the number of LEF1-TurboID protein interactors, the yellow circle representing the TurboID-LEF1 protein interaction and the overlapping region representing the proteins interacting with both C- and N-terminus FL-LEF1 TurboID fusions. ***(B, C, and D)*** GO Molecular Function enrichment analysis (Enrichr) of high confidence unique and common proteins interacting with C- and N-terminus FL-LEF1 TurboID fusions. Top 10 enriched terms are columns ranked by enrichment score (P-value; top ten ≤ 0.001), and the input genes enriched in top 10 terms are the rows. Rows in C shows only 100 top rows.

4.2.4. FL-LEF1-Specific Interactors in the Presence of CHIR

As the number of FL-LEF1 interactors in the presence of CHIR is high (920 hits), an overview of the FL-LEF1 interactors is represented in a network generated by Cytoscape (3.10.0) (***Figure 4-7***). We focused on transcription factors and cofactors, epigenetic regulators, and post-translational modifiers (***Figure 4-8***). Both C- and N-terminus FL-LEF1 TurboID fusions identified Wnt enhanceosome components, such as Ep300, Bcl9, and Hdac1. β-catenin is a well-established

LEF1-associated protein.[166] In our TurboID assay, β-catenin is among the most abundant biotinylated proteins detected in the presence of CHIR and only detected as a proximal protein to FL-LEF1. This helps validate our experimental system.

Transcription factors interacting only with FL-LEF1 (both C- and N-terminus FL-LEF1 TurboID fusion protein) with high confidence, such as ZNF462, TCF20, SALL4, and GATAD2B, are shown in ***Figure 4-8***. An interesting NLEF1 (TurboID-LEF1) and DNLef1 (ΔN-LEF1) interactor detected is Tcf7l1, with higher relative abundance detected with ΔN-LEF1 as bait. By contrast, Tcf7 has a higher relative abundance in FL-LEF1 versus ΔN-LEF1 bait samples, and its interaction with ΔN-LEF1 has much lower abundance and confidence (BFDR ≤ 0.05).

Other important preys identified as FL-LEF1 interactors in our data are epigenetic regulators such as DNA methyltransferases (DNMT1 and DNMT3L), histone methyltransferases (SHMT1, SHMT2, and EHMT1), histone lysine demethylases (Kdmt1a, Kdmt1b, Kdmt3a, Kdmt3b, and Kdmt6a), histone acetyltransferases (EP300 and ELP3), and histone deacetylases (HDAC1, HDAC2, and HDAC3). Of these epigenetic regulators, only EP300 interacts with ΔN-LEF1 with high confidence.

Post-translational modifiers interacting with FL-LEF1 include kinases (BAZ1B, PKM, HK2, and CHEK1), phosphatases (PPP2R1A and PPP2CB), and ubiquitin-protein ligases (UBR5, UBR4, CUL4B, NEDD4, CAND1, UBA2, UHRF1, TRIM33, TRIM24, and TRIM72), and one deubiquitylase OTUD4. Tripartite Motif-containing 33 (TRIM33) is the only ubiquitin-protein ligase detected with both FL-LEF1 and ΔN-LEF1.

We are particularly interested in two chromatin remodelling complexes: the BAF complex and the NuRD complex. Some components of BAF complex, such as SMARCE1, SMARCD1, and

BRD7, are among the high confident hits interacting with FL-LEF1, while almost all components of NuRD complex are identified as high confident Fl-LEF1 interactors.

Three of the BAF complex components detected in our TurboID data (the scaffold subunit SMARCC2 and two transcription factors binding subunits SMARCE1 and SMARCD1) are among the eight subunits shared by both BAF and PBAF complexes. PBAF unique subunits identified in our TurboID data are the DNA-binding subunit ARID2 (interacting with both FL-LEF1 and ΔN-LEF1) and two histone tail-binding subunits PBRM1 and BRD7 (both only interacting with FL-LEF1). We detected MTA2, MTA3, MBD3, HDAC1/2, RBBP4/7, SALL4 and GATAD2B as significant interactors of FL- LEF1, while only MBD3 and GATA2A interact with ΔN-LEF1 (***Figure 4-8***).

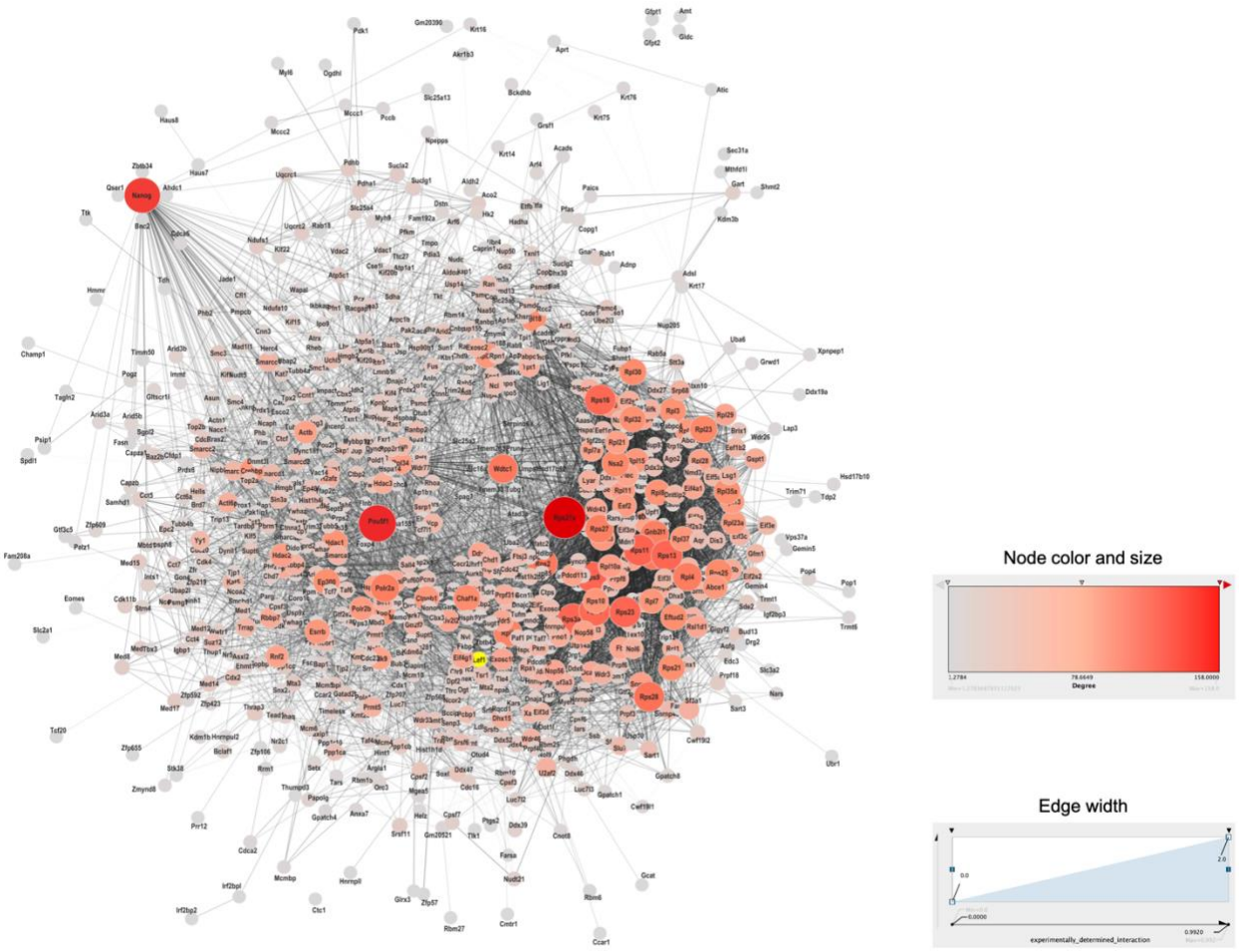

Figure 4-7. Network Representation of FL-LEF1 Protein Interactors in the Presence of CHIR.

High confidence FL-LEF1 interactors (BFDR ≤ 0.01; 920 hits) were selected from our SAINT analysis and were filtered further with ProHits.viz. The gene list passing the ProHits.viz filter was mapped to the search tool for retrieval of interacting genes (STRING) to acquire a PPI network. The resulting network was filtered further for physical subnetworks and network edges based on confidence (minimal required interaction score (confidence): 0.4. The network was clustered by K-means clustering to 25 clusters (functional clusters). A list of 844 genes was visualized by Cytoscape. Connectivity significance assessed by STRING was $p < 0.1e\text{-}16$. Node size and coloring is based on degree (i.e., the number of connections of a node to other nodes in the network); degree is determined by three interaction sources: textmining, experiments, and databases. The minimum degree is 1.28 (grey and small circles), medium degree 78.66 (orange and medium size circles), and maximum 158 (red and large circles). Edge thickness (width) is based on the experimentally determined confidence in the connection between nodes (minimum confidence 0 and maximum confidence 0.99).

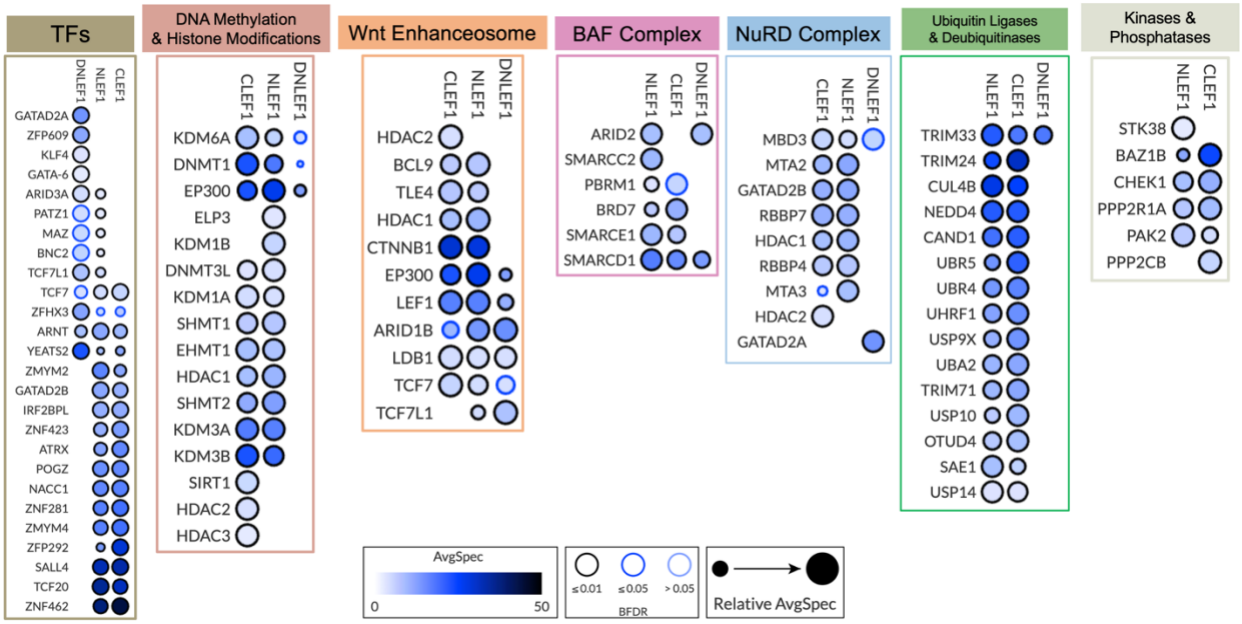

Figure 4-8. Dot Plot of Relative Average Spectral Count of Hits from TurboID in Differentiating mESCs in the Presence of CHIR.

Preys are grouped according to PANTHER (Pantherdb.org) and Enrichr gene ontology (GO) terms.

4.3. Aim 2: Determining the Genomic Occupancy of LEF-1 and ΔN-LEF1 *in vitro* in Differentiating mESCs

4.3.1. Generating *Lef1*$^{-/-}$ Knockout Clones in mESCs

To fully understand the function of FL-LEF1 and ΔN-LEF1, we need to characterize their genomic distribution. ΔN-LEF1 isoform is known as a dominant-negative (inhibitor), as it has been observed to oppose the function of FL-LEF1 in overexpression studies by supressing Wnt target gene activation.[91] Data about the differences in genomic distribution of LEF1 isoforms expressed at endogenous levels in a biologically relevant model are lacking. Furthermore, the results of our TurboID experiments presented in *Aim 1* determined that FL-LEF1 and ΔN-LEF1 interact with different proteins, and this could suggest that they interact with distinct genomic loci that place the LEF1 isoforms in the proximity of different proteins.

CRISPR/Cas9 methodology was used to generate full-length *Lef1*$^{-/-}$ knockout clones in mESCs with the aim of exploring the specific genomic distribution of the N-terminally truncated LEF1 isoform. As described in Material and Methods (*Section 3.5.4.*), WT mESCs were transfected with sgRNA-guided Cas9 endonuclease vector (PX459 V2.0) targeting the first exon of the *Lef1* gene. To obtain clonal populations from the resulting cells, a subsequent dilution plating method plus puromycin (final concentration: 2 µg/mL) was employed. A total of 30 clones were expanded, and protein extraction was conducted. Western blot analysis was then performed on all 30 clones to evaluate the abundance of LEF1 protein. Furthermore, DNA was extracted from 8 putative knockout clones, which were subsequently submitted for Sanger sequencing to identify potential allele-specific edits, such as nucleotide insertions or deletions.

4.3.2. Identifying *Lef1* Edits

To amplify and validate the successful amplification of the 961 bp region surrounding the target locus, PCR and DNA gel electrophoresis were employed (as described in Appendix B). The amplified PCR products were then sent for bidirectional Sanger sequencing by Eurofins Genomics LLC (Louisville, United States), and the obtained results were analyzed using Benchling's sequence alignments tool to determine the size and location of CRISPR/Cas9-induced modifications. Consequently, a total of three independent homozygous knockout clones, namely Lef1$^{-/-}$ A, Lef1$^{-/-}$ B, and Lef1$^{-/-}$ C, were generated and employed in CUT&RUN experiments. Lef1$^{-/-}$ A exhibited a 2 bp deletion, Lef1$^{-/-}$ B had a 4 bp deletion, and Lef1$^{-/-}$ C contained a 38 bp insertion (as shown in ***Figure 4-9C***). To predict the impact of frame-shift deletions and insertions on the resulting amino acid substitutions, the sequences were subjected to in silico translation using Benchling. The analysis aimed to identify potential premature stop codons (as shown in ***Figure 4-9D***), which aligned with the findings obtained from our western blot analyses (***Figure 4-9A***). Western blot data shows that FL-LEF1 protein is not detectable in the clones A, B, M, and C, whereas in WT mESCs both FL-LEF1 and ΔN-LEF1 are detectable. QKO clone, which lack all full- length TCF/LEFs, in ***Figure 4-9B*** was used as a positive control for ΔN-LEF1. Overall, these findings validated the suitability of Lef1$^{-/-}$ A, Lef1$^{-/-}$ B, and Lef1$^{-/-}$ C clones as appropriate models for investigating the genomic distribution of ΔN-LEF1 (in the absence of FL-LEF1) via the CUT&RUN assay.

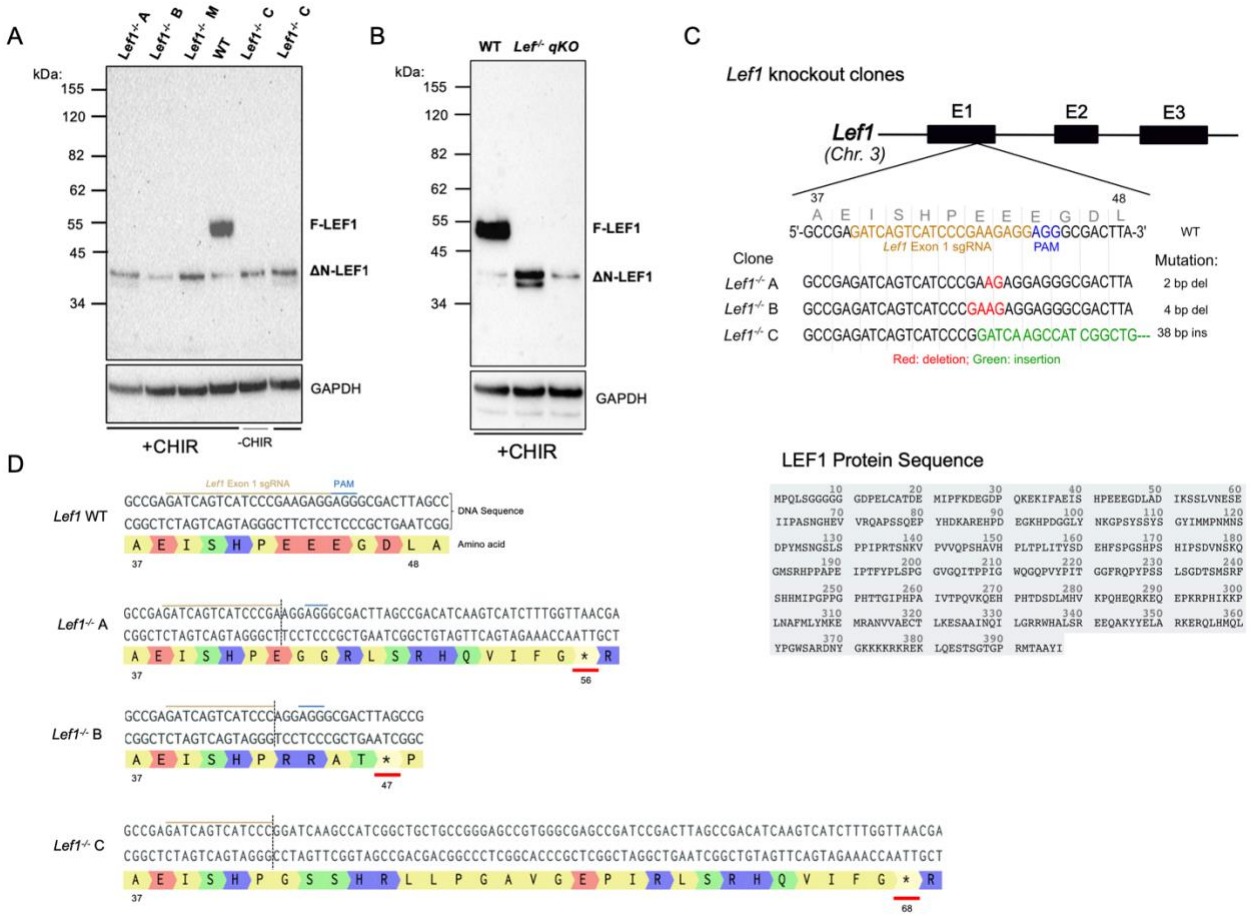

Figure 4-9. DNA Sequencing Analysis Determined *Lef1*$^{-/-}$ Knockout Clones in mESCs.
(A) Western bots illustrating the absence of FL-LEF1 expression relative to WT mESCs. GAPDH is the loading control. ***(B)*** Western blot analysis validating the anti-LEF1 antibody as being capable of detecting both FL-LEF1 and ΔN-LEF1 isoforms. qKO mESCs generated in Doble lab lack full-length TCF/LEF expression but express the ΔN-LEF1. qKO; quadruple knockout. ***(C)*** The DNA sequencing results of exon 1 of *Lef1* in three *Lef1*$^{-/-}$ clones were compared to the reference sequence (ENSMUSG00000027985). Deletions are represented in red, while insertions are shown in green. PAM; protospacer adjacent motif (AGG). ***(C)*** The resulting amino acid changes caused by the editing mediated by CRISPR/Cas9 technology. The numbers correspond to the amino acid positions in the LEF1 reference sequence (P27782-1; right sequence). All frameshift mutations introduce premature stop codons (* and red underline), which are predicted to trigger the degradation of mRNA though nonsense-mediated decay.

4.3.3. Bioanalyzer Profile of Amplified Total LEF1 and ΔN-LEF1 CUT&RUN Libraries

The CUT&RUN libraries of three biological replicates were generated, and their quality (fragment distribution) was assessed by a DNA fragment analyzer (Bioanalyzer DNA Analysis). Wild type (WT) mESCs express both FL-LEF1 and ΔN-LEF1 proteins (thus called "total LEF1"), while *Lef1*-KO (three biological $LEF1^{-/-}$ clones: A5, D5, and F1) mESCs do not express FL-LEF1 protein. Therefore, WT samples are for total LEF1 CUT&RUN libraries while $LEF1^{-/-}$ samples are for ΔN-LEF1 CUT&RUN libraries. The reason for having two LEF1-KO samples (A5 and F1) Bioanalyzed separately is that the i5.7 indexing primer tube provided by the library preparation kit manufacturer (Epicypher) was empty; therefore, these two samples were analyzed later (subsequent to two additional bead clean-ups after Indexing PCR) when new i5.7 primers were received. The bioanalyzer results are from samples obtained after after Indexing PCR and two bead cleanup steps to remove adaptor dimers (except for the input samples, which do not require the additional cleanup steps). Gel-like images are shown in ***Figure 4-10***, and the electropherograms are shown in ***Figure 4-11 A-D***.

Libraries were pooled based on calculations compatible with the Genome Québec sequencing facility, as mentioned in Material and Methods (*Section 3.6.7.*). The quantity and quality of the pooled library was checked via Qubit Fluorometer and Bioanalyzer, respectively, before sending for sequencing (***Figure 12A***). The quantity and the quality of the pooled library was checked again in the Genome Québec sequencing facility using qPCR and LabChip (more sensitive than Bioanalyzer), respectively (***Figure 12B***).

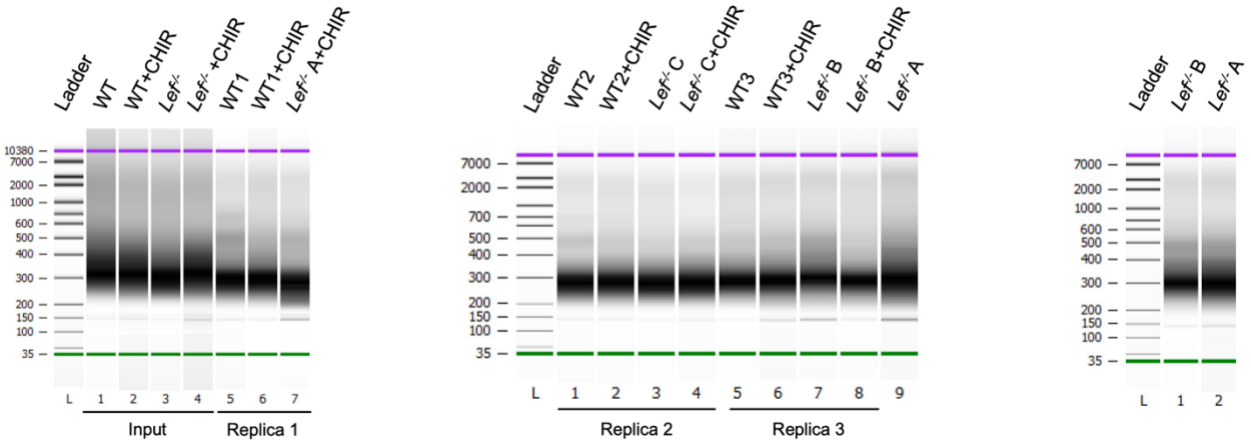

Figure 4-10. Gel-Like Fragment Analysis of CUT&RUN Libraries for FL-LEF1 and ΔN-LEF1 versus ΔN-LEF1.

Analysis was performed on 1 µL sample (~10 ng/µL), except *Lef1*$^{-/-}$ A and B (-CHIR) which were ~6 ng/µL due to two additional bead clean-up steps. Replica: replicate

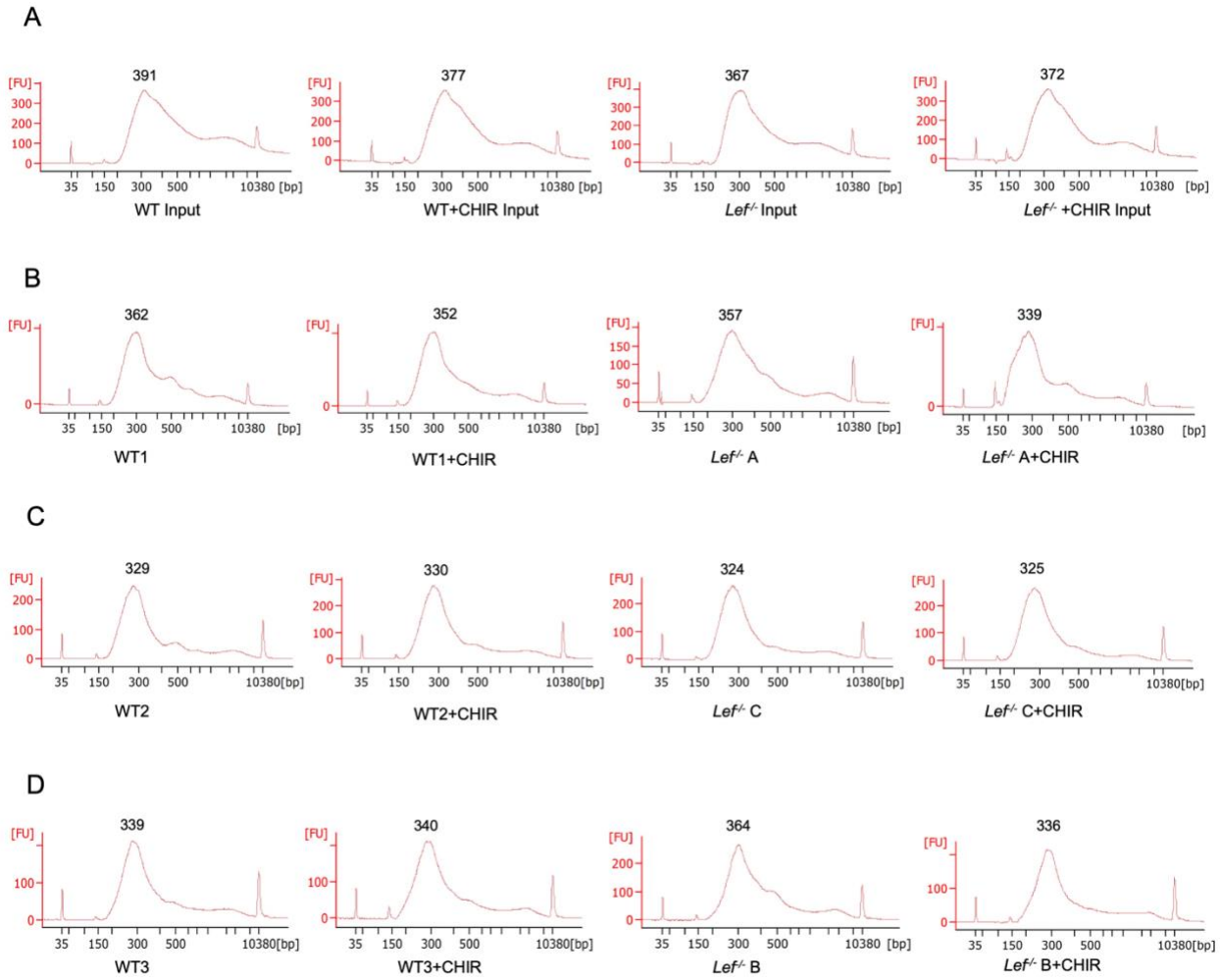

Figure 4-11. Electropherogram (Fragment Analysis) of Total LEF1 and ΔN-LEF1 CUT&RUN Libraries.

The average fragment size represented by the main middle peak in each electropherogram is based on the Bioanalyzer analysis. Markers (M): lower = 35 bp and upper =10380 bp. Y-axis shows the sample fluorescence intensity and X-axis shows fragment size (bp). Analysis was performed on 1 µL sample (~10 ng/µL), except *Lef1*$^{-/-}$ A and B (-CHIR) which were ~6 ng/µL due to two additional bead-cleanup steps. ***(A)*** Bioanalyzer DNA analysis of Input sample libraries. ***(B, C, and D)*** Bioanalyzer DNA analysis of three biological replicates of WT and *Lef1*$^{-/-}$ mESC CUT&RUN libraries. WT; wild type.

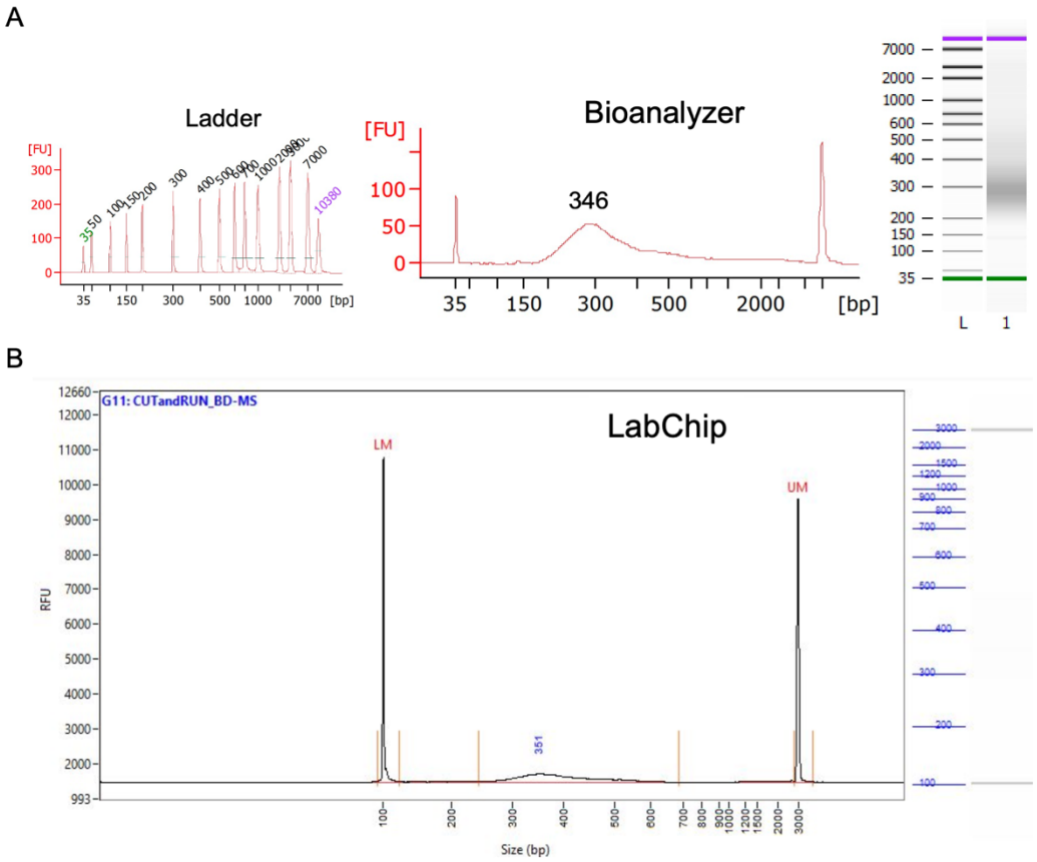

Figure 4-12. Fragment Analysis of the Pooled Library.
(A) 1 µL (10 ng/µL) of the ladder and pooled library were analysed via Bioanalyzer as shown in the electropherogram. A gel-like image of the ladder and the pooled library is shown on left. The average fragment size peak was 346 bp. ***(B)*** The quality of the pooled library was checked via LabChip DNA analysis. The average fragment size peak was 351 bp. Gel-like image of the ladder and the pooled library is shown on left.

4.3.4. LEF1 Genomic Occupancy in the Presence and Absence of CHIR

Our primary objective in using the CUT&RUN technique was to elucidate and differentiate the genome-wide distribution patterns of FL-LEF1 and ΔN-LEF1. This investigation aligns with our overarching goal of characterizing the functionality of these two isoforms. As detailed in the Materials and Methods section, we cultured WT and *Lef1* KO mESCs (where *Lef1* KO refers to the absence of FL-LEF1 and presence of ΔN-LEF1) in mEB medium, with and without the presence of CHIR. CHIR was administered to inhibit GSK-3 activity, thereby promoting β-catenin

stability and its translocation to the nucleus. While the CUT&RUN sequencing data requires further analysis, our preliminary findings indicate strong *Lef1* peaks upstream and downstream of the *Lef1* transcription start site in both WT and *Lef1* KO samples, regardless of the presence or absence of CHIR. These initial results bolster the reliability of the data obtained from the CUT&RUN experiment (***Figure 4-13A***). Intriguingly, our investigation revealed a noteworthy observation of LEF1 (both FL-LEF1 and ΔN-LEF1) binding to the *Zscan10* gene, primarily localized within the exon regions (***Figure 4-13B***).

Please note that the CUT&RUN sequencing results will undergo more in-depth scrutiny to extract additional insights. In our subsequent analyses, we aim to explore the differential and shared peaks in the genomic distribution of total LEF1 (FL-LEF1 and ΔN-LEF1 in WT samples) and ΔN-LEF1 (in *Lef1* KO samples) in both Wnt/β-catenin active and inactive state.

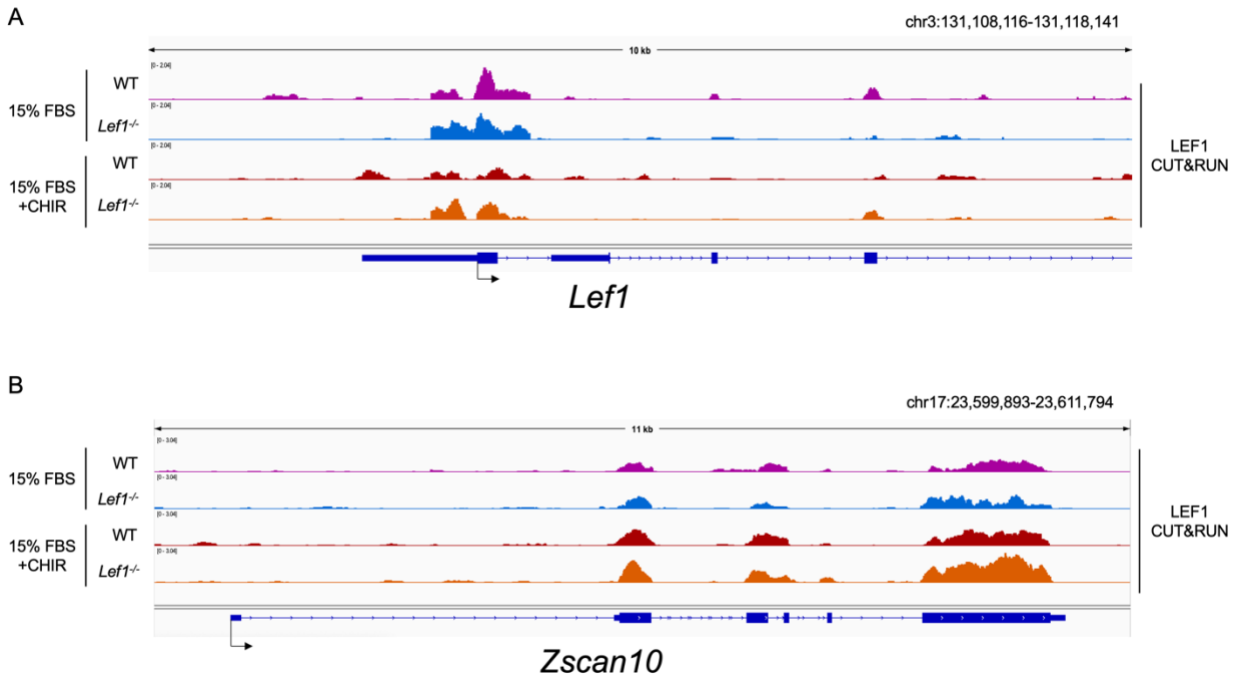

Figure 4-13. LEF1 Binding to *Lef1* and *Zscan10* Loci.
(A and B) CUT&RUN tracks displaying the binding of total LEF1 (in WT samples) and ΔN-LEF1 (in *Lef1* KO samples) under conditions with and without CHIR. The gene structure and transcription direction are indicated below the respective tracks. The track illustrates LEF1 binding activity within specific genomic regions, with subfigure ***A*** focusing on *Lef1* loci and subfigure ***B*** highlighting *Zscan10* loci.

CHAPTER 5. SUMMARY, DISCUSSION, CONCLUSION & SIGNIFICANCE

5.1. SUMMARY

To investigate the distinct molecular functions of FL-LEF1 and ΔN-LEF1 isoforms during mESC differentiation, two techniques were employed in the research described in this thesis. In *Aim 1,* the TurboID proximity-labeling technique was employed to detect proteins located in very close proximity to, and potentially interacting with, FL-LEF1 and ΔN-LEF1, whereas in *Aim 2*, the CUT&RUN technique was employed to map the genome-wide distribution of both FL-LEF1 and ΔN-LEF1 in WT mESCs and ΔN-LEF1 in FL-LEF1 KO mESCs. In our TurboID experiment, we not only identified well-established LEF1 interactors, which helped validate our data, but we also identified novel isoform-specific interactors of FL-LEF1 and ΔN-LEF1, such as NuRD complex and ZSCAN10. In the CUT&RUN data, which requires further analysis, we detected LEF1 binding to its own gene. Additionally, and surprisingly, we also detected binding of LEF1 to *Zscan10,* a protein that was identified as preferentially interacting with ΔN-LEF1 in our TurboID study.

5.2. DISCUSSION

5.2.1. Preys Interacting with LEF1 Both in the Presence and Absence of CHIR Exhibit Isoform Preference

We identified 12 proteins that interacted with LEF1 both in the presence and absence of CHIR conditions. However, these common interactors displayed a preference for different LEF1 isoforms (FL-LEF1 vs ΔN-LEF1) under these two conditions (-/+CHIR) (***Figure 4-4A***, page 63).

Among the 12 proteins interacting with FL-LEF1 and/or ΔN-LEF1 in the presence and absence of CHIR, certain preys, including ELAVL1, MYBBP1A, NOP58, RPL34, KMT2C, favored FL-LEF1 association in both conditions, whereas other preys, such as ARID2, JPT2, and RPS13 favored FL-LEF1 in the presence of CHIR and favored ΔN-LEF1 in the absence of CHIR (***Figure 4-4B***, page 63). ELAVL1, is an RNA-binding protein, known for its role in mRNA stabilization and its involvement in regulating β-catenin translation and Wnt/β-catenin-mediated transcriptional activity.[167,168] MYBBP1A, NOP58, and RPL34 are associated with ribosome biogenesis,[169,170] which might be linked to the selective and cell-type-specific translational regulation of developmental signaling networks during mESC differentiation. For instance, it has been demonstrated that RPL10A exerts its function on gastrulation and mesoderm formation via regulating translation of Wnt pathway components.[171,172] Furthermore, KMT2C, also known as MLL3, has been shown to play an essential role in regulating H3K4 mono-methylation at enhancers and tri-methylation at promoters.[173] ARID2, as a PBAF-specific subunit complex, is associated with chromatin regulation.[174] Given the changes in BAF complex composition in a temporal and lineage-specific manner in mESCs, detecting ARID2 as an isoform-specific LEF1 interactor suggests that ARID2 plays a functionally specific role in regulating the function of LEF1 isoforms and Wnt/β-catenin signaling pathway.[174-176] Additionally, ribosomal protein S13

(RPS13) was found to interact with E130012A19Rik (E13) as regulators of gene expression in mESCs.[177]

By contrast, KMT2D and TBX3 favored both FL-LEF1 and ΔN-LEF1 in the presence of CHIR, while only favoring FL-LEF1 in the absence of CHIR. TBX3, a member of the T-box family of transcription factors and a "tissue-specific component of Wnt/β-catenin enhanceosome", plays a role in sustaining mESC self-renewal and is activated by canonical Wnt signaling.[178] It also negatively regulates the transition of mESCs into mesodermal lineages by repressing genes associated with mesoderm and the Wnt pathway.[179,180] Furthermore, experimental evidence has revealed that TBX3 exerts direct repression on *Lef1* transcription by binding directly to the *Lef1* promoter.[181] Similarly, KMT2D, also known as MLL4, is a major mammalian H3K4 monomethyltransferase involved in lineage-selective gene expression during mESC differentiation. It predominantly localizes on enhancers and colocalizes with lineage-determining TFs on active enhancers during mESC differentiation.[182-184] Therefore, it appears that Wnt/β-catenin pathway activation alters TBX3 and KMT2D interactions with LEF1 isoforms, likely affecting TBX3, KMT2D, and LEF1 functions in mESCs.

Our results suggest that additional investigations should be conducted to explore specific functions of the identified interactors, such as ARID2, KMT2C, and TBX3, in modulating the activity of FL-LEF1 and ΔN-LEF1. This involves investigating these interactions in both their undifferentiated and differentiating states, or alternatively, comparing these associations as human or mouse embryonic stem cells undergo differentiation into each of the three germ layers. This could yield valuable insights into the intricate mechanisms that govern the biological functions of LEF1 isoforms.

5.2.2. Distinct Isoform-Specific LEF1 Preys Are Identified in the Absence of CHIR

We identified five preys interacting with LEF1 only in the absence of CHIR, with GNL2, PTBP1, and TLE3 exclusively favoring FL-LEF1, and ZSCAN10 and RPS18 favoring ΔN-LEF1 (***Figure 4-4B***, page 63). GNL2 is a nucleolar GTP-binding protein involved in ribosome biogenesis.[185] In our data, GNL2 is only interacting with LEF1-TurboID (CLEF1) and neither with TurboID-LEF1 (NLEF1) or ΔN-LEF1. PTBP1 protein, known as hnRNP I, is an RNA-binding protein involved in mRNA processing, including alternative splicing, in embryonic stem cells.[186-188] Consistent with the literature, TLE3 is interacting with LEF1 only in Wnt/β-catenin "Off" (the absence of CHIR) state, as in Wnt/β-catenin "On" state TLE3 is displaced by β-catenin.[189] Our data provides additional information detecting TLE3 isoform-specific interaction with FL-LEF1 in mESCs. This finding aligns with the role of LEF1 in repressing Wnt target genes in Wnt/β-catenin "Off" state.[190,191] By contrast, although ΔN-LEF1 is known to function as the dominant-negative LEF1 isoform inhibiting LEF1 function, in our TurboID experiment, ΔN-LEF1, surprisingly, is not interacting with any of TLE corepressor family proteins.[85,90,192] The literature indicates that the TLE/Groucho-binding sites on LEF1 are located in the context regulatory domain and HMG domain,[193] which are downstream of the region truncated in ΔN-LEF1, thus we were expecting to observe TLE as a significant interactor that would be detected by using ΔN-LEF1 as a bait. Our finding that this was not the case warrants further investigation.

Two unique ΔN-LEF1 interactors in the absence of CHIR are ZSCAN10 and RSP18. ZSCAN10, also known as ZFP206, is zinc finger transcription factor and an "embryonic stem-cell specific transcription factor" required for pluripotency maintenance. It has been demonstrated that both LEF1 and ZSCAN10 are upregulated upon mouse embryo implantation, however the levels of ZSCAN10 transcripts undergo a dramatic reduction during ES cell differentiation.[194-197] The

ZSCAN10 consensus DNA-binding motif is GCGCATGCGC, and it has been shown that ZSCAN10 is involved in regulating the expression of a handful of ESC-specific transcripts, such as *Thoc4, Tcstv1, eIF-1A,* and *Zscan4*.[196,198,199] Further, Jiang *et al.* showed that ZSCAN10 appears to be involved in regulating critical molecules, including β-catenin, in Wnt/β-catenin signaling in glioma.[200] RPS18 is another unique interactor specific to ΔN-LEF1 in the absence of CHIR.

The identification of exclusive interactors for each LEF1 isoform in the absence of CHIR, such as GNL2 favoring FL-LEF1 and ZSCAN10 favoring ΔN-LEF1, suggests LEF1 isoform-specific roles in ribosome biogenesis and pluripotency, respectively, providing new insights into the regulation of LEF1 functions in different cellular contexts. Therefore, it would offer valuable insight to confirm these interactions in other contexts, such as employing human embryonic stem cells. This includes investigating these interactions in both their undifferentiated and differentiating states or even comparing these associations as human embryonic stem cells are forced to differentiate toward each of the three germ layers. By doing so, a more specific biological function of these interactions could be attained.

5.2.3. Enrichment of RNA Binding, DNA Binding, and Chromatin Remodellers with TurboID-FL-LEF1 as Bait in the Presence of CHIR Suggests Potential RNA/LEF1/DNA Complex Formation

The dramatic increase in the number of proteins detected with TurboID fused to FL-LEF1 in the presence of CHIR (Wnt/β-catenin "On") underscores the huge impact of Wnt/β-catenin pathway activation on the LEF1 protein interaction network, with an approximate 50-fold increase in the number of network nodes detected. In the presence of CHIR (Wnt/β-catenin "On"), most interactors (over 95%) displayed isoform specificity, with FL-LEF1 interactions being favored by 94.8%, ΔN-LEF1 being favored by 1.4%, and approximately 2% being isoform-independent

(***Figure 4-5A and B***, page 66). The unique interactors of FL-LEF1 in the presence of CHIR were predominantly enriched in RNA- and mRNA-binding proteins, while the unique interactors of ΔN-LEF1 were primarily enriched in transcription factors or epigenetic regulators (***Figure 4-6B***, page 67). These observations suggest that FL-LEF1 may function in the regulation of RNA processing in addition to its DNA-binding role, particularly in the presence of CHIR. Therefore, further investigating the interaction of LEF1 with RNA-binding proteins and associated RNA molecules, focusing on potential differences between Wnt/β-catenin "On" and "Off" states, would be of interest.

The increase in FL-LEF1 interactors in the presence of CHIR condition could be attributed to the upregulation of pioneering TFs, such as OCT4 (also known as POUF5F1), and stabilization of β-catenin, thereby increasing chromatin accessibility and recruiting other proteins (including TFs, epigenetic regulators, and RNA-binding proteins), respectively.[201,202] It is likely that the increased number of proteins associated with FL-LEF1 are also, in part, due to non-Wnt/β-catenin pathway mediated effects of GSK-3 inhibition, as GSK-3-mediated phosphorylation is associated with destabilization of proteins other than β-catenin,[203] although previous studies from the Doble lab have shown that mESCs with complete GSK-3 ablation are largely rescued by knocking down β-catenin levels to those in WT mESCs with inactive Wnt/β-catenin signaling.[204] Furthermore, as revealed in the GO analysis for FL-LEF1 unique interactors in the presence of CHIR, we observed an enrichment of terms associated with RNA-binding and metabolism, DNA-binding, and chromatin modification. This finding implies that FL-LEF1 exhibits the capacity to interact concurrently with both RNA and DNA, resulting in the formation of a ternary RNA/LEF1/DNA complex. This conclusion is supported by previous experimental evidence demonstrating the indirect interaction between LEF1 and RNA *in vitro*.[205,206]

5.2.4. Additional Regulatory Mechanisms Contribute to Wnt Enhanceosome Activity

Our TurboID analyses revealed numerous novel interactors that are not in the current model of the Wnt enhanceosome in its active or inactive states. Apart from the BAF chromatin remodeling complex, which is the primary remodeler in the current model of the Wnt enhanceosome active state, we also identified components of other chromatin modifier complexes, such as the nucleosome remodeling and deacetylation complex (NuRD), and PBAF complex (***Figure 4-8***, page 71). GATAD2B and GATAD2A are considered mutually exclusive in the NuRD complex.[207] Intriguingly, in our data, GATAD2A is a unique interactor of ΔN-LEF1, while GATAD2B interacts with both C- and N-terminal TurboID fusions of FL-LEF1. BRG1, a member of the BAF complex, was not identified as an interactor with FL-LEF1 or ΔN-LEF1 in our data. On the other hand, MBD3, a member of the NuRD complex, interacted with all three LEF1 bait variants. BRG1 and MBD3 have been found to exert opposing regulatory effects on a shared set of genes by modulating the occupancy of nucleosomes at the gene promoters. This antagonistic chromatin regulation has been suggested as a common regulatory mechanism in ES cells.[208] Previous studies have demonstrated that BRG1 is involved in the activation of pluripotency genes in mESCs through LIF-STAT3 signaling pathway in response to the LIF signal.[209] Interestingly, this finding aligns with our experimental setup, as we cultured mESCs in mEB medium lacking LIF. Considering that all NuRD complex components, except GATAD2A, exclusively interact with FL-LEF1 and only in the presence of CHIR, it can be inferred that NuRD complex recruitment to the Wnt target genes and its interaction with FL-LEF1 is β-catenin dependent. This also implies that GATAD2A is an isoform-specific LEF1 interactor (only interacting with ΔN-LEF1) in the Wnt/β-catenin "On" state which might have an NuRD-independent function in the context of our study. The NuRD-independent activity of GATAD2A has been suggested and supported by

previous studies evaluating the effect of *Gatad2a* KO compared to other NuRD components, such as *Mbd3*, on embryonic development.[210-212]

Only three SWI/SNF complex components, a transcription factor-binding subunit SMARCD1 and two DNA-binding subunit ARID1B and ARID2 are interacting with ΔN-LEF1 in the presence of CHIR. This would suggest that some SWI/SNF components possess differential affinity for LEF1 isoforms in mESCs. In this study, we observed that SWI/SNF components interacted with FL-LEF1 and ΔN-LEF1 only in the presence of CHIR. Notably, we detected unique components of the PBAF complex, such as ARID2, BRD7, and PBRM1, which suggests that the PBAF complex interacts with LEF1 isoforms either simultaneously or independently of the BAF complex.

Another interesting finding is the enrichment of TLE4 interacting with FL-LEF1 in the presence of CHIR, while TLE3 was enriched as a unique FL-LEF1 interactor in the absence of CHIR. Not finding other TLE proteins, such as TLE1, is consistent with previous findings observing that, for instance, TLE1 does not bind to LEF1.[17] TLE3 and TLE4, which are abundant TLE proteins in both undifferentiated and differentiating mESCs, play a crucial role in mediating the process of mESCs differentiation.[213] In differentiating mESCs, the expression of TLE4 has been observed to increase upon LIF removal, while TLE3 exhibits high levels both in the presence and absence of LIF.[213] Therefore, based on our TurboID data, it is evident that there is a discernible switch or change in the TLE component (TLE4 and/or TLE3) of the Wnt enhanceosome between the presence and absence of CHIR.

Moreover, detecting various well-established and novel enzymes involved in histone and DNA modification, including CBP/EP300, KMD3, EHMT1, HDAC1, and DNMT1, in our TurboID data

serves as further evidence supporting the involvement of these enzymes, at least CBP/EP300 and HDAC1, in helping β-catenin to modify chromatin accessibility in response to CHIR.[140,214]

Our data suggest a more dynamic interplay of proteins involved in the Wnt enhanceosome than is currently acknowledged. This could arise from changes in the Wnt enhanceosome composition and/or conformation; although it has been suggested, by Melissa & Bienz, 2018, that "the composition of the Wnt enhanceosome does not change upon Wnt signaling; rather, the components of this complex undergo a conformational rearrangement upon binding of β-catenin."[6]

5.2.5. Post Translational Modifiers Interacting Specifically with FL-LEF1 in the Presence of CHIR

LEF1 protein levels and activity are regulated through diverse post-translational modifications, such as phosphorylation and ubiquitination. LEF1 has been found to interact with multiple E3 ubiquitin ligases, which facilitate its ubiquitination and subsequent degradation.[215] One of the ubiquitin ligases identified as a high confident FL-LEF1 interactor in the presence of CHIR is NEDD4 (***Figure 4-8***, page 71), which is an E3-ubiquitin ligase that has been shown to negatively regulate Wnt/β-catenin signaling in intestinal stem cells.[216] Furthermore, it has been demonstrated that *Nedd4* knockout increased LEF1 (full-length) expression in intestinal tumors.[217] It would be informative to determine the impact of *Nedd4* knockdown on LEF1 levels in mESCs. There is the possibility that LEF1 is a direct target of NEDD4 ubiquitylation, thus affecting the LEF1 stability and protein level, which would be an exciting new finding. Additionally, to find UBR5 as a FL-LEF1 interactor only in the presence of CHIR validates the involvement of UBR5 in the ubiquitination of TLE3 induced by the activation of Wnt/β-catenin signaling.[141,190]

In the presence of CHIR, we detected three members of the tripartite motif (TRIM) proteins with the various LEF1-TurboID baits. Two of these proteins, TRIM24 and TRIM71, only interact

with FL-LEF1. TRIM24 is known as a transcriptional co-regulator exerting its influence on mESCs by modulating the expression of genes involved in cell proliferation and differentiation processes.[218] TRIM71, known as a "stem cell-specific protein", is strongly expressed during the initial stages of mouse embryonic development. Furthermore, mESCs are known to be TRIM71 positive.[219,220] An interesting ubiquitin ligase detected proximal to both FL-LEF1 and ΔN-LEF1 in our TurboID data is TRIM33, which is a chromatin reader playing a crucial role in mammalian mesendodermal differentiation via regulating signaling pathways such as Nodal and Wnt/β-catenin.[221,222] An interaction between TRIM33 and β-catenin has already been reported in mESCs.[221]

Our TurboID screen also detected the OTU deubiquitinase (OTUD) 4, a member of the ovarian tumour (OTU) deubiquitinase family, which has been shown to regulate pathways such as transforming growth factor-β via both catalytic and noncatalytic activities.[223,224] In a recent study, it was discovered that OTUD7B,[225] another member of OTU family, interacts with LEF1 and enhances its transcriptional activity. However, intriguingly, OTUD7B does not deubiquitinate LEF1 or contribute to its stabilization. Instead, it facilitates the nuclear translocation of LEF1.[225] These findings suggest that further exploring the potential involvement of the OTU-family members in the regulation of LEF1 expression and function is warranted.

We also identified other deubiquitinases, such as the universal stress protein (USP) 10, USP9X, and USP14, which are known to have diverse functions in ESCs, contributing to pluripotency maintenance, differentiation, and genomic stability.[226] For instance, direct interaction between USP9X and transcription factors such as ESRRB, SALL4, and SOX2, have been demonstrated in mESCs, playing crucial role in regulating the delicate balance between pluripotency and the differentiation potential of mESCs.[227] Moreover, it has been revealed that in human breast cancer

cells, USP9X mediates the deubiquitylation of DVL2 and BCL9, which is a crucial step necessary for the activation of Wnt/β-catenin signaling pathway.[228,229]

These findings provide a sound rationale for pursuing comprehensive studies to elucidate the functional role of both the identified ligases (such as NEDD4, TRIM24, TRIM33, and UBR5) and deubiquitinases (OTUD4, USP10, USP9X, and USP14) in modulating LEF1 stability, activity, and its impact on downstream Wnt/β-catenin signaling across different biological contexts.

Regarding the kinases identified in our TurboID screen, we found BAZ1B, which is a tyrosine-protein kinase, interacting with FL-LEF1 only in the presence of CHIR. BAZ1B, also known as Williams syndrome transcription factor, has been shown to be involved in cell fate determination and EMT via chromatin remodeling in conjunction with SMARCA1/5 (ISWI or Imitation Switch proteins).[230,231] Among the phosphatases identified with high confidence as being proximal to FL-LEF, PPP2R1A stands out as an intriguing enzyme due to its role as a scaffolding subunit within PP2A complex, a large multi-subunit assembly. Research has shown that PP2A plays a crucial role in Wnt signaling by dephosphorylating β-catenin, thereby facilitating its stabilization, activation, and translocation to the nucleus.[232] Interestingly, PPP2R1A has been shown to be responsible for gastrulation defects and impaired embryonic patterning due to the disruption of WNT and Nodal signaling pathways in mice carrying the t^{w18} homozygous-lethal *t* haplotype, one of four "*t* haplotypes" found on mouse chromosome 17.[233,234]

5.2.6. LEF1 Binds to *Lef1* and *Zscan10* Gene Loci

Our LEF1 CUT&RUN sequencing data validates the binding of both total LEF1 (including FL-LEF1 and ΔN-LEF1) and ΔN-LEF1 to *Lef1* and *Zscan10* loci (***Figure 4-13***; page 80). In a study that employed CUT&RUN to map the genomic binding patterns of LEF1 and CTNNB1 in chicken neural crest cells, researchers observed a significantly higher number of binding events for

CTNNB1 (9,892) compared to LEF1 (1,878), indicating approximately five times more binding occurrences. Moreover, they identified 1,398 regions where both LEF1 and CTNNB1 were found to co-occupy the genome. Notably, the majority of these regions were located in intergenic regions, suggesting their potential as active enhancers.[235] A study conducted by Feder et. al (2020) examined the genomic occupancy of two isoforms of LEF1 (FL-LEF1 and ΔN-LEF1) in both normal mice and those with acute myeloid leukemia using ChIP-seq.[236] Through their analysis, they discovered distinct ΔN-LEF1 DNA-binding sites that were enriched for proteins associated with stemness, including OCT4-SOX2-TCF-NANOG, and CDX2. These findings align with the observed differences in ΔN-LEF1 expression in normal hematopoietic stem cells compared to the expression of FL-LEF1 in leukemic stem cells, which contribute to the development of acute myeloid leukemia.[236] These findings are consistent with our hypothesis that ΔN-LEF1 is involved in regulating self-renewal and preserving pluripotency in mESCs. Furthermore, the identification of ZSCAN10 as an interacting partner of ΔN-LEF1 in the absence of CHIR (Wnt/ β-catenin "Off"), whose gene is also targeted by LEF1 isoforms, provides further support for our hypothesis.

5.3. FUTURE DIRECTIONS

In this study, we identified numerous novel FL-LEF1 and ΔN-LEF1 protein interactors, including ZSCAN10, and detected the *Zscan10* gene as a new LEF1 target. We also evaluated the effect of Wnt/β-catenin signaling pathway on LEF1 isoform-specific PPIs and their genomic occupancy. Some of these novel protein interactors, such as NuRD complex components, have the potential to be involved in the Wnt enhanceosome, thereby have roles in Wnt target gene regulation in the context of differentiating mESCs. These findings provide new insights into our understanding of the complexity of underlying mechanisms through which LEF1 isoforms

function in regulating gene expression and warrants new studies characterizing the function of other TCF/LEF isoforms.

5.3.1. Confirming FL-LEF1 Interaction with NuRD Complex Components

The NuRD complex plays a pivotal role in governing pluripotency and priming mESCs for differentiation. We identified almost all components of the NuRD complex interacting with FL-LEF1 in the presence of CHIR, which was also found associated with TCF7L1 in the absence of CHIR in a BioID study conducted by former PhD student, Dr. Steven Moreira, in the Doble lab.[237]

A study investigating the heterogeneity within wild-type ES cells cultured in standard ES medium revealed that the NuRD complex is responsible for this variation, and NuRD-mutant ES cells lost their capability for spontaneous differentiation.[122] Furthermore, it has been demonstrated that the NuRD complex regulates nucleosome positioning across regulatory elements, thereby regulating the accessibility of DNA to DNA-binding proteins. This process leads to the transcriptional activation of transcriptionally inactive genes.[238,239]

Considering the important role of the NuRD complex in ESC biology and that the Doble lab has detected the NuRD complex as a high confident LEF1 and TCF7L1 interactor in the presence and absence of CHIR, respectively, warrant further investigation to confirm their interaction via co-immunoprecipitation (coIP) and/or reciprocal TurboID to support their putative interaction.

5.3.2. Confirming ZSCAN10 Interaction with LEF1

ZSCAN10 has received limited attention to date, despite playing a crucial role in mESCs pluripotency and cell fate determination.[195,240,241] It has been shown that ZSCAN10 physically interacts with other pluripotency-regulating transcription factors such as OCT4 and SOX2, and directly regulates *Oc4* expression via binding to *Oct4* gene in mESCs.[242] While ZSCAN10 exhibits widespread expression during embryogenesis, it is characterized by relatively low levels in certain

adult human tissues, such as cerebellum, pituitary gland, skin, bone marrow, spleen, and skeletal muscle,[243] as can be observed in a search conducted in the human protein atlas (https://www.proteinatlas.org/).

The identification of ZSCAN10 as a high-confidence TurboID hit, specifically with ΔN-LEF1 bait in the absence of CHIR, suggests that ΔN-LEF1 may play a unique role in maintaining mESCs self-renewal via its interaction with proteins such as ZSCAN10. Moreover, the detection of *Zscan10* as a target of LEF1 in our CUT&RUN data strengthen the possibility of a functional relationship between LEF1 and ZSCAN10. Therefore, it is important to explore the physical interaction between LEF1 isoforms and ZSCAN10 via reciprocal TurboID (ZSCAN10-TurboID) as well as coIP. Conducting coIP experiments with overexpressed FL-LEF1 and ZSCAN10 in one experiment, and ΔN-LEF1 and ZSCAN10 in another experiment, can help to confirm whether ZSCAN10 exhibits distinct interactions with FL-LEF1 and ΔN-LEF1.

5.4. CONCLUSION AND SIGNIFICANCE

In this study, our findings have provided valuable insights into known and putative protein interactors of LEF1 isoforms and their genomic occupancy, shedding light on the intricate mechanisms underlying LEF1-mediated gene regulation and its role in mESC fate determination. We not only validated the interaction of LEF1 with well-established interactors, but also identified novel isoform-specific interactors, such as ZSCAN10. Furthermore, our findings offer a comprehensive and dynamic perspective on Wnt enhanceosome interactors. These interactors encompass a range of functional categories, including more temporally sustained interactors like the NuRD complex, as well as transient interactors such as ubiquitin ligases, deubiquitinase, kinases and phosphatases, and histone and DNA modifiers. Notably, the enrichment of RNA-binding proteins and chromatin modifiers among FL-LEF1 interactors in the presence of CHIR

indicates the potential formation of RNA/LEF1/DNA complexes and underscores the impact of Wnt/β-catenin signaling on LEF1 function. The CUT&RUN data confirmed the binding of LEF1 isoforms to the *Lef1* gene and intriguingly revealed an interaction with *Zscan10*. These findings highlight the potential for discovering additional novel LEF1 targets through further analysis of our LEF1 CUT&RUN data, which could unveil new insights into the regulatory mechanisms governed by LEF1 isoforms in mESCs.

Overall, this research lays the foundation for a broad spectrum of future studies that will delve deeper into the precise role of LEF1 isoforms, their interactors, and their impact on Wnt target gene regulation. These findings may ultimately lead to the identification of novel targets and the elucidation of additional regulatory mechanisms involved in LEF1 function. There are numerous links between LEF1 and various human diseases, most notably cancer, and a better understanding of LEF1 regulation and function has the potential to reveal new strategies to treat LEF1-associated diseases.

CHAPTER 6. REFERENCES

1 Miller, J. R. The wnts. *Genome biology* **3**, 1-15 (2001).

2 Wiese Katrin, E. & Roel, N. van Amerongen Renée. Wnt signalling: conquering complexity. *Development* **145**, 12 (2018).

3 Schenkelaars, Q. *et al.* Animal multicellularity and polarity without Wnt signaling. *Sci Rep* **7**, 15383 (2017).

4 Routledge, D. & Scholpp, S. Mechanisms of intercellular Wnt transport. *Development* **146**, dev176073 (2019).

5 Shah, K. & Kazi, J. U. Phosphorylation-dependent regulation of WNT/Beta-catenin signaling. *Frontiers in Oncology* **12**, 858782 (2022).

6 Gammons, M. & Bienz, M. Multiprotein complexes governing Wnt signal transduction. *Current opinion in cell biology* **51**, 42-49 (2018).

7 Brock, R. The uptake of arginine-rich cell-penetrating peptides: putting the puzzle together. *Bioconjugate chemistry* **25**, 863-868 (2014).

8 Ko, S.-B. *et al.* Functional role of the Frizzled linker domain in the Wnt signaling pathway. *Communications Biology* **5**, 421 (2022).

9 Chen, W. *et al.* Dishevelled 2 recruits ß-arrestin 2 to mediate Wnt5A-stimulated endocytosis of Frizzled 4. *Science* **301**, 1391-1394 (2003).

10 Clevers, H. Wnt/β-catenin signaling in development and disease. *Cell* **127**, 469-480 (2006).

11 Kang, K., Shi, Q., Wang, X. & Chen, Y.-G. Dishevelled phase separation promotes Wnt signalosome assembly and destruction complex disassembly. *Journal of Cell Biology* **221**, e202205069 (2022).

12 MacDonald, B. K, T., & X, H.(2009). Wnt/beta-catenin signaling: components, mechanisms, and diseases. *Developmental cell* **17**

13 Ye, B., Li, L., Xu, H., Chen, Y. & Li, F. Opposing roles of TCF7/LEF1 and TCF7L2 in cyclin D2 and Bmp4 expression and cardiomyocyte cell cycle control during late heart development. *Laboratory Investigation* **99**, 807-818 (2019).

14 Klingel, S. *et al.* Subfunctionalization and neofunctionalization of vertebrate Lef/Tcf transcription factors. *Developmental biology* **368**, 44-53 (2012).

15 Arce, L., Yokoyama, N. & Waterman, M. Diversity of LEF/TCF action in development and disease. *Oncogene* **25**, 7492-7504 (2006).

16 Mao, C. D. & Byers, S. W. Cell-context dependent TCF/LEF expression and function: alternative tales of repression, de-repression and activation potentials. *Critical Reviews™ in Eukaryotic Gene Expression* **21** (2011).

17 Chodaparambil, J. V. *et al.* Molecular functions of the TLE tetramerization domain in W nt target gene repression. *The EMBO Journal* **33**, 719-731 (2014).

18 Moreira, S. *et al.* A single TCF transcription factor, regardless of its activation capacity, is sufficient for effective trilineage differentiation of ESCs. *Cell reports* **20**, 2424-2438 (2017).

19 Shy, B. R. *et al.* Regulation of Tcf7l1 DNA binding and protein stability as principal mechanisms of Wnt/β-catenin signaling. *Cell reports* **4**, 1-9 (2013).

20 Morrison, G., Scognamiglio, R., Trumpp, A. & Smith, A. Convergence of cMyc and β-catenin on Tcf7l1 enables endoderm specification. *The EMBO journal* **35**, 356-368 (2016).

21 Kikuchi, A., Kishida, S. & Yamamoto, H. Regulation of Wnt signaling by protein-protein interaction and post-translational modifications. *Experimental & molecular medicine* **38**, 1-10 (2006).
22 Hrckulak, D., Kolar, M., Strnad, H. & Korinek, V. TCF/LEF transcription factors: an update from the internet resources. *Cancers* **8**, 70 (2016).
23 Wang, C. *et al.* pVHL regulates protein stability of TCF/LEF transcription factor family via ubiquitin-independent proteasomal degradation. *bioRxiv*, 2021.2010. 2014.464355 (2021).
24 Han, W., Lee, H. & Han, J.-K. Ubiquitin C-terminal hydrolase37 regulates Tcf7 DNA binding for the activation of Wnt signalling. *Sci Rep* **7**, 1-13 (2017).
25 Roose, J. *et al.* Synergy between tumor suppressor APC and the β-catenin-Tcf4 target Tcf1. *Science* **285**, 1923-1926 (1999).
26 Koenig, S. F. *et al.* Autoregulation of XTcf-4 depends on a Lef/Tcf site on the XTcf-4 promoter. *Genesis* **46**, 81-86 (2008).
27 Zhu, Y., Wang, W. & Wang, X. Roles of transcriptional factor 7 in production of inflammatory factors for lung diseases. *Journal of translational medicine* **13**, 1-13 (2015).
28 Chhabra, S., Liu, L., Goh, R., Kong, X. & Warmflash, A. Dissecting the dynamics of signaling events in the BMP, WNT, and NODAL cascade during self-organized fate patterning in human gastruloids. *PLoS Biology* **17**, e3000498 (2019).
29 Pond, K. W., Doubrovinski, K. & Thorne, C. A. Wnt/β-catenin signaling in tissue self-organization. *Genes* **11**, 939 (2020).
30 Brons, I. G. M. *et al.* Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**, 191-195 (2007).
31 Czechanski, A. *et al.* Derivation and characterization of mouse embryonic stem cells from permissive and nonpermissive strains. *Nature protocols* **9**, 559-574 (2014).
32 Boroviak, T., Loos, R., Bertone, P., Smith, A. & Nichols, J. The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. *Nature cell biology* **16**, 513-525 (2014).
33 Wang, X. *et al.* Formative pluripotent stem cells show features of epiblast cells poised for gastrulation. *Cell research* **31**, 526-541 (2021).
34 Kinoshita, M. *et al.* Capture of mouse and human stem cells with features of formative pluripotency. *Cell stem cell* **28**, 453-471. e458 (2021).
35 Rivera-Pérez, J. A. & Hadjantonakis, A.-K. The dynamics of morphogenesis in the early mouse embryo. *Cold Spring Harbor perspectives in biology* **7**, a015867 (2015).
36 Brennan, J. *et al.* Nodal signalling in the epiblast patterns the early mouse embryo. *Nature* **411**, 965-969 (2001).
37 Ben-Haim, N. *et al.* The nodal precursor acting via activin receptors induces mesoderm by maintaining a source of its convertases and BMP4. *Developmental cell* **11**, 313-323 (2006).
38 Power, M.-A. & Tam, P. Onset of gastrulation, morphogenesis and somitogenesis in mouse embryos displaying compensatory growth. *Anatomy and embryology* **187**, 493-504 (1993).
39 Probst, S. & Sagar, T. J., Schwan, C., Grün, D., and Arnold, SJ (2020). Spatiotemporal sequence of mesoderm and endoderm lineage segregation during mouse gastrulation. *Development* **148**
40 Sheng, G., Martinez Arias, A. & Sutherland, A. The primitive streak and cellular principles of building an amniote body through gastrulation. *Science* **374**, abg1727 (2021).

41 Shahbazi, M. N. & Zernicka-Goetz, M. Deconstructing and reconstructing the mouse and human early embryo. *Nature cell biology* **20**, 878-887 (2018).
42 Hamidi, S. *et al.* Mesenchymal-epithelial transition regulates initiation of pluripotency exit before gastrulation. *Development* **147**, dev184960 (2020).
43 Gadue, P., Huber, T. L., Nostro, M. C., Kattman, S. & Keller, G. M. Germ layer induction from embryonic stem cells. *Experimental hematology* **33**, 955-964 (2005).
44 Marconi, G. D. *et al.* Epithelial-mesenchymal transition (EMT): the type-2 EMT in wound healing, tissue regeneration and organ fibrosis. *Cells* **10**, 1587 (2021).
45 Choi, S. S. & Diehl, A. M. Epithelial-to-mesenchymal transitions in the liver. *Hepatology* **50**, 2007-2013 (2009).
46 Zhou, P. *et al.* The epithelial to mesenchymal transition (EMT) and cancer stem cells: implication for treatment resistance in pancreatic cancer. *Molecular cancer* **16**, 1-11 (2017).
47 Mittal, V. Epithelial mesenchymal transition in tumor metastasis. *Annual Review of Pathology: Mechanisms of Disease* **13**, 395-412 (2018).
48 Scheibner, K. *et al.* Epithelial cell plasticity drives endoderm formation during gastrulation. *Nature Cell Biology* **23**, 692-703 (2021).
49 Kobayashi, W. & Ozawa, M. The epithelial-mesenchymal transition induced by transcription factor LEF-1 is independent of β-catenin. *Biochemistry and Biophysics Reports* **15**, 13-18 (2018).
50 Aguilera-Castrejon, A. *et al.* Ex utero mouse embryogenesis from pre-gastrulation to late organogenesis. *Nature* **593**, 119-124 (2021).
51 Macklon, N. S., Geraedts, J. P. & Fauser, B. C. Conception to ongoing pregnancy: the 'black box'of early pregnancy loss. *Human reproduction update* **8**, 333-343 (2002).
52 ten Berge, D. *et al.* Wnt signaling mediates self-organization and axis formation in embryoid bodies. *Cell stem cell* **3**, 508-518 (2008).
53 Gadue, P., Huber, T. L., Paddison, P. J. & Keller, G. M. Wnt and TGF-β signaling are required for the induction of an in vitro model of primitive streak formation using embryonic stem cells. *Proceedings of the National Academy of Sciences* **103**, 16806-16811 (2006).
54 Hart, M. J., De Los Santos, R., Albert, I. N., Rubinfeld, B. & Polakis, P. Downregulation of β-catenin by human Axin and its association with the APC tumor suppressor, β-catenin and GSK3β. *Curr. Biol.* **8**, 573-581 (1998).
55 Patel, S., Doble, B. & Woodgett, J. Glycogen synthase kinase-3 in insulin and Wnt signalling: a double-edged sword? *Biochemical Society Transactions* **32**, 803-808 (2004).
56 Neal, J. W. & Clipstone, N. A. Glycogen synthase kinase-3 inhibits the DNA binding activity of NFATc. *Journal of Biological Chemistry* **276**, 3666-3673 (2001).
57 Price, M. A. & Kalderon, D. Proteolysis of the Hedgehog signaling effector Cubitus interruptus requires phosphorylation by Glycogen Synthase Kinase 3 and Casein Kinase 1. *Cell* **108**, 823-835 (2002).
58 Chidiac, R. & Angers, S. Wnt signaling in stem cells during development and cell lineage specification. *Current Topics in Developmental Biology* **153**, 121-143 (2023).
59 Singh, A. M., Bechard, M., Smith, K. & Dalton, S. Reconciling the different roles of Gsk3β in "naïve" and "primed" pluripotent stem cells. *Cell Cycle* **11**, 2991-2996 (2012).

60 Ye, S. *et al.* Pleiotropy of glycogen synthase kinase-3 inhibition by CHIR99021 promotes self-renewal of embryonic stem cells from refractory mouse strains. *PLoS One* **7**, e35892 (2012).
61 Yu, X. *et al.* Notch signaling activation in human embryonic stem cells is required for embryonic, but not trophoblastic, lineage commitment. *Cell Stem Cell* **2**, 461-471 (2008).
62 Williams, R. L. *et al.* Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature* **336**, 684-687 (1988).
63 Nicola, N. A. & Babon, J. J. Leukemia inhibitory factor (LIF). *Cytokine & growth factor reviews* **26**, 533-544 (2015).
64 Kunisada, K. Hirota H, Fujio Y, Matsui H, Tani Y, Yamauchi-Takihara K, Kishimoto T. *Activation of JAK-STAT and MAP kinases by leukemia inhibitory factor through gp130 in cardiac myocytes. Circulation* **94**, 2626-2632 (1996).
65 Cartwright, P. *et al.* LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism. (2005).
66 Savatier, P., Lapillonne, H., Van Grunsven, L., Rudkin, B. & Samarut, J. Withdrawal of differentiation inhibitory activity/leukemia inhibitory factor up-regulates D-type cyclins and cyclin-dependent kinase inhibitors in mouse embryonic stem cells. *Oncogene* **12**, 309-322 (1996).
67 Burdon, T., Smith, A. & Savatier, P. Signalling, cell cycle and pluripotency in embryonic stem cells. *Trends in cell biology* **12**, 432-438 (2002).
68 Burdon, T. Smith A, and Savatier P. *Signalling, cell cycle and pluripotency in embryonic stem cells. Trends Cell Biol* **12**, 432-438 (2002).
69 Aulicino, F. *et al.* Canonical wnt pathway controls mESC Self-renewal through inhibition of spontaneous differentiation via β-Catenin/TCF/LEF functions. *Stem Cell Reports* **15**, 646-661 (2020).
70 Kim, S. *et al.* The Distinct Role of Tcfs and Lef1 in the Self-Renewal or Differentiation of Mouse Embryonic Stem Cells. *International Journal of Stem Cells* **13**, 192-201 (2020). https://doi.org:10.15283/ijsc20044
71 Pereira, L., Yi, F. & Merrill, B. J. Repression of Nanog gene transcription by Tcf3 limits embryonic stem cell self-renewal. *Molecular and cellular biology* **26**, 7479-7491 (2006).
72 Ye, S. *et al.* Depletion of Tcf3 and Lef1 maintains mouse embryonic stem cell self-renewal. *Biology open* **6**, 511-517 (2017).
73 Athanasouli, P. *et al.* The Wnt/TCF7L1 transcriptional repressor axis drives primitive endoderm formation by antagonizing naive and formative pluripotency. *Nature Communications* **14**, 1210 (2023).
74 De Jaime-Soguero, A. *et al.* Wnt/Tcf1 pathway restricts embryonic stem cell cycle through activation of the Ink4/Arf locus. *PLoS Genetics* **13**, e1006682 (2017).
75 Liang, R. & Liu, Y. Tcf7l1 directly regulates cardiomyocyte differentiation in embryonic stem cells. *Stem Cell Research & Therapy* **9**, 1-7 (2018).
76 Merrill, B. J., Gat, U., DasGupta, R. & Fuchs, E. Tcf3 and Lef1 regulate lineage differentiation of multipotent stem cells in skin. *Genes & development* **15**, 1688-1705 (2001).
77 Chatterjee, S. S. *et al.* Inhibition of β-catenin–TCF1 interaction delays differentiation of mouse embryonic stem cells. *Journal of Cell Biology* **211**, 39-51 (2015).

78 Verbeek, S., Izon, D., Hofhuis, F. & Robanus-Maandag, E. to Riele H, van de Wetering M, Oosterwegel M, Wilson A, MacDonald HR, Clevers H. An HMG-box-containing T-cell factor required for thymocyte differentiation. *Nature* **374**, 70-74 (1995).
79 Schilham, M. W. *et al.* Critical involvement of Tcf-1 in expansion of thymocytes. *The Journal of Immunology* **161**, 3984-3991 (1998).
80 Yi, K. *et al.* Dynamic maintenance of asymmetric meiotic spindle position through Arp2/3-complex-driven cytoplasmic streaming in mouse oocytes. *Nature cell biology* **13**, 1252-1258 (2011).
81 Korinek, V. Barker N, Moerer P, van Donselaar E, Huls G, Peters PJ, Clevers H. *Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. Nat Genet* **19**, 379-383 (1998).
82 Huang, C. & Qin, D. Role of Lef1 in sustaining self-renewal in mouse embryonic stem cells. *Journal of Genetics and Genomics* **37**, 441-449 (2010).
83 van Genderen, C. *et al.* Development of several organs that require inductive epithelial-mesenchymal interactions is impaired in LEF-1-deficient mice. *Genes & development* **8**, 2691-2703 (1994).
84 Li, T. W.-H. *et al.* Wnt activation and alternative promoter repression of LEF1 in colon cancer. *Molecular and cellular biology* **26**, 5284-5299 (2006).
85 Hovanes, K. *et al.* β-catenin–sensitive isoforms of lymphoid enhancer factor-1 are selectively expressed in colon cancer. *Nature genetics* **28**, 53-57 (2001).
86 Jimenez, J., Jang, G. M., Semler, B. L. & Waterman, M. L. An internal ribosome entry site mediates translation of lymphoid enhancer factor-1. *Rna* **11**, 1385-1399 (2005).
87 Yuan, L., Roy, B., Ratna, P., Uhler, C. & Shivashankar, G. Lateral confined growth of cells activates Lef1 dependent pathways to regulate cell-state transitions. *Sci Rep* **12**, 17318 (2022).
88 Ayoubi, T. A. & Van De Yen, W. J. Regulation of gene expression by alternative promoters. *The FASEB Journal* **10**, 453-460 (1996).
89 Willinger, T. *et al.* Human naive CD8 T cells down-regulate expression of the WNT pathway transcription factors lymphoid enhancer binding factor 1 and transcription factor 7 (T cell factor-1) following antigen encounter in vitro and in vivo. *The Journal of Immunology* **176**, 1439-1446 (2006).
90 Hovanes, K., Li, T. & Waterman, M. The human LEF-1 gene contains a promoter preferentially active in lymphocytes and encodes multiple isoforms derived from alternative splicing. *Nucleic acids research* **28**, 1994-2003 (2000).
91 Yokoyama, N. N., Pate, K. T., Sprowl, S. & Waterman, M. L. A role for YY1 in repression of dominant negative LEF-1 expression in colon cancer. *Nucleic acids research* **38**, 6375-6388 (2010).
92 Hoeppner, L. H. *et al.* Runx2 and bone morphogenic protein 2 regulate the expression of an alternative Lef1 transcript during osteoblast maturation. *Journal of cellular physiology* **221**, 480-489 (2009).
93 Nawshad, A. & Hay, E. D. TGFβ3 signaling activates transcription of the LEF1 gene to induce epithelial mesenchymal transformation during mouse palate development. *The Journal of cell biology* **163**, 1291-1301 (2003).
94 Hao, Y.-H. *et al.* Induction of LEF1 by MYC activates the WNT pathway and maintains cell proliferation. *Cell Communication and Signaling* **17**, 1-16 (2019).

95 Yang, M. *et al.* Lef1 is transcriptionally activated by Klf4 and suppresses hyperoxia-induced alveolar epithelial cell injury. *Experimental Lung Research* **48**, 213-223 (2022).
96 Zhao, Y. *et al.* Prognostic value of association of OCT4 with LEF1 expression in esophageal squamous cell carcinoma and their impact on epithelial-mesenchymal transition, invasion, and migration. *Cancer medicine* **7**, 3977-3987 (2018).
97 Wang, C.-J. *et al.* Interaction between CASP8AP2 and ZEB2-CtBP2 Regulates the Expression of LEF1. *Pediatric Hematology and Oncology* **39**, 549-560 (2022).
98 Giese, K., Amsterdam, A. & Grosschedl, R. DNA-binding properties of the HMG domain of the lymphoid-specific transcriptional regulator LEF-1. *Genes & development* **5**, 2567-2578 (1991).
99 Zhao, Y. *et al.* The transcription factor LEF1 promotes tumorigenicity and activates the TGF-β signaling pathway in esophageal squamous cell carcinoma. *Journal of Experimental & Clinical Cancer Research* **38**, 1-16 (2019).
100 Wang, X. *et al.* LEF1/Id3/HRAS axis promotes the tumorigenesis and progression of esophageal squamous cell carcinoma. *International Journal of Biological Sciences* **16**, 2392 (2020).
101 Zhang, M., Jin, M., Gao, Z., Yu, W. & Zhang, W. High COL10A1 expression potentially contributes to poor outcomes in gastric cancer with the help of LEF1 and Wnt2. *Journal of Clinical Laboratory Analysis* **36**, e24612 (2022).
102 Dolezal, D., Zhang, X. & Harigopal, M. Increased Expression of LEF1 and β-Catenin in Invasive Micropapillary Carcinoma of the Breast is Associated With Lymphovascular Invasion and Lymph Node Metastasis. *Applied Immunohistochemistry & Molecular Morphology* **30**, 557-565 (2022).
103 Guo, X. *et al.* Characterization of LEF1 high expression and novel mutations in adult acute lymphoblastic leukemia. *PLoS One* **10**, e0125429 (2015).
104 Lin, A. Y. *et al.* Comparative profiling of primary colorectal carcinomas and liver metastases identifies LEF1 as a prognostic biomarker. *PLoS One* **6**, e16636 (2011).
105 Wu, W. *et al.* High LEF1 expression predicts adverse prognosis in chronic lymphocytic leukemia and may be targeted by ethacrynic acid. *Oncotarget* **7**, 21631 (2016).
106 Wang, W.-J. *et al.* Increased LEF1 expression and decreased Notch2 expression are strong predictors of poor outcomes in colorectal cancer patients. *Disease markers* **35**, 395-405 (2013).
107 Erbilgin, Y. *et al.* Prognostic evidence of LEF1 isoforms in childhood acute lymphoblastic leukemia. *International journal of laboratory hematology* **43**, 1093-1103 (2021).
108 Li, Y. *et al.* LEF1 in androgen-independent prostate cancer: regulation of androgen receptor expression, prostate cancer growth, and invasion. *Cancer research* **69**, 3332-3338 (2009).
109 Wu, L. *et al.* ERG Is a Critical Regulator of Wnt/LEF1 Signaling in Prostate CancerTMPRSS2–ERG Gene Fusions Activate Wnt/LEF1 Signaling. *Cancer research* **73**, 6068-6079 (2013).
110 Gehrke, I., Gandhirajan, R. K. & Kreuzer, K.-A. Targeting the WNT/β-catenin/TCF/LEF1 axis in solid and haematological cancers: multiplicity of therapeutic options. *European Journal of Cancer* **45**, 2759-2767 (2009).
111 K Gandhirajan, R., J Poll-Wolbeck, S., Gehrke, I. & Kreuzer, K.-A. Wnt/β-catenin/LEF-1 signaling in chronic lymphocytic leukemia (CLL): a target for current and potential therapeutic options. *Current cancer drug targets* **10**, 716-727 (2010).

112 Fakhr, E., Zare, F., Azadmanesh, K. & Teimoori-Toolabi, L. LEF1 silencing sensitizes colorectal cancer cells to oxaliplatin, 5-FU, and irinotecan. *Biomedicine & Pharmacotherapy* **143**, 112091 (2021).
113 Hsieh, T.-H. *et al.* A novel cell-penetrating peptide suppresses breast tumorigenesis by inhibiting β-catenin/LEF-1 signaling. *Sci Rep* **6**, 19156 (2016).
114 Zambanini, G., Nordin, A., Jonasson, M., Pagella, P. & Cantù, C. A new CUT&RUN low volume-urea (LoV-U) protocol optimized for transcriptional co-factors uncovers Wnt/β-catenin tissue-specific genomic targets. *Development* **149**, dev201124 (2022).
115 García-Sánchez, A. & Marqués-García, F. Review of methods to study gene expression regulation applied to asthma. *Molecular Genetics of Asthma*, 71-89 (2016).
116 Göös, H. *et al.* Human transcription factor protein interaction networks. *Nature Communications* **13**, 766 (2022).
117 Andrews, G. *et al.* Mammalian evolution of human cis-regulatory elements and transcription factor binding sites. *Science* **380**, eabn7930 (2023).
118 Dann, G. P. *et al.* ISWI chromatin remodellers sense nucleosome modifications to determine substrate preference. *Nature* **548**, 607-611 (2017).
119 Klemm, S. L., Shipony, Z. & Greenleaf, W. J. Chromatin accessibility and the regulatory epigenome. *Nature Reviews Genetics* **20**, 207-220 (2019).
120 Clapier, C. R. & Cairns, B. R. The biology of chromatin remodeling complexes. *Annual review of biochemistry* **78**, 273-304 (2009).
121 Hargreaves, D. C. & Crabtree, G. R. ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. *Cell research* **21**, 396-420 (2011).
122 Signolet, J. & Hendrich, B. The function of chromatin modifiers in lineage commitment and cell fate specification. *The FEBS Journal* **282**, 1692-1702 (2015).
123 Fazzio, T. G. & Panning, B. Control of embryonic stem cell identity by nucleosome remodeling enzymes. *Current opinion in genetics & development* **20**, 500-504 (2010).
124 Ye, Y., Chen, X. & Zhang, W. Mammalian SWI/SNF chromatin remodeling complexes in embryonic stem cells: Regulating the balance between pluripotency and differentiation. *Frontiers in Cell and Developmental Biology* **8**, 626383 (2021).
125 Bode, D., Yu, L., Tate, P., Pardo, M. & Choudhary, J. Characterization of two distinct nucleosome remodeling and deacetylase (NuRD) complex assemblies in embryonic stem cells. *Molecular & Cellular Proteomics* **15**, 878-891 (2016).
126 Smith, J. J. *et al.* The SWI/SNF chromatin remodeling assemblies BAF and PBAF differentially regulate cell cycle exit and cellular invasion in vivo. *PLoS genetics* **18**, e1009981 (2022).
127 Panwalkar, P. *et al.* SWI/SNF complex heterogeneity is related to polyphenotypic differentiation, prognosis, and immune response in rhabdoid tumors. *Neuro-oncology* **22**, 785-796 (2020).
128 Pulice, J. L. & Kadoch, C. in *Cold Spring Harbor symposia on quantitative biology.* 53-60 (Cold Spring Harbor Laboratory Press).
129 Yan, Z. *et al.* PBAF chromatin-remodeling complex requires a novel specificity subunit, BAF200, to regulate expression of selective interferon-responsive genes. *Genes & development* **19**, 1662-1667 (2005).
130 Alver, B. H. *et al.* The SWI/SNF chromatin remodelling complex is required for maintenance of lineage specific enhancers. *Nature communications* **8**, 14648 (2017).

131 Osman, S. Reshuffling chromatin: how the human chromatin remodeler PBAF recognizes nucleosomes. *Nature Structural & Molecular Biology* **29**, 419-419 (2022).
132 Yuan, J., Chen, K., Zhang, W. & Chen, Z. Structure of human chromatin-remodelling PBAF complex bound to a nucleosome. *Nature* **605**, 166-171 (2022).
133 Leighton, G. & Williams Jr, D. C. The methyl-CpG–Binding domain 2 and 3 proteins and formation of the nucleosome remodeling and deacetylase complex. *Journal of molecular biology* **432**, 1624-1639 (2020).
134 Arvindekar, S. *et al.* Molecular architecture of nucleosome remodeling and deacetylase sub-complexes by integrative structure determination. *Protein Science* **31**, e4387 (2022).
135 Hirota, A., Nakajima-Koyama, M., Ashida, Y. & Nishida, E. The nucleosome remodeling and deacetylase complex protein CHD4 regulates neural differentiation of mouse embryonic stem cells by down-regulating p53. *Journal of Biological Chemistry* **294**, 195-209 (2019).
136 Wang, H., Bienz, M., Yan, X.-X. & Xu, W. Structural basis of the interaction between BCL9-Pygo and LDB-SSBP complexes in assembling the Wnt enhanceosome. *Nature Communications* **14**, 3702 (2023).
137 Park, J.-H., Park, E.-J., Hur, S.-K., Kim, S. & Kwon, J. Mammalian SWI/SNF chromatin remodeling complexes are required to prevent apoptosis after DNA damage. *DNA repair* **8**, 29-39 (2009).
138 Renko, M. *et al.* Rotational symmetry of the structured Chip/LDB-SSDP core module of the Wnt enhanceosome. *Proceedings of the National Academy of Sciences* **116**, 20977-20983 (2019).
139 Fiedler, M. *et al.* An ancient Pygo-dependent Wnt enhanceosome integrated by Chip/LDB-SSDP. *Elife* **4**, e09073 (2015).
140 Pagella, P. *et al.* Time-resolved analysis of Wnt-signaling reveals β-catenin temporal genomic repositioning and cell type-specific plastic or elastic chromatin responses. *bioRxiv*, 2022.2008. 2005.502932 (2022).
141 Flack, J. E., Mieszczanek, J., Novcic, N. & Bienz, M. Wnt-dependent inactivation of the Groucho/TLE co-repressor by the HECT E3 ubiquitin ligase Hyd/UBR5. *Molecular cell* **67**, 181-193. e185 (2017).
142 Pollard, K. S. *et al.* Forces shaping the fastest evolving regions in the human genome. *PLoS genetics* **2**, e168 (2006).
143 Isbel, L., Grand, R. S. & Schübeler, D. Generating specificity in genome regulation through transcription factor sensitivity to chromatin. *Nature Reviews Genetics* **23**, 728-740 (2022).
144 Akiva, E., Friedlander, G., Itzhaki, Z. & Margalit, H. A dynamic view of domain-motif interactions. *PLoS Comput. Biol.* **8**, e1002341 (2012).
145 Kim, Y., Min, B. & Yi, G.-S. in *Proteome science.* 1-9 (BioMed Central).
146 Xing, S., Wallmeroth, N., Berendzen, K. W. & Grefen, C. Techniques for the analysis of protein-protein interactions in vivo. *Plant physiology* **171**, 727-758 (2016).
147 Zhou, M., Li, Q. & Wang, R. Current experimental methods for characterizing protein–protein interactions. *ChemMedChem* **11**, 738-756 (2016).
148 Prieto, D. A. *et al.* Mass spectrometry in cancer biomarker research: a case for immunodepletion of abundant blood-derived proteins from clinical tissue specimens. *Biomarkers in medicine* **8**, 269-286 (2014).
149 Branon, T. C. *et al.* Efficient proximity labeling in living cells and organisms with TurboID. *Nature biotechnology* **36**, 880-887 (2018).

150 Cho, K. F. *et al.* Proximity labeling in mammalian cells with TurboID and split-TurboID. *Nature Protocols* **15**, 3971-3999 (2020).
151 Roux, K. J., Kim, D. I., Burke, B. & May, D. G. BioID: a screen for protein-protein interactions. *Current protocols in protein science* **91**, 19.23. 11-19.23. 15 (2018).
152 Skene, P. J., Henikoff, J. G. & Henikoff, S. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nature protocols* **13**, 1006-1019 (2018).
153 Choi, H. *et al.* SAINT: probabilistic scoring of affinity purification–mass spectrometry data. *Nature methods* **8**, 70-73 (2011).
154 Madden, K. *et al.* Mycobacterium tuberculosis infection triggers epigenetic changes that are enriched in a type I IFN signature. *Microlife* **4**, uqad006 (2023).
155 Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic acids research* **44**, W160-W165 (2016).
156 Zerbino, D. R., Johnson, N., Juettemann, T., Wilder, S. P. & Flicek, P. WiggleTools: parallel processing of large collections of genome-wide datasets for visualization and statistical analysis. *Bioinformatics* **30**, 1008-1009 (2014).
157 Ter Huurne, M., Chappell, J., Dalton, S. & Stunnenberg, H. G. Distinct cell-cycle control in two different states of mouse pluripotency. *Cell stem cell* **21**, 449-455. e444 (2017).
158 Ummethum, H. & Hamperl, S. Proximity labeling techniques to study chromatin. *Frontiers in Genetics* **11**, 450 (2020).
159 Liu, Z. *et al.* Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Sci Rep* **7**, 2193 (2017). https://doi.org:10.1038/s41598-017-02460-2
160 Karasawa, S., Araki, T., Yamamoto-Hino, M. & Miyawaki, A. A green-emitting fluorescent protein from Galaxeidae coral and its monomeric version for use in fluorescent labeling. *J Biol Chem* **278**, 34167-34171 (2003). https://doi.org:10.1074/jbc.M304063200
161 Liu, Z. *et al.* Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Sci Rep* **7**, 2193 (2017).
162 Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nature protocols* **8**, 2281-2308 (2013).
163 Hersbach, B. A. *et al.* Probing cell identity hierarchies by fate titration and collision during direct reprogramming. *Molecular Systems Biology* **18**, e11129 (2022).
164 Hulsen, T., de Vlieg, J. & Alkema, W. BioVenn–a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC genomics* **9**, 1-6 (2008).
165 Oliveros, J. C. VENNY. An interactive tool for comparing lists with Venn Diagrams. *http://bioinfogp. cnb. csic. es/tools/venny/index. html* (2007).
166 Behrens, J. *et al.* Functional interaction of β-catenin with the transcription factor LEF-1. *Nature* **382**, 638-642 (1996).
167 Kim, I., Hur, J. & Jeong, S. HuR represses Wnt/β-catenin-mediated transcriptional activity by promoting cytoplasmic localization of β-catenin. *Biochemical and biophysical research communications* **457**, 65-70 (2015).
168 Hutchins, E. J., Gandhi, S., Chacon, J., Piacentino, M. & Bronner, M. E. RNA-binding protein Elavl1/HuR is required for maintenance of cranial neural crest specification. *Elife* **11**, e63600 (2022).
169 Hochstatter, J. *et al.* Myb-binding protein 1a (Mybbp1a) regulates levels and processing of pre-ribosomal RNA. *Journal of Biological Chemistry* **287**, 24365-24377 (2012).

170 Cervantes, M. *et al.* BMAL1 associates with NOP58 in the nucleolus and contributes to pre-rRNA processing. *Iscience* **23**, 101151 (2020).
171 Genuth, N. R. *et al.* A stem cell roadmap of ribosome heterogeneity reveals a function for RPL10A in mesoderm production. *Nature communications* **13**, 5491 (2022).
172 Pfister, A. S. & Kühl, M. Of Wnts and ribosomes. *Progress in molecular biology and translational science* **153**, 131-155 (2018).
173 Maurya, S. *et al.* Loss of KMT2C reprograms the epigenomic landscape in hPSCs resulting in NODAL overexpression and a failure of hemogenic endothelium specification. *Epigenetics* **17**, 220-238 (2022).
174 Xu, F., Flowers, S. & Moran, E. Essential role of ARID2 protein-containing SWI/SNF complex in tissue-specific gene expression. *Journal of Biological Chemistry* **287**, 5033-5041 (2012).
175 Gatchalian, J. *et al.* A non-canonical BRD9-containing BAF chromatin remodeling complex regulates naive pluripotency in mouse embryonic stem cells. *Nature communications* **9**, 5139 (2018).
176 Kaeser, M. D., Aslanian, A., Dong, M.-Q., Yates, J. R. & Emerson, B. M. BRD7, a novel PBAF-specific SWI/SNF subunit, is required for target gene activation and repression in embryonic stem cells. *Journal of Biological Chemistry* **283**, 32254-32263 (2008).
177 De Cegli, R. *et al.* Reverse engineering a mouse embryonic stem cell-specific transcriptional network reveals a new modulator of neuronal differentiation. *Nucleic acids research* **41**, 711-726 (2013).
178 Zimmerli, D. *et al.* TBX3 acts as tissue-specific component of the Wnt/β-catenin enhanceosome. *bioRxiv*, 2020.2004. 2022.053561 (2020).
179 Weidgang, C. E. *et al.* TBX3 directs cell-fate decision toward mesendoderm. *Stem cell reports* **1**, 248-265 (2013).
180 Waghray, A. *et al.* Tbx3 controls Dppa3 levels and exit from pluripotency toward mesoderm. *Stem cell reports* **5**, 97-110 (2015).
181 Melzer, M. K. *et al.* TBX3 is dynamically expressed in pancreatic organogenesis and fine-tunes regeneration. *BMC biology* **21**, 55 (2023).
182 Lee, J.-E. *et al.* H3K4 mono-and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. *elife* **2**, e01503 (2013).
183 Cao, K. *et al.* An Mll4/COMPASS-Lsd1 epigenetic axis governs enhancer function and pluripotency transition in embryonic stem cells. *Science advances* **4**, eaap8747 (2018).
184 Xie, G. *et al.* MLL3/MLL4 methyltransferase activities control early embryonic development and embryonic stem cell differentiation in a lineage-selective manner. *Nature Genetics*, 1-13 (2023).
185 Iuchi, S. & Paulo, J. A. RNAmetasome network for macromolecule biogenesis in human cells. *Communications Biology* **4**, 1399 (2021).
186 Pina, J. M., Hernandez, L. A. & Keppetipola, N. M. Polypyrimidine tract binding proteins PTBP1 and PTBP2 interact with distinct proteins under splicing conditions. *PLoS One* **17**, e0263287 (2022).
187 Linares, A. J. *et al.* The splicing regulator PTBP1 controls the activity of the transcription factor Pbx1 during neuronal differentiation. *Elife* **4**, e09268 (2015).
188 Iannone, C. *et al.* PTBP1-activated co-transcriptional splicing controls epigenetic status of pluripotent stem cells. *Molecular Cell* (2023).

189 Daniels, D. L. & Weis, W. I. β-catenin directly displaces Groucho/TLE repressors from Tcf/Lef in Wnt-mediated transcription activation. *Nature structural & molecular biology* **12**, 364-371 (2005).
190 Ramakrishnan, A. B., Sinha, A., Fan, V. B. & Cadigan, K. M. The Wnt transcriptional switch: TLE removal or inactivation? *Bioessays* **40**, 1700162 (2018).
191 Billin, A. N., Thirlwell, H. & Ayer, D. E. β-catenin–histone deacetylase interactions regulate the transition of LEF1 from a transcriptional repressor to an activator. *Molecular and cellular biology* **20**, 6882-6890 (2000).
192 Lau, W. d. & Clevers, H. LEF1 turns over a new leaf. *Nature Genetics* **28**, 3-4 (2001).
193 Arce, L., Pate, K. T. & Waterman, M. L. Groucho binds two conserved regions of LEF-1 for HDAC-dependent repression. *BMC Cancer* **9**, 159 (2009). https://doi.org:10.1186/1471-2407-9-159
194 Boroviak, T. *et al.* Lineage-specific profiling delineates the emergence and progression of naive pluripotency in mammalian embryogenesis. *Developmental cell* **35**, 366-382 (2015).
195 Yamane, M., Fujii, S., Ohtsuka, S. & Niwa, H. Zscan10 is dispensable for maintenance of pluripotency in mouse embryonic stem cells. *Biochemical and Biophysical Research Communications* **468**, 826-831 (2015).
196 Zhang, W. *et al.* Zfp206 regulates ES cell gene expression and differentiation. *Nucleic acids research* **34**, 4780-4790 (2006).
197 Frankenberg, S. R. Different species choose their own paths to pluripotency. *Developmental Cell* **35**, 267-268 (2015).
198 Zhang, Q. *et al.* Tcstv1 and Tcstv3 elongate telomeres of mouse ES cells. *Sci Rep* **6**, 1-11 (2016).
199 Wang, L. *et al.* The THO complex regulates pluripotency gene mRNA export and controls embryonic stem cell self-renewal and somatic cell reprogramming. *Cell stem cell* **13**, 676-690 (2013).
200 Jiang, Y. *et al.* ZSCAN10 promotes cell proliferation, upregulates OCT4 expression, and activates Wnt/β-catenin signaling in glioma. *International Journal of Clinical and Experimental Pathology* **12**, 700 (2019).
201 Theka, I. *et al.* Wnt/β-catenin signaling pathway safeguards epigenetic stability and homeostasis of mouse embryonic stem cells. *Sci Rep* **9**, 948 (2019).
202 Vodnala, M., Choi, E.-B. & Fong, Y. W. Low complexity domains, condensates, and stem cell pluripotency. *World Journal of Stem Cells* **13**, 416 (2021).
203 Sutherland, C. What Are the bona fide GSK3 Substrates? *Int J Alzheimers Dis* **2011**, 505607 (2011). https://doi.org:10.4061/2011/505607
204 Kelly, K. F. *et al.* β-catenin enhances Oct-4 activity and reinforces pluripotency through a TCF-independent mechanism. *Cell Stem Cell* **8**, 214-227 (2011). https://doi.org:10.1016/j.stem.2010.12.010
205 Hamilton, D. J., Hein, A. E., Holmes, Z. E., Wuttke, D. S. & Batey, R. T. The DNA-binding high-mobility group box domain of sox family proteins directly interacts with RNA in vitro. *Biochemistry* **61**, 943-951 (2022).
206 Hamilton, D. J., Hein, A. E., Wuttke, D. S. & Batey, R. T. The DNA binding high mobility group box protein family functionally binds RNA. *Wiley Interdisciplinary Reviews: RNA*, e1778

207 Spruijt, C. G. *et al.* ZMYND8 co-localizes with NuRD on target genes and regulates poly (ADP-ribose)-dependent recruitment of GATAD2A/NuRD to sites of DNA damage. *Cell reports* **17**, 783-798 (2016).
208 Yildirim, O. *et al.* Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. *Cell* **147**, 1498-1510 (2011).
209 Ho, L. *et al.* esBAF facilitates pluripotency by conditioning the genome for LIF/STAT3 signalling and by regulating polycomb function. *Nature cell biology* **13**, 903-913 (2011).
210 Mor, N. *et al.* Neutralizing Gatad2a-Chd4-Mbd3/NuRD complex facilitates deterministic induction of naive pluripotency. *Cell stem cell* **23**, 412-425. e410 (2018).
211 Marino, S. & Nusse, R. Mutants in the mouse NuRD/Mi2 component P66α are embryonic lethal. *PLoS One* **2**, e519 (2007).
212 Kaji, K. *et al.* The NuRD component Mbd3 is required for pluripotency of embryonic stem cells. *Nature cell biology* **8**, 285-292 (2006).
213 Laing, A. F., Lowell, S. & Brickman, J. M. Gro/TLE enables embryonic stem cell differentiation by repressing pluripotent gene expression. *Developmental biology* **397**, 56-66 (2015).
214 Kurimoto, K. *et al.* Quantitative dynamics of chromatin remodeling during germ cell specification from mouse embryonic stem cells. *Cell stem cell* **16**, 517-532 (2015).
215 Yamada, M. *et al.* NARF, an nemo-like kinase (NLK)-associated ring finger protein regulates the ubiquitylation and degradation of T cell factor/lymphoid enhancer factor (TCF/LEF). *Journal of Biological Chemistry* **281**, 20749-20760 (2006).
216 Novellasdemunt, L. *et al.* NEDD4 and NEDD4L regulate Wnt signalling and intestinal stem cell priming by degrading LGR5 receptor. *The EMBO Journal* **39**, e102771 (2020).
217 Lu, C. *et al.* Intestinal knockout of Nedd4 enhances growth of Apcmin tumors. *Oncogene* **35**, 5839-5849 (2016).
218 Rafiee, M.-R., Girardot, C., Sigismondo, G. & Krijgsveld, J. Expanding the circuitry of pluripotency by selective isolation of chromatin-associated proteins. *Molecular cell* **64**, 624-635 (2016).
219 Cuevas, E., Rybak-Wolf, A., Rohde, A. M., Nguyen, D. T. & Wulczyn, F. G. Lin41/Trim71 is essential for mouse development and specifically expressed in postnatal ependymal cells of the brain. *Frontiers in Cell and Developmental Biology* **3**, 20 (2015).
220 Torres-Fernández, L. A. *et al.* The stem cell–specific protein TRIM71 inhibits maturation and activity of the prodifferentiation miRNA let-7 via two independent molecular mechanisms. *RNA* **27**, 805-828 (2021).
221 Xia, X. *et al.* Role of TRIM33 in Wnt signaling during mesendoderm differentiation. *Science China Life Sciences* **60**, 1142-1149 (2017).
222 Sun, H. *et al.* Recruitment of Trim33 to Cell-Context specific PML nuclear bodies regulates nodal signaling in mESCs. *The EMBO Journal* **42**, e112058 (2023).
223 Jaynes, P. W. *et al.* OTUD4 enhances TGFβ signalling through regulation of the TGFβ receptor complex. *Sci Rep* **10**, 15725 (2020).
224 Zhao, X. *et al.* OTUD4: A potential prognosis biomarker for multiple human cancers. *Cancer Management and Research*, 1503-1512 (2020).
225 Lee, Y., Piao, H.-l. & Kim, J. OTUD7B Activates Wnt Signaling Pathway through the Interaction with LEF1. *Biomolecules* **13**, 1001 (2023).

226 Suresh, B., Lee, J., Kim, H. & Ramakrishna, S. Regulation of pluripotency and differentiation by deubiquitinating enzymes. *Cell Death & Differentiation* **23**, 1257-1264 (2016).
227 de Dieuleveult, M. *et al.* USP9X deubiquitinase couples the pluripotency network and cell metabolism to regulate ESC differentiation potential. *Biorxiv*, 2020.2001. 2013.904904 (2020).
228 Shang, Z. *et al.* USP9X-mediated deubiquitination of B-cell CLL/lymphoma 9 potentiates Wnt signaling and promotes breast carcinogenesis. *Journal of Biological Chemistry* **294**, 9844-9857 (2019).
229 Nielsen, C. P., Jernigan, K. K., Diggins, N. L., Webb, D. J. & MacGurn, J. A. USP9X deubiquitylates DVL2 to regulate WNT pathway specification. *Cell reports* **28**, 1074-1089. e1075 (2019).
230 Goodwin, L. R. & Picketts, D. J. The role of ISWI chromatin remodeling complexes in brain development and neurodevelopmental disorders. *Molecular and Cellular Neuroscience* **87**, 55-64 (2018).
231 Meng, J. *et al.* WSTF promotes proliferation and invasion of lung cancer cells by inducing EMT via PI3K/Akt and IL-6/STAT3 signaling pathways. *Cellular Signalling* **28**, 1673-1682 (2016).
232 Zhang, W. *et al.* PR55α, a regulatory subunit of PP2A, specifically regulates PP2A-mediated β-catenin dephosphorylation. *Journal of Biological Chemistry* **284**, 22649-22656 (2009).
233 Sugimoto, M. Developmental genetics of the mouse t-complex. *Genes Genet Syst* **89**, 109-120 (2014). https://doi.org:10.1266/ggs.89.109
234 Lange, L. *et al.* Patterning and gastrulation defects caused by the tw18 lethal are due to loss of Ppp2r1a. *Biology open* **6**, 752-764 (2017).
235 Azambuja, A. P. & Simoes-Costa, M. The connectome of neural crest enhancers reveals regulatory features of signaling systems. *Developmental cell* **56**, 1268-1282. e1266 (2021).
236 Feder, K. *et al.* Differences in expression and function of LEF1 isoforms in normal versus leukemic hematopoiesis. *Leukemia* **34**, 1027-1037 (2020).
237 Moreira, S. *et al.* Endogenous BioID elucidates TCF7L1 interactome modulation upon GSK-3 inhibition in mouse ESCs. *bioRxiv*, 431023 (2018).
238 Reynolds, N. *et al.* NuRD suppresses pluripotency gene expression to promote transcriptional heterogeneity and lineage commitment. *Cell stem cell* **10**, 583-594 (2012).
239 Bornelöv, S. *et al.* The nucleosome remodeling and deacetylation complex modulates chromatin structure at sites of active transcription to fine-tune gene expression. *Molecular Cell* **71**, 56-72. e54 (2018).
240 Kraus, P. *et al.* Pleiotropic functions for transcription factor zscan10. *PLoS One* **9**, e104568 (2014).
241 Yanagihara, Y. *et al.* Zscan10 suppresses osteoclast differentiation by regulating expression of Haptoglobin. *Bone* **122**, 93-100 (2019).
242 Yu, H.-b., Kunarso, G., Hong, F. H. & Stanton, L. W. Zfp206, Oct4, and Sox2 are integrated components of a transcriptional regulatory network in embryonic stem cells. *Journal of Biological Chemistry* **284**, 31327-31335 (2009).
243 Wang, Z.-X. *et al.* Zfp206 is a transcription factor that controls pluripotency of embryonic stem cells. *Stem cells* **25**, 2173-2182 (2007).

APPENDIX A: SOLUTIONS

CELL CULTURE

Mouse Embryonic Stem Cell Media

Name	Amount
DMEM/High Glucose Media (Sigma)	500.0 mL
Fetal Bovine Serum (Sigma)	75 mL
Non-Essential Amino Acids (Sigma; 100 mM)	5 mL
Sodium Pyruvate (Sigma; 100 mM)	5 mL
GlutaMax I (or L-Glutamine) (Sigma; 200 mM)	5 mL
2-Mercaptoethanol (Gibco; 55 mM)	500.0 µL
Leukemia Inhibitory Factor (Amsbio; 10 ng/mL)	120.0 µL
Total Volume	~590.0 mL

WESTERN BLOT

2x RIPA Buffer

Name	Amount
1 M Tris-pH 8.0	10.0 mL
5 M NaCl	6 mL
SDS (10% [w/v])	2 mL
NP40 (10% [w/v])	20.0 mL
Sodium Deoxycholate (10% [w/v])	10.0 mL
500 mM EDTA	400.0 µL
Milli-Q Water	Up to 100.0 mL
Total Volume	100.0 mL

1x RIPA Buffer

Name	Amount
2x RIPA Buffer	500.0 µL
100x Protease/Phosphatase Inhibitor cocktail (CST)	10.0 µL
Milli-Q Water	490.0 µL
Total Volume	1 mL

-CST: Cell Signaling Technology

1x Running Buffer

Name	Amount
MES SDS Running Buffer (20X)	25 mL
Milli-Q Water	475 mL
Total Volume	500.0 mL

1xSample Loading Buffer

Name	Amount
4x SDS Sample Loading Buffer (Invitrogen)	50.0 µL
Bond Breaker (Invitrogen)	10.0 µL
Protein Sample	Varies
Milli-Q Water	Up to 200.0 µL
Total Volume	200.0 µL

10x Tris Buffered Saline (TBS) pH 7.6

Name	Amount
Tris-Base	24 g
NaCl	88 g
HCl	Varies
Milli-Q Water	Up to 1.0 L
Total Volume	1.0 L

1x TBS-Tween20 (TBST)

Name	Amount
10x TBS	50 .0 mL
Tween-20 (Bioshop)	500.0 µL
Milli-Q Water	449.5 mL
Total Volume	500.0 mL

TURBOID

10 mM Biotin

Name	Amount
Biotin (Sigma)	122 mg
DMEM/F12 (Gibco)	50.0 mL

Wash Buffer

Name	Amount
Tris-HCl pH 7.4	600.0 mg
Urea (Sigma)	48 g
Milli-Q Water	Up to 100.0 mL
Total Volume	100.0 mL

Storage Buffer

Name	Amount
50 mM Ammonium Bicarbonate	9.5 mL
1 mM Biotin	500.0 µL
Total volume	10.0 mL

PLASMID CONSTRUCTION

1000x Carbenicillin (100 mg/mL)

Name	Amount
Carbenicillin	1 g
Milli-Q Water	10.0 mL
Total Volume	~10.0 mL

-Store at -20 °C

Luria-Bertani (LB) Agar Plates (containing 100 µg/mL Carbenicillin)

Name	Amount
LB Broth with Agar (Lennox)	32 g
Milli-Q Water	Up to 1.0 L
Carbenicillin (100 mg/mL)	1.0 mL
Total Volume	1.0 L

-Combine LB Agar and water in a bottle for autoclave
-Autoclave for agar dissolution and sterilization
-Mix carbenicillin into the solution while it is still warm (at around 40 °C)
-Pour 20 mL of the solution into 10 cm plates (~20 mL/plate), then let the agar cool and solidify
-store at 4 °C

LB Broth (containing 100 µg/mL Carbenicillin)

Name	Amount
LB Broth Base	20.0 g
Milli-Q Water	Up to 1.0 L
Carbenicillin (100 mg/mL)	1.0 mL
Total Volume	1.0 L

-Combine LB Agar and water in a bottle for autoclave
-Autoclave for agar dissolution and sterilization
-Mix carbenicillin into the solution while it is still warm (at around 40 °C)
-store at 4 °C

10x Tris-Acetate-EDTA (TAE) Buffer

Name	Amount
Tris	48.4 g
Acetic Acid	11.4 mL
EDTA	3.72 g
Milli-Q Water	Up to 1.0 L
Total Volume	1.0 L

1x TAE Buffer

Name	Amount
10x TAE Buffer	50.0 mL
Milli-Q Water	450.0 mL
Total Volume	500.0 mL

1% Agarose

Name	Amount
Agarose (Invitrogen)	1.0 g
Milli-Q Water	100.0 mL
Total Volume	100.0 mL

APPENDIX B: SUPPLEMENTARY TABLES

Table S 1. Primer and Oligo Sequences

Type	Name	Sequence	Construct
Primer			
	TurboID ΔN-LEF1	Short Lef1-Fwd: CTGCGGTCTGCCGAAAAGGGCGGATCTGGAGGCATGATG CCCAATATGAACAGC	Short LEF1 (ΔN-LEF1; 852 bp)
		Short Lef1-Rev: ATCAGCGAGCTCTAGATCATTTAGTAGGCAGCTGTCATTC	
	TurboID ΔN-LEF1 Sequencing	F-mid-Turbo-Seq: GTGCGAGTCAAATGGCCCAATGAC	
		R-pA-Seq: GTGAGCGCAACGCAATTAATGTGAG	
	FL-LEF1 TurboID	LEF1 FL Fwd: GAGAACCCTGGACCTGGTACATGCCCCAACTTTCCGGAG	1188 bp
		LEF1 FL Rev: GTAGGCAGCTGTCATTCTGG	
		3XHA TurboID FWD: ATGTACCCGTATGATGTTCC	1050 bp
		3X TurboID Rev: ATCAGCGAGCTCTAGATCATTCACTTTTCGGCAGACCG	
		Long linker*	105 bp
	FL-LEF1 TurboID Sequencing	F-mAGend-Seq: CAACGACTACAACAAGGTGAAGCTG	
		F-mid-Turbo-Seq: GTGCGAGTCAAATGGCCCAATGAC	
		R-pA-Seq: GTGAGCGCAACGCAATTAATGTGAG	
		LEF1-F: GAGAACCCTGGACCTGGTACCATGCCCCAACTTTCCGG	
		LEF1-R: GTAGGCAGCTGTCATTCTG	
		6-dBBBDe-FWD: ATAGGGAGAGCGGCCGCATGATGCCCAATATGAACAGCG ACCCGTAC	
		3b-dl34-FWDH: TCCAGACACCCTCCATTTTCCCATCATATGATTCCTGGTC	
	TurboID-NLS	3XHA-TurboID-NLS_fwd: GAGAACCCTGGAACCATGTACCCGTATGATGTT	1188 bp
		3XHA-TurboID-NLS_rev: ATCAGCGAGCTCTAGATCATCGATTCACACCTTCCTCTTC	
	Lef1 gRNA-T7E1 digestion	T7E1-Cas9 Fwd: CCAAATCACCGAAATCAGCATC	665 bp
		T7E1-Cas9 Rev: TGCTGGCTGGGATGATTT	
	Putative *Lef1* knockout PCR	T7E1-Cas9 Fwd: CCAAATCACCGAAATCAGCATC	961 bp
		Lef1 KO-Rev: GTTTCGGTGGCTCTGAGGGACGCGGTGCATTAA	
Custom oligos	3X GGGGS Linker	GGTACAGGTCCCAGAATGACAGCTGCCTACGGCGGAGGA GGGAGCGGGGGAGGCGGATCTGGCGGAGGCGGATCCAT GTACCCGTATGATGTTCCGGATTACGCT	Long linker*
	Lef1 sgRNA	Lef1 sgRNA-Fwd: 5-CACCGGATCAGTCATCCCGAAGAGG-3	
		Lef1 sgRNA-Rev: 5-AAACCCTCTTCGGGATGACTGATCC-3	