Dissecting MicroRNA Requirements in Early Human Embryonic Development

A Dissertation
Presented to the Faculty of the Weill Cornell Graduate School
of Medical Sciences
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

by
Maria Virginia Teijeiro
August 2017

© 2017 Maria Virginia Teijeiro

Dissecting MicroRNA Requirements in Early Human Embryonic Development

Maria Virginia Teijeiro, Ph.D.

Cornell University 2017

DICER1 is the enzyme responsible for cleaving double-stranded RNAs (dsRNAs) into functionally mature ~20-24 nucleotide (nt) long microRNAs (miRNAs). miRNAs are the effectors of an RNA-induced gene silencing system that functions at the post-transcriptional and post-translational levels, and are essential for mouse development. Due to maternal Dicer1 contribution in the early stages of embryonic development, the exact timing for miRNA requirement has not yet been established. Furthermore, precise miRNA functional studies in early embryonic development have been lacking since miRNA knockout studies are hindered due to redundancy in the miRNA network. Thus, miRNA function in early embryonic development remains elusive. To address this knowledge gap, we set out to knockout *DICER1* in human ESCs (hESCs) and assess human-specific miRNA requirements in the primed state. We report that DICER1 is essential in hESCs unlike in mouse embryonic stem cells (mESCs), and that this likely reflects a unique requirement for DICER1 in primed versus naïve pluripotency. Additionally, we designed an inducible DICER1 system to bypass the lethality of DICER1 loss in hESCs and enable the generation of homozygous mutants. A targeted mature miRNA rescue screen identified members of the miRNA-302-367 and miRNA-371-373 clusters, but surprisingly not miRNA-17-92, as having pro-survival functions in hESCs. Since it bypasses the common issue of redundancy in the miRNA network, our screening platform is particularly suited

to dissect the roles of individual miRNAs and miRNA clusters in early human development and hESC differentiation.

BIOGRAPHICAL SKETCH

Maria Virginia Teijeiro was born in Washington DC where her family was temporarily residing and soon after moved to Argentina where she is from. She grew up in the suburbs of Buenos Aires, and when she finished high school she moved to California in pursuit of a scientific career. She attended Stanford University and graduated in 2009 with a B.S. in Biological Sciences. After spending a year studying the effects of DNA lesions on gene transcription in the laboratory of Phil Hanawalt at Stanford, she moved to New York City to pursue her Ph.D. in Cell & Developmental Biology at Cornell University's medical school. At the program's retreat, she met Danwei Huangfu, an incoming Assistant Professor at Memorial Sloan Kettering Cancer Center, and enthusiastically joined her lab to study human development using human embryonic stem cells (hESCs) as a platform. Here, she studied the role of *DICER1* in pluripotency, and discovered that DICER1 is essential in human ESCs, unlike in mouse ESCs (mESCs), and that this discrepancy is likely due to the different developmental stages that mESCs and hESCs represent. Furthermore, she set up a platform to study the contribution of individual and clusters of microRNAs in hESCs to gain insight into their function in early human development. Armed with great scientific training and three new powerful letters after her name, she is now embarking on a career in the healthcare/biotech industry.

Para mamá y papá,

por su apoyo incondicional,
por las infinitas oportunidades que me dieron
y por lo mucho que creyeron en mí.

Les dedico mi tesis con muchísimo amor y agradecimiento.

ACKNOWLEDGEMENTS

Completing a Ph.D. is a trying feat, and one that I could not have concluded without the unwavering support of many wonderful people in my life. I want to thank them all profoundly.

Thanks to my mentor, Danwei Huangfu, for her optimism, her enthusiasm with science and discovery, and for building a supportive lab, full of amazing people I call my friends. As the great developmental biologist she is, she knew how to elicit the right signals at the right time to trigger my development into the scientist I am today, and for that, I am truly grateful.

I want to thank my committee members Eric Lai, Lorenz Studer, and Kat Hadjantonakis for their advice and support throughout my years as a graduate student. I am particularly grateful for Eric's insight in my project, and for enabling an essential and fruitful collaboration with his lab.

Big thanks to my labmates Nipun, Chew-Li, Qing, Kihyun, ZD, Zhu, Daniela, and Fede for their daily moral support in matters of science as well as matters of life. I have spent most of my time with you for the past seven years and I am truly happy for the deep friendships we have forged. I also want to thank Sonali Majumdar who has been an instrumental collaborator in this project and a great friend. I have learnt so much from all of you, and I could not have imagined better labmates and friends for this rollercoaster ride.

Heartfelt thanks to my oldest friends: Angeles, Emilia, Ximena, Valentina, Brenda, Gabriela, Sabrina, Rocio and Luz for being a constant source of support and happiness, and for keeping me grounded to what is most important in life. Many thanks to my college friends Cristina, Isabella, Maria Victoria, and all the MKs for your steady support and inspiration, and for

being a source of fun and lightness in my life. A heartfelt thanks to Yaya, for her caring soul and for sharing (and surviving) many crucial moments in our twenties together. Special thanks to Kat and Tharu, for being my home away from home - for celebrating birthdays, engagements, graduations, and many more joyful events together these past seven years. I love you all! Having such wonderful friends to rely on in good and bad times has made my graduate student life so much sweeter.

I want to especially thank my science partner-in-crime, Iris, for being more like a sister than a friend - the person I go to when I'm happy, when I'm sad, or when I need advice. It is a blessing to have shared so many profound and life-changing experiences with you: our mutual desire to study in the USA, our passion for biology, moving away from home when we were only eighteen (and nobody else around us was doing that), and more recently, the ups and downs of pursuing a Ph.D. Because of you I have never felt alone in this ride - and that has been so crucial. Thank you for your kindheartedness, and for your illuminating presence in my life.

The most profound thanks to my beautiful family. To my siblings Agustina, Jimena, Rodrigo, and Alvaro, for being my role models and for inspiring me since I was a kid. Thank you for being the rock in my life. To my siblings-in-law, Marcos, Patricia, and Josefina, thank you for your friendship and support throughout the years. To Elvi, for loving me so much.

To my parents, Silvia and Mario, the most generous souls I know - thank you for supporting me and allowing me to pursue my dreams even when it meant living far away from each other. I know how hard it must have been to leave your eighteen year old daughter in a foreign country, but I am truly grateful that you had faith in me and selflessly supported me. Thank you for

pushing me to be the best I can be, and for rooting for me since day one – I could not have done it without your unwavering love and support.

Finally, the most heartfelt thanks to my fiancé, Andrés, for encouraging me and believing in me always, especially during the rough patches of graduate school. Thank you for moving back to New York when I needed you the most - I could not have done it without your constant support and your healing humor. Coming home to you has made this chapter in my life so much more meaningful because it has become a chapter in *our* lives. I can't wait to spend the rest of our time building a life together. You are the reason for my happiness, and I am eternally grateful for you.

TABLE OF CONTENTS

BIOGRAPHICAL SKETCH iii
ACKNOWLEDGEMENTS v
LIST OF FIGURES xi
LIST OF TABLES xii
CHAPTER 1: Introduction 1
1.1 OUR MODEL SYSTEM: HUMAN EMBRYONIC STEM CELLS 1
1.1.1 Human Embryonic Stem Cells as a Platform to Study Human Development 1
1.1.2 A New Genetic Toolset to Study Gene Function in hESCs: CRISPR/Cas9 and TALENs 4
1.2 MICRORNA's & DICER1 6
1.2.1 MicroRNA Biogenesis & Function 6
1.2.2 Dicer1 Protein 9
1.2.3 MicroRNA Expression in Naïve and Primed ESCs 11
1.3 MICRORNA FUNCTION IN DEVELOPMENT 13
1.3.1 Transcriptional Control of MicroRNA Expression in ESCs 13
1.3.2 Global microRNA Function in Mouse Development 14
1.3.3 Global microRNA Function in Naïve and Primed Cultured ESCs 17
1.3.4 Function of miRNA-290-295 Cluster (miRNA-371-373 in human) in ESCs 19
1.3.5 Function of miRNA-302-367 Cluster in ESCs 22
1.3.6 Function of miRNA-17-92 Cluster in ESCs 24
1.3.7 miRNAs in Reprogramming 26
1.4 THESIS AIMS 28
CHAPTER 2: DICER1 Requirement and miRNA Function in hESCs 30
2.1 INTRODUCTION 30
2.2 MATERIALS & METHODS 32
2.2.1 hESC Culture 32
2.2.2 Generation of gRNAs for CRISPR/Cas9 Targeting 33
2.2.3 Establishment of Clonal Mutant Lines Using CRISPR/Cas9 33
2.2.4 Construction of pIND_DICER1_ΔPAM-Cr4 and -Cr6 35
2.2.5 Generation of TRE-DICER1* Lines 36
2.2.6 RNA Processing and Quantitative qRT-PCR 38
2.2.7 Western Blotting 39
2.2.8 Northern Blotting 39
2.2.9 Alkaline Phosphatase and Immunofluorescence Staining 40
2.2.10 miRNA Mimic Rescue Screen and Quantification 40
2.2.11 Small RNA-Sequencing Analysis 41

2.2.12 RNA-Sequencing....42
2.2.13 Naïve Stem Cell Culture....42
2.2.14 Flow Analysis....43
2.2.15 Teratoma Assay....44
2.2.16 Colony Forming Assay....44
2.2.17 Suspension Culture....45
2.2.18 Neuroectoderm Differentiation....45
2.2.19 Proliferation Assay....46
2.2.20 Statistical Analysis....46
2.3 RESULTS....47
2.3.1 Failure to Generate True DICER1 Null hESCs using iCRISPR....47
2.3.2 DICER1 Hypomorphs Maintain Pluripotency Marker Expression....54
2.3.3 DICER1 Hypomorphs Retain Characteristic Cell Cycle Properties and Normal Apoptosis Levels....56
2.3.4 DICER1 Hypomorphs Show Impaired Colony Forming Ability....58
2.3.5 DICER1 Hypomorphs Dysregulate Differentiation Markers but Retain the Ability to Differentiate....60
2.3.6 DICER1 Hypomorphs Downregulate Argonautes 2-4....66
2.3.7 Successful Generation of DICER1 Knockout hESCs using a TRE-DICER1* Rescue Strategy....68
2.3.8 DICER1 is Essential for Self-Renewal in hESCs....72
2.3.9 Loss of DICER1 in hESCs Causes Caspase-3-Mediated Apoptosis....77
2.3.10 Five Days of Doxycycline Removal Results in Few Changes in Transcriptome....77
2.3.11 ESCC miRNA Clusters 371-373, 302-367 but not 17-92 Rescue DICER1 Knockout hESCs....80
2.3.12 DICER1 Seems Not to be Essential in Naïve hESCs....84
2.4 DISCUSSION....86
CHAPTER 3: Conclusions & Perspectives....91
3.1 SUMMARY....91
3.2 FUTURE DIRECTIONS....93
3.2.1 Detailed characterization of DICER1 Knockout Phenotype in hESCs....93
3.2.2 Mechanism of action of miRNA-302-367 and miRNA-371-373 in hESCs....93
3.2.3 Genome-wide miRNA Screen to Identify Novel miRNA Players in hESCs Pluripotency and Differentiation....94
3.2.4 Identify Naïve Versus Primed miRNA Requirements in hESCs....95
3.3 CONCLUSIONS....95
APPENDIX....97
Appendix 1....97

Appendix 2. 98
Appendix 3. 99
Appendix 4. 100
REFERENCES 101

LIST OF FIGURES

Figure 1.1……2
Figure 1.2……8
Figure 1.3……16
Figure 1.4……17
Figure 1.5……29
Figure 2.1……48
Figure 2.2……51
Figure 2.3……53
Figure 2.4……55
Figure 2.5……57
Figure 2.6……59
Figure 2.7……61
Figure 2.8……62
Figure 2.9……65
Figure 2.10……67
Figure 2.11……70
Figure 2.12……73
Figure 2.13……75
Figure 2.14……76
Figure 2.15……78
Figure 2.16……79
Figure 2.17……82
Figure 2.18……85

LIST OF TABLES

Table 2.1……34
Table 2.2……36
Table 2.3……38
Table 2.4……40
Table 2.5……47
Table 2.6……54
Table 2.7……69
Table 2.8……71

CHAPTER 1: Introduction

1.1 OUR MODEL SYSTEM: HUMAN EMBRYONIC STEM CELLS

1.1.1 Human Embryonic Stem Cells as a Platform to Study Human Development

Understanding our basic biology as human beings is at the core of our innate curiosity about life, but also at the core of our necessity to find cures for developmental defects and cancer development. It has long been a priority to understand how genes function and how they can provoke abnormalities in disease states by performing controlled human genetic studies. Due to the obvious lack of material to study human development, gene function has been traditionally studied in model systems such as *Drosophila* and mouse, which share substantial functional conservation with human genes (99%) and major similarities in development, physiology and anatomy (Zhu and Huangfu, 2013). It is reassuring that many gene homologs in mouse and human behave similarly, as is the case for the Hox genes responsible for controlling body plan configuration in the developing embryo (Mallo et al., 2010). Considerable insight into human genetics has been gained from these mouse models, particularly in aspects pertaining to cell fate specification and tissue morphogenesis.

However, studying the specifics of human development poses a unique challenge, for it is more complex and not entirely conserved from mouse. For example, there are obvious differences in gestation period, morphology, and the spatial-temporal regulation of gene expression (Zhu and Huangfu, 2013). Additionally, around 1% of human genes have no recognizable mouse homologs. Thus, it is possible to miss the crucial function of a gene in human

development even after extensive study of the mouse homolog. For example, patients deficient for the hypoxanthine-guanine phosphoribosyl transferase (Hprt) gene develop Lesch-Nyhan syndrome, a disease characterized by increased uric acid production and neurological dysfunction. However, the Hprt knockout mice do not reflect these symptoms (Finger et al., 1988). Thus, mouse models do not always reflect human development or disease.

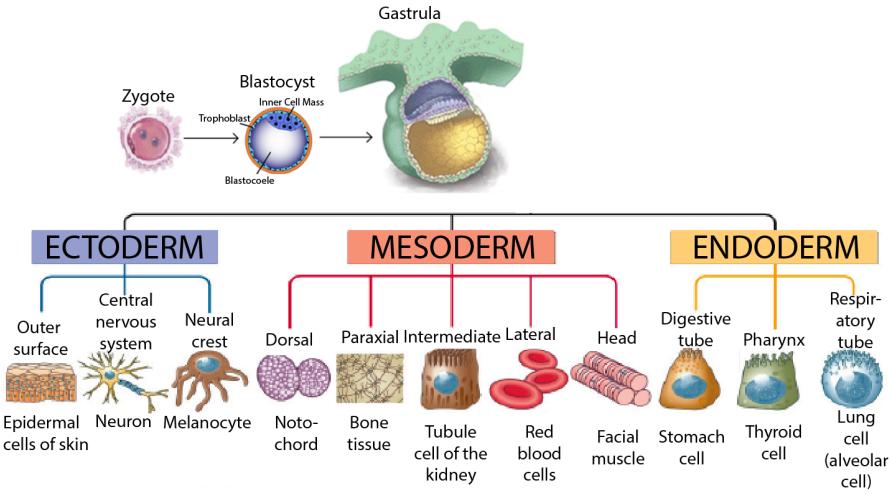


Figure 1.1. Human Development. Development of a complex organism is an intricate process that requires many signals and control checks. An organism begins to form from a single fertilized cell called the zygote. This zygote then divides and forms the blastocyst which contains the embryonic stem cells that will go on to form all three germ layers (ectoderm, mesoderm, and endoderm), and subsequently every single cell in our bodies.

The need for a better model system that recapitulates human development more intimately is critical. Human pluripotent stem cells have two properties that render them ideal for human genetic studies: 1) they have self-renewing ability, which provides us with an unlimited amount of material to work with, and 2) they are pluripotent, which means that they can be coaxed into any cell type of the body. The applications of hESCs are vast; studying human embryonic development ‘in a dish’, disease modeling, and cell

replacement therapies being the main ones. We will focus our study on human embryonic stem cells (hESCs) but it is worth mentioning that induced pluripotent stem cells (iPSCs), that are pluripotent stem cells derived from somatic cells of an adult patient (Takahashi and Yamanaka, 2006), can also be used for human genetic studies and developmental modeling.

To this end, the ability to differentiate hESCs into cell types of interest is of crucial importance. In recent years, much progress has been achieved in differentiating hESCs into different cells types such as pancreatic β-cells, neuroectoderm, and specialized neuronal cells, amongst many other cell types (Fattahi et al., 2016; Qi et al., 2017; Rezania et al., 2014; Williams et al., 2012). These directed differentiation protocols have been informed by mouse studies on signaling pathways in mouse development with much success, albeit some limitations (Zhu and Huangfu, 2013). For example, the formation of mature cell types still poses a challenge, as well as the variable ability of different hESC lines to differentiate into distinct cell types. Nevertheless, considerable progress is being made with directed differentiation protocols.

To take full advantage of the hESC model system one must be able to manipulate it genetically to study the underlying mechanisms of human development and of human patient phenotypes. Traditional homologous recombination (HR) used in mouse genetics is not efficient in hESCs since the rate of HR is extremely low (Hockemeyer and Jaenisch, 2010). Some studies have shown success with traditional gene targeting in hESCs (Zwaka and Thomson, 2003), but these studies were extremely time consuming and required the generation of very large constructs which were many times unsuccessful (Hockemeyer and Jaenisch, 2016). This challenge has recently been overcome by the advent of site-specific nucleases (SSN) that

dramatically increase the rate of HR (Rouet et al., 1994), namely Transcription Activator-Like Effector Nuclease (TALENs) and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) technologies.

1.1.2 A New Genetic Toolset to Study Gene Function in hESCs: CRISPR/Cas9 and TALENs

At the intersection of human embryonic stem cells (hESCs) and genome editing lies the unique opportunity to study human development and disease in the lab. In hESCs we are now able to alter the genome effectively and precisely using SSN's ('DNA scissors'), which are enzymes that cut the backbone of DNA and cause either small insertions or deletions by Non-Homologous End Joining (NHEJ), or precise genome editing by Homology Directed Repair (HDR) using a ssDNA donor template (Doudna and Charpentier, 2014). These technologies have allowed researchers to perform gain-of-function and loss-of-function studies in hESCs efficiently and accurately, and boosted confidence in the use of hESCs as a human genetic model system.

Our laboratory has pioneered a rapid, multiplexable, and highly efficient genome-engineering platform in hESCs, called iCRISPR (inducible CRISPR), which takes advantage of both TALEN and CRISPR/Cas9 technologies (Gonzalez et al., 2014). We generated an inducible Cas9-expressing cell line by integrating a Cas9 expression cassette into the AAVS1 locus using TALENs. The AAVS1 locus is a safe harbor locus like the ROSA26 locus in mice (Smith et al., 2008). In our system, Cas9 is induced by expression of doxycycline so that genome editing only requires the expression or delivery of the single guide RNA (sgRNA), enhancing the efficiency of CRISPR targeting.

Our iCas9-HUES8 cell line allows the generation of biallelic knockout hESCs for loss-of-function studies, as well as homozygous knockin hESCs that harbor specific nucleotide alterations for disease modeling.

Work from our lab has uncovered human-specific phenotypes for pancreatic differentiation genes such as NGN3 and GATA6 using the iCRISPR platform (Shi et al., 2017; Zhu et al., 2016). For example, NGN3 null hESC mutants differentiated towards the pancreatic lineage consistently give rise to ~0.5% of insulin-positive cells, whereas Ngn3 has been shown to be essential for β-cell development in the mouse and no insulin-positive cells are detected *in vivo* (Gradwohl et al., 2000; Zhu et al., 2016). This phenotypic difference raises the possibility of an unidentified NGN3-independent pro-endocrine pathway(s), and highlights the importance of probing gene function in human ESCs.

Another advantage of using hESCs and gene editing tools is that it allows the study of haploinsufficient human disease genes. Modeling a haploinsufficient disease gene is very difficult in the mouse since it usually requires inactivating both alleles of the mouse ortholog to recapitulate the disease phenotype. Our lab studied the contribution of a heterozygous GATA6 inactivating mutation that has been observed in many patients with pancreatic agenesis, a rare birth defect characterized by the lack of a pancreas or a small one (Lango Allen et al., 2011). Inactivation of only one allele of GATA6 in the mouse does not cause apparent pancreatic defects (Morrisey et al., 1998), but our lab showed that heterozygous GATA6 mutants are compromised in definitive endoderm formation, as well as formation of pancreatic progenitor and glucose-responsive β-cells (Shi et al., 2017). Thus, our hESC model system is amenable to studying human-specific genetic phenotypes.

Other labs have also taken advantage of gene editing in hESCs. Recently, knockout studies on both DNMT1 and EZH2, which are genes that have been extensively studied in the mouse, rendered surprisingly different phenotypes in hESCs. In the case of DNMT1, it was found to be required for hESC viability (Liao et al., 2015), although not being essential for mouse ESCs (Tsumura et al., 2006). Similarly, EZH2 knockout hESCs are not viable (Collinson et al., 2016) despite being nonessential in mESCs (Shen et al., 2008). These studies highlight the value of an hESC model system that allows human-specific requirements to be studied.

Powerful gene editing techniques in hESCs have opened the doors into unforeseen territory, and will deepen our understanding of human development and disease states. A model system tends to be as powerful as its genetic tools, and hESCs are now amenable to powerful genome editing tools.

1.2 MICRORNA's & DICER1

The hESC platform has become an attractive model system for studying human genetics. For example, it has enabled the study of DICER1, the RNase III protein responsible for processing microRNAs into their functionally mature forms, in human primed ESCs.

1.2.1 MicroRNA Biogenesis & Function

MicroRNAs are endogenous short noncoding RNAs 19-25nt long that regulate gene expression at the post-transcriptional level (Bartel, 2004). Since their discovery in *C. elegans* by Victor Ambros (Lee et al., 1993), thousands of microRNAs have been identified in a wide range of organisms from plants to

mammals. They are evolutionarily conserved and are thus recognized as one of the essential regulators in the control of many biological processes including development, metabolism, and homeostasis (Bartel, 2009). Aberrant microRNA expression is associated with human disease, including cancer (Chang and Mendell, 2007).

MicroRNAs represent ~4% of the human genome and regulate the expression of more than a third of the protein-coding genes at the post-transcriptional level (Bentwich et al., 2005). About half of the microRNAs are in intergenic regions controlled individually or as a cluster by their own promoters, and the other half lie within protein-coding genes and are co-transcribed with their host genes or from miRNA-specific promoters (Bartel, 2004, 2009; Rodriguez et al., 2004).

MicroRNAs are transcribed by RNA-polymerase II into long transcripts called primary miRNA (pri-miRNA) (Lee et al., 2004a). These transcripts undergo multiple sequential endonucleolytic cleavage steps that ultimately result in ~20nt long mature hairpin molecules. First they are processed by the RNase III enzyme Drosha (Lee et al., 2003) and the RNA-binding protein Dgcr8 in the nucleus into ~60-70nt long hairpin precursor miRNAs (pre-miRNAs) (Denli et al., 2004; Gregory et al., 2004; Han et al., 2004; Landthaler et al., 2004). Then they are actively transported from the nucleus into the cytoplasm through the export receptor Exportin-5 (Lund et al., 2004; Yi et al., 2003), where the pre-miRNAs are cleaved by the RNAse III enzyme Dicer1 into 19-25 nt double-stranded mature miRNAs (Hammond et al., 2000). Finally, either the 3p or 5p arm of the double stranded RNA is bound by an RNA-binding protein of the Argonaute family (Ago1-4) forming the RNA-induced silencing complex (RISC) (Khvorova et al., 2003; Schwarz et al.,

2003). Endogenous small interfering RNAs (endo-siRNAs) are also cleaved by Dicer1 and loaded into the RISC complex but are derived from long double-stranded RNAs from different genomic loci and bypass Dgcr8/Drosha cleavage (Elbashir et al., 2001; Zamore et al., 2000).

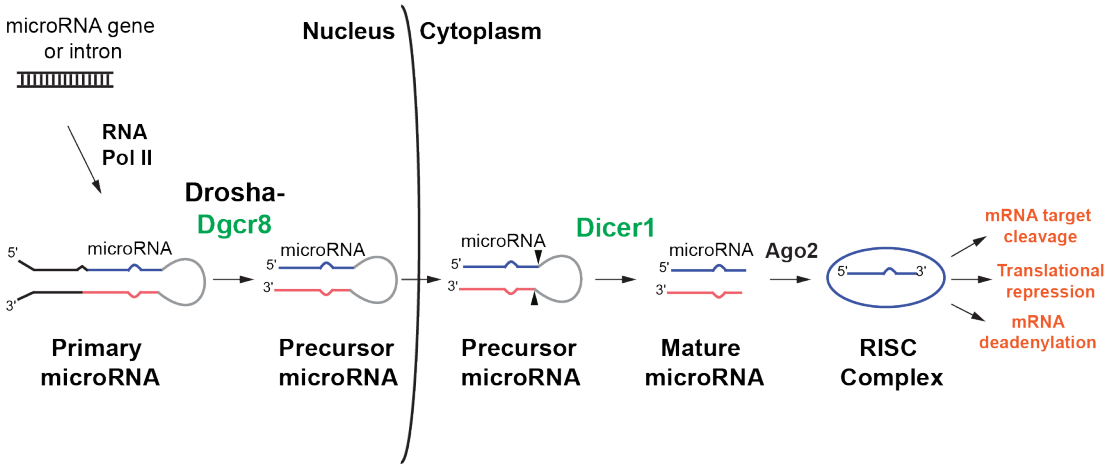


Figure 1.2. MicroRNA Biogenesis Pathway. In green, Dgcr8 and Dicer1, two RNaseIII enzymes required for functional microRNA maturation.

The RISC complex bound to the miRNA localizes to the 3' untranslated region (UTR) of target mRNAs through complementary base pairing (Bartel, 2004). Nucleotides 2-8 from the 5'end of the microRNA (seed region) are crucial for target identification. Whether the RISC favors degradation, destabilization, or translation inhibition depends on the free energy of the base pairing between the small miRNA and the mRNA target (Baek et al., 2008; Djuranovic et al., 2012; Guo et al., 2010). Perfect complementarity between miRNAs and their target mRNAs result in Ago2 degradation of the target (Brennecke et al., 2005; Lewis et al., 2005). Important to note that only Ago2 has slicer activity and therefore is the only Argonaute that can degrade target mRNA (Liu et al., 2004). Partial complementarity leads to repression of the mRNA translation at the initiation step, or to sequestration of the target

mRNAs into cytoplasmic processing bodies where they undergo polyA-nuclease degradation via deadenylation pathways (Guo et al., 2010). miRNAs function primarily through translational inhibition and target destabilization, whereas siRNAs result in target degradation. However, there are rare exceptions to this rule where mature miRNAs cleave targets directly (Yekta et al., 2004), and siRNAs behave like miRNAs (Doench et al., 2003).

1.2.2 Dicer1 Protein

Dicer1, an RNase III endonuclease, was first discovered for its role in generating the small interfering RNAs (siRNAs) that mediate RNA interference (RNAi) (Bernstein et al., 2001). Since then, at least three distinct functions of Dicer1 in the RNAi pathway have been delineated. First, as previously mentioned, its role in processing gene-encoded miRNA precursor hairpins (pre-miRs) into mature miRNAs (Grishok et al., 2001; Hutvagner et al., 2001; Ketting et al., 2001), and long dsRNAs into small interfering RNAs (siRNAs) (Elbashir et al., 2001; Knight and Bass, 2001; Zamore et al., 2000). Second, its role in loading small RNAs onto Argonaute (Ago) proteins to form the RNA-induced silencing complex (RISC), a function called RISC-loading (Maniataki and Mourelatos, 2005; Pham and Sontheimer, 2005). And third, Dicer1's role in serving as a scaffold in protein-protein interactions between RNAi cofactors in the RISC loading complex (RLC), the RISC itself, and other complexes involved in endo-RNAi mechanisms (Duchaine et al., 2006; Lee et al., 2004b).

Although Dicer1's most well-known function is in microRNA biogenesis, it is also involved in the endogenous siRNA pathway (endo-siRNA). EndoRNAi is elicited by transcribed genomic sequences that fold onto themselves or base-pair with other transcripts to generate double-stranded RNA molecules

that are processed by Dicer1 into endogenous siRNAs (endo-siRNAs) (Flamand and Duchaine, 2012; Okamura and Lai, 2008). This pathway then affects gene expression and genome organization at the transcriptional or post-transcriptional levels and, although similar to miRNAs, they can also modify histone marks and DNA methylation. Endo-siRNA targets in mammalian cells include a wide variety of repeated sequences, such as centromeric repeats, transposable elements and Alu repeats, but also include protein-coding genes (Fukagawa et al., 2004; Reinhart and Bartel, 2002; White and Allshire, 2004).

Higher order organisms like vertebrates that evolved alternate antiviral strategies retained only one Dicer gene and lost Dicer2, which is presumably involved in host defense against RNA viruses in *Drosophila* (Kurzynska-Kokorniak et al., 2015; Mukherjee et al., 2013). In humans, *DICER1* is located on chromosome 14q32.13, it contains 26 protein coding exons, and it is made up of 1,922 aminoacids. It is a 220kD multidomain enzyme harboring an amino (N)-terminal helicase domain (homologous to DExD/H-box helicases), a DUF283 domain (domain of unknown function), a PAZ (Piwi-Argonaute-Zwille) domain, two RNase III domains (RNase IIIa and IIIb) and a dsRNA-binding domain (dsRBD) (Macrae et al., 2006a; Macrae et al., 2006b; Zhang et al., 2002; Zhang et al., 2004). From this *DICER1* gene, multiple transcripts have been identified as a result of alternative transcription initiation sites and alternative splicing. To date, four full-length DICER1 proteins have been identified comprising all 1,922 aminoacid residues. These four variants differ only in their 5’ and 3’ non protein-coding sequences as they maintain the coding regions. Numerous shorter alternative splice variants have been identified as well. Some of them encode short proteins retaining only the N- or

C-terminus of DICER1, and other variants do not code for protein. The function of these shorter variants has not been elucidated yet (Kurzynska-Kokorniak et al., 2015).

1.2.3 MicroRNA Expression in Naïve and Primed ESCs

Embryonic stem cell (ESC) lines can be derived from mouse and human embryos. These include mouse ESCs (mESCs), mouse epiblast-like stem cells (mEpiSCs), and human ESCs (hESCs) (Greve et al., 2013). These stable cell lines can be passaged indefinitely in culture and differentiate into all three embryonic germ layers (endoderm, ectoderm and mesoderm). mESCs are derived from the inner cell mass (ICM) of embryos in the blastocyst stage (Evans and Kaufman, 1981; Martin, 1981). These mESCs are characterized by a specific self-reinforcing circuitry of transcription factors that allows the long term self-renewal and maintenance of pluripotency *in vitro*. As mESCs remain pluripotent, they integrate and contribute to the embryo proper when they are injected into blastocysts, generating chimeras.

Interestingly, although mESCs and hESCs are derived from the blastocyst stage of pre-implantation embryos, they seem to differ in their transcriptional program and appear to represent different developmental stages in embryonic development (Brons et al., 2007; Tesar et al., 2007; Thomson et al., 1998). Molecular profiling, X-chromosome activation status, morphological observations and growth conditions indicate that hESCs resemble mEpiSCs more than mESCs (Pera and Tam, 2010). mEpiSCs are derived from post-implantation mouse embryos or by differentiating mESCs in the presence of Activin and Fgf. mEpiSCs, unlike mESCs, are unable to contribute to chimeras and upregulate early differentiation markers, but they

can still form teratomas (Pera and Tam, 2010). The discovery of these two different pluripotent states led to the idea that hESCs and mEpiSCs share a ‘primed’ pluripotent state and mESCs represent a ‘naïve’ pluripotent state (Brons et al., 2007; Tesar et al., 2007; Weinberger et al., 2016).

The microRNA profile of pluripotent stem cells (PSCs) has been extensively studied (Bar et al., 2008; Calabrese et al., 2007; Houbaviy et al., 2003; Marson et al., 2008; Mathieu and Ruohola-Baker, 2013; Morin et al., 2008; Stadler et al., 2010). Interestingly, a single microRNA family sharing an AAGUGC seed sequence and known as the ESC-specific cell cycle-regulating (ESCC) family of miRNAs dominates the miRNA landscape in PSCs (Wang et al., 2007). This is in stark contrast to most somatic cells which express a diversity of miRNA molecules. Individual ESCC miRNAs are expressed from different clusters, mainly the miR-290-295 (miR-371-373 in human), miR-302-367, and miR-17-92b clusters, and these clusters contain both ESCC, ESCC-like (shifted by one base) and non-ESCC miRNAs. Consequently, although expressed from distinct loci, this ESCC miRNA superfamily comprises between 20% and 50% of all miRNAs in PSC lines (Greve et al., 2013).

The differences between mESCs, mEpiSCs, and hESCs are also distinguishable in their microRNA expression profiles. In mESCs, the clusters miR-290-295 (human analog miR-371-373) and miR-17-92 are the predominant contributors to the miRNA profile (Marson et al., 2008), whereas hESCs and mEpiSCs mainly express the miR-302-367 cluster (Bar et al., 2008; Jouneau et al., 2012), and the miR-290-295 (=miR-371-373) and miR-17-92b are expressed at lower levels. During mouse development, miR-290 is ubiquitously expressed in the zygote, but later become restricted to the extraembryonic structures such as the placenta and the yolk sac (Parchem et

al., 2015; Parchem et al., 2014; Tang et al., 2007). However, miR-302 expression starts after implantation ~(E) 5.5 in the epiblast and is co-expressed with miR-290 until embryonic day (E) 7.5, when miR-290 is downregulated and miR-302 expression persists localized to anterior structures until ~(E) 10.5 (Card et al., 2008; Houbaviy et al., 2003; Parchem et al., 2015; Parchem et al., 2014; Tang et al., 2007).

Consistent with their primed pluripotent state, hESCs and mEpiSCs upregulate miRNAs commonly found in differentiated tissues (Chiang et al., 2010; Jouneau et al., 2012). Upon differentiation, the ESCC miRNA-containing clusters are silenced (Suh et al., 2004), demonstrating that these clusters are specific to early embryonic tissues.

1.3 MICRORNA FUNCTION IN DEVELOPMENT

1.3.1 Transcriptional Control of MicroRNA Expression in ESCs

Maintaining appropriate levels of miRNAs in ESCs is crucial to maintain self-renewal and to deterr differentiation, but also to 'prime' the cells for differentiation. The well-defined pluripotency circuitry of Oct4, Nanog, and Sox2 in the inner cell mass (ICM) and in ESCs (Avilion et al., 2003; Chambers et al., 2003; Mitsui et al., 2003; Nichols et al., 1998; Niwa et al., 2000) co-occupy promoters of many miRNAs and protein-coding genes in ESCs. They bind and activate promoters of pluripotency-associated genes, including their own promoters (Boyer et al., 2006; Marson et al., 2008), thus maintaining a state of self-renewal. The Oct4/Nanog/Sox2 axis also represses genes involved in lineage commitment by co-occupying these promoters with

Polycomb-group protein (PcG), priming them for differentiation (Bernstein et al., 2006; Boyer et al., 2006).

Three research groups in 2008 found that Oct4, Nanog, and Sox2 positively associated with the ESCC miRNA cluster promoters (mmu-miR-290-295, hsa-miR-371-373, miR-302-367, and miR-106a-363 clusters), as well as the promoters of differentiation-inducing miRNAs such as the let-7 family (Barroso-delJesus et al., 2008; Card et al., 2008; Marson et al., 2008). The expression of pri-let-7 family of miRNAs is elevated in mESCs (Thomson et al., 2006) but processing of these primary transcripts into their functional forms is inhibited by Lin28 (Thornton and Gregory, 2012). Lin28 is also under transcriptional control of Oct4/Nanog/Sox2 and it is downregulated upon differentiation (Marson et al., 2008; Moss and Tang, 2003). High levels of immature let-7 transcripts in ESCs thus function as a priming mechanism for rapid differentiation upon Lin28 downregulation that doesn't require immediate transcription of the let-7 genes (Greve et al., 2013; Marson et al., 2008).

Furthermore, the ESCC microRNAs inhibit the let-7 family from silencing the pluripotency program of ESCs, thus promoting their self-renewal (Melton et al., 2010). Contrary to the ESCC miRNAs, let-7 is a suppressor of cell-cycle progression (Johnson et al., 2007) and causes ESCs to accumulate in the G1 phase. Blelloch's group identified an additional 5 microRNAs that promote differentiation by inducing accumulation at the G1 phase, and that are antagonized by the ESCC miRNAs: miR-26a, miR-99b, miR-193, miR-199a-5p, and miR-218 (Wang et al., 2013b).

1.3.2 Global microRNA Function in Mouse Development

As previously discussed, Dicer1 is required for the last endonucleolytic cleavage step that precursor miRNAs undergo before forming functionally mature miRNAs. Thus, to study global microRNA function *Dicer1* is a good candidate to knockout (Greve et al., 2013). *Dicer1* knockout in the mouse was established by Emily Bernstein and colleagues and demonstrated that Dicer1 is essential for early mouse development (Bernstein et al., 2003). Specifically, they replaced exon 21 (RNase IIIa domain) with a neomycin resistant gene, and then carried out timed heterozygous matings to determine at which developmental stage *Dicer1* is required. *Dicer1*-/- embryos collected at days (E) 9.5, 10.5, and 11.5 displayed empty or necrotic epithelium. At E8.5 they also found empty or necrotic epithelium but it was surrounded by a normal yolk sac. At E7.5 the embryos appeared small and morphologically abnormal, although embryonic and extraembryonic regions were clearly demarcated. At this stage, *Dicer1*-/- embryos did not stain for the primitive streak marker Brachyury, which means that Dicer1 is required before the body plan is configured during gastrulation. Moreover, about half of the embryos predicted from mendelian inheritance were observed at this developmental stage, presumably because a fraction of the embryos die at an earlier timepoint (Bernstein et al., 2003).

Contrary to the finding that Dicer1 is required for gastrulation, a more recent study showed that gastrulation is initiated in the epiblast of *Dicer1*-/- mice but with a day of delay causing smaller embryos compared to wildtype (Spruce et al., 2010). Upon further investigation, they found this delay is caused not by decreased proliferation but rather from increased apoptosis. These findings were corroborated in null mutant embryos and in embryos where *Dicer1* has been removed specifically from the epiblast. MicroRNAs,

therefore, are responsible for controlling cell survival at E5.5. It is possible that differences in genetic background of the embryos analyzed in these two studies could account for the increased severity of the phenotype observed in the work of Bernstein and colleagues (Spruce et al., 2010).

Nevertheless, both studies observe embryonic lethality of *Dicer1* null embryos and conclude that Dicer1 is essential for mammalian development (Bernstein et al., 2003; Spruce et al., 2010). One confounding factor in these *Dicer1* null studies is the fact that mature miRNAs are still present in *Dicer1*-/- mice in preimplantation stages due to perdurance of maternally-contributed Dicer1. In fact, three members of the miR-290 family were still expressed in *Dicer1*-/- blastocyts (Spruce et al., 2010). Thus, the exact timing for Dicer1 requirement during development remains elusive, so is Dicer1's requirement in preimplantation ESCs.

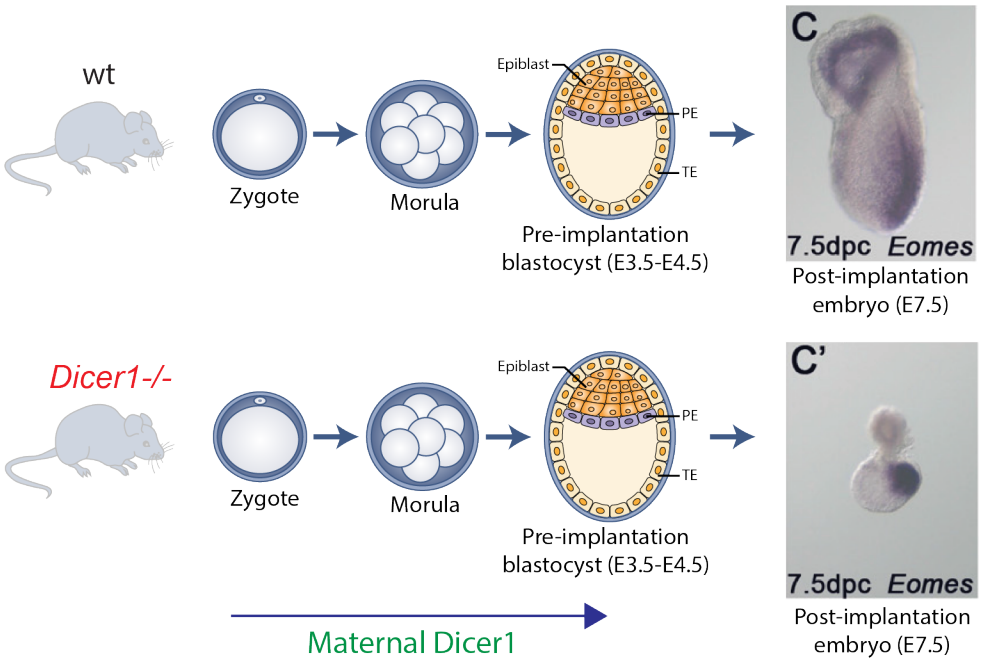


Figure 1.3. Dicer1-/- Mice Are Embryonic Lethal and Embryos Show Abnormalities at E7.5. Figure adapted from Spruce, T, et al., Dev. Cell, 2010. Due to maternal Dicer1 presence in early embryogensis, Dicer1 cannot be ruled out to be required in the pre-implantation stage.

1.3.3 Global microRNA Function in Naïve and Primed Cultured ESCs

To bypass the confounding factor of maternally contributed Dicer1 in preimplantation stages, researchers set out to study Dicer1's role in mESCs, which are derived from E3.5 blastocysts and are said to be in a "naïve" pluripotency state (Kanellopoulou et al., 2005; Murchison et al., 2005; Weinberger et al., 2016). *Dicer1* knockout mESCs are viable, display normal ES morphology and express the pluripotency circuitry. However, *Dicer1* knockout cells proliferate slower than wildtypes and accumulate in the G1 phase of the cell cycle, unlike the ICM in *Dicer1* knockout embryos (Spruce et al., 2010). Upon induction of embryo body formation, *Dicer1* knockout mESCs fail to upregulate differentiation markers and downregulate pluripotency markers (Kanellopoulou et al., 2005). In line with these findings, *Dicer1* knockout mESCs failed to generate teratomas upon subcutaneous injection in mice. Together, these reports suggest that miRNAs are not required for the survival or self-renewal of ESCs, but are instead important for proliferation and cell cycle structure, and are essential for proper differentiation and silencing of the pluripotency network.

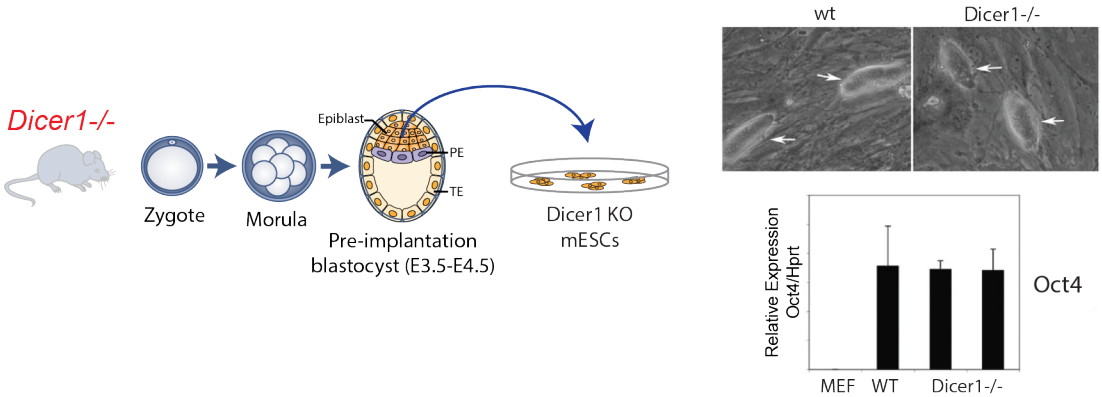


Figure 1.4. MicroRNAs Are Not Essential for Mouse Embryonic Stem Cell Survival. Figure adapted from Kanellopoulou, C, et al, Genes & Dev., 2005.

Blelloch's group took a slightly different approach. Since *Dicer1* is required for both miRNA and siRNA pathways, Blelloch's group opted to knockout Dgcr8, which is only required in the miRNA pathway for pri-miRNA processing. Dgcr8 knockout mESCs present comparable phenotypes when compared with *Dicer1* knockout mESCs, albeit less severe (Wang et al., 2007). For example, upon differentiation, Dgcr8 knockout mESCs show reduced levels rather than a complete lack of differentiation markers. These phenotypic differences may suggest additional functional roles for Dicer1 in mESCs (Greve et al., 2013), or may suggest a role for mirtrons, which are a class of intronic miRNAs that bypass Drosha/Dgcr8 processing but require Dicer1 (Okamura et al., 2007; Ruby et al., 2007).

However, these mESC studies do not explain the post-implantation cell death observed *in vivo* (Spruce et al., 2010). To this end, mEpiSCs and hESCs, which are cells primed for differentiation and resemble the epiblast of the post-implantation embryo (Nichols and Smith, 2009), are better suited to study. In fact, *Dicer1* deletion in mEpiSCs mimicks the *in vivo* observation and leads to increased apoptosis as measured by Annexin and Cleaved Caspase-3 (Pernaute et al., 2014). Specifically, miRNAs seem to be required to control apoptosis via the mitochondrial pathway by regulating the expression of the proapototic protein BIM in the primed pluripotent state. Changes in expression of pluripotency or lineage-specific markers were not observed, neither were changes in the activation of signaling pathways required for EpiSC maintenance (FGF and Activin), demonstrating that regulation of cell survival is the primary role of miRNAs in the primed pluripotent state in mice.

Although *DICER1* knockout studies have not yet been performed in hESCs, Ruohola-Baker's group attempted shRNA knockdown of DICER1 and

DROSHA in hESCs (Qi et al., 2009). Reduced levels of miRNAs caused proliferation defects, but surprisingly, had no effect on survival of the cells resembling the *Dicer1* knockout mESCs. There are several possible explanations for these primed hESCs behaving differently than their mouse counterparts, mEpiSCs. One possibility is that the DICER1 and DROSHA knockdowns did not reach sufficient levels to elicit the proper phenotype. The other explanation is that human and mouse primed ESCs may behave differently. Proper characterization of *DICER1* knockout hESCs is required to discern between these possibilities.

1.3.4 Function of miRNA-290-295 Cluster (miRNA-371-373 in human) in ESCs

The miRNA-290-295 cluster, and its human homolog miRNA-371-373, is amongst the most abundant miRNA classes in ESCs (Calabrese et al., 2007; Marson et al., 2008; Suh et al., 2004). It is a cluster specific to ESCs, as its expression increases during preimplantation development, remains elevated in undifferentiated ESCs, and decreases upon differentiation (Houbaviy et al., 2003; Tang et al., 2007).

In mESCs, individual miRNAs from these clusters directly target the cell-cycle inhibitors p21 and LATS2, which normally repress the G1-S phase transition (Wang et al., 2008). mESCs have a unique cell cycle structure characterized by a short G1 phase, and thus it is believed that the miR-290 cluster encourages the cell cycle transition at the G1-S phase, shortening it and thus allowing faster cycling of the cells (Wang et al., 2008; White and Dalton, 2005).

In addition to its role in cell proliferation, this cluster plays a role in differentiation of mESCs by promoting *de novo* DNA methylation (Benetti et al., 2008; Sinkkonen et al., 2008). miR-290 miRNAs repress the expression of Rbl2, which is a transcriptional repressor of *de novo* DNA methyltransferases. *Dicer1*-/- mESCs downregulate *de novo* DNA methyltransferases like Dnmt1, Dnmt3a, Dnmt3b and Dnmt3L and are unable to *de novo* methylate the Oct3/4 locus, which is at least partially caused by miR-290 loss-of-function (Benetti et al., 2008; Sinkkonen et al., 2008). The inability of mESCs to differentiate in *Dicer1* and *Dgcr8* knockout backgrounds is likely attributed to the absence of the miR-290 cluster, causing defective *de novo* DNA methylation during differentiation especially at the Oct3/4 locus (Sinkkonen et al., 2008). Another group found that miR-290 cluster miRNAs repress the Wnt pathway inhibitor Dkk1 and thus prevent ESC differentiation to mesoderm and germ cells *in vitro* (Zovoilis et al., 2009).

Furthermore, Kanellopoulou and colleagues found that bivalent genes, specifically of the Hox family, were upregulated in *Dicer1*-/- mESCs, and that miR-290 is responsible for Hox gene repression by inactivating Ash1l (Kanellopoulou et al., 2015), which normally evicts Polycomb during differentiation (Miyazaki et al., 2013; Tanaka et al., 2011). Another group found that *Dicer1* is required for bivalent gene repression and polycomb group binding at bivalent genes, which is rescued by miR-294, a member of the miR-290-205 cluster (Graham et al., 2016). Thus, the miRNA-290-295 cluster is required for repression of bivalent genes in ESCs to prevent differentiation.

The miRNA-290 cluster indirectly antagonizes the maturation of the let-7 family of miRNAs, which are strong inducers of differentiation, to maintain pluripotency and prevent differentiation in mESCs (Melton et al., 2010). By an

unknown mechanism, miR-290 maintains the expression of Lin28, an RNA-binding protein that blocks the maturation of let-7. The miRNA-290 cluster and the let-7 family show a mutually exclusive expression pattern: let-7 is not expressed in ESCs, but upon induction of differentiation, the loss of pluripotency transcription factors causes miR-290 levels to decrease, which downregulates Lin28. Consequently, let-7 maturation is reestablished increasing its expression and further promoting differentiation by repressing pluripotency factors that normally increase miR-290 levels (Melton et al., 2010). The mutually exclusive expression pattern and opposing functions of these clusters form a regulatory feedback loop that regulates the balance between self-renewal and differentiation.

A more recent study links the miR-290 cluster with a pro-survival function in mESCs (Zheng et al., 2011). Using gain-of-function and loss-of-function studies, the authors confirmed that the miR-290 cluster prevents apoptosis during exposure to genotoxic stress by targeting Caspase 2 and Ei24, both known pro-apoptotic factors activated by p53 (Zheng et al., 2011). This anti-apoptotic function of miR-290 in mESCs can be particularly relevant during physiological stress in embryonic development, and could partially explain the increased apoptosis observed in post-implantation *Dicer1*-/- embryos (Spruce et al., 2010).

miR-290 cluster knockout studies have also been performed. Null mESCs are viable and indistinguishable from wildtypes, suggesting that the miR-290 cluster is dispensable for maintaining the pluripotent state in ESCs (Medeiros et al., 2011). Nevertheless, miR-290 null mice resulted in partially penetrant embryonic lethality (Medeiros et al., 2011). Three quarters of the miR-290 cluster nulls are lost during development before birth, specifically

between E11.5 and E18.5, but embryo abnormalities are observed as early as E8.5. Abnormalities include embryos localized outside the yolk sac or developmentally delayed embryos (Medeiros et al., 2011).

Overall, the miRNA-290-295 cluster is important for development and studies from mESCs have highlighted its role in cell cycle, differentiation, bivalent gene repression, and apoptosis. However, the human homolog 371-373 has not been fully elucidated in the context of ES cell biology, particularly in the primed state, i.e. hESCs.

1.3.5 Function of miRNA-302-367 Cluster in ESCs

The miRNA-302-367 cluster belongs to the ESCC family of miRNAs that are highly expressed in ESCs and that function, at least partially, to promote proliferation in ESCs through a characteristically short G1 phase (Wang et al., 2008). Although expressed in mESCs, it is the most abundant miRNA cluster in primed hESCs and mEpiSCs (Bar et al., 2008; Jouneau et al., 2012; Suh et al., 2004).

Studies have shown that this cluster functions in the cell cycle to promote the G1-S transition. For example, transfection of miR-302b, miR-302c, and miR-302d in Dgcr8 deficient mESCs decreases the percentage of cells in the G1 phase to wildtype levels (Wang et al., 2008). This effect is orchestrated through suppression of a cyclinE-Cdk2 complex inhibitor, Cdkn1a (Wang et al., 2008). Another group found that ectopic expression of miR-302a in primary and transformed cell lines leads to a decrease in the percentage of cells in the G1 phase by translational inhibition of the CyclinD1-Cdk4 complex, which is a G1-2 transition regulator (Card et al., 2008). Similarly, studies in human embryonal carcinoma cells (hECCs) show that miR-302b indirectly

regulates the expression of OCT3/4 and directly regulates expression of Cyclin D2 protein, which is a developmental regulator during gastrulation (Lee et al., 2008).

Besides its role in cell proliferation, the miR-302 cluster has been implicated in differentiation. Prior reports have shown that the TGF-β/Nodal signaling pathway is critical in the maintenance of pluripotency in hESCs (James et al., 2005; Vallier et al., 2005). miR-302 has been shown to target the TGF-β/Nodal inhibitors LEFTY1/2 in hESCs and thus promoting TGF-β/Nodal signaling (Rosa et al., 2009; Tabibzadeh and Hemmati-Brivanlou, 2006). As an upstream regulator of the TGF-β/Nodal signaling pathway, miR-302 leads to a delay in early hESC differentiation and facilitates the maintenance of pluripotency (Barroso-delJesus et al., 2011). Another group showed that by indirectly raising Nodal activity, miR-302 promotes the mesendodermal lineage at the expense of neuroectoderm differentiation, thus preventing the default neural induction of hESCs (Rosa et al., 2009).

miR-302 is also implicated in bone morphogenetic protein (BMP) signaling. BMP signaling activation promotes differentiation towards mesoderm and trophectoderm lineages in hESCs at the expense of neuroectoderm (Xu et al., 2002; Zhang et al., 2008). miR-302 was shown to inhibit neural differentiation and enhance trophectodermal fate by promoting BMP signaling (Lipchina et al., 2011). Since neural induction is initiated by inhibition of the TGF-β and BMP signaling pathways, by targeting inhibitors of both pathways, and therefore activating them, miR-302-367 both suppresses neural induction and promotes pluripotency (Lipchina et al., 2011). This strategy makes sense as in the absence of cell-cell signaling, hESCs are fated

for neural differentiation at the expense of mesendodermal and trophectodermal lineages (Munoz-Sanjuan and Brivanlou, 2002).

Consistent with miR-302's role in suppressing neural induction, mouse knockout ESCs for miR-302a-d show an expansion in neuroepithelium at (E) 9.5 caused by early increases in proliferation and later decreases in apoptosis (Parchem et al., 2015). miR-302a-d knockout mice are embryonic lethal with a fully penetrant phenotype, unlike the partial penetrance of miR-290 knockout mice (Medeiros et al., 2011; Parchem et al., 2015). Although expression of the miR-302 cluster begins ~(E) 5.5, the phenotype is only evident at (E) 9.5, presumably because of overlapping expression and function with the miR-290 cluster between (E) 5.5 and (E) 7.5. In fact, double knockout mice for miR-290 and miR-302 clusters led to an earlier phenotype prior to neurulation, which could not be appropriately staged because of the highly abnormal morphology of the mutant embryos. Nevertheless, this demonstrated a role for miR-302 before neurogenesis that is redundant with miR-290, but that has not been studied yet (Parchem et al., 2015).

To conclude, the miR-302 cluster is essential for embryonic development. Studies have shown it plays important roles in the cell cycle, differentiation, and maintenance of pluripotency. Furthermore, there are indications of an earlier role for miR-302 during gastrulation which has not yet been elucidated.

1.3.6 Function of miRNA-17-92 Cluster in ESCs

The miR-17-92 cluster of miRNAs is enriched in ESCs (Houbaviy et al., 2003; Suh et al., 2004) and contains miRNAs with similar seed sequences to those of the ESCC family of miRNAs (Marson et al., 2008). Unlike other ESCC

miRNAs, this cluster is expressed in adult tissues as well, and is therefore not specific to the epiblast. This cluster forms a family with two paralogues: the miR-106b-25 and the miR-106a-363 clusters, which also share similar seed sequences as the ESCC miRNAs (Mogilyansky and Rigoutsos, 2013). During mouse development, miR-17-92 and miR-106b-25 clusters, but not miR-106a-363, are expressed in mESCs and mid-gestation embryos (Ventura et al., 2008).

In line with other ESCC family miRNA, five miRNAs from the miR-17-92 family (miR-19a, miR-20a, miR-20b, miR-93, and miR-106a) were able to partially rescue the proliferation defects observed in Dgcr8-/- mESCs (Wang et al., 2008). Thus, miR-17-92, like other ESCC miRNAs, play an activating role in proliferation of mESCs.

Deletion of miR-17-92 in mice leads to fully penetrant neonatal lethality and specific defects in heart, lung and B cell development (Ventura et al., 2008). Redundancy with miR-106b-25 was confirmed by a double knockout which showed severe cardiac defects, increased apoptosis and embryonic death by mid-gestation - much earlier than miR-17-92 deletion alone (Ventura et al., 2008). Important to note, miR-17-92 is the only cluster in this family that is essential for proper development. The miR-106b-25 and miR-106a-363 single knockout mice were normal and viable, and it was only in the context of miR-17-92 deletion that miR-106b-25 caused a phenotype (Ventura et al., 2008). No phenotype was observed in embryos earlier than mid-gestation, presumably due to overlapping functions of the ESCC miRNAs which are abundantly expressed in early development.

Of particular interest, Tyler Jacks' group found that miR-17-92 negatively regulates the expression of the proapoptotic gene Bim during

development (Ventura et al., 2008), which leads to apoptosis upon overexpression (Bouillet and Strasser, 2002).Thus, the miR-17-92 cluster seems to have a protective role against apoptosis at least during B cell development and lymphomagenesis.

It is no wonder that roles in accelerating proliferation and diminishing apoptosis may be co-opted in cancer. miR17-92, also called OncomiR-1, has been extensively studied in the context of cancer and is known to be upregulated in this setting (Mendell, 2008), but will not be discussed in depth here.

miR17-92 has been implicated in cancer, as well as in proliferation and apoptosis in the developing embryo. However, its function in ESCs, particularly during early embryonic development, has not yet been elucidated (Mallanna and Rizzino, 2010), presumably due to redundancy issues.

1.3.7 miRNAs in Reprogramming

The *in vitro* derivation of induced pluripotent stem cells (iPSCs) from somatic cells, a process denominated reprogramming (Takahashi and Yamanaka, 2006), has been an area of extensive research due to its potential clinical applications. Briefly, a set of defined transcription factors (such as Oct4, Sox2, Klf4, and c-Myc, collectively dubbed OSKM) is introduced into somatic cells in DNA vectors, and these factors induce a dedifferentiation process in some cells that adopt the morphology and molecular profile of ESCs, reestablishing both self-renewal and pluripotency (Greve et al., 2013). Since miRNAs are important regulators of the pluripotent state and can be easily overexpressed or suppressed, many groups investigated the potential of miRNAs in the generation of iPSCs.

Studies of gene expression patterns during reprogramming have identified several miRs from the miR-17-92, miR-106a-363, and miR-106b-25 clusters as being activated early in OSKM reprogramming of mouse embryonic fibroblasts (MEFs) (Chen et al., 2012; Li et al., 2011; Polo et al., 2012). This activation is induced by direct c-Myc binding to their promoters (Liao et al., 2011). On the other hand, the miR-290 cluster reactivates late in reprogramming, together with other core pluripotency factors (Chen et al., 2012; Judson et al., 2009; Polo et al., 2012). Although the promoter of this cluster is bound by all OSKM factors, it is not activated until epigenetic remodeling occurs (Judson et al., 2009; Marson et al., 2008). Other microRNA clusters with known roles in PSCs, such as the miR-302 and miR-200 clusters, have been reported to be activated during reprogramming (Chen et al., 2012; Lee et al., 2013; Li et al., 2011; Liao et al., 2011; Polo et al., 2012; Samavarchi-Tehrani et al., 2010; Wang et al., 2011). The timing of their activation has been inconsistent, which probably reflects the stochasticity of the conversion process.

Exogenous overexpression of a myriad of miRNAs has been shown to enhance reprogramming of MEFs (Greve et al., 2013). The first study focused on the miR-290 and miR-302 clusters due to their important roles in pluripotency. Only the ESCC seed-containing miRNAs in these clusters were able to enhance iPSC colony formation in a dose-dependent manner (Judson et al., 2009). These miRNAs not only increased reprogramming efficiency up to 70%, but also reduced the generation of partially reprogrammed or transformed colonies (Judson et al., 2009). In fact, all ESCC-containing miRNAs, even ESCC-like miRNAs, enhanced reprogramming, showing the

robustness of ESCC miRNA-complemented reprogramming (Li et al., 2011; Liao et al., 2011), even in human fibroblasts (Subramanyam et al., 2011).

Reprogramming of fibroblasts is initiated by a mesenchymal-to-epithelial transition (MET) (Samavarchi-Tehrani et al., 2010). ESCC miRNAs have been shown to cause an acceleration of MET (Liao et al., 2011; Subramanyam et al., 2011), which likely enhances reprogramming. The miR-200 family, also highly expressed in PSCs, is also well-known for its role in promoting MET and, as expected, enhances MET during reprogramming (Samavarchi-Tehrani et al., 2010). However, it does not increase the number of ultimate iPSC colonies generated, suggesting that the ESCC miRNAs must play additional roles in reprogramming (Liao et al., 2011; Samavarchi-Tehrani et al., 2010). In fact, it has been shown in human that the ESCC miRNAs reduce cellular senescence by directly targeting P21 and P130, activate Oct4 expression through targeting of NR2C2, and inhibit a range of epigenetic modifiers (Banito et al., 2009; Hu et al., 2013; Lee et al., 2013; Lin et al., 2011; Subramanyam et al., 2011). It is unknown whether these are the only targets of ESCC miRNAs during reprogramming or whether there are more.

1.4 THESIS AIMS

While Dicer1 is undoubtedly essential for mammalian development, the exact timing for its requirement has not yet been established due to the confounding factor of maternal Dicer1 in early mouse development. Since *Dicer1*-/- mESCs are viable, it seems like the requirement for *Dicer1* occurs after implantation. Individual miRNA function at these early stages has not been thoroughly assessed since miRNA knockout studies are hindered due to redundancy in the miRNA network. Thus, precise miRNA functional studies in

early embryonic development have been lacking. To address this knowledge gap, we set out to knockout *DICER1* in human ESCs, which model stem cells of the post-implantation epiblast, to assess global miRNA requirement in the primed state. Also, we set out to specifically dissect human miRNA requirements at this stage. We report that *DICER1* is essential in hESCs unlike in mESCs, and that this likely reflects a unique requirement for *DICER1* in primed versus naïve pluripotency. Additionally, we set up a system to interrogate individual and miRNA clusters that bypasses redundancy problems and allows proper assessment of their functions in human primed pluripotency.

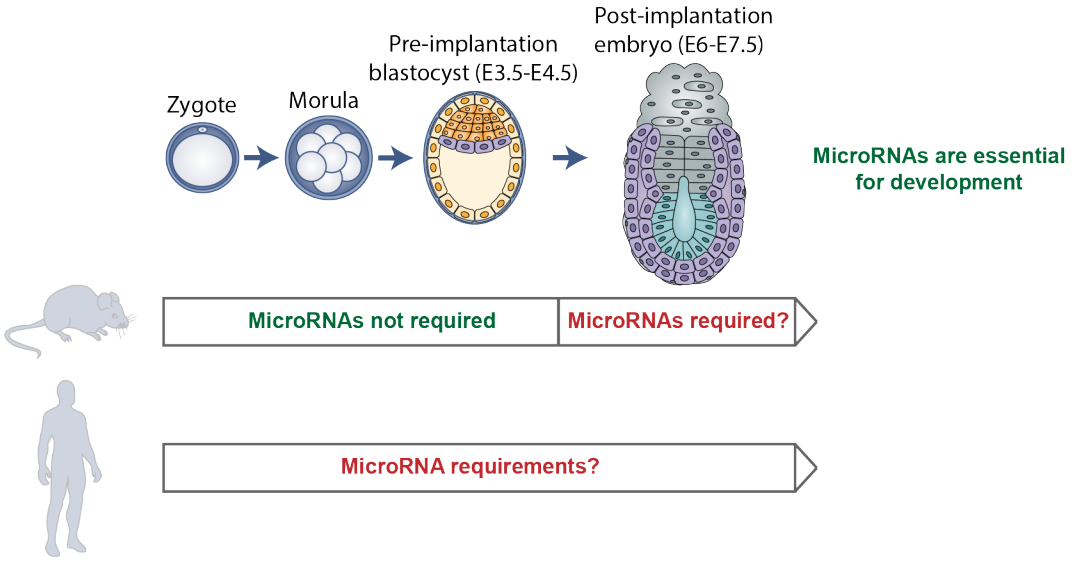


Figure 1.5. Schematic of Early Human Development Showing the Acquired and Missing Knowledge on MicroRNA Requirements at Different Stages. In green, known facts; in red, knowledge that is lacking.

CHAPTER 2: DICER1 Requirement and miRNA Function in hESCs

2.1 INTRODUCTION

MicroRNAs (miRNAs) are ~20-24 nucleotides (nt) long noncoding RNAs that bind target messenger RNAs (mRNAs) through partial complementarity with their 3'UTRs to destabilize or inhibit translation of the transcripts (Ebert and Sharp, 2012). This gene regulatory system is largely enabled by the RNase III enzyme Dicer1 that is required for the biogenesis of canonical miRNAs into their functional forms (Kurzynska-Kokorniak et al., 2015). Genetic deletion of Dicer in the mouse is embryonic lethal, and embryos are grossly abnormal by E7.5 during the postimplantation stage (Bernstein et al., 2003; Spruce et al., 2010), emphasizing the importance of microRNAs in early mouse development.

Nevertheless, *Dicer1*-/- mESCs exhibit normal mESC morphology and express the proper pluripotency markers corroborating the *in vivo* observation that miRNAs are required not for the establishment of pluripotent cells but rather for differentiation (Kanellopoulou et al., 2005; Spruce et al., 2010). However, *Dicer1*-/- and Dgcr8-/- mESCs exhibit a proliferation and cell cycle defect not observed *in vivo*, and accumulate in the G1 phase of the cell cycle (Kanellopoulou et al., 2005; Murchison et al., 2005; Wang et al., 2007). miRNA rescue experiments in Dgcr8-/- mESCs have shown that the cell cycle of mESCs is regulated by members of the miR-290-295 (its human analog miR-371-373), miR-302, and miR-17-92b clusters, commonly referred to as the ESC-specific cell cycle-regulating (ESCC) family of miRNAs (Wang et al., 2008). However, miR-290-295 and miR-17-92b cluster mouse knockouts do not mimic *Dicer1* embryonic lethal phenotype (Medeiros et al., 2011; Ventura

et al., 2008). Thus, the *Dicer1*-null lethality remains unexplained due to the redundant nature of miRNAs, which poses a technical challenge since multiple clusters would have to be deleted in order to recapitulate the global knockout phenotype (Greve et al., 2013).

Although the field has inferred some microRNA function from studying mESCs, it is not clear which miRNAs are responsible for the *Dicer1* knockout phenotype, and more importantly, human ESC (hESC) microRNA function is still lacking (Greve et al., 2013). To the best of our knowledge, *DICER1* has not been knocked out in hESCs yet, despite several indications that we should study hESC-specific phenotypes and not rely solely on mESC data. mESCs and hESCs differ in their microRNA expression profiles. Firstly, in mESCs the miR-290-295 and miR-17-92b clusters are the dominant contributors to the miRNA profile, whereas hESCs express miRNAs from the miR-302-367 cluster predominantly (Bar et al., 2008; Jouneau et al., 2012). Different miRNA profiles might indicate different functional roles for miRNAs in these different cell types. Secondly, recent knockout studies in hESCs have shown significant differences in the regulation of the stem cell state between mouse and human ESCs. For example, Dnmt3b and EZH2 are unexpectedly essential in hESCs but not in mESCs (Collinson et al., 2016; Liao et al., 2015), highlighting the need to study hESC-specific knockout phenotypes. Thus, it is crucial that we study the effects of *DICER1* deletion to uncover novel and human-specific miRNA functional roles in the context of human development.

Here we report the generation and characterization of DICER1-deficient human ESCs. Our study shows that *DICER1* is essential for hESC self-renewal, revealing its crucial function in early human development. We also identify an unexpected difference in the requirement for *DICER1* between

human and mouse ESCs, since *DICER1* is not essential in mESCs. Furthermore, our study provides a powerful tool for the discovery of specific interactions and functions of individual and families of miRNAs in regulating hESCs during early human embryonic development.

2.2 MATERIALS & METHODS

2.2.1 hESC Culture

All experiments were performed on HUES8 iCas9 lines (Gonzalez et al., 2014). HUES8 iCas9 lines, and the lines derived from these (like TRE-DICER1*), were cultured on irradiated mouse embryonic fibroblast (iMEFs) feeder layers in DMEM/F12 (without HEPES) supplemented with 20% KnockOut Serum Replacement, 1X Non-Essential Amino Acids, 1X GlutaMAX, 100 U/ml Penicillin/100 µg/ml Streptomycin (Gemini), 0.055mM 2-mercaptoethanol, and 10ng/mL recombinant human basic FGF. *DICER1* knockout lines derived from the TRE-DICER1* line (B2, C2, D11, F2) were maintained on doxycycline (2 µg/ml) unless otherwise specified. Cultures were passaged at a 1:12 split ratio every 4-6 days using TrypLE. Upon passaging of the cells, 5µM Rho-associated protein kinase (ROCK) inhibitor Y-27632 (Selleck Chemicals, S1049) was added to the culture media. All cell culture reagents were purchased from Life Technologies. We routinely checked for mycoplasma contamination in our hESCs cultures to ensure that the cells were not infected. The hESC culture work was approved by the Embryonic Stem Cell Research Committee (ESCRO) and conducted following NIH guidelines.

2.2.2 Generation of gRNAs for CRISPR/Cas9 Targeting

To generate CRISPR/Cas9 expression vectors targeting specific genomic loci, 20 base pairs (bp) of sequence located 5' of the PAM sequence was cloned into piCRg Entry (Gonzalez et al., 2014) following an established protocol (Cong et al., 2013). In summary, piCRg Entry was digested with BbsI, dephosphorylated and gel purified. A pair of oligos including the 20bp homology (**Table 1**) were annealed and phosphorylated, generating BbsI overhangs that can be cloned into the BbsI-digested and dephosphorylated vector. For production of gRNAs, we first generated a T7-gRNA IVT template by adding the T7 promoter to the gRNA sequence in the piCRg Entry vector through PCR amplification using CRISPR-specific forward primers and a universal reverse primer (**Table 1**). Alternatively, for Cr7-9 gRNAs we designed a 120-nt oligo including the T7 promoter and the full-length gRNA sequence. This oligo was used as a template for PCR amplification using T7 and gRNA universal primers (**Table 1**). T7-gRNA PCR products were used as templates for IVT using the MEGAshortscript T7 kit (Life Technologies). The resulting gRNAs were purified using the MEGAclear kit (Life Technologies), eluted in RNase-free water and stored at −80°C until use.

2.2.3 Establishment of Clonal Mutant Lines Using CRISPR/Cas9

A detailed description on how to establish clonal mutant lines from HUES8 iCas9 hESCs has been described previously (Gonzalez et al., 2014; Zhu et al., 2014). To generate mutant lines, iCas9 hESCs were treated with doxycycline (2 μg/ml) for 1 day before gRNA transfection, or in the case of the TRE-DICER1* lines, they were kept on doxycycline continuously. Cells were dissociated using TrypLE Select, replated onto iMEF-coated plates and

Table 1. CRISPR gRNA sequences, T7E1 primers, and Sanger sequencing primers.

CRISPR gRNA	DICER1 Region Targeted	CRISPR gRNA Target Sequence (5'→ 3')	Oligo Sequence for Cloning into piCRg Entry Vector (5'→ 3')
Cr1	PAZ domain	GAGAAGTCTGAAGCTCGCAT	F: CACCGAGAAGTCTGAAGCTCGCAT R: AAACATGCGAGCTTCAGACTTCTC
Cr2	PAZ domain	TTCCATTTAAATACCTACCT	F: CACCGTCCATTTAAATACCTACCT R: AAACAGGTAGGTATTTAAATGGAC
Cr3	RNase IIIa domain	GCTAACAGAGACTTTTGCCA	F: CACCGCTAACAGAGACTTTTGCCA R: AAACTGGCAAAAGTCTCTGTTAGC
Cr4	RNase IIIa domain	GATCTGCTGAAACTTCAACG	F: CACCGATCTGCTGAAACTTCAACG R: AAACCGTTGAAGTTTCAGCAGATC
Cr5	RNase IIIa domain	CCTGTGATGGCCGTAATGCC	F: CACCGCTGTGATGGCCGTAATGCC R: AAACGGCATTACGGCCATCACAGC
Cr6	RNase IIIa domain	CAGGAGAGTACATTCATCGC	F: CACCGAGGAGAGTACATTCATCGC R: AAACGCGATGAATGTACTCTCCTC
Cr7	RNase IIIa domain	ACTTACCCTGATGCGCATGA	N/A
Cr8	RNase IIIa domain	ACTCTGTCAAACGCTAGTGA	N/A
Cr9	RNase IIIa domain	CGCTAGTGATGGATTTAACC	N/A

Oligonucleotides used for *in vitro* transcription form piCRg Entry templates

CRISPR gRNA	Oligo (5'→3')
Cr1	TAATACGACTCACTATAGGGAGAAGTCTGAAGCTCGCAT
Cr2	TAATACGACTCACTATAGGTTCCATTTAAATACCTACCT
Cr3	TAATACGACTCACTATAGGGCTAACAGAGACTTTTGCCA
Cr4	TAATACGACTCACTATAGGGATCTGCTGAAACTTCAACG
Cr5	TAATACGACTCACTATAGGCCTGTGATGGCCGTAATGCC
Cr6	TAATACGACTCACTATAGGGAGGAGAGTACATTCATCGC
Universal - RNA-tracr R	AAAAGCACCGACTCGGTGCC

Oligonucleotides used for *in vitro* transcription from ssDNA templates:

CRISPR gRNA	110nt Oligo Template for IVT (5'→3')
Cr7	TAATACGACTCACTATAGGGACTTACCCTGATGCGCATGAGTTTTAGAGCTAGAAA TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGT
Cr8	TAATACGACTCACTATAGGGACTCTGTCAAACGCTAGTGAGTTTTAGAGCTAGAAA TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGT
Cr9	TAATACGACTCACTATAGGGCGCTAGTGATGGATTTAACCGTTTTAGAGCTAGAAA TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGT
Universal - T7 F	TAATACGACTCACTATAGGG
Universal - gRNA R	AAAAGCACCGACTCGGTGCC

PCR Primers for T7E1 and Sanger Sequencing:

CRISPR locus	Forward Primer (5'→ 3')	Reverse Primer (5' →3')	Sequencing Primer (5' →3')
PAZ domain (Cr1-2)	TCAACTTTAGA AGGCGGAAGCTC	TACAATGCTAAAATCACAGCCCAC	GATCGAGGTGCCTCTTCTATT
RNase IIIa (Cr3-6)	CGGGTGGAAAA AATCTATTGACAG	CTCATATATGAAAGGCGGCCC	CGGGTGGAAAAAATCTATTGACAG
RNase IIIa (Cr4&6 for TRE-DICER1*)	CGGGTGGAAAA AATCTATTGACAG	GCATGATACGTTCTCATCCTC	CGGGTGGAAAAAATCTATTGACAG
RNase IIIa (Cr7-9)	GCGATGAATGTA CTCTCCTGA	CAATACTCATCAACTGCCAGG	GCGATGAATGTACTCTCCTGA

transfected in suspension with gRNA using Lipofectamine RNAiMAX (Life Technologies) per manufacturer's instructions. gRNAs were added at a final concentration of 10nM. A second transfection was performed a day later. Two days after the last gRNA transfection, hESCs were dissociated into single cells, passed through a 40μm cell strainer, and replated at a low density (2,000 cells/10cm dish). Genomic DNA was also extracted at this stage to perform T7E1 assay to verify that the gRNAs had generated indels. Cells were grown until clonal colonies became visible (~12-14 days). 48-96 colonies were picked when colonies were in excess. Colonies were mechanically disaggregated and replated into individual wells of 96-well plates. Clonal lines were expanded and analyzed by Sanger sequencing to identify mutant clones. PCR and sequencing primers are listed in **Table 1**. Selected clonal lines carrying desired mutations were further expanded and frozen down.

2.2.4 Construction of pIND_DICER1_∆PAM-Cr4 and -Cr6

Construction of low copy pBR_ENT vector: pCR-Blunt II-TOPO vector (Thermo Fisher Cat. K280002) was partially digested with PvuII and BspHI to replace the high copy pUC origin of replication and the Zeocin resistance gene by a PvuII-BspHI fragment obtained from pMC1403 vector (http://seq.yeastgenome.org/vectordb) containing the low copy pBR322 origin of replication. The resulting vector was then opened with ApaI and SacI in order to insert an ApaI-SacI digested (in bold) GenScript-synthetized DNA fragment (Piscataway, NJ, USA):

TTTAGGGCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTT
GCAACAAATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAA
AGCAGGCTCCGCGGCCGCCAGTGTGATGGATATCTGCAGAATTCAAGGG

Table 2. Primers for site-directed mutagenesis of DICER1*.

Cloning Construct	Forward Primer (5'→3')	Reverse Primer (5'→3')
Cr4	GCGAGTCCCCTGGTAAGCTACACGTT GAAGTTTCAGC	GCTGAAACTTCAACGTGTAGCTTACCA GGGGACTCGC
Cr6	CGAGAACCAGCCCCAGCCGAGCGAT GAATGTACTCTCCTG	CAGGAGAGTACATTCATCGCTCGGCTG GGGCTGGTTCTCG

TGGGCGCGCCGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAAAGC ATTGCTTATCAATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAA TCATTATTTGCCATCCAGATCCGAGCTCCCTT. This fragment contains attL1 and attL2 Gateway cloning sequences flanking a polylinker including SacII, NotI, EcoRV, PstI and EcoRI unique restriction sites.

Construction of pBR_ENT hsDICER vector: pBR_ENT was opened with EcoRV and dephoshorylated to clone the Xho-Acc65I digested, Klenow blunt ended Flag-tagged hsDICER cDNA from pCAGGS-Flag-hsDICER (a gift from Phil Sharp, Addgene plasmid #41584)(Gurtan et al., 2012)

Construction of pBR_ENT hsDICER∆PAM-Cr4 and -Cr6: To introduce silent mutations in the amino acid sequence of hsDICER disrupting the PAM sequence of RNACr4 and Cr6 we used the QuikChange II Site-Directed Mutagenesis Kit (Agilent Cat. #20052) following manufacturers guidelines, using pBR_ENT hsDICER as template and the primers specified on **Table 2**.

Construction of pIND_hsDICER ∆PAM-Cr4 and -Cr6: We used Gateway LR Clonase (Invitrogen Cat. #11791-020) to transfer hsDICER ∆PAM-Cr4, and hsDICER ∆PAM-Cr6 attL1-attL2 flanked inserts into the pINDUCER21 (a gift from Stephen Elledge & Thomas Westbrook, Addgene plasmid #46948) (Meerbrey et al., 2011).

2.2.5 Generation of TRE-DICER1 Lines*

Preparation of pIND_ΔPAM-Cr4 and -Cr6 virus: Plated 14 million 293T cells in 15cm dishes previously coated with collagen. Allowed cells to grow in the absence of Pen/Strep for two days. Thirty minutes before transfection, fresh media was added to the cells. Transfection was performed using JetPRIME Polyplus transfection reagent (VWR, Cat. No. 114-01) as per manufacturer's recommendations. To summarize, 20.4µg of pIND_hsDICER ΔPAM lentiviral vector was mixed with 8.16µg of psPAX2 and 2.04µg of pCMV-VSVG. This mix was then diluted in 1500µl of JetPRIME buffer, followed by the addition of 60µl of JetPRIME transfection reagent. The mix was incubated at room temperature for 10 minutes, and then added dropwise to one 15cm dish. The media was changed the following day, and the supernatant was harvested on days 2,3, and 4 post transfection. The three harvests were pooled together and passed through a 0.45µm filter. Virus was concentrated by ultracentrifugation for 1.5 hours at 4°C and 250,000 rpm, and resuspended in 400µl of hESC medium overnight at 4°C. The virus was then aliquoted and stored at -80°C.

Infection of HUES8 iCas9 lines: HUES8 iCas9 cells were grown to 60-70% confluency, and then disaggregated using TrypLE, counted, and resuspended at 1million/ml in hESC medium with 5µm of Rock inhibitor (Selleck Chemicals, S1049) and 10µg/ml of Protamine Sulfate (MP Biomedicals). 1ml of cells were plated per well of a 6well dish that was previously seeded with iMEFs. Infection was carried out overnight using 150µl of virus. Media was changed everyday until the cells were confluent and ready for FACS sorting of GFP+ colonies. GFP+ cells were plated at low density (2,500 cells) in 10cm dishes previously seeded with iMEFs, and grown for ~2 weeks until colonies were

visible and ready to pick. Five colonies were picked per line, and one was chosen to carry out experiments.

2.2.6 RNA Processing and Quantitative qRT-PCR

MicroRNA and total RNA were isolated using miRNeasy Mini Kit (Qiagen, Cat. No. 217004). For mRNA qRT-PCR, this was followed by cDNA synthesis using the High Capacity cDNA Reverse Transcription Kit (Life Technologies), and qRT-PCR was performed with QPCR SYBR Green low ROX Kit (Thermo Scientific, AB4322B). For microRNA qRT-PCR, we used Taqman MicroRNA Assays (Thermo Fisher Scientific, Cat. No. 4427975) per manufacturer's instructions. RNU-44 was used for normalization. We used the Applied Biosystems 7500 Real-Time PCR System for all qRT-PCRs. Primer

Table 3. qRT-PCR primers.

Gene	Forward Primer (5'→3')	Reverse Primer (5'→3')
DICER1 – total (DICER1_total_transcript_ (ex-ex) F & R)	AGCATGCCATCACCACATATC	CCAAGGCGATACAGATTACAG
DICER1* – exogenous (Flag_F & DICER_Flag_R=128nt)	TGGACTACAAAGACGATGACG	TTGTTGCCATGGCAGTCCAAA
OCT4	TGGTCCGAGTGTGGTTCTGTAA	TGTGCATAGTCGCTGCTTGAT
NANOG	GCTGGTTGCCTCATGTTATTATGC	CCATGGAGGAAGGAAGAGGAGAGA
SOX2	GGCAGCTACAGCATGATGCAGGAGC	CTGGTCATGGAGTTGTACTGCAGG
KLF4	TATGACCCACACTGCCAGAA	TGGGAACTTGACCATGATTG
ECAD	GCTGAGCTGGACAGGGAGGA	ATGGGGGCGTTGTCATTCAC
NCAD	CCACCTTAAAATCTGCAGGC	GTGCATGAAGGACAGCCTCT
ZEB1	GCACCTGAAGAGGACCAGAG	TGCATCTGGTGTTCCATTTT
ZEB2	AACAACGAGATTCTACAAGCCTC	TCGCGTTCCTCCAGTTTTCTT
TWIST1	TCCATTTTCTCCTTCTCTGGAA	GTCCGCGTCCCACTAGC
SNAIL	AGGTTGGAGCGGTCAGC	CCTTCTCTAGGCCCTGGCT
SLUG	AGATGCATATTCGGACCCAC	CCTCATGTTTGTGCAGGAGA
GATA4	AAAGAGGGGATCCAAACCAG	TTGCTGGAGTTGCTGGAAG
GATA6	GTGCCCAGACCACTTGCTAT	TGGAATTATTGCTATTACCAGAGC
SOX7	CATGCAGGACTACCCCAACT	ACTCACCCCTGTCCTCCTTC
FOXA2	GGGAGCGGTGAAGATGGA	TCATGTTGCTCACGGAGGAGTA
SOX17	GGCGCAGCAGAATCCAGA	CCACGACTTGCCCAGCAT
SOX1	AACACTTGAAGCCCAGATGGA	GCAGGCTGAATTCGGTTCTC
PAX6	TGGGCAGGTATTACGAGACTG	ACTCCCGCTTATACTGGGCTA
OTX2	CATGCAGAGGTCCTATCCCAT	AAGCTGGGGACTGATTGAGAT
GSC	AACGCGGAGAAGTGGAACAAG	CTGTCCGAGTCCAAATCGC
BRACHY	ACCCAGTTCATAGCGGTGAC	CCATTGGGAGTACCCAGGTT
GAPDH	GGAGCCAAACGGGTCATCATCTC	GAGGGGCCATCCACAGTCTTCT

sequences are specified in **Table 3**.

2.2.7 Western Blotting

Protein samples were collected from cell lysate homogenized in RIPA buffer in the presence of proteinase inhibitor cocktail (Cell Signaling Technology) and stored in -80ºC until use. Proteins were separated on a NuPAGE Novex Bis-Tris 3-8% Tris-Acetate protein gel (Life Technologies), and transferred to a PVDF membrane (Life Technologies) overnight at 4ºC. Blocking was performed with 5% milk in TBST. Membranes were incubated with primary antibodies overnight at 4ºC and HRP-conjugated secondary antibodies for 1 hour at RT. ECL reagents (Pierce) were obtained from GE Healthcare. Primary antibodies used were DICER1 (1:1000, Cell Signaling, D5F2 #5325), Flag (1:5000, Sigma Aldrich, Cat. No. F1804), AGO2 (1:1000, borrowed from Eric Lai's lab, monoclonal mouse Ab) and GAPDH (1:5000, Cell signaling, D16H11 #5174).

2.2.8 Northern Blotting

Total RNA were isolated using the miRNeasy Mini Kit as per manufacturer's recommendation (Qiagen, Cat. No. 217004). 20 µg of total RNA per lane were separated on 15% polyacrylamide 7M urea gels and transferred onto GeneScreen Plus membrane (Perkin Elmer) using A Trans-Blot SD Semi-Dry Cell (Biorad). The blots were UV crosslinked (Stratagene), baked at 80°C for 1hour and probes with γ-[^{32}P]-ATP labeled DNA oligonucleotide probes. The probe sequences are listed in **Table 4**. The blots were stripped and re-probed for multiple miRNAs and loading control (RNU44). Decade Marker RNA

(Thermo Fisher) was labeled with γ-[^{32}P]-ATP and used as size standard (10-100 bases).

Table 4. Northern Blot primers.

miRNA	Probe (5'→3')
miR-302c-3p	CCACTGAAACATGGAAGCACTTA
miR-200c-3p	TCCATCATTACCCGGCAGTATTA
miR-92a-1-5p	AGCATTGCAACCGATCCCAACCT
miR-92a-1-3p	ACAGGCCGGGACAAGTGCAATA
miR-367-3p	TCACCATTGCTAAAGTGCAATT
miR-17-5p	CTACCTGCACTGTAAGCACTTTG
DICER1(RNase III)	**F:** CTGGAGAGGTTACCATATCCA ***** **R:** GCAGACTTTCCCATTTGGCTT
RNU44	AGTTAGAGCTAATTAAGACCT

*****the PCR product of the F&R primers was used as the probe.

2.2.9 Alkaline Phosphatase and Immunofluorescence Staining

Cells were fixed with 4% paraformaldehyde in PBS for ≤10 minutes. For nuclear immunostaining, we washed once with PBS, and then permeabilized in PBS with 0.1% Triton (PBS-T) for 15min. Blocking was done using donkey blocking solution at RT for 5 min (5% donkey serum in PBS-T). Primary and secondary antibodies were diluted in blocking solution, and incubated at RT for 1 hour. The following primary antibodies were used: Nanog (1:100, Cosmobio Japan, Cat. No. REC-RCAB0004P-F), Oct3/4 (1:100, Santa Cruz, sc-8628), Sox2 Y-17 (1:100, Santa Cruz, sc-17320), Ki67 (1:1000, Vector Laboratories, VP-K451), PH3 (1:100, Cell Signaling, Ser10 #9701), and Csp3 (1:400, Cell Signaling, Asp175 #9661). AP staining was performed using Vector Red Alkaline Phosphatase Substrate Kit following manufacturer's guidelines (Vector Laboratories, SK-5100).

2.2.10 miRNA Mimic Rescue Screen and Quantification

For the miRNA screen, *DICER1* knockout hESCs (B2, D11) were dissociated with TrypLE and counted. They were replated at a concentration of 125,000 cells per ml in 24-well plates, and transfected in suspension with microRNA mimics (*mir*Vana® miRNA mimics, Thermo Fisher, Cat. No. 4464066) using Lipofectamine RNAiMAX (Life Technologies) per manufacturer's instructions. 5μM ROCK inhibitor Y-27632 (Selleck Chemicals, S1049) was added upon replating. The final concentration of each small RNA was 10nM, as used in (Wang et al., 2008). Cells were allowed to grow without Doxycycline for 6 days, at which point the cells were either fixed and stained for Alkaline Phosphatase, or counted with a cell counter.

2.2.11 Small RNA-Sequencing Analysis

For small RNA-Sequencing total RNA was isolated with the miRNeasy Mini Kit (Qiagen, Cat. No. 217004), and gel purified to collect RNAs <200nt. We cloned small RNA libraries from *DICER1*-Hypomorph (B4-1) and wildtype (A5) hESC lines in two biological replicates. Single pair 50bp, 70 million reads. After removing 5' and 3' adaptor sequences with cutadapt (Martin, 2011), we mapped small RNA-Seq reads from these datasets to UCSC *Homo sapiens* (hg38) genome assembly using Bowtie. Unmapped reads were iteratively trimmed one nucleotide each iteration retaining a read length of >=17nt, and then mapped to the genome using Bowtie with no mismatches, up to 30 iterations. The reads were required to match to the microRNA mature sequences with at least 15 nt overlap, and within 2 nt of the 5' end. We normalized small RNA-Seq by total microRNA counts in each library by the trimmed mean of M-values (TMM) normalization method in the edgeR/Limma Bioconductor library (Oshlack et al., 2010). We used the voom method of

Limma (Smyth, 2005) to correct for the Poisson noise due to the discrete counts of small RNA-seq. Differentially expressed miRNAs between Dicer-Hypomorph versus wild type were identified by a moderated t-test, and FDR (Benjamini- Hochberg) was estimated, using the Limma library (Smyth, 2005) in Bioconductor. The genes with FDR $\leq$ 0.1 and FC $\geq$ 2-fold were considered to be differentially expressed. To examine microRNA mature and star strand change in *DICER1*-Hypomorph, a proportion of microRNA 5p vs. 3p was calculated by 5p/(5p+3p) for each microRNA in each sample and proportion difference was calculated between *DICER1*-Hypomorph and wildtype.

2.2.12 RNA-Sequencing

For RNA-Sequencing total RNA was isolated with the miRNeasy Mini Kit (Qiagen, Cat. No. 217004) from HUES8-iCas9 (A5), *DICER1*-Hypo (B4-1, B4-4), *DICER1*-KO (D11+Dox) and *DICER1*-KO (D11-Dox) hESCs (n=2 for A5, B4-1 and B4-2, and n=3 for D11+Dox and D11-Dox each). RNA samples were submitted to the MSKCC Integrated Genomics Core for library prep and sequencing. Paired-end 50bp, 20 million reads each.

2.2.13 Naïve Stem Cell Culture

hESCs were cultured on irradiated mouse embryonic fibroblast (iMEFs) feeder layers as described in “hESC Culture” above, and then passaged and cultured using the Stem Cell Technologies RSet Medium (Cat. # 05970) as per manufacturer’s recommendation. The naïve cells were maintained on iMEFs and the cells were kept in a hypoxic chamber at 5% O_2 levels, except for the ~30min when the cells were passaged.

2.2.14 Flow Analysis

Intracellular marker. hESCs were disaggregated with TrypLE for 5 minutes and washed with cold FACS buffer: 5% Fetal Bovine Serum (Invitrogen, 261400799) in PBS. Cells were pelleted by centrifugation and washed again with FACS buffer. Each sample was resuspended in FACS buffer and incubated with a Live/Dead Fixable stain (Life Technologies L34964, 1:1000) for 30 minutes at RT. Afterwards cells were washed with FACS buffer and resuspended in fixation solution (eBioscience, 00-5523-00) for 1 hour at RT. Cells were washed with permeabilization buffer (eBioscience, 00-5523-00) and then resuspended in permeabilization buffer and the appropriate antibody: OCT4 (Santa Cruz, Cat. # sc-8628), DICER1 (Cell Signaling, Cat. # 5325S), Caspase-3 (Cell Signaling Asp175, Cat. # 9602S), Caspase-2 (EMD Millipore, Cat. # MAB3501), Caspase-8 (Cell Signaling, Cat. # 9746), AIF (Abcam, Cat. # ab32516), and BIM [Y36] (Abcam, Cat. # ab32158) for 1 hour at RT. Cells were washed with FACS buffer and then incubated with the appropriate Molecular Probes Alexa Fluor dye conjugated secondary antibody (Life Technologies, 1:500) for 1 hour at RT. Cells were then washed with FACS buffer and analyzed by FACS.

Surface marker. Confluent hESCs were disaggregated with TrypLE for 5 minutes and washed with cold FACS buffer. Cells were pelleted by centrifugation and washed again with FACS buffer. Each sample was resuspended in FACS buffer with the appropriate conjugated antibody: Tra1-60-FITC (BD Pharminogen, Cat. # 560380), Tra1-81-FITC (BD Pharminogen, Cat. # 560194), or SSEA3-FITC (BD Pharmingen, Cat. # 561145). Cells were incubated in FACS buffer with the antibody for 30 minutes on ice. After staining cells were washed two times with FACS buffer and resuspended in

FACS buffer with DAPI and analyzed by FACS.

Cell Cycle Analysis. hESCs were harvested in FACS buffer (5% FBS in PBS) at a concentration of 500,000 cells/ml. 1ml of cells was placed on a 15ml Falcon tube on ice and allowed to cool. Then, 3ml of cold (-20°C) ethanol was added dropwise while vortexing. Cells were fixed overnight at -20°C. Cells were washed twice with PBS, centrifuged at 3,500 rpm for 5 minutes at 4°C each time. Supernatant was then removed, and samples were resuspended in 200ul of RNase-Propidium Iodide (PI) solution (188ul of PBS and 2ul of RNase (5mg/ml) and 10ul of PI (1mg/ml). Samples were left in the dark overnight, then passed on a cell strainer, and analyzed by flow.

2.2.15 Teratoma Assay

Confluent hESCs grown on a MEF feeder layer were collected in 1mg/mL Collagenase type IV (Life Technologies, 17104-019) for 10min, then the collagenase was removed and the cells were scraped off and resuspended in hESC medium. Cells were spun for 5min and then resuspended in 400ul of PBS per 10cm dish. One quarter of the cells (100ul) from a confluent 10cm dish were injected subcutaneously to the dorsal flank of a SCID mouse. Palpable tumors were typically observed 1-2 months after injection. Tumor samples were usually collected in 2-3 months, fixed in 4% paraformaldehyde and processed for paraffin embedding and hematoxylin and eosin staining following standard procedures.

2.2.16 Colony Forming Assay

hESC cells were washed with PBS, treated with TrypLE for 5min at 37°C, then collected with fresh hESC medium, passed through a 40um cell strainer to

make sure cells were single cells, and then counted. Two-thousand cells were plated in a 10cm dish previously coated with iMEF and either treated with Rock inhibitor or not for the first 4 days. Cell were grown for two weeks, then stained with alkaline phosphatase (AP) and counted.

2.2.17 Suspension Culture

Cells were adapted to the E8 feeder-free condition, and after two passages, were transported into low attachment U-bottom 6-well plates, at a density of 5 million cells per well. Rock inhibitor was added on the first day of plating, and media was changed every day. The plates were maintained in rocking conditions at 95rpm in the incubator.

2.2.18 Neuroectoderm Differentiation

hESCs were plated on matrigel-coated (BD, 354234) dishes in hESC media with ROCK-inhibitor at a density of 180,000–200,000 cells/cm^2. After 12 hours, differentiation into neuroectoderm was initiated by switching to knockout serum replacement (KSR) media with 10 μM TGF-β inhibitor (SB431542, Tocris 161410) and 100 nM LDN (Axonmedchem, 1509). On day 1 and day 2 of differentiation, the media was removed and fresh KSR with 10 μM TGF-β inhibitor and 100 nM LDN was added. Starting on day 4 of differentiation an increasing amount of N2 media was added to the KSR media every two days, while maintaining 10 μM TGF-β inhibitor and 100 nM LDN. On day 4 a 3:1 mixture of KSR/N2 media was added. On day 6 a 1:1 mixture of KSR/N2 media was added, and on day 8, a 1:3 mixture of KSR/N2 media was added. The cells were isolated for analysis on days 4,6,8, and 10 of differentiation. KSR media contains Knockout DMEM (Invitrogen, 10829018), Knockout

Serum Replacement (Invitrogen, 10828028), 1X MEM Non-Essential Amino Acids (Life Technologies, 11140050), 1X GlutaMAX (Life Technologies, 35050079), and 2-mercaptoethanol (Life Technologies, 21985023). N2 media contains DMEM/F12 medium (Life Technologies, 12500-062), glucose (Sigma, G8270), sodium bicarbonate (Sigma, S5761), putrescine (Sigma, P5780), progesterone (Sigma, P8783), Sodium selenite (Sigma, S5261), apo-transferrin (Sigma, T1147), and insulin (Sigma, I2643).

2.2.19 Proliferation Assay

hESCs were washed with PBS, counted, and then 50,000 cells were plated per well of a 24-well plate previously coated with iMEF. Two replicates of each sample were plated per timepoint. Every 24hs the cells were disaggregated and counted, up to day 6.

2.2.20 Statistical Analysis

All values are shown as mean +/- SD. Data were analyzed using Student's t-test in GraphPad Prism. p-values of <0.05 (*) were considered significant.

2.3 RESULTS

2.3.1 Failure to Generate True DICER1 Null hESCs using iCRISPR

To investigate the role of DICER1 in human primed pluripotency, we used CRISPR/Cas9 to disrupt the catalytic domains of *DICER1* in hESCs. Nine guide RNAs (gRNAs) designed to target the PAZ and RNase IIIa functional domains of DICER1 were identified as resulting in high indel efficiencies by the T7 Endonuclease 1 (T7E1) assay, which recognizes and cleaves heteroduplexes formed from the hybridization between wildtype and mutant sequences (Mashal et al., 1995) (Figure 2.1A and 2.1B). These gRNAs were individually transfected into iCRISPR HUES8 hESC line in the presence of doxycycline (DOX), and colonies were then picked, expanded, and analyzed by Sanger DNA sequencing (Gonzalez et al., 2014). The efficiency of disrupting the target sequence within the *DICER1* coding region with

Table 5. Results of CRISPR gRNA targeting of DICER1 in iCas9 HUES8 hESCs. if, in frame allele; fs, frameshift allele.

			Mutant Alleles Per Clone						
			1		2				
					Compoud Heterozygotes			Homozygous	
Gene	Domain	CRISPR	In frame	Frameshift	if/if	fs/fs	if/fs	In frame	Frameshift
DICER1	PAZ	Cr1	4/36	3/36	2/36	0/36	0/36	0/36	0/36
		Cr2	2/9	0/9	0/9	0/9	0/9	0/9	0/9
	RNase IIIa	Cr3	2/21	4/21	0/21	0/21	1/21	0/21	0/21
		Cr4	4/96	2/96	11/96	0/96	38/96	0/96	2/96
		Cr5	6/48	7/48	0/48	0/48	0/48	0/48	0/48
		Cr6	4/79	15/79	0/79	0/79	2/79	0/79	1/79
		Cr7	1/59	4/59	0/59	0/59	0/59	0/59	0/59
		Cr8	1/86	8/86	0/86	0/86	1/86	0/86	0/86
		Cr9	0/30	1/30	0/30	0/30	1/30	0/30	0/30

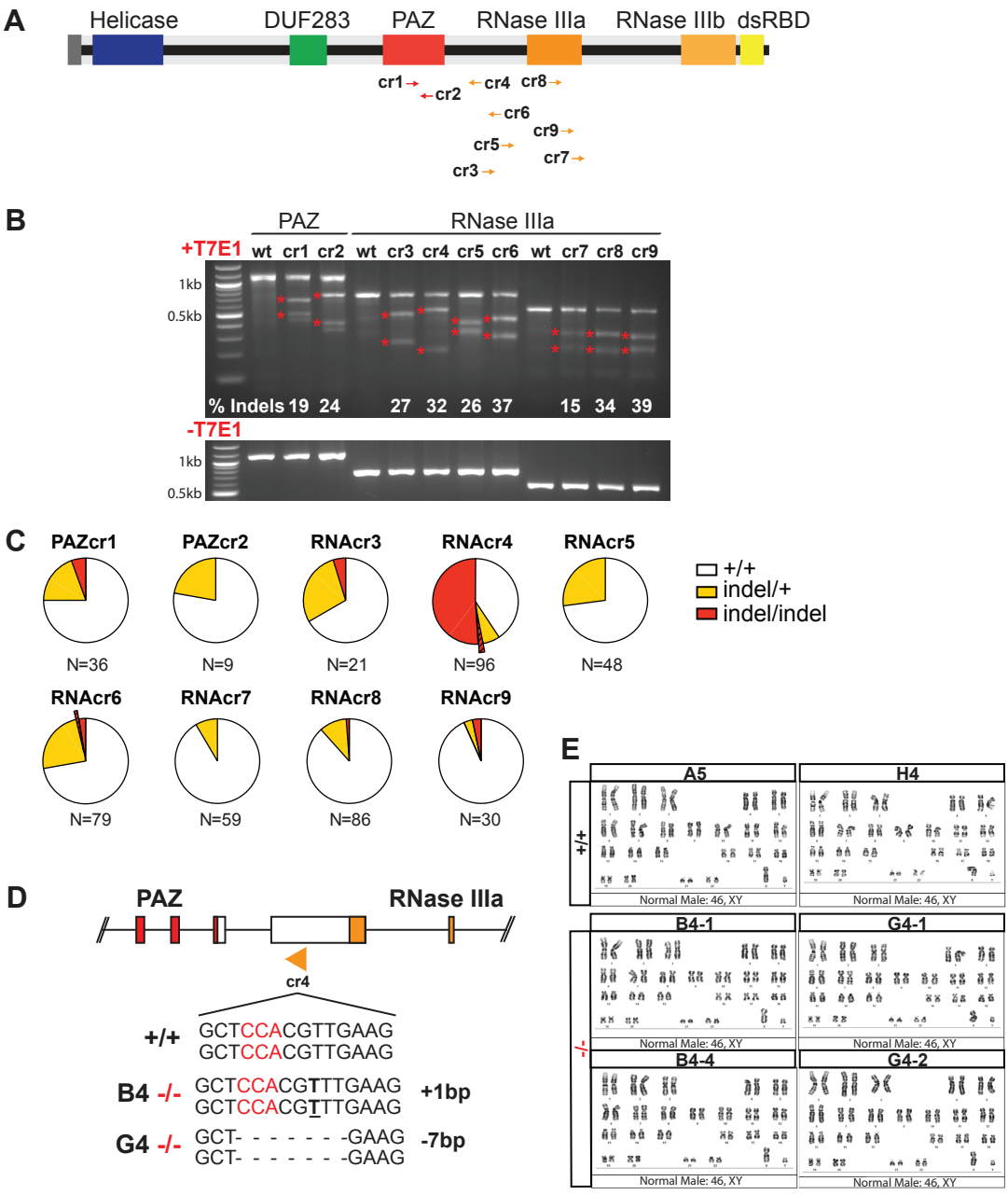


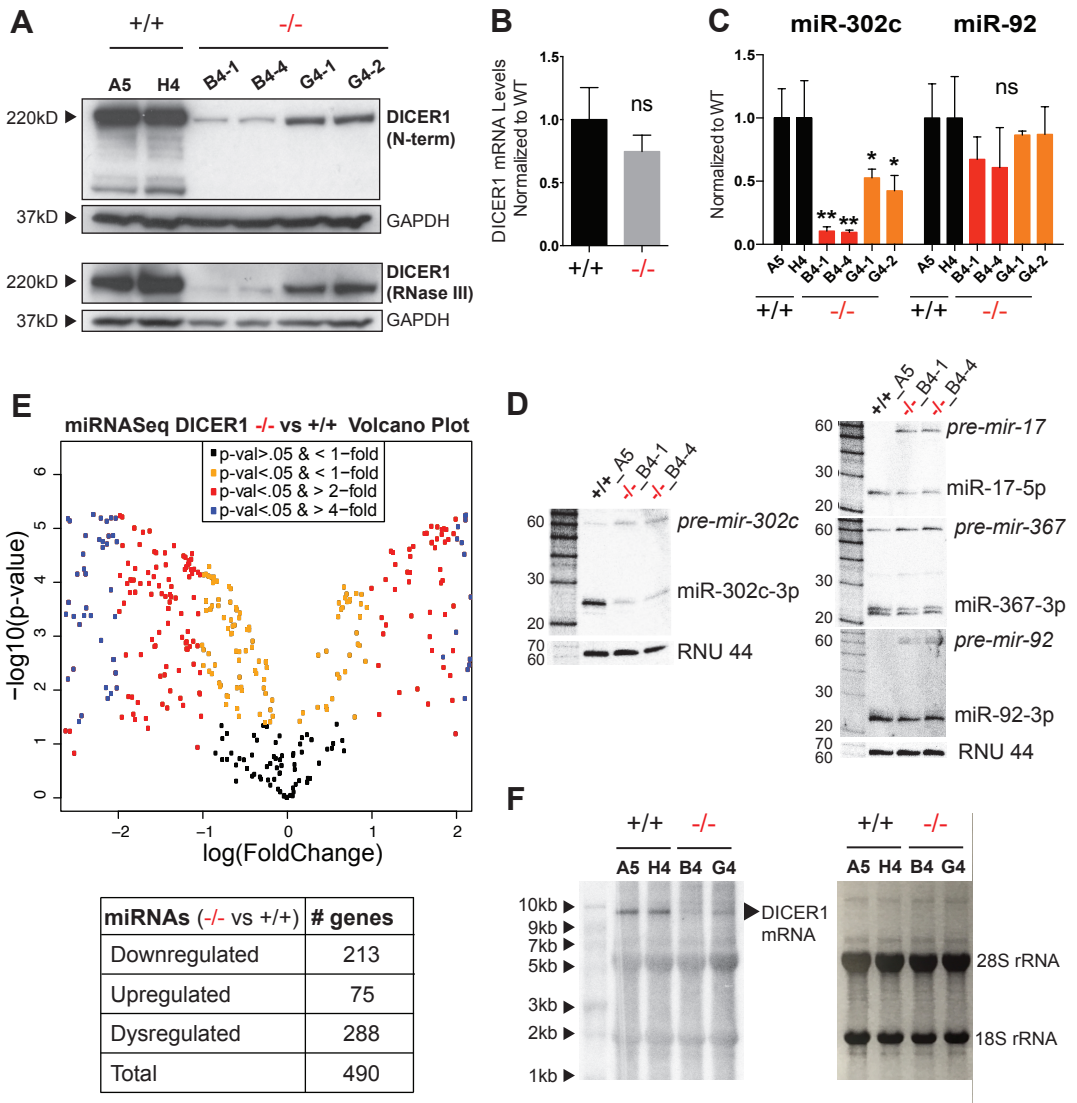
Figure 2.1. Generation of DICER1-/- hESCs using iCRISPR HUES8 hESCs. (A) Schematic representation of DICER1 protein with domains. Red and orange arrows represent gRNA-targeting loci. Red arrows represent gRNAs targeting the PAZ domain, and orange arrows represent gRNAs targeting the RNase IIIa domain. cr, CRISPR; DUF, Domain of Unknown Function; dsRBD, double-stranded RNA-Binding Domain. (B) T7E1 assay in HUES8 iCas9 cells transfected with nine different gRNAs targeting the functional domains of PAZ and RNase IIIa. cr, CRISPR; T7E1, T7 endonuclease 1. (C) CRISPR/Cas9 targeting results of HUES8 iCas9 lines. White, wildtype lines; Yellow, monoallelic mutant lines; Red, biallelic mutant lines. Indel: insertion/deletion. N= number of clones analyzed. (D) Sequences of two DICER1 biallelic mutant lines and wild type line generated using Cr4. (E) Normal karyotype of two clones of B4 and G4 DICER1-/- and their passage-matched wild-types.

different gRNAs was variable, with ~3% - 29% of clonal lines containing a mutation on one allele (*DICER1*$^{-/+}$) (Figure 2.1C and Table 5). This efficiency lies within the expected percentages for other iCRISPR targetings in hESCs (Gonzalez et al., 2014). However, only 3 out of 464 screened lines had biallelic frameshift mutations, and one of them did not survive (Figure 2.1C and Table 5). The two-other surviving *DICER1*$^{-/-}$ lines were generated using Crispr 4 (cr4), and one harbors a single (T) base pair insertion and the other one harbors a seven base pair deletion (Figure 2.1D). The lines were subcloned to ensure they were clonal and free from wildtype contamination, and this was confirmed by TA cloning. Two subcloned lines for each mutant were chosen for further studies (B4-1, B4-4, G4-1, and G4-2), and two passage-matched wildtype lines (A5 and H4) that underwent targeting but were not mutated were also picked and subcloned, and used as controls. These six lines were all karyotypically normal (Figure 2.1E).

Next, we set out to validate the *DICER1*$^{-/-}$ lines. First, we tested DICER1 protein expression using a DICER1 antibody that recognized the N-terminal domain of the protein. Surprisingly, we observed expression of DICER1 in both mutant lines (B4, G4), albeit much reduced compared to wildtype lines (Figure 2.2A). G4 *DICER1* mutant line appeared to produce ~4 times as much protein as the B4 line. To confirm these results, we used a second DICER1 antibody that recognizes the RNase IIIa functional domain of the protein, and again we saw reduced, but not completely ablated, DICER1 protein expression (Figure 2.2A). This is surprising since the genomic mutations are expected to render null alleles. The mutant mRNA transcript is expressed, and thus does not seem to undergo nonsense-mediated mRNA decay (NMR) (Figure 2.2B).

To functionally validate the *DICER1*$^{-/-}$ lines, we tested their ability to 'dice' primary microRNAs (pri-miRNAs) into mature microRNAs that are normally expressed in hESCs (Suh et al., 2004). We checked mature miRNA levels of miR-302c, miR-92, miR-17, and miR-367 by qRT-PCR and Northern blot, and both assays showed compromised but still functional DICER1 activity (Figure 2.2C and 2.2D). Northern blots were performed by Sonali Majumdar in Eric Lai's lab at MSKCC. Interestingly, the reduction in DICER1 activity seems to be miRNA-dependent since miR-302c showed about a 10-fold reduction, and miR-92 showed no significant reduction compared to wildtype. miR-302c was more reduced in the B4 line (10-fold reduction) compared to the G4 line (2-fold reduction) (Figure 2.2C). As expected of DICER1-compromised cells, pri-miRNAs accumulated in the *DICER1*$^{-/-}$ line (B4) in all four miRNAs tested, even when their mature miRNA reduction did not reach significant levels by qRT-PCR (Figure 2.2C and 2.2D).

To investigate the global depletion of mature microRNAs in our *DICER1*$^{-/-}$ lines, we first ran size-fragmented RNA samples (<150nt long species) from wildtype (A5) and *DICER1* mutant (B4-1 and B4-4) lines through the Agilent Bioanalyzer 2100 to find out the percentage of the global mature miRNA population in our samples (Appendix 1). Sonali Majumdar from the Lai lab performed the RNA gel extraction. We found a global reduction in the percentage of mature miRNA populations in the *DICER1* mutants lines (15-21%) compared to the wildtype (35%). To precisely determine these miRNA populations, we next performed miRNA-deep sequencing (miRNASeq) of *DICER1*$^{-/-}$ (B4) and wildtype (A5) lines followed by analysis performed by Jiayu Wen in the Lai lab. We found 288 mature miRNAs out of a total of ~500 mature miRNAs significantly dysregulated (above a 1-fold change, and p-value



miRNAs (-/- vs +/+)	# genes
Downregulated	213
Upregulated	75
Dysregulated	288
Total	490

Figure 2.2. DICER1-/- lines are hypomorphic. (A) Western blots of DICER1 in two wildtype lines (A5, H4) and 4 DICER1-/- (B4-1, B4-4, G4-1, G4-2) using antibodies recognizing the N-terminus and the RNase III domain. kD, kiloDalton. (B) mRNA levels of DICER1 transcript in wild type and DICER1 mutant by qRT-PCR. n=3. (C) qRT-PCR of mature miRNAs 302c and 92 in two wildtype lines (A5, H4) and 4 DICER1-/- (B4-1, B4-4, G4-1, G4-2). n=3. (D) Northern blots of precursor and mature miRNAs 302c, 17, 367, and 92 in one wildtype line (A5) and two DICER1 mutant lines (B4-1, B4-4). RNU44 was used as loading control. pre-mir, precursor miRNA; miR, mature miRNA. (E) miRNA-Seq of wild-type (A5) and DICER1-/- (B4-1) lines presented as a volcano plot with the log Fold Change on the x-axis and the -log10 of the p-value on the y-axis. In black, miRNAs with a p-value>0.05 and less than 1-fold difference comparing DICER1 -/- to wildtype; in yellow, miRNAs with a p-value<0.05 and less than 1-fold difference comparing DICER1 -/- to wildtype; in red, miRNAs with a p-value<0.05 and over 1-fold difference comparing DICER1 -/- to wildtype; in blue, miRNAs with a p-value<0.05 and more than 2-fold difference comparing DICER1 -/- to wildtype. (F) Northern blot of DICER1 mature mRNA in two wildtype lines (A5, H4) and two DICER1 -/- lines (B4, G4). The blot on the right is the loading control showing the 28S and 18S rRNA bands.

<0.05) in the *DICER1*$^{-/-}$ line compared to wildtype; of which 213 were downregulated as expected, and 75 were upregulated (Figure 2.2E). However, around 50% of the mature miRNAs were not significantly dysregulated, and this demonstrated that there are microRNAs' expressions that remain the same in these *DICER1*$^{-/-}$ cells. Additionally, some microRNAs' expression diminished but are not abolished, suggesting low levels of DICER1 activity are still functional.

We next sought to investigate whether the mature mRNA of *DICER1* in mutant (B4-1 and G4-1) versus wildtype lines (A5 and H4) was still being processed. We designed an oligo that recognizes the exon-exon junction in the RNase IIIa domain of *DICER1* and performed a Northern blot (Figure 2.2F). The Northern blot was ran by Sonali Majumdar in the Lai lab. We found that around ~5-10% of the mature mRNA is still being processed in the mutant lines, presumably allowing some extent of miRNA maturation for some but not all miRNAs (Figure 2.2F).

Thus, after extensive characterization, we were not able to validate these *DICER1*$^{-/-}$ lines as true nulls, but rather as hypomorphic alleles. Therefore, we were not able to generate true *DICER1* nulls using iCRISPR. This hypothesis was further supported by the failure to create *DICER1* frameshift homozygous-mutants by targeting *DICER1*$^{-/+}$ lines instead of *DICER1*$^{+/+}$ lines (Figure 2.3A-C). We chose two karyotypically normal *DICER1* heterozygous lines (with one frameshifted allele) generated by targeting with Crispr 3 (Cr3) and Crispr 6 (Cr6) gRNA (Figure 2.3A,B). We selected lines that contained deletions spanning the PAM sequence in the mutant allele so that it would not be targeted again by the gRNA. We subjected these lines to another round of targeting with either Cr3 or Cr6, and compared the results to the

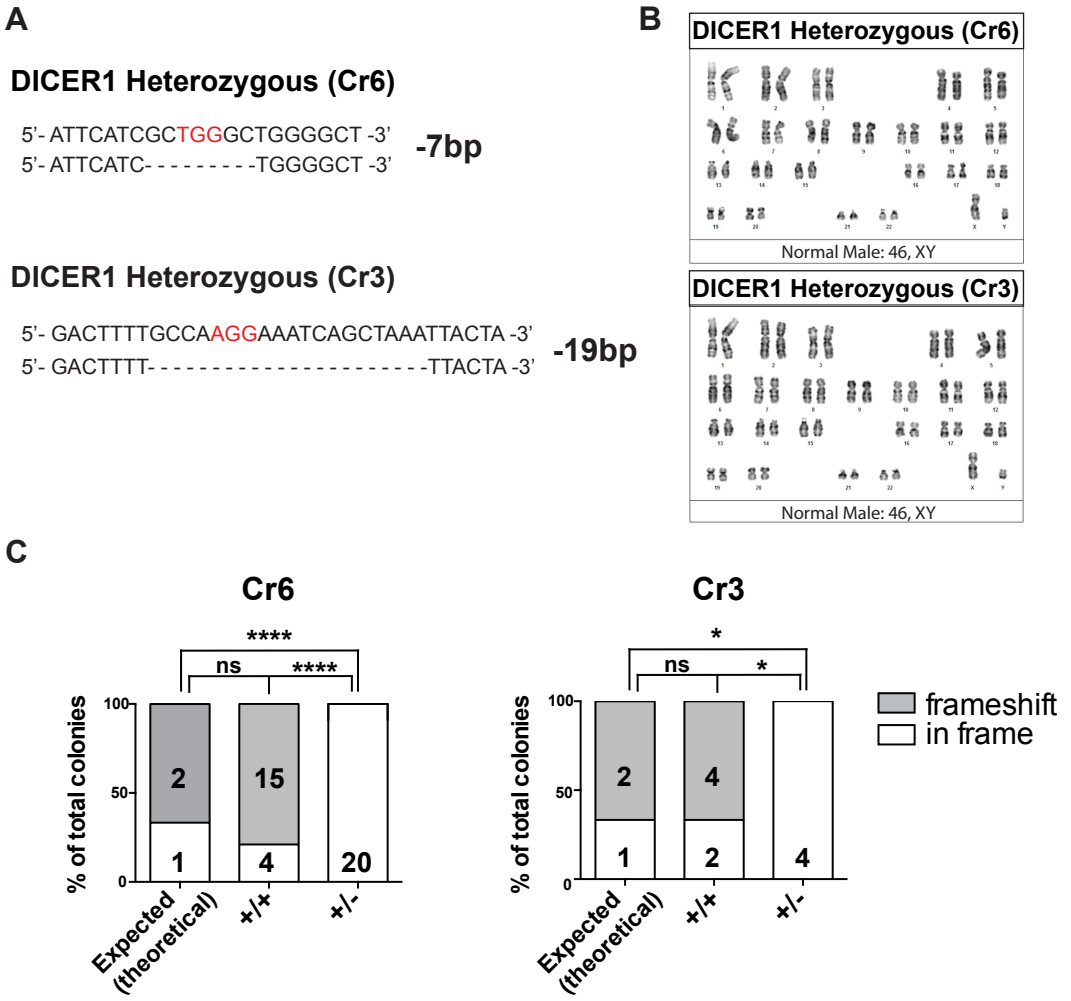


Figure 2.3. Inability to Generate DICER1 Null Alleles Suggests DICER1 is Essential in hESCs. (A) Sequences of DICER1 heterozygous lines chosen for a second round of targeting using the same gRNAs (Cr3 and Cr6) they were initially targeted with. Chosen lines harbor a deletion spanning the NGG on one allele. In red, NGG sequence. (B) Normal karyotype of chosen DICER1 heterozygous lines. (C) Frameshift and in frame allele frequency results of Cr3 and Cr6 targetings in wildtype lines (+/+) versus heterozygous lines (+/-), and compared with the theoretical expectation. Chi-test was used for statistical analysis.

expected (theoretical) ratio of two frameshifts per one in frame (Table 6). While the results for targeting a wildtype line fell in line with this theoretical expectation, targeting a heterozygous line did not as no frameshift homozygous lines were generated (Figure 2.3C and Table 6). These results suggest that DICER1-deficient hESCs are not viable.

Table 6. Results of CRISPR gRNA targeting of DICER1+/- and DICER1+/+ hESCs. if, in frame allele; fs, frameshift allele.

			Mutant Alleles Per Clone						
			1		2				
					Compoud Heterozygotes			Homozygous	
Gene	hESC Line	CRISPR	In frame	Frameshift	if/if	fs/fs	if/fs	In frame	Frameshift
DICER1	DICER1+/-	Cr3	4/69	0/69	←		N/A	→	
	DICER1+/+	Cr3	2/21	4/21	0/21	0/21	1/21	0/21	0/21
	DICER1+/-	Cr6	20/172	0/172	←		N/A	→	
	DICER1+/+	Cr6	4/79	15/79	0/79	0/79	2/79	0/79	1/79

2.3.2 DICER1 Hypomorphs Maintain Pluripotency Marker Expression

Since our *DICER1* hypomorphs downregulate DICER1 about 10-fold or more, we decided to study them further as they may be useful in identifying *DICER1* knockout phenotypes. First, we checked the transcript levels of four pluripotency factors (OCT4, NANOG, SOX2, and KLF4) in B4-1 and B4-4 lines by qRT-PCR, and found that these genes were all expressed in the mutants as expected from mouse *DICER1* knockouts (Spruce et al., 2010). However, OCT4 was significantly downregulated by ~25% in the B4-4 line and NANOG was downregulated by 25% or ~50% in the B4-1 and B4-4 lines respectively (Figure 2.4A). These results contrast the shRNA-mediated downregulation of DICER1 in hESCs in which they found upregulation of OCT4 (Qi et al., 2009). Furthermore, we found that OCT4, NANOG, and SOX2 protein were still expressed at comparable levels (Figure 2.4C). This finding suggests that one or more miRNAs might control pluripotency factor transcription in an as-of-yet

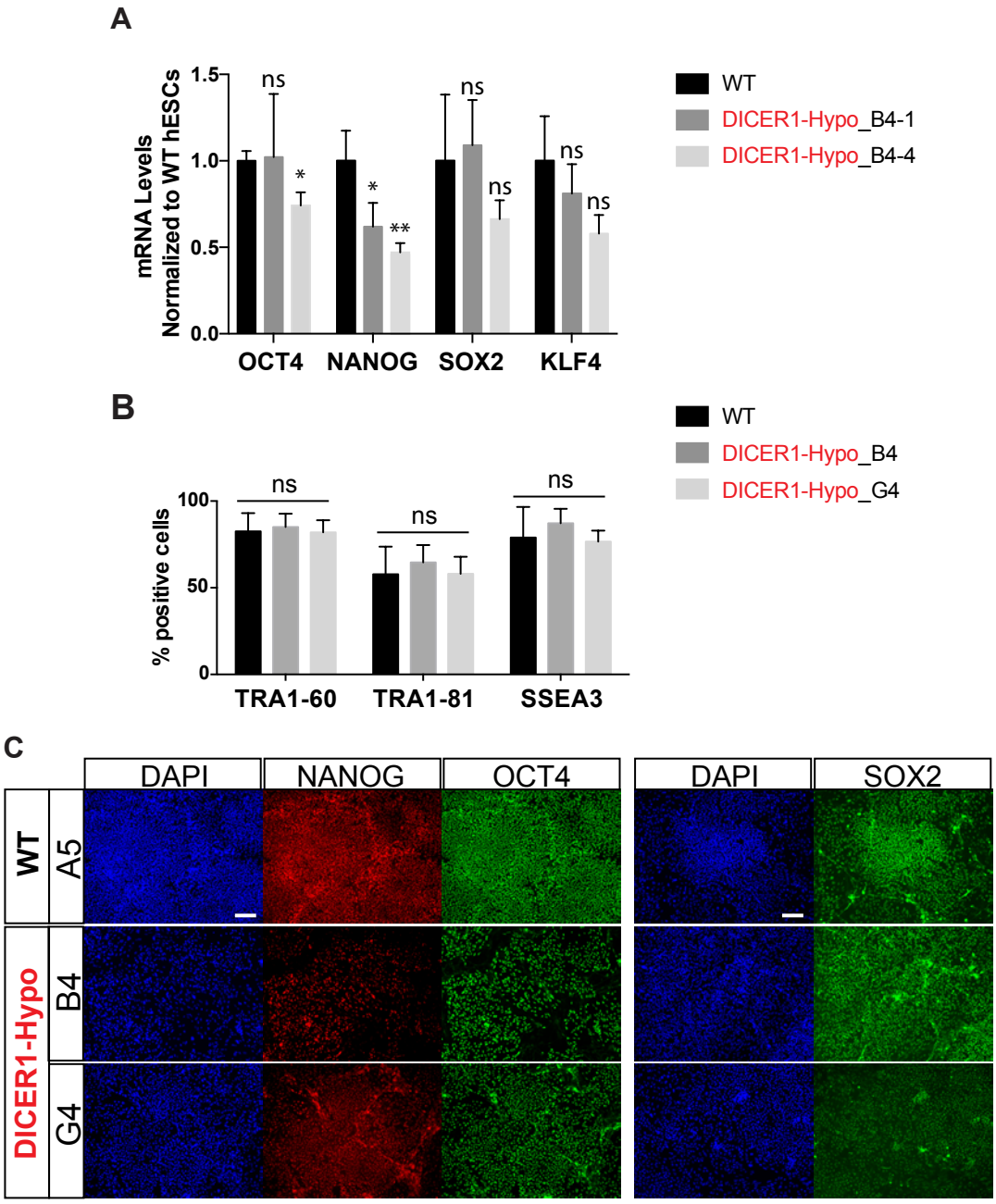


Figure 2.4. DICER1 Hypomorphs Maintain Expression of Pluripotency Markers.
(A) qRT-PCR of pluripotency markers OCT4, NANOG, SOX2 and KLF4. mRNA levels are normalized to wildtype. WT, wildtype; ns, not significant; *, p-value <0.05 **, p-value <0.01. (B) Flow cytometry of pluripotency-associated cell-surface markers TRA1-60, TRA1-81 and SSEA3. Hypo, hypomorph. (C) Immunofluorescence of pluripotency markers NANOG, OCT4 and SOX2 in DICER1 hypomorph and wildtype lines. Scale bar, 100μm.

unknown pathway, although more studies are needed to test this hypothesis. Finally, we tested pluripotency surface marker expression (TRA1-60, TRA1-81, and SSEA3) by flow cytometry and found no difference in expression between the wildtype (A5) and knockout lines (B4 and G4) (Figure 2.4B). Together, these results suggest that pluripotency markers are maintained, although OCT4 and NANOG transcript levels are downregulated.

2.3.3 DICER1 Hypomorphs Retain Characteristic Cell Cycle Properties and Normal Apoptosis Levels

Next, we were interested in analyzing the cell cycle properties of the *DICER1* hESC hypomorphs, since *Dicer1* mESC knockouts and *DICER1* hESC knockdowns show a slower proliferation rate and an accumulation of cells in the G1 phase compared to wildtypes (Qi et al., 2009; Wang et al., 2008). Thus, we first counted cell number every-day for 6 days and found that the wildtype (A5) and the *DICER1* Hypomorph (B4) proliferate at comparable rates (Figure 2.5A). We then looked at proliferation marker Ki67 and mitotic marker Phospho-Histone3 (PH3) by flow cytometry during days 0, 2, 4, and 6 and found that the levels in *DICER1* Hypomorphs remain the same as in the wildtype controls (Figure 2.5B). These results were corroborated by immunofluorescence staining as well (Figure 2.5C). Next, we sought to investigate the proportion of cells in the G1, S, or G2 stages of the cell cycle by Propidium Iodide (PI) staining and flow cytometry. Again, we found no significant accumulation in G1 of the cell cycle in *DICER1* hypomorphs (Figure 2.5D, E). Although our results show that *DICER1* hypomorphic hESCs do not have an altered cell cycle structure unlike found in other studies (Qi et al., 2009), these results are in line with the *in vivo* observations of *Dicer1* knockout

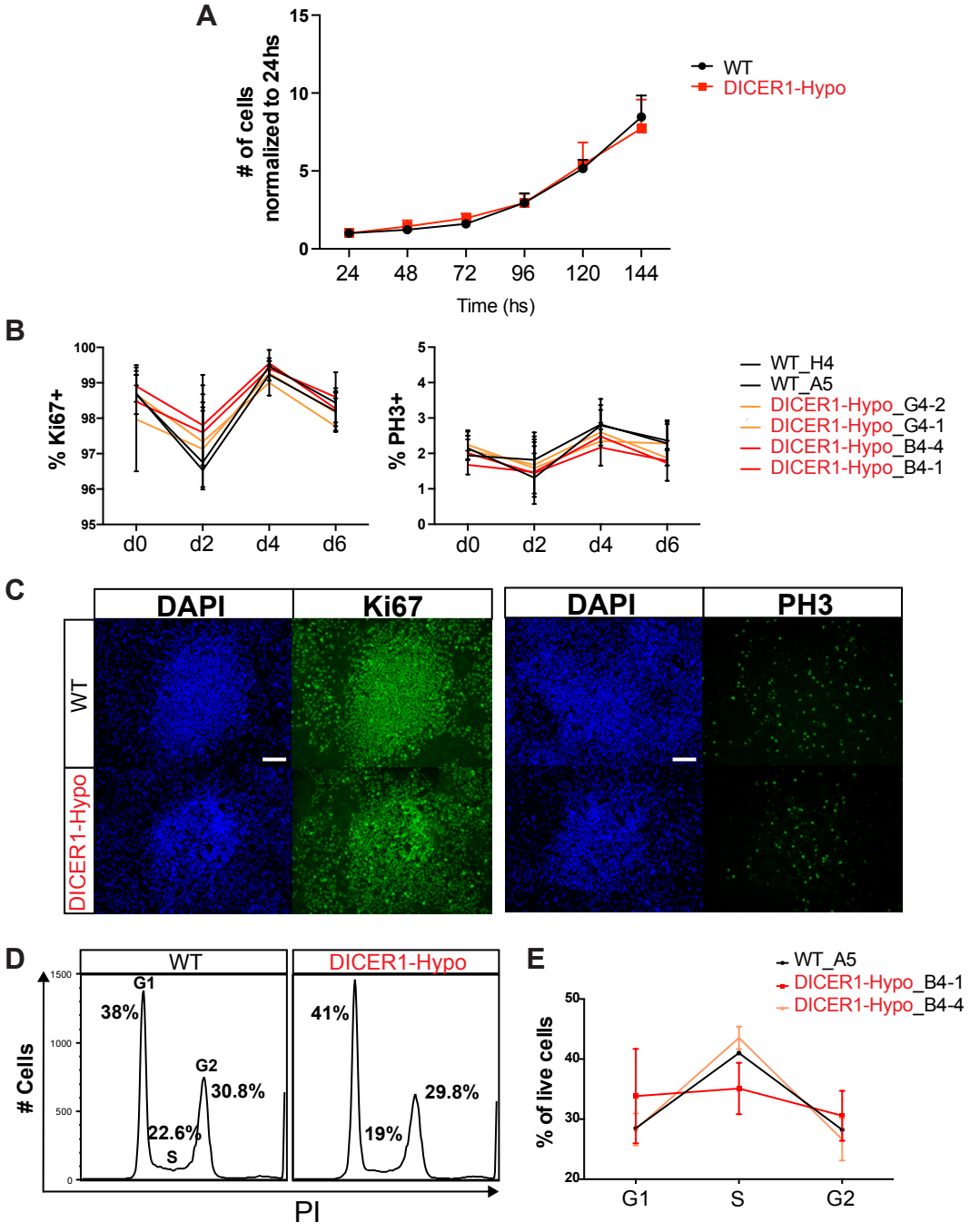


Figure 2.5. DICER1 Hypomorphs Maintain Proliferation and Cell Cycle Properties. (A) Proliferation assay of DICER1-Hypomorph and wildtype lines by counting total cell number every 24hs for 6 days. Results are normalized to total cell number for each line at 24hs. (B) Flow cytometry analysis of prolfieration markers Ki67 and PH3 in wildtype (A5, H4) and DICER1-Hypomorphs (B4-1,B4-4, G4-1, G4-2) every 48hs for 6 days. n=3. (C) Immunofluorescence of Ki67 and PH3 proliferation markers in DICER1-Hypomorph (B4) and wildtype (A5) lines. (D) Representative propidium iodide flow cytometry histograms showing the percentage of cells in the G1, S or G2 stags of the cell cycle in wildtype (A5) and DICER1-Hypomorph (B4-4) lines. (E) Quantification of propidium iodide flow cytometry. n=3.

during mouse development (Spruce et al., 2010). It remains to be seen whether these differences are cause because the *DICER1* hypomorphs do not reach sufficient abolition of DICER1 protein levels, or if these differences display a mouse to human difference in *DICER1* function.

We also checked levels of apoptosis in the *DICER1* hypomorphs compared to wildtypes. We performed flow cytometry to check for Cleaved Caspase-3 staining at days 0, 2, 4, and 6 of plating and observed no significant differences in apoptosis levels at any of the days analyzed (Figure 2.6A). These results were corroborated with immunostaining for Cleaved Caspase-3 (Figure 2.6B). Thus, apoptosis levels remain unchanged in *DICER1* hypomorphs.

2.3.4 DICER1 Hypomorphs Show Impaired Colony Forming Ability

Even though proliferation and apoptosis rates remained the same in *DICER1* hypomorphs, we did notice that the total number of cells after 6 days of growth without normalizing to 24 hours was lower in *DICER1* hypomorphs compared to wildtype. Thus, we investigated the ability of the hypomorphs to self-renew by subjecting them to colony forming assay, hypothesizing that they may have impaired colony forming ability, which could account for the difference in population at day six. Therefore, we plated *DICER1* hypomorph (B4) or wildtype (A5) cells at low density in 10cm dishes, and counted the cell number after two weeks. The results showed that the *DICER1* hypomorphs were indeed impaired in colony forming ability, resulting in half as many colonies when cultured without Rock inhibitor, and a quarter less when cultured in the presence of Rock inhibitor (Figure 2.7A). Addition of Rock inhibitor partially rescued the colony forming ability (Figure 2.7A).

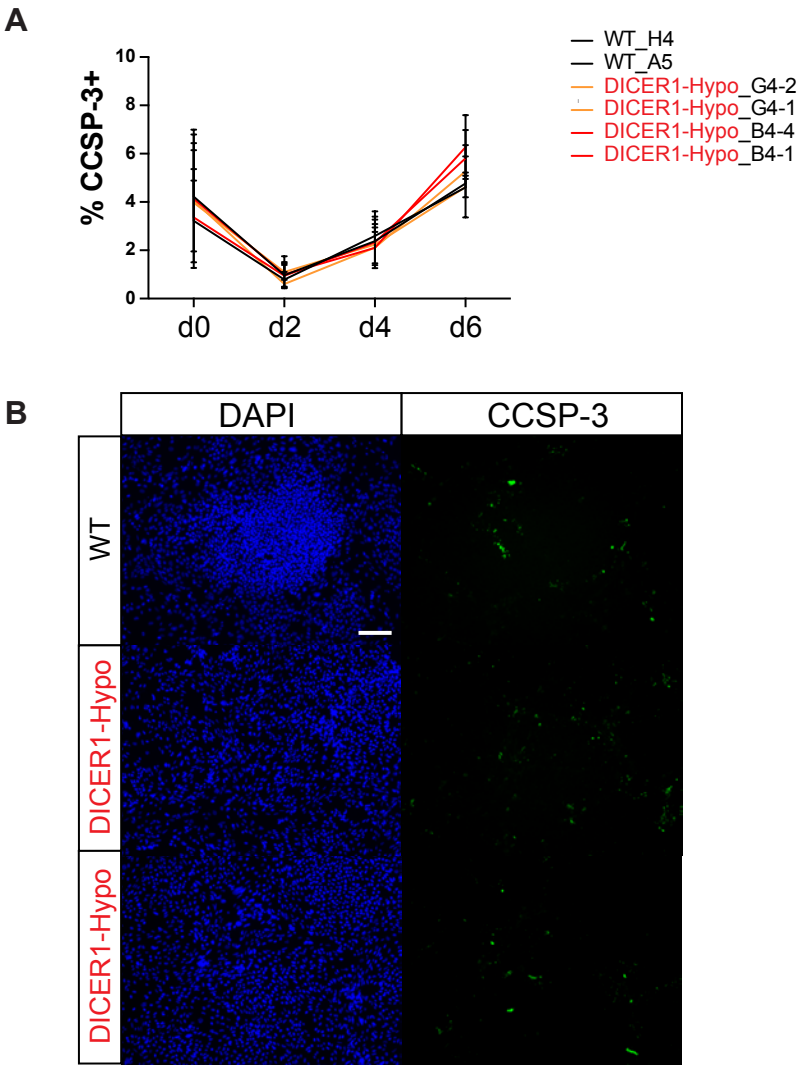


Figure 2.6. Cell Death is Not Observed in DICER1 Hypomorphs.
(A) Flow cytometry analysis of apoptotic marker cleaved caspase-3 (CCSP-3) in wildtype (A5, H4) and DICER1-Hypomorphs (B4-1,B4-4, G4-1, G4-2) every 48hs for 6 days. n=3. (B) Immunofluorescence images of cleaved caspase-3 (CCSP-3) in WT (A5) and DICER1-Hypomorphs (B4, G4) lines.

We next checked the levels of epithelial-to-mesenchymal transition (EMT) markers in *DICER1* hypomorph (B4) compared to wildtype (A5). Loss of E-Cadherin is considered a fundamental event in EMT (Kim et al., 2014), and we found that E-Cadherin levels were significantly downregulated in the *DICER1* hypomorph by about a half (Figure 2.7B). Consistent with these results, ZEB1, an EMT-promoting transcription factor, was highly upregulated by around 5-fold in the *DICER1* hypomorph (Figure 2.7B). These findings could be explained by the reduction of the hESC-expressed miR-200 family in the *DICER1* hypomorphs since this family has been shown to prevent EMT in hESCs and target the E-Cadherin inhibitors ZEB1 and ZEB2 directly (Gill et al., 2011; Zhang and Ma, 2012). Additional studies are needed to test this hypothesis.

2.3.5 DICER1 Hypomorphs Dysregulate Differentiation Markers but Retain the Ability to Differentiate

Since *DICER1* knockout mESCs are unable to differentiate as evidenced by their inability to generate teratomas, to contribute to the embryo proper, and to upregulate differentiation markers upon embryoid body formation (Bernstein et al., 2003; Kanellopoulou et al., 2005; Murchison et al., 2005), we set out to investigate whether this also holds true in the *DICER1* hESC hypomorphs. First, we performed qRT-PCR of transcription factors associated with the three germ layers (endoderm, ectoderm, and mesoderm) and the extraembryonic primitive endoderm in *DICER1* hypomorphs (B4-1 and B4-4) and wildtype (A5). The results showed a 5 to 15-fold significant increase in FOXA2 expression and a 30 to 100-fold significant increase in SOX17 expression in the *DICER1* hypomorphs (Figure 2.8A). Ectoderm markers

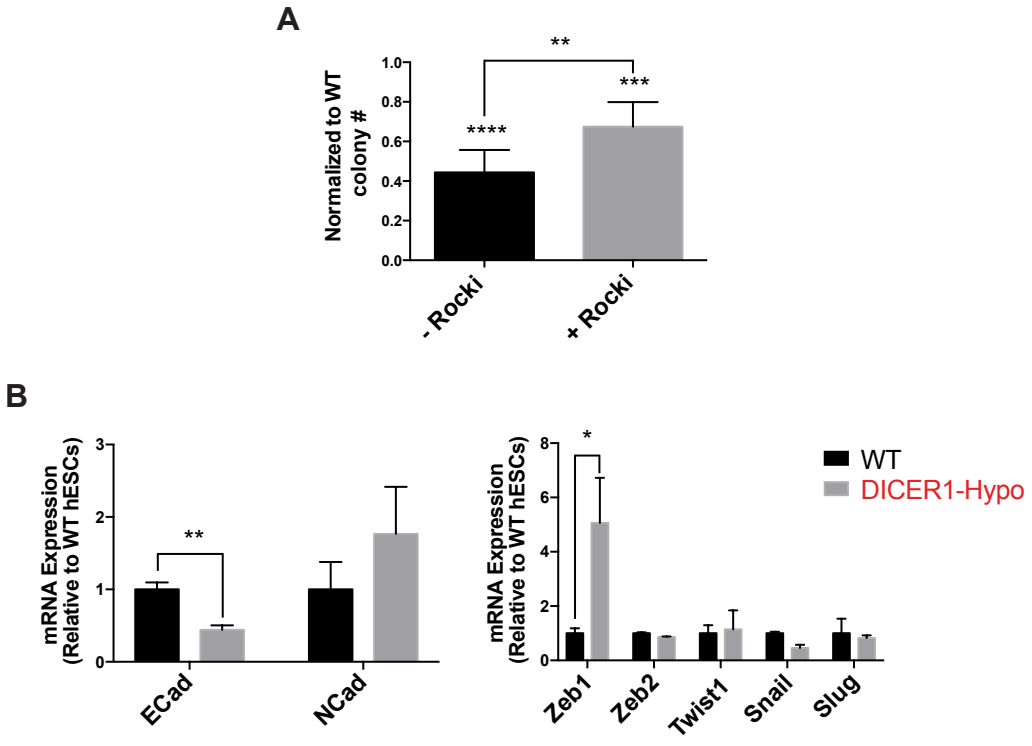
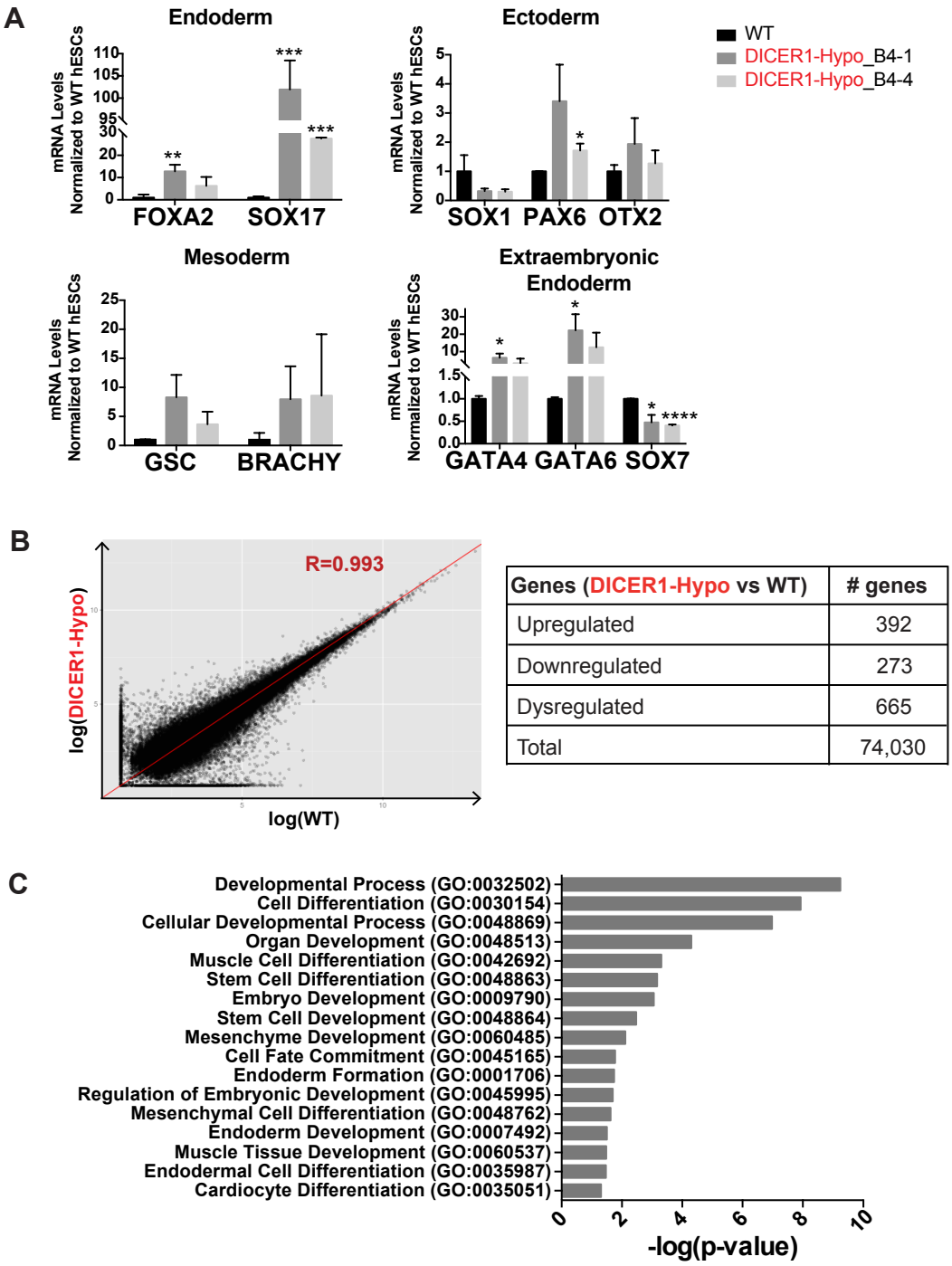


Figure 2.7. DICER1 Hypomorphs Show Impaired Colony Forming Ability.
(A) Colony forming assay showing number of colonies after 2 weeks with or without Rock inhibitor for wildtype and DICER1-Hypomorph (B4) lines. Results are normalized to wildtype number of colonies. In black, colony forming assay for without Rock inhibitor; in grey, results with Rock inhibitor. n=3. (B) qRT-PCR of ECAD, NCAD and EMT-promoting transcription factors ZEB1, ZEB2, TWIST1, SNAIL and SLUG in DICER1-Hypomorph lines normalized to wildtype. In black, wildtype line; in grey, DICER1-Hypomorph line. n=3.



Genes (DICER1-Hypo vs WT)	# genes
Upregulated	392
Downregulated	273
Dysregulated	665
Total	74,030

Figure 2.8. DICER1 Hypomorphs Dysregulate Developmental Transcription Factors. (A) qRT-PCR of transcription factors associated with the three germ layers (endoderm, ectoderm, mesoderm) and the extraembryonic endoderm in DICER1-Hypomorph lines and normalized to WT hESCs. n=3. (B) RNA-Sequencing plot of DICER1-Hypomorph and WT hESC lines. n=2. R, correlation coefficient. (C) Gene ontology analysis of genes upregulated in DICER1-Hypomorph compared to WT line. GO, gene ontology.

SOX1 and OTX2 seem not to change significantly, but PAX6 does increase by about 2-fold (Figure 2.8A). Mesoderm markers Goosecoid (GSC) and Brachyury (BRACHY) seem to increase, albeit the changes detected were not statistically significant. Extraembryonic primitive endoderm markers like GATA4 and GATA6 were significantly upregulated 7 and 20-fold respectively, and SOX7 was significantly downregulated by a half (Figure 2.8A). Thus, we did find significant changes in differentiation marker expression.

To investigate the global transcript levels in the *DICER1* hypomorph (B4) compared to wildtype (A5), we performed RNA-sequencing followed by analysis performed by Solomon Shenker from the Lai lab at MSKCC. Overall, we found that global transcript levels remained fairly the same with a correlation coefficient of R=0.993 (Figure 2.8B). This is not surprising given that we did find some level of functional DICER1 in these hypomorphic lines. However, out of a total of 74,030 transcripts, 392 were significantly upregulated and 273 were significantly downregulated. Gene ontology (GO) analysis on the upregulated gene list, presumably direct targets of the microRNAS, rendered GO terms related to differentiation and development. The top three terms were “Developmental Process”, “Cell Differentiation”, and Cellular Developmental Process” (Figure 2.8C). Contrary to *DICER1* knockout mESC data, our *DICER1* hypomorph hESCs upregulated differentiation genes.

Thus, we were interested in figuring out whether our *DICER1* hypomorphs could form teratomas when injected under the skin of SCID-mice. We injected eight mice per line (A5, B4-1 and B4-4), out of which 3 mice injected with wildtype hESCs generated teratomas, 4 mice injected with B4-1 and 4 mice injected with B4-4 also generated teratomas (Figure 2.9A). The

teratomas formed were comparable in size between wildtype and mutant (Figure 2.9A). All teratomas derived from wildtype and *DICER1* hypomorph hESCs were stained with hematoxylin & eosin (H&E) and then visualized under the microscope for endoderm, mesoderm and ectoderm structures. *DICER1* hypomorph-derived teratomas displayed structures from all three germ layers, like wildtype-derived ones, showing that *DICER1* hypomorph hESCs can differentiate, unlike *Dicer1* knockout mESCs.

To investigate whether *DICER1* hypomorph hESCs are amenable to directed differentiation, we opted to differentiate these mutants into the neuroectodermal lineage. Studies performed in hESCs have shown that miR-302, the most abundant miRNA cluster in hESCs, inhibits neural differentiation by promoting BMP signaling, and thus we investigated whether the *DICER1* hypomorph, which have reduced levels of miR-302, might be precociously differentiating into neuroectoderm (Barroso-delJesus et al., 2011; Lipchina et al., 2011). Thus, we induced neuroectoderm differentiation for ten days and analyzed the expression of pluripotency markers (OCT4 and NANOG) and neuroectoderm markers (PAX6 and SOX1) by qRT-PCR on days 0, 6, and 10 of differentiation (Figure 2.9C). We found that pluripotency markers OCT4 and NANOG were properly downregulated by day 6 in *DICER1* hypomorphs (B4 and G4), with a concomitant increase in PAX6 and SOX1 transcripts at similar levels to wildtype (A5) (Figure 2.9D). Then, we checked the percentage of cells that expressed PAX6 protein, which is a transcription factor that is both necessary and sufficient for neuroectoderm formation (Zhang et al., 2010), by flow cytometry on days 4, 6, 8 and 10 of differentiation (Figure 2.9E, F). The percentage of *DICER1* hypomorph cells expressing PAX6 at days 4, 6, 8 and 10 were comparable to wildtype cells (Figure 2.9E). After 10 days of

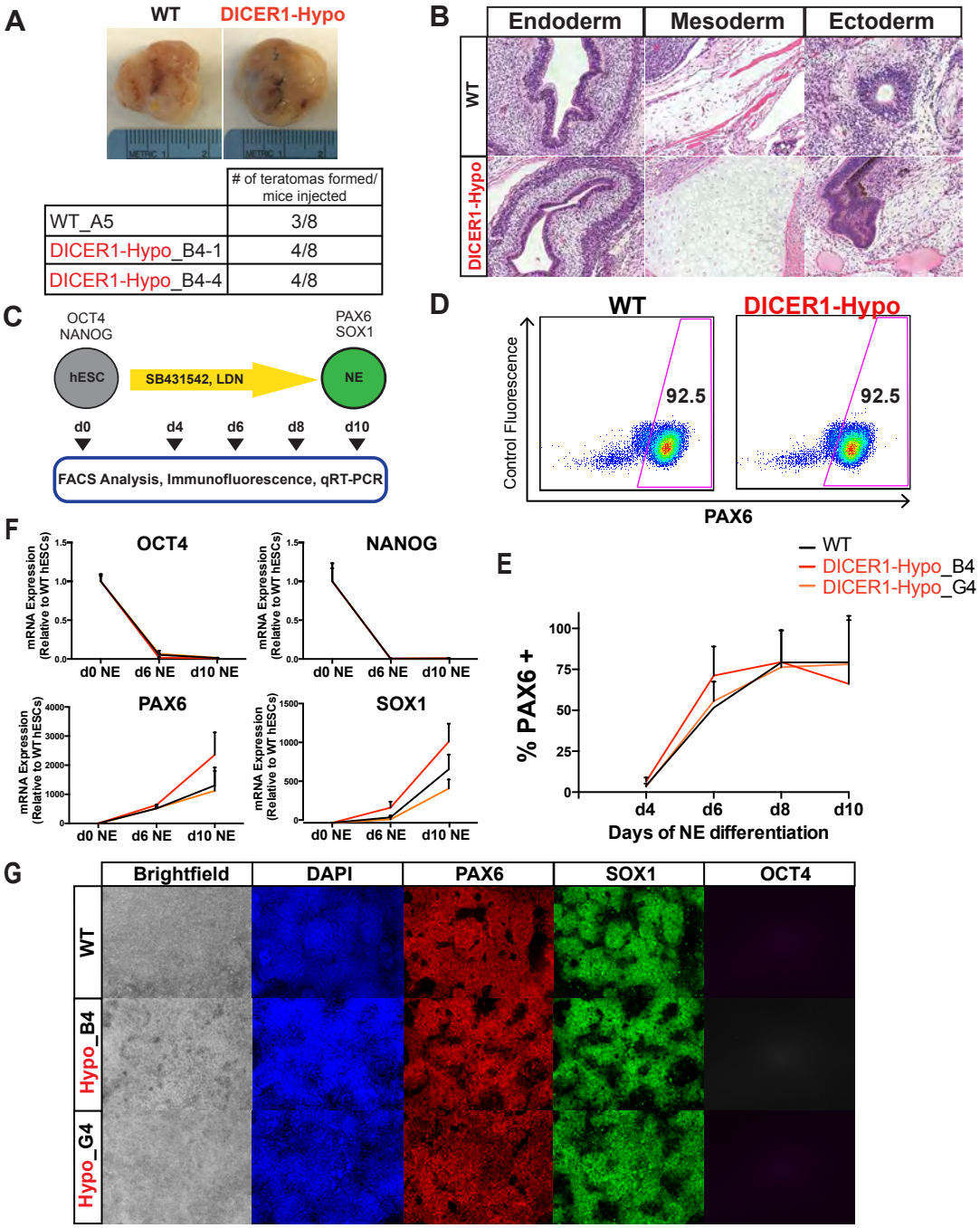


Figure 2.9. DICER1 Hypomorphs Retain the Ability to Differentiate into All Three Germ Layers. (A) Teratoma assay of DICER1-Hypomorph (B4-1, B4-4) and WT (A5) lines by subcutaneous injections into 8 immunodeficient SCID mice per line. Table shows number of injected mice that developed a teratoma. (B) Hematoxylin & Eosin staining of teratoma cross sections displaying all 3 germ layers (endoderm, ectoderm, mesoderm). (C) Schematic representation of neuroectoderm differentation experimental procedure. (D) Flow cytometry plots of PAX6 positive cells on day 10 of neuroectoderm differentiation in DICER1-Hypomorph and WT lines. (E) Flow cytometry results of PAX6 positive cells at days 4,6,8,10 of neuroectoderm differentiation in DICER1-Hypomorphs (B4, G4) and WT lines. (F) qRT-PCR of pluripotency markers (OCT4 and NANOG) and neuroectoderm markers (PAX6 and SOX1) at days 0, 6 and 10 of neuroectoderm differentiation. (G) Immunofluorescence staining of PAX6, SOX1 and OCT4 at day 10 of neuroectoderm differentiation in DICER1-Hypomorph and wildtype lines.

differentiation, ~90% of the neuroectoderm cells derived from either *DICER1* hypomorph or wildtype hESCs expressed PAX6 protein at comparable levels (Figure 2.9D, E). Immunostaining of PAX6, SOX1 and OCT4 confirmed these results (Figure 2.9G). Therefore, this means that *DICER1* hypomorph hESCs can differentiate to neuroectoderm and do not appear to differentiate precociously.

In sum, *DICER1* hypomorph hESCs can differentiate efficiently and downregulate pluripotency factors at an appropriate pace, unlike *Dicer1* knockout mESCs.

2.3.6 DICER1 Hypomorphs Downregulate Argonautes 2-4

Out of the four mammalian Argonautes (AGO1-4), only Argonaute 2 (AGO2) has catalytic activity (Liu et al., 2004). We were interested to investigate whether Argonaute 2 might be upregulated in the *DICER1* hypomorphs to aid in 'slicer' activity of microRNAs. To this end, we performed qRT-PCR of Argonautes 1 through 4 (AGO1-4) in *DICER1* hypomorph (B4-1 and B4-4) and wildtype (A5), and observed a significant downregulation for Argonautes 2, 3 and 4, but not 1 (Figure 2.10A). Western blotting of AGO2 was performed by Sonali Majumdar of the Lai lab and confirmed that protein levels were also reduced in these *DICER1* hypomorphs (Figure 2.10B). Even though the results don't align with our hypothesis, these results make sense in the context of a previous study, which found that the stability of AGO declines in *Dicer1*-knockout mouse cells, and that Argonaute levels are finely tuned in part by the cellular availability of mature miRNAs (Smibert et al., 2013). Nevertheless, whether the remaining AGO2 levels are playing a role in slicing

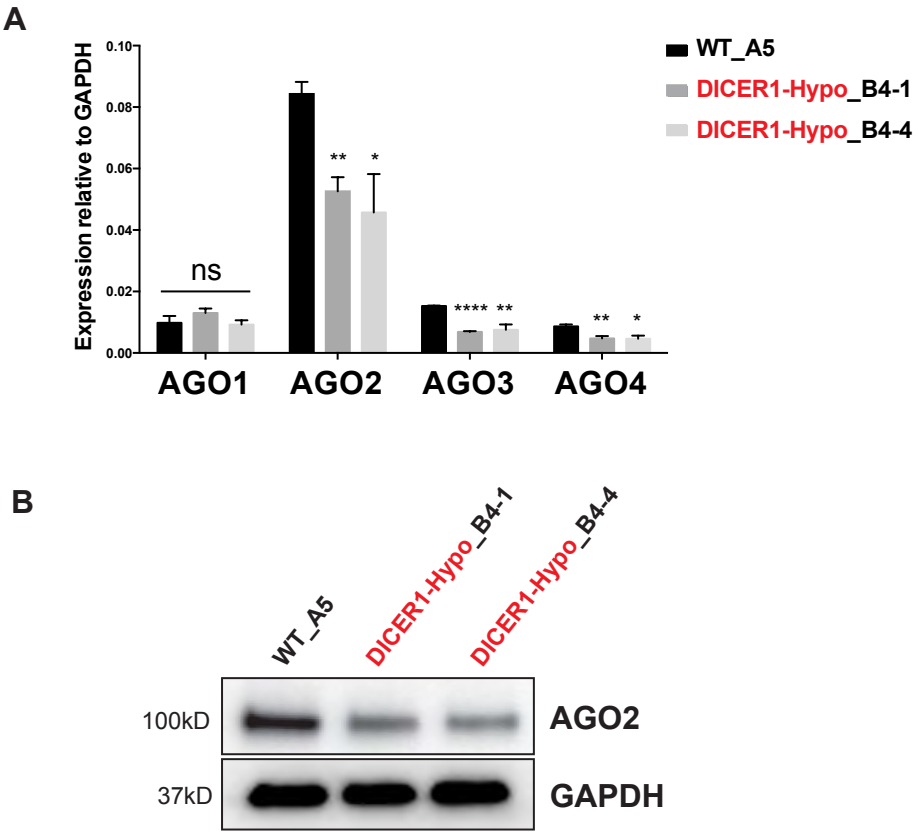


Figure 2.10. Argonaute Proteins Are Downregulated in DICER1 Hypomorphs.
(A) qRT-PCR of the four human Argonaute proteins (AGO1, AGO2, AGO3, AGO4) in DICER1 Hypomorph lines (B4-1, B4-4) and a wildtype line. mRNA levels are normalized to GAPDH. AGO, Argonaute. n=3. (B) Western blot of Argonaute-2 in DICER1 Hypomorph lines (B4-1, B4-4) and a wildtype line. GAPDH was used as loading control.

pre-miRNAs in the *DICER1* hypomorph context is still unknown and needs further exploration.

2.3.7 Successful Generation of DICER1 Knockout hESCs using a TRE-DICER1 Rescue Strategy*

To overcome the hypothesized *DICER1*$^{-/-}$ defect, we designed a rescue strategy to enable the generation of homozygous-mutant lines without the apparent lethality of DICER1 loss, similar to the strategy used for knocking out DNMT3b in hESCs (Liao et al., 2015). We introduced a lentivirus containing a doxycycline-inducible TRE-DICER1* transgene into an iCas9 HUES8 line, creating a Tet-On DICER1* cell line to be used for subsequent targeting (Figure 2.11A and Appendix 2.A). The TRE-DICER1* vector was cloned by Federico Gonzalez in our lab. We selected cr4 and cr6 for further targeting due to their ability to generate indels and biallelic mutations (Figure 2.1C, Figure 2.11B and Table 5). We generated two Tet-On DICER1* lines each with a silent mutation in the protospaceradjacent motif (PAM) sequence of cr4 or cr6 to prevent the gRNA from disrupting exogenous DICER1* (Appendix 2.B). Expression of DICER1* transgene was confirmed by addition of doxycycline followed by qRT-PCR, which confirmed tight expression of the transgene (Appendix 2.C). We next used the CRISPR/Cas9 approach to knock out endogenous *DICER1* while maintaining expression of exogenous DICER1* by doxycycline addition (Figure 1A). We now obtained several homozygous mutant lines using cr4 and cr6 gRNAs, suggesting that our rescue strategy maintains sufficient levels of DICER1 activity for survival (Figure 2.11C, Table 7 and Table 8). We picked four lines (B2, C2, D11, F2),

two from each gRNA targeting, and confirmed that they were karyotypically normal to perform more detailed characterization (Figure 2.11D, E). It is worth noting that the C2 line harbors the same biallelic base pair insertion as the B4 *DICER1* hypomorphic line generated initially (Figure 2.1D).

Table 7. Results of CRISPR gRNA targeting of DICER1 in TRE-DICER1* and iCas9 HUES8 hESCs. if, in frame allele; fs, frameshift allele.

			Mutant Alleles Per Clone						
			1		2				
					Compoud Heterozygotes			Homozygous	
Gene	hESC Line	CRISPR	In frame	Frameshift	if/if	fs/fs	if/fs	In frame	Frameshift
DICER1	TRE-DICER1*	Cr4	0/93	7/93	0/93	13/93	11/93	0/93	18/93
	iCas9 HUES8	Cr4	4/96	2/96	11/96	0/96	38/96	0/96	2/96
	TRE-DICER1*	Cr6	6/94	24/94	0/94	1/94	4/94	1/94	1/94
	iCas9 HUES8	Cr6	6/79	7/79	0/79	0/79	0/79	0/79	0/79

DICER1 homozygous knockout hESCs proliferated normally and could be maintained over several passages as long as exogenous DICER1* was expressed. Removal of doxycycline for 5 days caused downregulation of DICER1* mRNA levels as well as DICER1* and DICER1 protein levels, which were barely detectable (Figures 2.12A, B, C). Additionally, we observed significant mature microRNA reduction by qRT-PCR and Northern blot at this stage with concomitant increase in precursor miRNA in the case of miRNA-92a, showing that these lines are compromised in microRNA processing as expected (Figures 2.12D, E). However, 5 days were not sufficient for microRNAs to be fully ablated.

As an alternate strategy to the lentiviral infection with the TRE-DICER1* transgene, we generated stable HUES8 hESC lines that harbor the TRE-DICER1* transgene with either the PAM of cr4 or cr6 mutated on one allele of the AAVS1 locus, and either an M2rtTA cassette alone or an M2rtTA couple with Cas9 cassette on the other allele (Appendix 3A). Results of the three

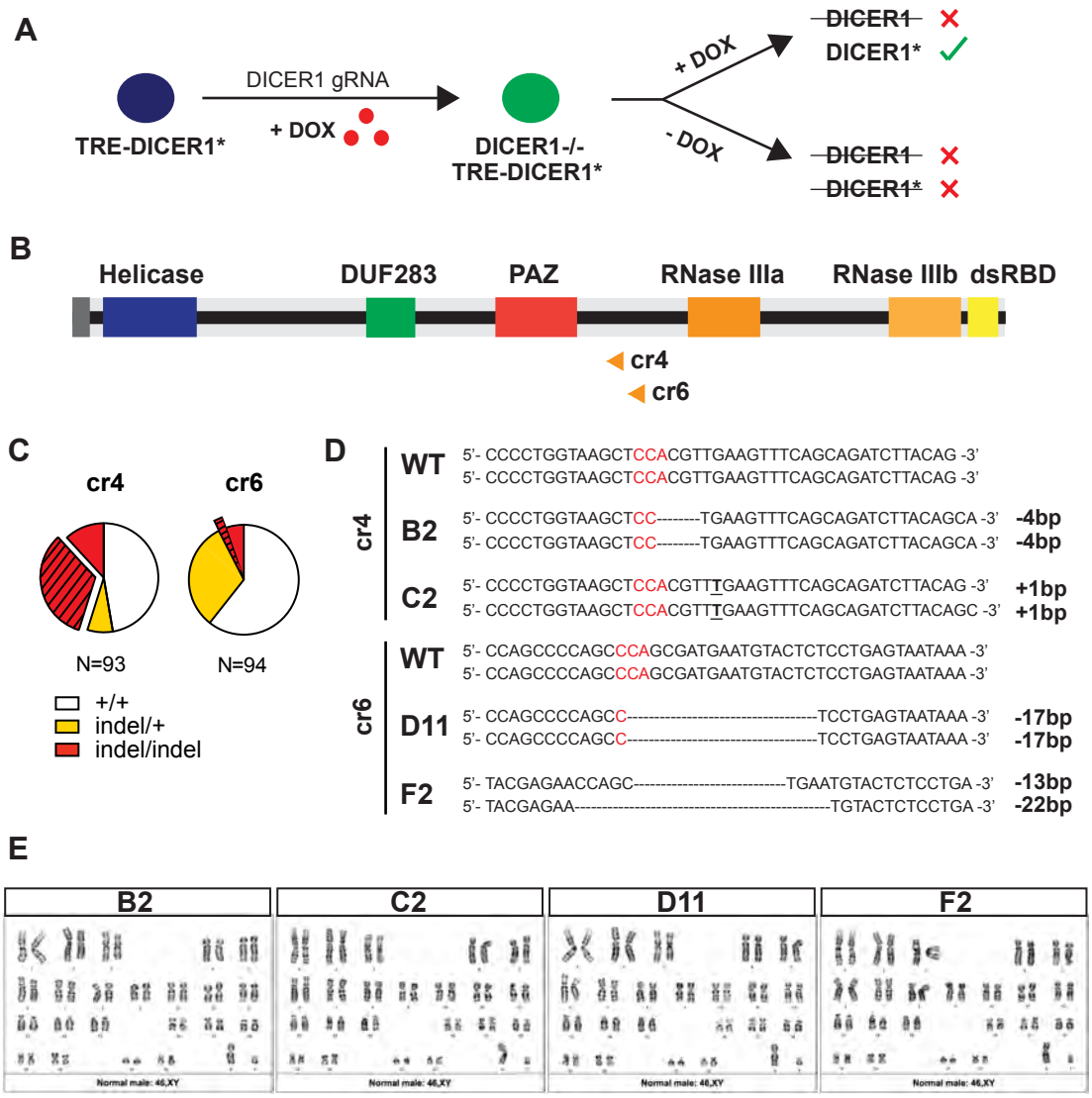


Figure 2.11. TRE-DICER1* Rescue Strategy Allows the Generation of DICER1-/- hESC Lines. (A) Schematic of the procedure followed to generate DICER1 knockout (-/-) hESCs. CRISPR/Cas9 was used to generate DICER1 -/- mutations in cells infected with a gRNA-immune doxycycline-inducible TRE-DICER1* transgene. Cas9 protein and DICER1* protein are induced simultaneously by doxycycline treatment. DICER1 -/- mutants were maintained in doxycycline until the start of experiments. Asterisk (*) is used to distinguish the exogenous modified DICER1* from endogenous DICER1. DOX, doxycycline (red dots). (B) Schematic representation of DICER1 protein with domains. Orange arrows represent gRNA-targeting loci upstream of the RNase IIIa functional domain. cr, CRISPR; DUF, Domain of Unknown Function; dsRBD, double-stranded RNA-Binding Domain. (C) CRISPR/Cas9 targeting results of TRE-DICER1* hESC lines. White, wildtype lines; Yellow, monoallelic mutant lines; Red, biallelic mutant lines. Stripes indicate DICER1 -/- lines. Indel: insertion/deletion. (D) Representative sequences of four selected DICER1 -/- clones with PAM sequences labeled in red. (E) Normal karyotype results of DICER1 -/- lines (B2 and C2 from cr4; D11 and F2 from cr6).

Table 8. Sequences of DICER1-/- form CRISPR targeting of TRE-DICER1* HUES8 hESCs.

TRE-DICER1* Cr6

A1	TTTATACAGTTACGAGAACCAGCCCCAGCCCAGCGATGAATGTACTCTCCTGAGTAATAAATACCTTGAT	**WT**
	TTTATACAGTTACGAGAACCAGCCCCAGCCCAGCGATGAATGTACTCTCCTGAGTAATAAATACCTTGAT	
F2	TTTATACAGTTACGAGAACCAGC----------TGAATGTACTCTCCTGAGTAATAAATACCTTGAT	**-13bp**
	TTTATACAGTTACGAGAA----------TGTACTCTCCTGAGTAATAAATACCTTGAT	**-22bp**
D11	TTTATACAGTTACGAGAACCAGCCCCAGC----------TCCTGAGTAATAAATACCTTGAT	**-17bp**
	TTTATACAGTTACGAGAACCAGCCCCAGC----------TCCTGAGTAATAAATACCTTGAT	**-17bp**

TRE-DICER1* Cr4

A1	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCACGTTGAAGTTTCAGCAGATCTTACAGCAATTAATGGTC	**WT**
	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCACGTTGAAGTTTCAGCAGATCTTACAGCAATTAATGGTC	
B2	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCC--------TGAAGTTTCAGCAGATCTTACAGCAATTAATGGTC	**-4bp**
	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCC--------TGAAGTTTCAGCAGATCTTACAGCAATTAATGGTC	**-4bp**
C2	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCACGTTTGAAGTTTCAGCAGATCTTACAGCAATTAATGGTC	**+1bp**
(16 total)	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCACGTTTGAAGTTTCAGCAGATCTTACAGCAATTAATGGTC	**+1bp**
G8	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCACGT--GAAGTTTCAGCAGATCTTACAGCAATTAATGGTC	**-1bp**
	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCACGT----------TTTCAGCAGATCTTACAGCAATTAATGGTC	**-5bp**
C3	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCACGTTTGAAGTTTCAGCAGATCTTACAGCAATTAATGGTC	**+1bp**
	CGTTGCTCAGCGAGTCCCCTGGTAAGC----------AGATCTTACAGCAATTAATGGTC	**-19bp**
E5	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCACGTTTGAAGTTTCAGCAGATCTTACAGCAATTAATGGTC	**+1bp**
	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCACGTT----------AGATCGTACAGCAATTAATGGTC	**-11bp**
	T>G	
G5	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCACGTTTGAAGTTTCAGCAGATCTTACAGCAATTAATGGTC	**+1bp**
	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCA----------GTTTCAGCAGATCTTACAGCAATTAATGGTC	**-7bp**
B6	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCACGTTTGAAGTTTCAGCAGATCTTACAGCAATTAATGGTC	**+1bp**
	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCACGT-GGAGTTGCA-GCAGATCTTACAGCAATTAATGGTC	**-10bp, +9bp**
	TGAAGTTTCA>GGAGTTGCA	
E7	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCACGTTTGAAGTTTCAGCAGATCTTACAGCAATTAATGGTC	**+1bp**
	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCACGT----------GCAGGTCTCACAGCAATTAATGGTC	**-10bp**
	A>G T>C	
F8	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCACGTTTGAAGTTTCAGCAGATCTTACAGCAATTAATGGTC	**+1bp**
	CGTTGCTCAGCGAGTCCC----------TGAAGTTTCAGCAGATCTTACAGCAATTAATGGTC	**-16bp**
E9	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCACGTTTGAAGTTTCAGCAGATCTTACAGCAATTAATGGTC	**+1bp**
	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCACGT-GGAGTTG-CAGCAGATCTTACAGCAATTAATGGTC	**-8bp, +7bp**
	TGAAGTTT>GGAGTTG	
H9	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCACGTTTGAAGTTTCAGCAGATCTTACAGCAATTAATGGTC	**+1bp**
	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCACGT--GAAGTTTCAGCAGATCTTACAGCAATTAATGGTC	**-1bp**
B10	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCACGTTTGAAGTTTCAGCAGATCTTACAGCAATTAATGGTC	**+1bp**
	CGTTGCTCAGCGAGTCCCCTGGTAAG----------TGAAGTTTCAGCAGATCTTACAGCAATTAATGGTC	**-8bp**
F10	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCACGTTTGAAGTTTCAGCAGATCTTACAGCAATTAATGGTC	**+1bp**
	CGTTGCTCAGCGAGTCCCCTGGTAAG----------TTTCAGCAGATCTTACAGCAATTAATGGTC	**-13bp**
G10	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCACGTTTGAAGTTTCAGCAGATCTTACAGCAATTAATGGTC	**+1bp**
	CGTTGCTCAGCGAGTCCC----------TGAAGTTTCAGCAGATCTTACAGCAATTAATGGTC	**-16bp**
D12	CGTTGCTCAGCGAGTCCCCTGGTAAGCTCCACGTTTGAAGTTTCAGCAGATCTTACAGCAATTAATGGTC	**+1bp**
	CGTTGCTCAGC----------AGATCTTACAGCAATTAATGGTC	**-35bp**

electroporations to generate these lines can be found in Appendix 3B. Since the lentiviral strategy described above worked we did not continue along this path, and thus further characterization of these cell lines by Southern blotting is required. However, these lines have been generated and could be used in future studies.

2.3.8 DICER1 is Essential for Self-Renewal in hESCs

To investigate whether these *DICER1* knockout hESC lines require DICER1* expression for survival, we removed doxycycline for 6 days (Figure 2.13A). Although displaying reduced levels of mature microRNAs at this stage, the *DICER1*-depleted hESCs survived and even seemed to grow faster as observed by alkaline phosphatase (AP) staining (Figure 2.13B). The apparent increased proliferation may be attributed to a slower proliferation of hESCs on doxycycline, but control experiments are needed to test this hypothesis. Moreover, the *DICER1*-depleted hESCs maintained comparable protein levels of pluripotency markers OCT4, NANOG, and SOX2 by immunofluorescence, albeit displaying slightly lower transcript levels by qRT-PCR (Figures 2.14A, B). This lower transcript level of pluripotency markers resembles what is observed in the *DICER1* hypomorphs (Figure 2.4A).

Since the majority of mature microRNAs are stable and have half-lives longer than 18hs, it is likely that a longer treatment without doxycycline was required to fully deplete these lines of mature microRNAs (Guo et al., 2015). Thus, we passaged these DICER1-depleted cells and maintained them in the absence of doxycycline for another 6 days (Figure 2.13A). After a total of 12 days of doxycycline removal there were very few DICER1-depleted hESCs remaining in the dish, except for the C2 line where clear colonies were

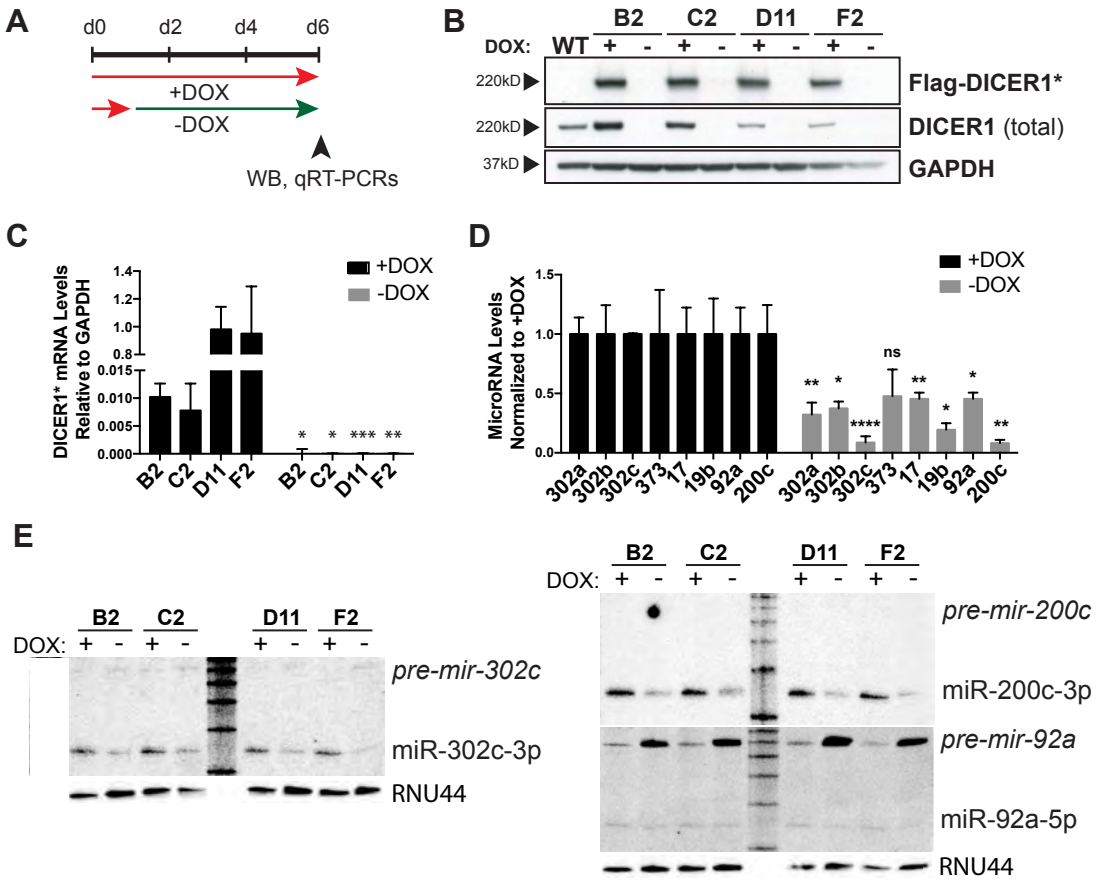


Figure 2.12. Validation of DICER1 -/- hESC Lines. (A) Schematic representation of the experimental procedure. (B) qRT-PCR analysis of DICER1* transgene expression in four DICER1 mutant lines with and without doxycycline treatment for 5 days. Levels relative to GAPDH. n=3. (C) Western blot of Flag-DICER1* and total DICER1 levels of four DICER1 mutant lines (B2, C2, D11, F2) with and without doxycycline treatment for 5 days. (D) Mature microRNA qRT-PCR analysis of eight hESC-expressed microRNAs in B2 DICER1 KO hESC line. Levels normalized to +DOX. n=3. (E) Northern blots of three hESC-expressed microRNAs (hsa-miR-302c-3p, hsa-miR-200c-3p, hsa-miR-92a-3p) in DICER1 -/- hESC lines treated with or without doxycycline for 5 days. RNU44 was used as loading control. Pre-mir, precursor microRNA; miR, mature microRNA.

growing (Figure 2.13C). The few remaining DICER1-depleted cells could not be propagated, but we could maintain the C2 line in culture. As previously mentioned, the C2 line is a *DICER1* hypomorph, and thus we were expecting this line to survive upon doxycycline removal. This argues against the hypothesis that adaptation of the *DICER1* hypomorph (B4) in culture causes its survival since this line survives despite immediate withdrawal of DICER1*. *DICER1* knockout hESC phenotype was partially rescued by inducing ectopic DICER1* expression upon replating (Figure 2.16B). These results confirmed that DICER1 is required for hESC self-renewal, unlike in mESCs.

To explore whether the *DICER1* knockout phenotype is caused by compromised adherence to the extracellular matrix (ECM), we followed the cells from day 7 to day 12 of doxycycline removal and checked for survival of pluripotent cells by AP staining and cell growth by cell counting (Figures 2.13D, E). Upon closer examination, we observed that the cells were properly attaching after replating, which means that at least initial adherence to the ECM remains intact. Along these lines, qRT-PCR of E-cadherin, N-cadherin and EMT-promoting transcription factors on day 6 shows that E-cadherin and N-cadherin levels remain the same as in wildtypes, albeit some downregulation of ZEB2, TWIST1, and SNAIL is observed, which indicates that epithelial-to-mesenchymal transition (EMT) is not occurring at this point (Figure 2.14C). Instead, the DICER1-deficient cells appear to die around days 9-10 of doxycycline removal (Figures 2.13D, E). This effect was also observed when plating four times as many cells, arguing against the possibility of a density-dependent phenotype (Figure 2.13F).

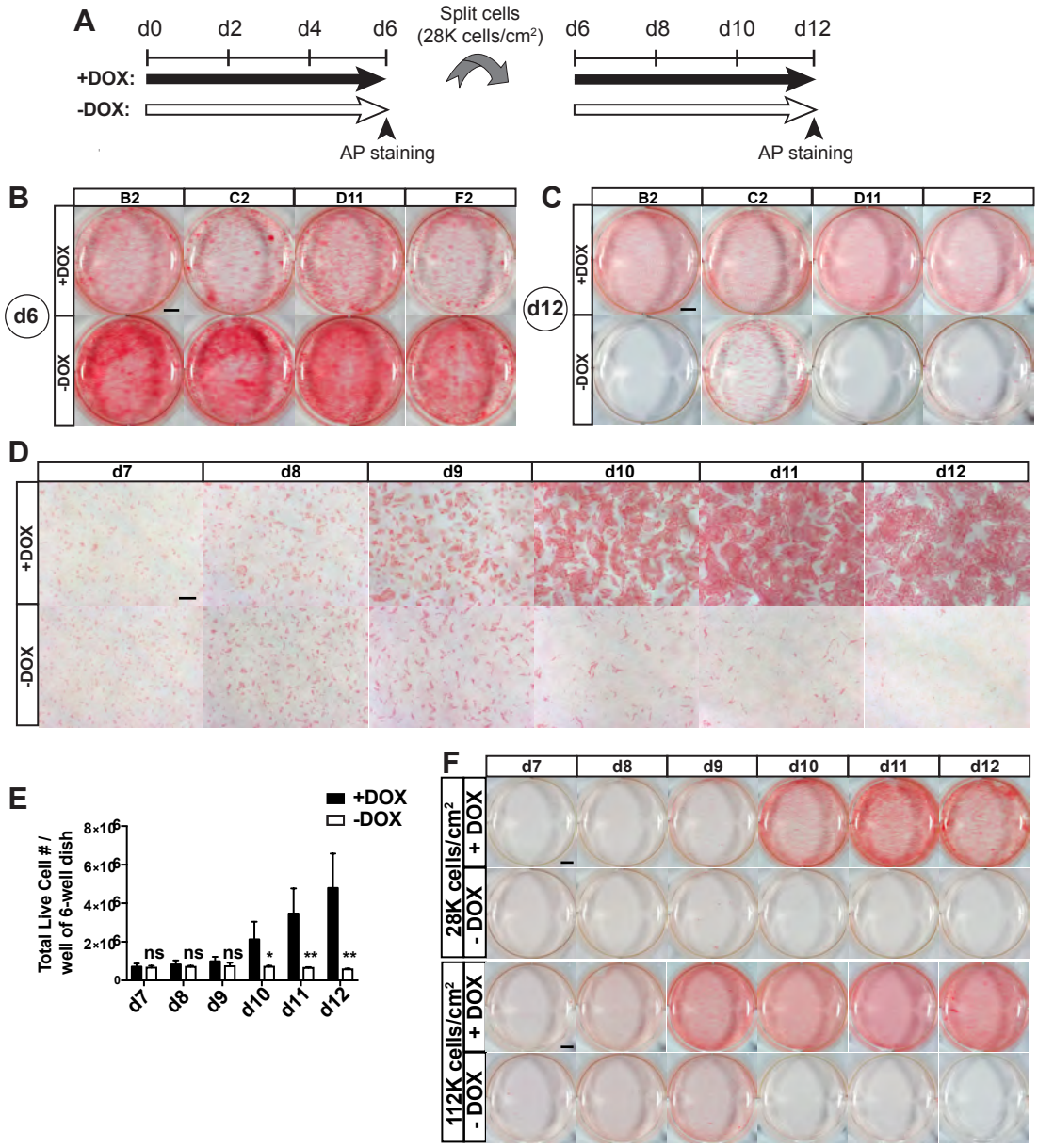


Figure 2.13. DICER1 is Essential in Human Embryonic Stem Cells. (A) Schematic representation of experimental procedure. AP, alkaline phosphatase; DOX, doxycycline. (B) Alkaline phosphatase (AP) staining of four DICER1 mutant lines with and without doxycycline treatment for 6 days. Scale bar represents 500mm. Levels relative to GAPDH. n=3. (C) Alkaline phosphatase (AP) staining of four DICER1 mutant lines treated with and without doxycycline treatment for 6 days, and then replated and grown with and without doxycycline for another 6 days. Scale bar represents 500mm. (D) Schematic representation of experimental procedure. Close-up pictures of alkaline phosphatase (AP) stained B2 DICER1 -/- hESC line with and without doxycycline after replating between days 7 and 12. Scale bar represents 500um.(E) Cell number quantification of B2 DICER1 -/- hESC line after replating between days 7 and 12. n=5 for d7-9, n=4 for d10, and n=3 for d11-12. (F) Whole-well view of AP-stained B2 DICER1 -/- after being replated at different densities (28,000 cells or 112,000 cells per cm^2) from days 7 to 12 with or without doxycycline treatment. Scale bar represents 500mm.

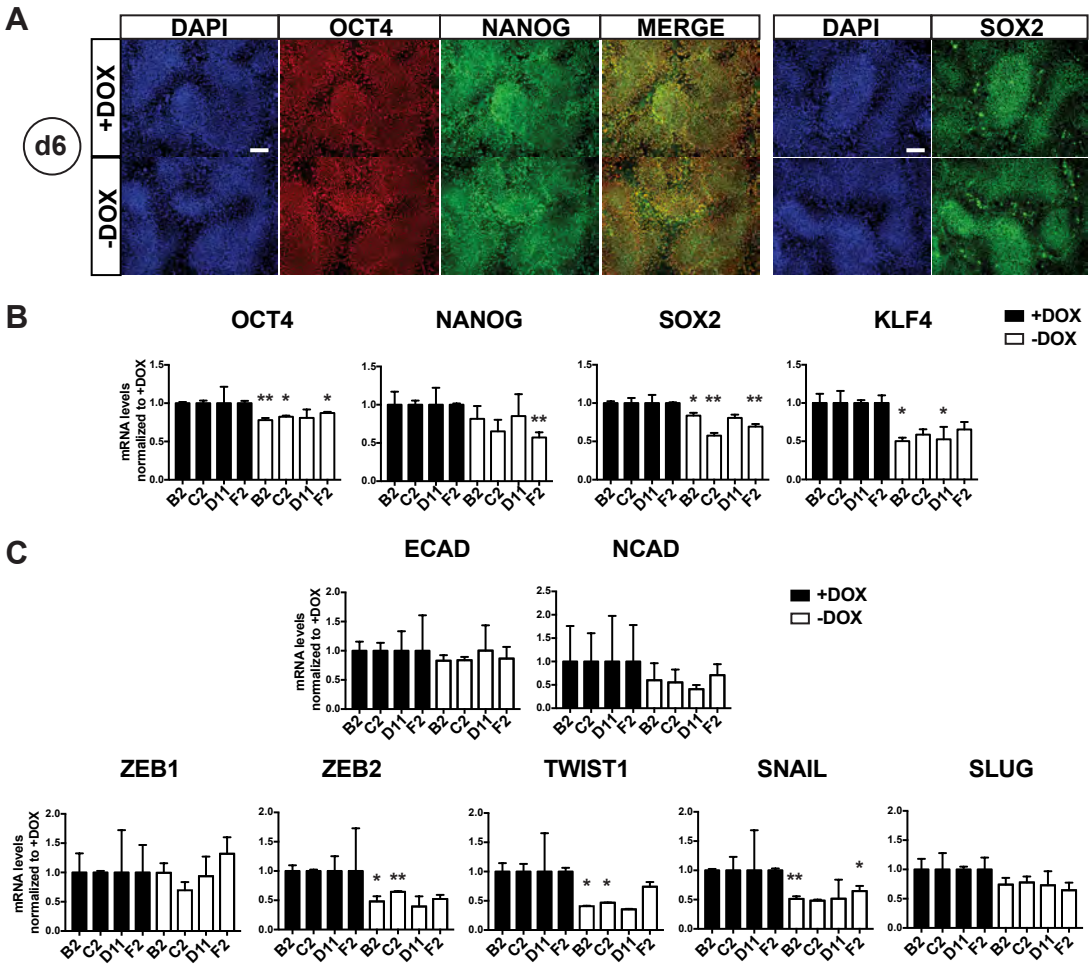


Figure 2.14. DICER1 -/- Lines Maintain Pluripotency Marker Expression and Do Not Seem to Undergo EMT Five Days After Doxycycline Removal. (A) Immunofluorescence staining of pluripotency markers (OCT4, NANOG, SOX2) in B2 DICER1 -/- with and without doxycycline treatment for 6 days. (B) qRT-PCR analysis of pluripotency marker expression (OCT4, NANOG, SOX2, and KLF4) in four DICER1-/- lines with and without doxycycline treatment for 5 days. Levels relative to GAPDH. n=3. (C) qRT-PCR analysis of E-Cadherin and N-Cadherin as well as EMT-promoting trancription factors ZEB1, ZEB2, TWIST1, SNAIL and SLUG in four DICER1.-/- lines with and without doxycycline treatment for 5 days. Levels relative to GAPDH. n=3.

2.3.9 Loss of DICER1 in hESCs Causes Caspase-3-Mediated Apoptosis

To explore whether hESCs require DICER1 for survival, we assessed Cleaved Caspase-3 levels by flow cytometry at days six through eleven after doxycycline withdrawal (Figure 2.15A). As a control, we kept the cells on doxycycline. We did not check levels at day twelve since there were not enough cells remaining at that timepoint. Flow cytometry showed that Cleaved Caspase-3 levels increase significantly when Doxycycline, and therefore DICER1*, is removed (Figure 2.15B). This increase is significant starting from day eight after Doxycycline removal and peaks at day ten (Figure 2.15B, C), which correlates with the alkaline phosphatase staining data (Figure 2.13D). In fact, at day ten there is a 6-fold increase in Cleaved Caspase-3 staining (Figure 2.15C). These results were confirmed by immunofluorescence staining of Cleaved Caspase-3 at days nine and ten when a clear increase in Cleaved Caspase-3 protein is observed (Figure 2.15D). Reassuringly, we find nuclei condensation and loss of OCT4 in Cleaved Caspase-3 positive cells (Figure 2.15D, inset).

2.3.10 Five Days of Doxycycline Removal Results in Few Changes in Transcriptome

To investigate global transcriptome changes in the *DICER1* hESC knockouts after five days of doxycycline withdrawal, we performed deep sequencing analysis of *DICER1* knockout (D11 cultured without doxycycline for 5 days) and wildtype cells (D11 cultured with doxycycline) followed by analysis performed by Professor Todd Evans. The results showed that after five days of doxycycline withdrawal the transcriptome remains almost unchanged (Figure 2.16A). As mentioned earlier, this may not be surprising

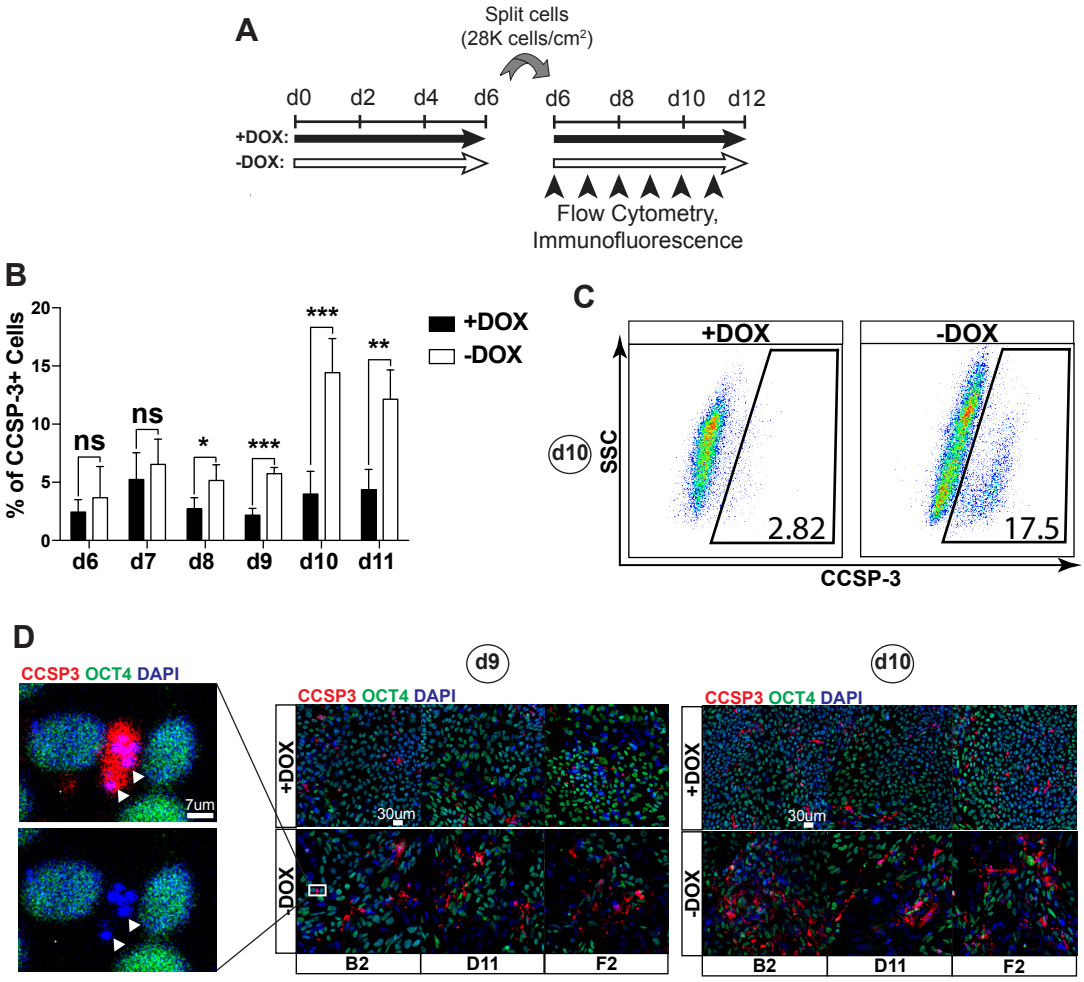
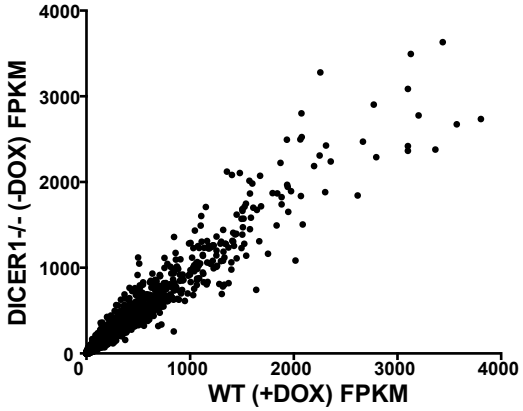


Figure 2.15. Loss of DICER1 in hESCs Causes Caspase-3 Mediated Cell Death. (A) Schematic representation of the experimental design. (B) Flow cytometry of Cleaved Caspase-3 from days 6 through 11 in D11 mutant hESCs with and without Doxycycline. CCSP-3, cleaved caspase-3; ns, not significant; DOX, doxycycline. n=3. (C) Representative flow plots of D11 mutant line with and without Doxycycline at day 10. CCSP-3, cleaved caspase-3; DOX, doxycycline. (D) Representative immunofluorescence pictures of cleaved Caspase-3 and OCT4 in D11 mutant hESCs with and without Doxycycline at days 9 and 10. Inset represents a close-up of showing nuclei fragmentation and loss of OCT4. Scale bar for closeup= 7um. Scale bar for main images= 30um. n=3.

A **Scatter Plot:** DICER1-/- vs. WT hESCs



Fold Change >2, p-val <0.05

Genes upregulated in DICER1-/-	44
Genes downregulated in DICER1-/-	38
Total dysregulated genes	82 / 16384

Fold Change >1.1, p-val <0.05

Genes upregulated in DICER1-/-	350
Genes downregulated in DICER1-/-	406
Total dysregulated genes	756 / 16384

B

Top 20 Upregulated Genes in DICER1-/-		Top 20 Downregulated Genes in DICER1-/-	
RP1-167F1.2	antisense	RP4-536B24.2	antisense
LDHA	protein_coding	RP11-685G9.2	antisense
IGF2	protein_coding	FERMT2	protein_coding
PRKAR1A	protein_coding	STC2	protein_coding
TYW3	protein_coding	AC004771.1	pseudogene
PLAUR	protein_coding	CTC-550B14.6	processed_transcript
SIM2	protein_coding	FLT1	protein_coding
RNF10	protein_coding	MT-TT	Mt_tRNA
PCDHA8	protein_coding	IFNWP19	pseudogene
EMC4	protein_coding	SLC7A11-AS1	processed_transcript
CTD-2325A15.5	antisense	PPP1R3B	protein_coding
PSME3	protein_coding	ARAP3	protein_coding
PDSS1	protein_coding	ADC	protein_coding
SPAG4	protein_coding	RP11-789C2.1	lincRNA
NEFH	protein_coding	PCK2	protein_coding
GPC3	protein_coding	GRB10	protein_coding
ATP1A2	protein_coding	KLHL18	protein_coding
NCAPG2	protein_coding	LINC00645	lincRNA
MT2P1	pseudogene	RP11-431J17.1	lincRNA
R3HDM4	protein_coding	CTD-3193O13.9	protein_coding

Figure 2.16. Deep Sequencing in DICER1-/- After Five Days of Doxycycline Removal Reveals Few Changes in Transcriptome. (A) Scatter plot of RNA-Sequencing results of DICER1-/- hESCs (D11 Minus Doxycycline) compared to WT hESCs (D11 Plus Doxycycline). Only eighty-two genes are significantly dysregulated by a factor of 2 or more, and seven hundred and fifty-six genes are significantly dysregulated by a factor of 1.1 or more. (B) A list of the top twenty upregulated and downregulated genes in DICER1-/-. Selection criteria involved genes that were significantly dysregulated (p-value <0.05) and by a factor of 2 or more. In green, FERMT2, a protein of interest.

given that most microRNAs are very stable and have half-lives longer than 18 hours (Guo et al., 2015), and we do see reduced but still present expression of all microRNAs tested (Figures 2.12D, E). Nevertheless, we found 44 genes with a 2-fold or more upregulation and a p-value <0.05, and 38 genes with 2-fold or more downregulation and a p-value <0.05 (Figure 2.16A). We found more genes, 756 to be exact, with subtle but significant changes above 1.1-fold (Figure 2.16A). Gene ontology (GO) analysis did not reveal any significantly enriched GO categories in the upregulated genes. However, we listed the top 20 upregulated and top 20 downregulated genes in the *DICER1* knockout cells that had been depleted of doxycycline for 5 days (Figure 2.16B). It will be informative to perform deep sequencing at ~day 8-9 when the cells start dying and have presumably downregulated the microRNAs enough to show a phenotype.

2.3.11 ESCC miRNA Clusters 371-373, 302-367 but not 17-92 Rescue DICER1 Knockout hESCs

To address whether hESC self-renewal requires miRNAs or other DICER1-dependent functions, we performed a rescue experiment with select microRNA families. We chose the ESCC-family of miRNAs (miR-302-367, miR-371-373, and miR-17-92) since they have been previously implicated in mouse ESC self-renewal (Greve et al., 2013; Wang et al., 2008). Additionally, we tested the miR-200 family because it has been shown to be important in reprogramming of human fibroblasts to iPSCs and because its role in hESCs has not yet been elucidated (Samavarchi-Tehrani et al., 2010; Suh et al., 2004; Wang et al., 2013a). Thus, we selected these four families to transfect into *DICER1* knockout hESCs and probe their function further. Specifically, we

cultured the cells without doxycycline for 6 days, and then passaged and transfected them with miRNA mimics, individually or all members of a cluster. Positive and negative controls were transfected with a control miRNA mimic that contains a non-specific sequence and that doesn't bind to any mRNAs. We allowed these cells to grow until day 12, at which point we either fixed them and stained for alkaline phosphatase to visualize pluripotent cells, or we counted live cells for quantification (Figure 2.17A). These experiments were performed in two *DICER1* knockout lines, D11 and B2 (Figures 2.17B, C and Appendix 4B, respectively). Criteria for rescuing the *DICER1* knockout hESC-phenotype was based on at least a 25% rescue as assessed by live cell counting by an individual miRNA or miRNA cluster.

The most striking rescues were observed for the miR-302-367 and miR-371-373 clusters by alkaline phosphatase (AP) staining and cell counting (Figure 2.17B, C). miR-302-367 rescued 45% of the number of cells in the positive control, whereas miR-371-373 rescued 36% (Figure 2.17C). Since ~60% of the cells are successfully transfected with miRNA mimics (Gonzalez et al., 2014), it is hard to predict whether these microRNAs are partially rescuing or fully rescuing the *DICER1* knockout hESC phenotype. Thus, the possibility that there are other microRNAs required to fully rescue the phenotype cannot be excluded at this time. The other two clusters, miR-17-92 and miR-200, had milder rescuing effects that lay under our 25% cutoff and therefore we did not consider them significant (Figure 2.17C). Notably, members of the miR-17-92 cluster that were shown to be sufficient to rescue the proliferation defect in miRNA-depleted mESCs (Wang et al., 2008), appear not to be required in miRNA-depleted hESCs. Also, although we did not observe high percentage rescue for the miR-200 cluster, the rescued cells

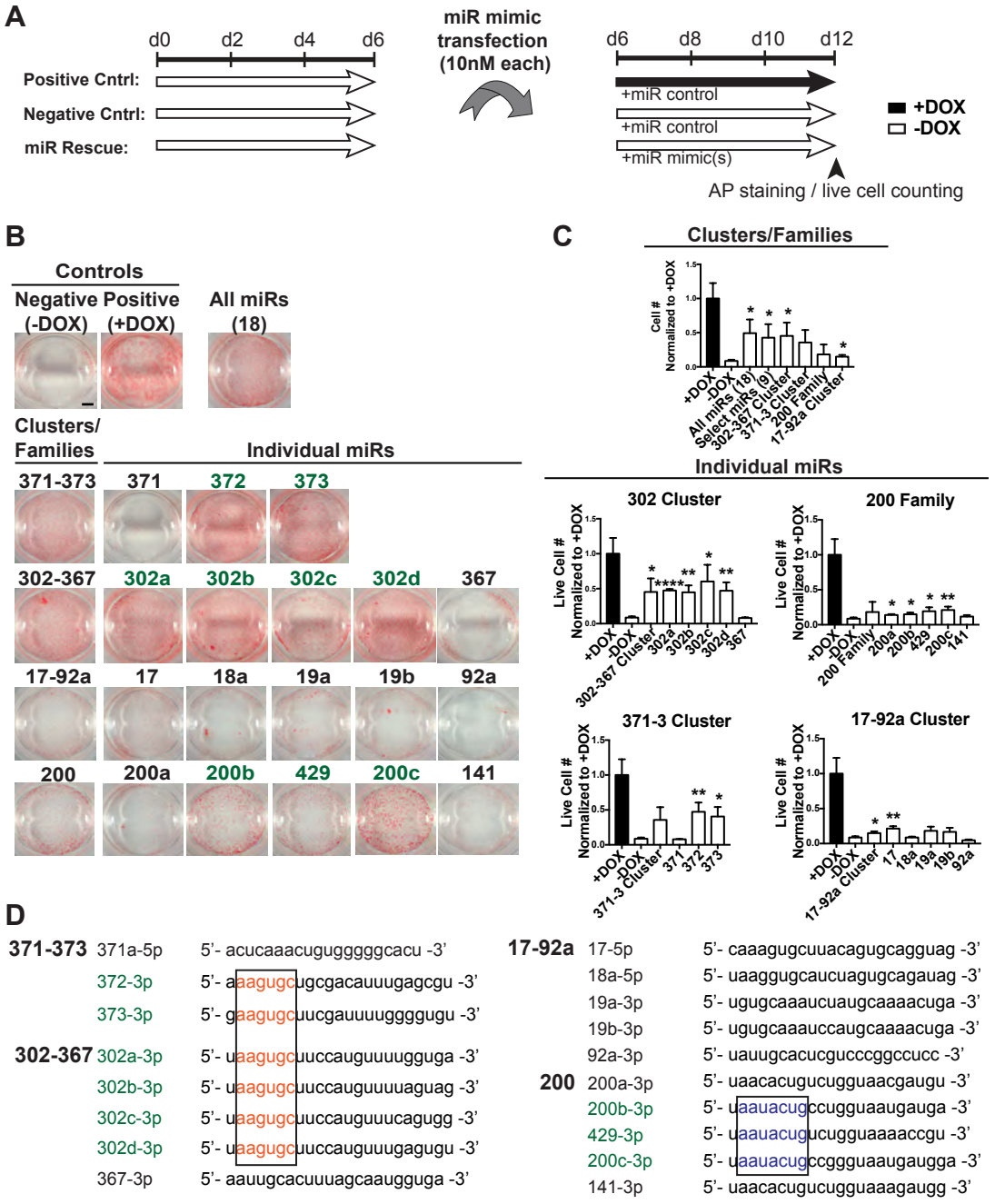


Figure 2.17. ESCC MicroRNAs 371-373 and 302-367, but not 17-92, Rescue DICER1-/- hESCs. (A) Schematic representation of microRNA mimic transfection procedure. Cntrl, control; miR, mature microRNA; DOX, doxycycline; AP, alkaline phosphatase. (B) Whole-well view of alkaline phosphatase (AP)-stained DICER1-/- hESCs on day twelve treated with or without microRNA transfection of individual microRNAs or groups of microRNAs. In green, nine microRNAs that rescue the DICER1-/- phenotype at varying degrees. Representative images of four independent experiments. (C) Quantification of miRNA rescue by cell counting of 24-well plates, and normalized to plus doxycycline control. n=3. (D) List of mature microRNAs (3p or 5p) transfected in D11 DICER1-/- hESCs. In green, mature microRNAs that significantly rescued the DICER1-/- phenotype. The boxes represent the common seed sequence amongst those mature microRNAs that rescue the DICER1-/- phenotype.

appeared to form refractive edges with an appearance resembling that of naïve cells, and they stained highly for alkaline phosphatase (Figure 2.17B). These observations were specific to miRNAs 200b-3p, 429-3p, and 200c-3p, which remarkably share the 'AAUACUG' seed sequence unlike the other two members of the family, miRNAs 200a-3p and 141-3p. It will be interesting to study whether this 'AAUACUG' seed sequence plays a role in the naïve stem cell state.

Interestingly, not all members of the miR-302-367 and miR-371-373 clusters rescued the *DICER1* knockout hESC phenotype. For example, although miRNAs 302a, 302b, 302c, and 302d significantly rescued between 44-60% of the cells in the positive control, miR-367 did not rescue (a non-significant 8%) (Figures 2.17B, C). Similarly, although miRNAs 372 and 373 rescued 47% and 41% of the cells in the positive control respectively, miRNA 371 failed to rescue (a non-significant 8%) (Figures 2.17B, C). Notably, the individual miRNAs that rescued in these two clusters contain the ESCC seed sequence 'AAGUGC', whereas miR-371a-5p and miR-367-3p do not. Since the seed sequence is predominantly responsible for mRNA target specificity (Bartel, 2009), these results emphasize the relevance of the ESCC seed sequence in survival of hESCs, and highlights the value of our system to identify individual key players.

To address whether the rescued cells expressed proper pluripotency markers, we analyzed expression of OCT4 and NANOG by immunofluorescence for each individual miRNA tested as well as the grouped miRNAs (Appendix 4 and data not shown). Individual, pooled and clusters of miRNAs transfected in *DICER1* knockout hESCs resulted in survival of hESCs

expressing the proper pluripotency markers, indicating that these miRNAs restored self-renewal of *DICER1* knockout hESCs.

2.3.12 DICER1 Seems Not to be Essential in Naïve hESCs

Although both derived from the pre-implantation blastocyst stage (Brons et al., 2007), mESCs and hESCs are believed to represent different embryonic developmental stages, in which hESCs represent a slightly later developmental stage, and thus resemble mEpiSCs more than mESCs (Tesar et al., 2007; Weinberger et al., 2016). Therefore, hESCs are classified as 'primed' ESCs, whereas mESCs are classified as 'naïve' ESCs (Weinberger et al., 2016). To investigate whether the requirement for DICER1 in hESCs is specific to the 'primed' state, we investigated whether our *DICER1* knockout hESCs would survive if they were converted to the naïve state. To this end, we used the Stem Cell Technologies RSeT Media, which is based on Jacob Hanna's protocol (Gafni et al., 2013), to convert the four *DICER1* mutant hESCs (B2, C2, D11, F2) to the naïve state under hypoxic conditions (5% O_2). This was carried out in the presence of doxycycline to maintain DICER1* levels in these *DICER1* knockout hESCs. By the third passage in naïve conditions, all four *DICER1* knockout hESCs under naïve conditions acquired typical naïve stem cell morphology with tightly-packed, dome-shaped colonies with refractive edges (Figures 2.18A, C). Molecularly, they activated key transcripts associated with naïve-like ESCs, such as STELLA and DNMT3L (Figure 2.18B). Thus, proper conversion from the primed to the naïve-like ESC state was achieved.

Next, we removed doxycycline from the culture to test whether the *DICER1* knockout hESCs could survive in naïve conditions without the

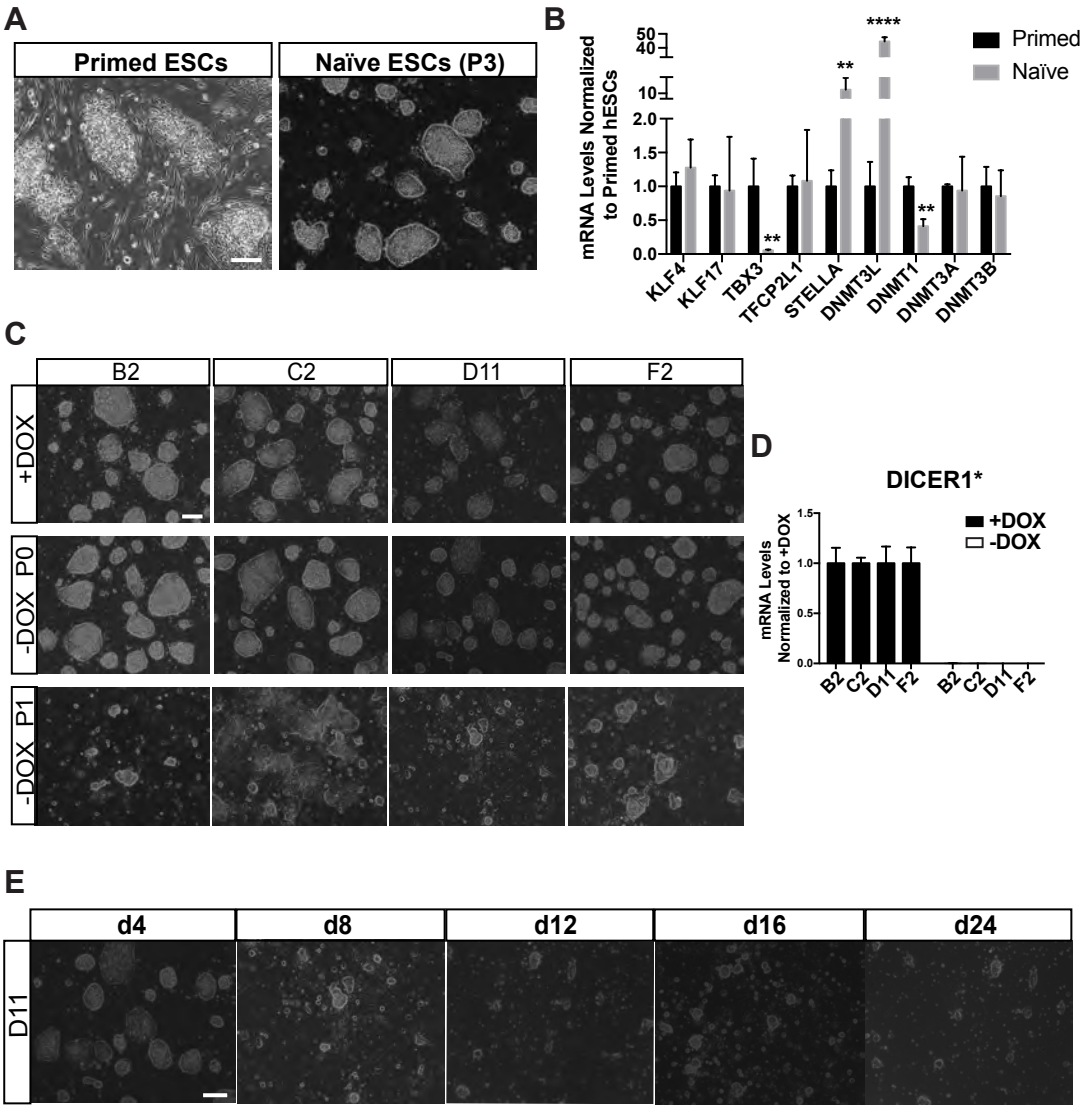


Figure 2.18. DICER1 is Apparently Not Essential in Naïve hESCs Lines. (A) Brightfield images of primed human embryonic stem cells (HUES8 line) and naive human embryonic stem cells at passage 3 (P3) of primed-to-naive conversion. Scale bar represents 100um. (B) qRT-PCR of markers associated with naive and primed stem cell states in primed and naive TRE-DICER1* lines (with doxycycline). mRNA levels are normalized to primed embryonic stem cells (HUES8 line). n=3 (for naive embryonic stemcells this corresponds to passages 3, 4 and 5 of primed-to-naive conversion). (C) Brightfield images of DICER1-/- (TRE-DICER1*) lines with doxycycline and without doxycycline for five days (P0, passage 0) and for ten days (P1, passage 1). Scale bar represents 100um. (D) qRT-PCR graph showing RNA levels of DICER1* transgene being downregulated upon doxycycline removal for five days. (E) Brightfield images of DICER1-/- naive human embryonic stem cells surviving despite long-term withdrawal of doxycycline. Five passages correspond to ~25 days of doxycycline withdrawal. Scale bar represents 100um.

exogenous DICER1* protein. The doxycycline was removed for five days and we confirmed the absence of DICER1* transgene expression (Figure 2.18D). As expected from our results in the primed state (Figure 2.13B), the cells did not die (Figure 18.C, “-DOX P0” panel). However, when these cells were replated again without doxycycline, the cells survived, albeit in decreased numbers (Figure 2.18C, “-DOX P1” panel). These naïve cells could be maintained in naïve conditions in the absence of doxycycline for at least five passages (Figure 2.18E). These results contrast what was observed for primed *DICER1* knockout hESCs that cannot be maintained when depleted of exogenous DICER1* (Figure 2.13C). Notably, these DICER1-depleted naïve-like cells proliferated very slowly and were hard to maintain as a line, similar to observations reported for *Dicer1* and *Dgcr8* knockout mESCs (Kanellopoulou et al., 2005; Murchison et al., 2005; Wang et al., 2007).

2.4 DISCUSSION

Analysis of *DICER1* knockout hESCs has allowed us to study miRNA requirements in the primed state of pluripotency. We have shown that DICER1 is essential for hESC survival, which is in contrast to its nonessential role in mouse ESCs. This is an interesting finding since DICER1 knockdown in hESCs had previously rendered viable cells, albeit displaying slower proliferation rates similar to the phenotype in mouse ESC knockouts (Murchison et al., 2005; Qi et al., 2009). We also generated two *DICER1* hypomorph hESC lines which were viable, did not show increased levels of apoptosis but, in contrast to the *DICER1* knockdown, seemed to proliferate normally. A possible explanation in the discrepancy between the *DICER1* knockout and the knockdown/hypomorph phenotypes might be that complete

lack of DICER1 is required to render the cells inviable, whereas reduced levels might impair certain functions but allow cell survival.

Notably, we found that the human *DICER1* knockout displays a different phenotype than the mouse knockout. There are two plausible explanations for this apparent discrepancy. The first is that the *Dicer1* knockout mESCs may have been selected for viability in culture and may not be viable in the *in vivo* context. Given that *in vivo* studies of *Dicer1* knockout at pre-implantation stages are compromised due to maternal Dicer1 contribution (Spruce et al., 2010), Dicer1 requirement at these early stages has not been elucidated. In fact, some groups have questioned whether the *Dicer1* knockout mouse ESCs are truly viable or result from a cell culture artifact (Greve et al., 2013; Murchison et al., 2005). In particular, Murchison and colleagues reported that most *Dicer1* knockout clones they generated were initially severely impaired in proliferation, and that after two weeks in culture, two lines eventually 'escaped' and achieved growth rates slower than but approaching that of Dicer wildtype lines, which could only then be maintained *in vitro* (Murchison et al., 2005). Our strategy to study *DICER1* knockouts in hESCs using a DICER1* transgene is specifically designed to bypass issues with cell culture adaptation and is thus particularly fit to study acute DICER1 depletion during development.

However, a more likely explanation for the discrepancy between the mouse and human requirement for Dicer1 might be the established notion that mouse and human ESCs represent different pluripotent states. Recently, knockout studies for *EZH2* and *DNMT3A* have shown essential phenotypes in human ESCs, contrary to their nonessential roles in mouse ESCs (Collinson et al., 2016; Liao et al., 2015; Shen et al., 2008; Tsumura et al., 2006). In fact,

our preliminary studies with naïve *DICER1* knockout hESCs suggest that this is the case since naïve *DICER1* knockout hESCs are able to survive, albeit with compromised proliferation rates. Since human ESCs model the post-implantation epiblast, this may indicate that this is the first developmental period when DICER1 becomes essential. In fact, a similar phenotype has been shown for mouse epiblast stem cells where *DICER1* deletion causes massive apoptosis (Pernaute et al., 2014). It will be interesting to perform more extensive characterization of *DICER1* knockouts in naïve versus primed hESCs to detail miRNA function in these contexts.

Dicer1 is not essential for survival in all cell types during development. For example, *Dicer1* conditional knockout studies of the adrenal gland showed that Dicer1 is dispensable in the specification and survival of the fetal adrenal cortex (Krill et al., 2013), and another study indicated that it is not required for survival in the γδ lineage cells during T cell development (Cobb et al., 2005). Similarly, it is not essential for survival in the colorectal cancer cell line HCT116 (Kim et al., 2016). However, Dicer1 is essential for survival in other developmental cell types such as the limb mesoderm at E10.5, the αβ lineage during T cell development, and in melanocyte stem cells and differentiated melanocytes (Cobb et al., 2005; Harfe et al., 2005; Levy et al., 2010). Thus, it seems like Dicer function is not universally required for survival but its requirement is instead context-dependent and is likely contingent on cell-type specific miRNA function.

In this study, we generated a platform to investigate cell-type specific miRNA requirements in primed and naïve pluripotency that can also be used to address miRNA requirements in differentiation of hESC-derived cell types. An advantage of our system is that it allows us to study the contribution of

individual microRNAs and microRNA clusters without encountering the problem of masked phenotypes due to redundancy in the miRNA network (Park et al., 2010). Of the more than forty reported genetic deletions of individual miRNAs, none have demonstrated embryonic defects prior to E14.5 (Kuhnert et al., 2008; Liu et al., 2008; Park et al., 2010; Ventura et al., 2008; Zhao et al., 2007). However, double knockouts for miRNA-302a-d and miRNA-290 clusters in mice exhibit an earlier phenotype than either knockout cluster alone, suggesting an earlier redundant function for these clusters (Medeiros et al., 2011). The contribution of individual miRNAs to this earlier function is very hard to study in mice, but our hESC platform is particularly suited to easily dissect miRNA function in this context. We found that the miRNA-302-367 and miRNA-371-373 clusters are necessary and sufficient on their own for survival of primed ESCs, which indicates that they act redundantly and that they are essential for survival during the post-implantation embryo. Additionally, we were able to distinguish which individual miRNAs within these clusters are functionally relevant. It will be interesting to study the exact mechanism of action in future studies.

Intriguingly, we did not find that the ESCC miRNA-17-92 cluster rescues *DICER1* knockout viability in hESCs. This result contrasts its role in miRNA-depleted mESCs where miRNA-19a was shown to rescue the proliferation defect (Wang et al., 2008), as well as the function of miRNA-92a in rescuing cell death in *Dicer1* knockout mEpiSCs (Pernaute et al., 2014). This difference in the role of the miRNA-17-92 cluster suggests that miRNA function differs between mouse and human, and underpins the importance of assessing human-specific miRNA requirements by performing studies in human ESCs.

Even though we identified two clusters that are necessary and sufficient to rescue survival in *DICER1* knockout hESCs, there may be other microRNAs that are required and sufficient in this context. Our platform is particularly suitable to investigate novel functions of miRNAs in the primed state, and it will be interesting to perform a blind miRNA rescue screen to uncover unknown miRNA functions in the future.

Taken together, our study provides a comprehensive examination of DICER1 and miRNA function in hESC pluripotency. Our study uncovers human to mouse differences in DICER1 requirement, and provides a platform to discover novel and essential miRNAs players in early human development.

CHAPTER 3: Conclusions & Perspectives

3.1 SUMMARY

Precise microRNA functional studies in the post-implantation stage of mouse development have been lacking due to redundancy issues in the microRNA network. Moreover, human specific miRNA requirements have not been assessed in primed ESCs. The purpose of my thesis work was to elucidate global and specific miRNA function in primed pluripotency as a proxy to the *in vivo* post-implantation stage. Thus, we set out to knockout *DICER1* in human ESCs for the first time and to evaluate miRNA-specific requirements in the primed state.

To our surprise, we were initially unable to generate *DICER1* knockout ESCs through CRISPR targeting. Nevertheless, we obtained two *DICER1* hypomorphic lines that we used for assessing the contribution of DICER1 in hESC pluripotency. *DICER1* hypomorphs maintained the overall levels of pluripotency factors, although we detected a slight decrease in the expression of OCT4 and NANOG transcript levels. Importantly, we did not observe a proliferation defect in the hypomorphs nor did we detect any changes in cell survival. However, we did find that the *DICER1* hypomorphs had impaired colony forming ability, and the cells downregulated E-CADHERIN and upregulated ZEB1, which are signs that the cells undergo epithelial-to-mesenchymal transition and that could, at least partially, explain the colony forming disability. Contrary to *Dicer1* mouse knockouts, our *DICER1* hypomorphs were able to differentiate into all three germ layers as evidenced by the formation of teratomas displaying structures of all three lineages. Moreover, they were able to properly upregulate differentiation markers and

downregulate pluripotency markers upon directed differentiation into neuroectoderm cells. In line with these results, RNA-Sequencing revealed an increase in developmental transcription factors in the hypomorphs. We tested the expression of mammalian Argonautes (AGOs) 1-4 in the hypomorphs and found the transcript levels of AGOs 2-4 decreased as well as the protein levels of AGO2, arguing against the possibility of AGO2 playing a role in processing the mature miRNAs found in the DICER1-compromised cells.

Our results suggested that DICER1 is essential in hESCs, and we confirmed this by generating a doxycycline-inducible line that expressed a gRNA-immune DICER1* transgene to enable the retrieval of *DICER1*$^{-/-}$ hESCs during CRISPR targeting. These knockout lines downregulated DICER1* upon doxycycline removal which caused depletion of miRNAs. miRNA-depleted hESCs did not survive passaging, demonstrating that DICER1 is indeed essential for survival in hESCs. Specifically, *DICER1* knockout hESCs died around days 9-10 of doxycycline withdrawal if they were passaged. Plating four times as many cells did not prevent the *DICER1* knockout cells from dying, suggesting that the phenotype is not density-dependent. Flow cytometry of Cleaved Caspase-3, an apoptotic marker, showed that hESCs in the absence of DICER1 undergo cell death. RNA-Sequencing studies after five days of doxycycline depletion revealed very few transcriptional changes, which might reflect the fact that miRNAs, which are known to be very stable, perdure at this timepoint. Nevertheless, we detected minor downregulation of pluripotency factors OCT4, NANOG, SOX2 and KLF4.

Additionally, we found that ESCC microRNA clusters 302-367 and 371-373, but not miRNA-17-92, rescue self-renewal in hESCs. Although we found that the miRNA-200 family did not significantly rescue the number of cells that

survived, we observed that the cells that did survive formed dome-shaped colonies with refractive edges that resembled naïve stem cells.

Finally, we converted our *DICER1* knockout hESCs into the naïve state and removed doxycycline. Contrary to the knockout phenotype in the primed state, we found that *DICER1* knockout naïve-like hESCs survived passaging, suggesting that DICER1 may not be required in the naïve state.

3.2 FUTURE DIRECTIONS

3.2.1 Detailed characterization of DICER1 Knockout Phenotype in hESCs

We identified a requirement for DICER1 in survival of hESCs. However, the exact mechanism by which the cells die remains elusive. Studies assessing Cleaved Caspase-3 by flow cytometry and morphological evidence of cell shrinkage point towards apoptosis as the culprit of the *DICER1* knockout phenotype in hESCs. We have not yet identified whether the apoptosis is caused by the extrinsic or intrinsic (mitochondrial) pathways of apoptosis. These observations correlate with the increased apoptosis observed in *DICER1* knockout in mouse EpiSCs (Pernaute et al., 2014). It will be important to investigate what causes the apoptosis in *DICER1* knockout hESCs.

3.2.2 Mechanism of action of miRNA-302-367 and miRNA-371-373 in hESCs

We uncovered a pro-survival function in hESCs for two clusters, miRNA-302-367 and miRNA-371-373. Both clusters have been previously

implicated in pro-survival functions, albeit in different contexts. For example, the miRNA-302-367 cluster is sufficient to partially rescue apoptosis in Dicer1-deficient mEpiSCs (Pernaute et al., 2014). In mESCs, the miRNA-290 cluster (miRNA-371-373 in human) prevents apoptosis during exposure to genotoxic stress, which might be relevant during physiological stress in the developing embryo (Zheng et al., 2011). It will be important to elucidate the mechanism of action for this pro-survival function in hESCs. An informative experiment will be to perform RNA-Sequencing at ~days 8-9 of doxycycline withdrawal in our *DICER1* knockout hESCs to understand which mRNA transcripts are upregulated at this stage, indicating possible miRNA targets and revealing potential mechanistic insights.

3.2.3 Genome-wide miRNA Screen to Identify Novel miRNA Players in hESCs Pluripotency and Differentiation

We performed a targeted screen with four clusters of miRNAs, and identified key miRNA players in hESC self-renewal. However, our study does not tell us whether there might be additional miRNAs that are necessary and sufficient to rescue the *DICER1* knockout phenotype. It might be interesting to perform double knockouts of miRNA-302-367 and miRNA-371-373 to test whether they phenocopy the *DICER1* knockout phenotype in hESCs, which would indicate whether there might be other microRNAs required in the primed state. Regardless, it will be informative to perform a genome-wide miRNA screen, or at least an ESC-specific miRNA library screen, to identify novel miRNAs that can fully or partially rescue aspects of the *DICER1* knockout phenotype.

Another exciting application for our inducible *DICER1* knockout hESC platform is its potential to uncover miRNA requirement during differentiation into any cell type of interest. As long as doxycycline does not impair differentiation, this platform will be extremely useful to discover miRNA requirements in hESC differentiation and in early stages of development.

3.2.4 Identify Naïve Versus Primed miRNA Requirements in hESCs

We converted our *DICER1* knockout hESCs to the naïve state, and found that *DICER1* knockout naïve cells can survive but show compromised proliferation abilities. It will be interesting to uncover human specific miRNA requirements in the naïve state, as well as miRNAs that promote conversion between primed and naïve ESCs identity. Our platform is suitable for both studies.

3.3 CONCLUSIONS

My doctoral thesis revealed an unknown requirement for DICER1 in human ESC survival, and showed that miRNA function is required as early as the post-implantation period during embryonic development.

Significantly, we uncovered differences between human and mouse global miRNA function in pluripotency. We found that the regulation of cell survival is the primary role of miRNAs in the primed pluripotent state, whereas in mESCs, microRNAs are primarily important for proliferation and differentiation. The differing roles of miRNAs in mESCs and hESCs can be attributed to the different developmental stages that mESCs and hESCs represent. Indeed, in preliminary studies, naïve knockout hESCs survived

without DICER1, but showed compromised proliferation rates mimicking the mESC phenotype.

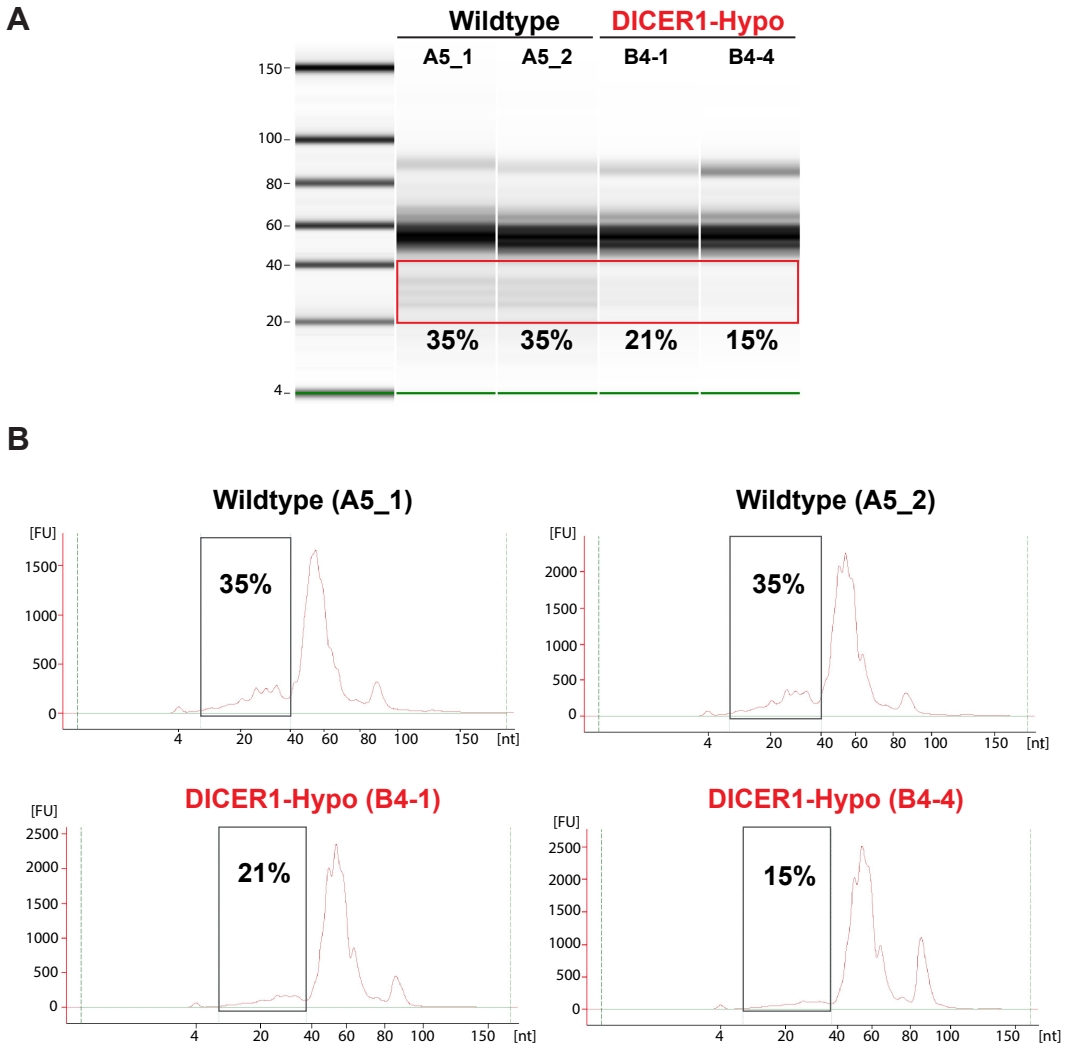
Additionally, we identified the ESCC-family members miRNA-302-367 and miRNA-371-373 as playing a pro-survival role in hESCs. Interestingly, these ESCC miRNAs are also required in mESCs although their functions differ, highlighting the importance of context in evaluating miRNA function. Unexpectedly, the miRNA-17-92 ESCC-family cluster does not prevent apoptosis in hESCs even though it is essential in the mouse development equivalent cell line, mEpiSCs. Our results emphasize the importance to study human-specific miRNA function in human ESCs rather than in mouse ESCs or mouse EpiSCs.

Finally, we generated an inducible platform that allows the assessment of acute *DICER1* knockout phenotypes in hESCs before adaptation can occur in cell culture. Our platform is particularly advantageous for studying miRNA requirements because it bypasses the problem of redundancy in the network. It is also versatile, allowing the study of miRNA function in both primed and naïve pluripotency, as well as in differentiation.

Thus, our study elucidated aspects of microRNA function during early human embryonic development and generated a valuable tool to uncover novel and human-specific miRNA function during human development and stem cell differentiation.

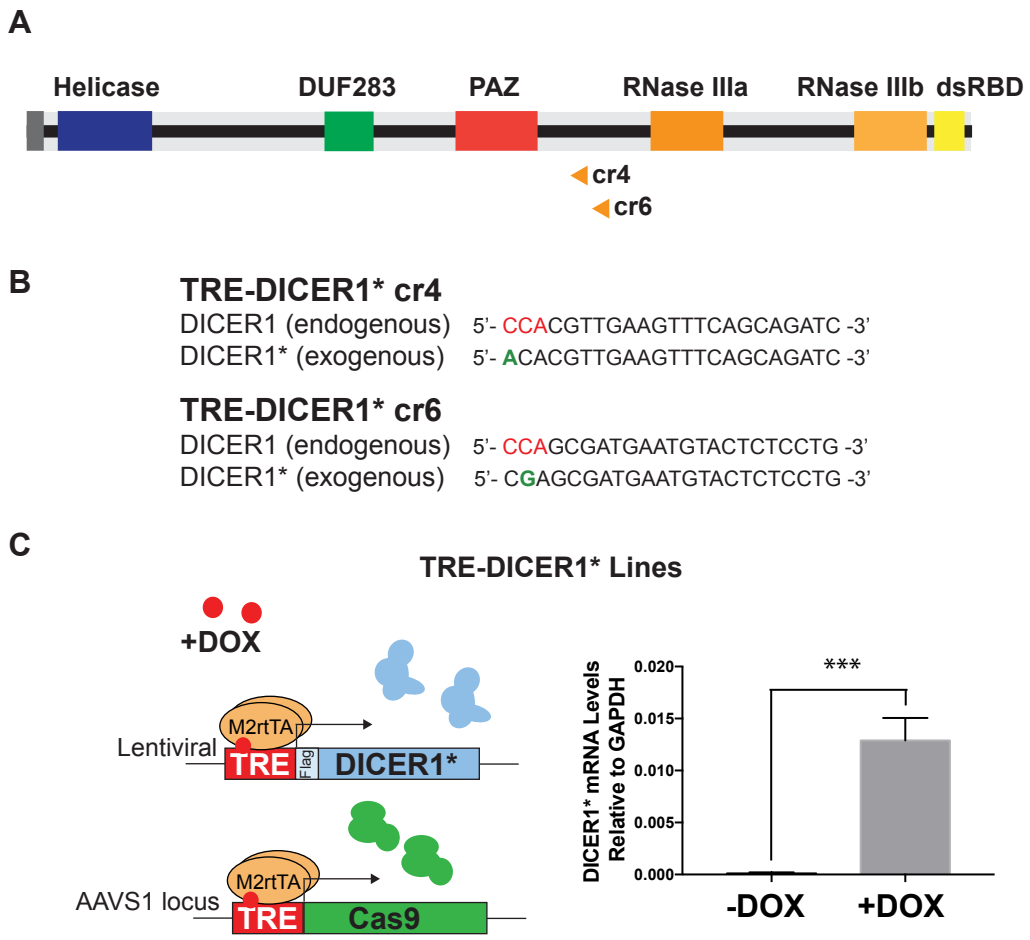
APPENDIX

Appendix 1.



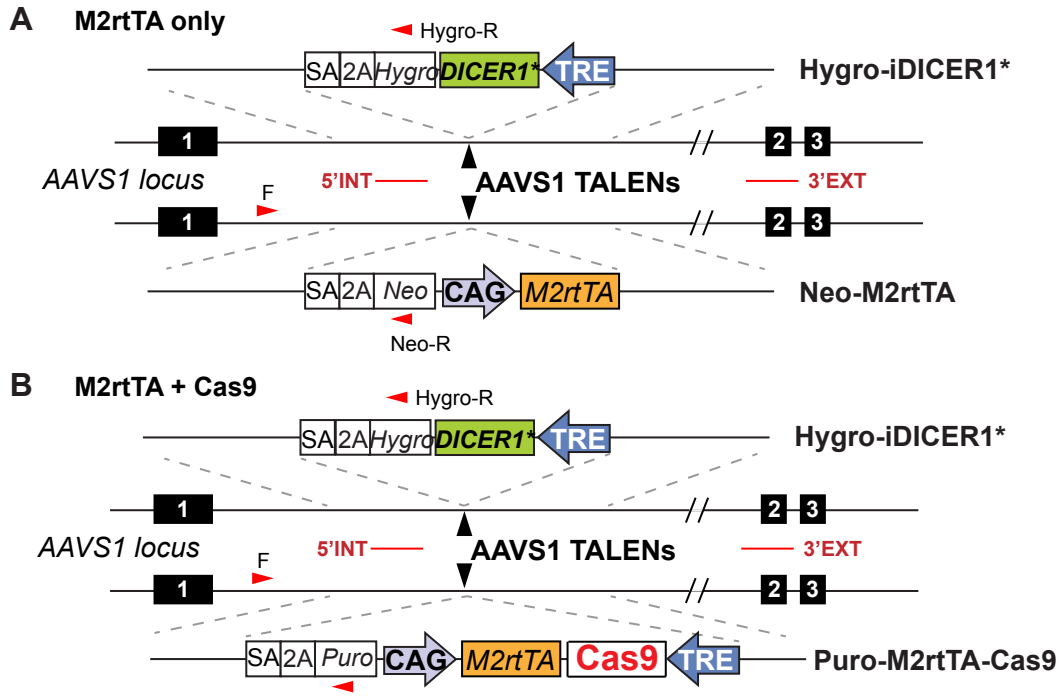
Appendix 1. DICER1 Hypomorphs Show Global Reduction of MicroRNA Species. Agilent Bioanalyzer 2100 small RNA electrophoresis (A) and electropherogram (B) results of size-fragmented RNA populations (<150nt long species) of two repeats of a wildtype hESC line (A5) and two DICER1 hypomorph hESC lines (B4-1 and B4-4). Percentages represent the relative microRNA to small RNA ratio in the samples. n=2.
Related to Figure 2.2.

Appendix 2.



Appendix 2. Generation of TRE-DICER1* Lines. (A) Schematic representation of DICER1 protein with domains. Orange arrows represent gRNA-targeting loci upstream of the RNase IIIa functional domain. cr, CRISPR; DUF, Domain of Unknown Function; dsRBD, double-stranded RNA-Binding Domain. (B) Sequences of the endogenous DICER1 and gRNA-immune exogenous DICER1* in TRE-DICER1*cr4 and TRE-DICER1*cr6 lines. NGG sequence is in red. Silent mutation base pair change is in green. (C) Diagram showing the TRE-DICER1* line expressing DICER1* and Cas9 upon doxycycline treatment. qRT-PCR data showing the expression of DICER1* upon doxycycline treatment. DOX, doxycycline. **Related to Figure 2.11.**

Appendix 3.

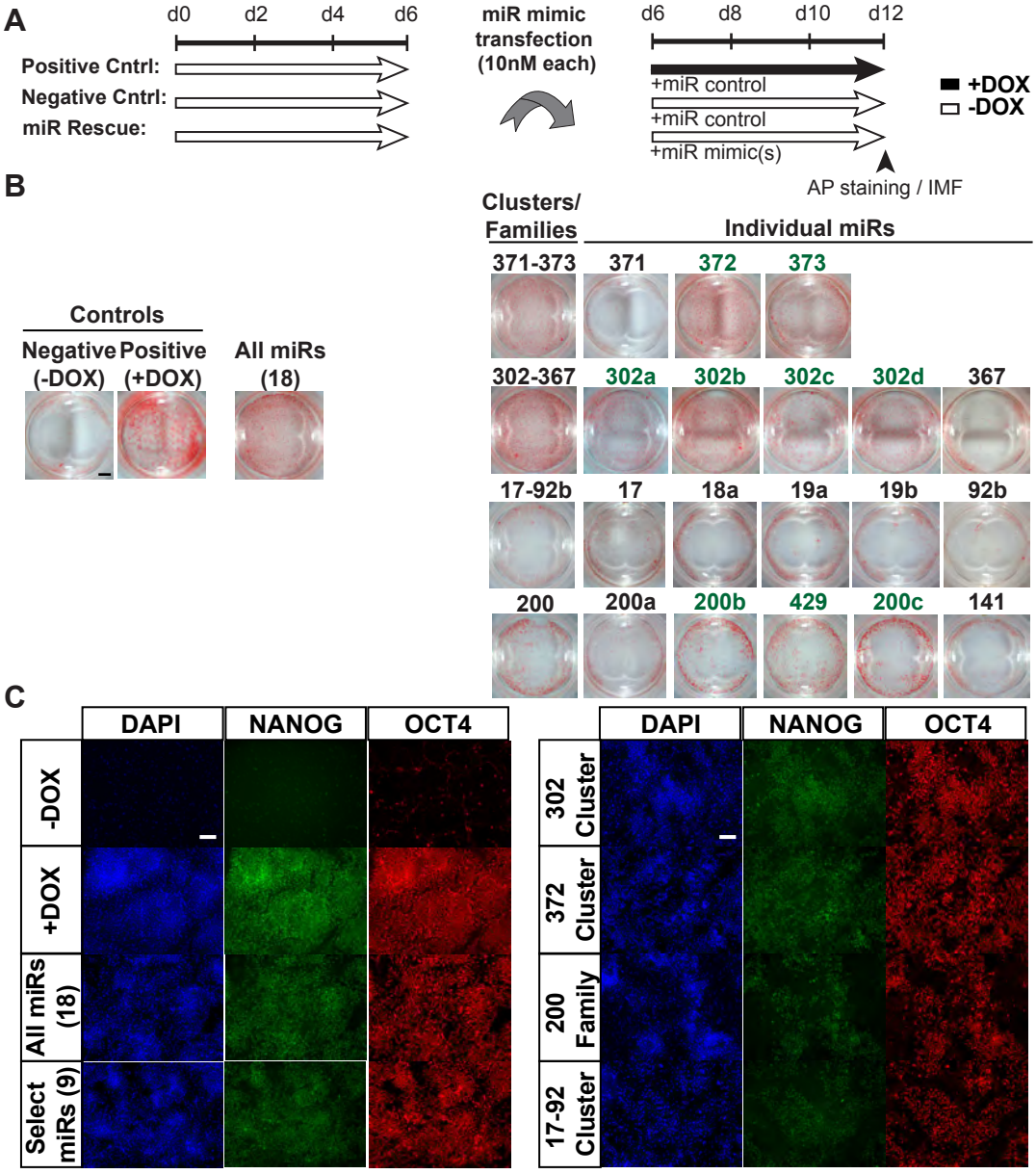


C

		# of colonies frozen / formed	
	Electroporation Round	M2rtTA only	M2rtTA + Cas9
AAVS1-TRE-DICER1-RNAcr4	1	6 / 24	1 / 1
	2	2 / 9	11 / 11
	3	10 / 17	5 / 5
AAVS1-TRE-DICER1-RNAcr6	1	6 / 24	2 / 2
	2	6 / 13	0 / 0
	3	8 / 16	3 / 4

Appendix 3. Generation of AAVS1-Inducible DICER1* Lines. (A) Schematic representation of AAVS1 locus targeting of a Hygro-iDICER1* cassette and a Neo-M2rtTA cassette. Small red arrows indicate PCR primers. Hygro, hygromycin; Neo, neomycin; INT, internal probe for Southern; EXT, external probe for Southern; SA, splice acceptor; 2A, 2A peptide; M2rtTA, M2 reverse tetracycline (tet)-controlled transactivator. These lines need to be transfected or electroporated with Cas9 plus gRNA. (B) Schematic representation of AAVS1 locus targeting of a Hygro-iDICER1* cassette and a Puro-M2rtTA-Cas9 cassette. Small red arrows indicate PCR primers. Hygro, hygromycin; Puro, neomycin; INT, internal probe for Southern; EXT, external probe for Southern; SA, splice acceptor; 2A, 2A peptide; M2rtTA, M2 reverse tetracycline (tet)-controlled transactivator. These lines need to be transfected with gRNA only. (C) Results of three electroporation rounds to generate four different lines (AAVS1-TRE-DICER1-M2rtTAOnly-RNAcr4 or 6 and AAVS1-TRE-DICER1-M2rtTA+Cas9-RNAcr4 or 6). Number of colonies frozen from colonies formed is shown. **Related to Figure 2.11.**

Appendix 4.



Appendix 4. MicroRNAs Rescue DICER1-/- hESCs Expressing Proper Pluripotency Markers. (A) Schematic representation of microRNA mimic transfection procedure in DICER1-/- hESCs (B2). Cntrl, control; miR, mature microRNA, DOX, doxycycline. Black, plus doxycycline; White, minus doxycycline. (B) Whole-well view of alkaline phosphatase (AP)-stained DICER1-/- hESCs (B2) with or without microRNA transfection of individual microRNAs or groups of microRNAs. Representative images of four experiments. In green, nine microRNAs that rescue the DICER1-/- phenotype. (C) Immunofluorescence staining of pluripotency markers NANOG and OCT4 in DICER1-/- (D11) transfected with select microRNAs or clusters/families. The positive and negative controls were transfected with control microRNAs, and the positive control contains doxycycline and the negative control does not contain doxycycline. Scale bar represents 500mm.
Related to Figure 2.17.

REFERENCES

Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev *17*, 126-140.

Baek, D., Villen, J., Shin, C., Camargo, F.D., Gygi, S.P., and Bartel, D.P. (2008). The impact of microRNAs on protein output. Nature *455*, 64-71.

Banito, A., Rashid, S.T., Acosta, J.C., Li, S., Pereira, C.F., Geti, I., Pinho, S., Silva, J.C., Azuara, V., Walsh, M.*, et al.* (2009). Senescence impairs successful reprogramming to pluripotent stem cells. Genes Dev *23*, 2134-2139.

Bar, M., Wyman, S.K., Fritz, B.R., Qi, J., Garg, K.S., Parkin, R.K., Kroh, E.M., Bendoraite, A., Mitchell, P.S., Nelson, A.M.*, et al.* (2008). MicroRNA discovery and profiling in human embryonic stem cells by deep sequencing of small RNA libraries. Stem Cells *26*, 2496-2505.

Barroso-delJesus, A., Lucena-Aguilar, G., Sanchez, L., Ligero, G., Gutierrez-Aranda, I., and Menendez, P. (2011). The Nodal inhibitor Lefty is negatively modulated by the microRNA miR-302 in human embryonic stem cells. FASEB J *25*, 1497-1508.

Barroso-delJesus, A., Romero-Lopez, C., Lucena-Aguilar, G., Melen, G.J., Sanchez, L., Ligero, G., Berzal-Herranz, A., and Menendez, P. (2008). Embryonic stem cell-specific miR302-367 cluster: human gene structure and functional characterization of its core promoter. Mol Cell Biol *28*, 6609-6619.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell *116*, 281-297.

Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell *136*, 215-233.

Benetti, R., Gonzalo, S., Jaco, I., Munoz, P., Gonzalez, S., Schoeftner, S., Murchison, E., Andl, T., Chen, T., Klatt, P.*, et al.* (2008). A mammalian microRNA cluster controls DNA methylation and telomere recombination via Rbl2-dependent regulation of DNA methyltransferases. Nat Struct Mol Biol *15*, 998.

Bentwich, I., Avniel, A., Karov, Y., Aharonov, R., Gilad, S., Barad, O., Barzilai, A., Einat, P., Einav, U., Meiri, E.*, et al.* (2005). Identification of hundreds of conserved and nonconserved human microRNAs. Nat Genet *37*, 766-770.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K.*, et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315-326.

Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature *409*, 363-366.

Bernstein, E., Kim, S.Y., Carmell, M.A., Murchison, E.P., Alcorn, H., Li, M.Z., Mills, A.A., Elledge, S.J., Anderson, K.V., and Hannon, G.J. (2003). Dicer is essential for mouse development. Nat Genet *35*, 215-217.

Bouillet, P., and Strasser, A. (2002). BH3-only proteins - evolutionarily conserved proapoptotic Bcl-2 family members essential for initiating programmed cell death. J Cell Sci *115*, 1567-1574.

Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K.*, et al.* (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature *441*, 349-353.

Brennecke, J., Stark, A., Russell, R.B., and Cohen, S.M. (2005). Principles of microRNA-target recognition. PLoS Biol *3*, e85.

Brons, I.G., Smithers, L.E., Trotter, M.W., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S.M., Howlett, S.K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R.A.*, et al.* (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature *448*, 191-195.

Calabrese, J.M., Seila, A.C., Yeo, G.W., and Sharp, P.A. (2007). RNA sequence analysis defines Dicer's role in mouse embryonic stem cells. Proc Natl Acad Sci U S A *104*, 18097-18102.

Card, D.A., Hebbar, P.B., Li, L., Trotter, K.W., Komatsu, Y., Mishina, Y., and Archer, T.K. (2008). Oct4/Sox2-regulated miR-302 targets cyclin D1 in human embryonic stem cells. Mol Cell Biol *28*, 6426-6438.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell *113*, 643-655.

Chang, T.C., and Mendell, J.T. (2007). microRNAs in vertebrate physiology and human disease. Annu Rev Genomics Hum Genet *8*, 215-239.

Chen, J., Wang, G., Lu, C., Guo, X., Hong, W., Kang, J., and Wang, J. (2012). Synergetic cooperation of microRNAs with transcription factors in iPS cell generation. PLoS One *7*, e40849.

Chiang, H.R., Schoenfeld, L.W., Ruby, J.G., Auyeung, V.C., Spies, N., Baek, D., Johnston, W.K., Russ, C., Luo, S., Babiarz, J.E.*, et al.* (2010). Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. Genes Dev *24*, 992-1009.

Cobb, B.S., Nesterova, T.B., Thompson, E., Hertweck, A., O'Connor, E., Godwin, J., Wilson, C.B., Brockdorff, N., Fisher, A.G., Smale, S.T.*, et al.* (2005). T cell lineage choice and differentiation in the absence of the RNase III enzyme Dicer. J Exp Med *201*, 1367-1373.

Collinson, A., Collier, A.J., Morgan, N.P., Sienerth, A.R., Chandra, T., Andrews, S., and Rugg-Gunn, P.J. (2016). Deletion of the Polycomb-Group Protein EZH2 Leads to Compromised Self-Renewal and Differentiation Defects in Human Embryonic Stem Cells. Cell Rep *17*, 2700-2714.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A.*, et al.* (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819-823.

Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F., and Hannon, G.J. (2004). Processing of primary microRNAs by the Microprocessor complex. Nature *432*, 231-235.

Djuranovic, S., Nahvi, A., and Green, R. (2012). miRNA-mediated gene silencing by translational repression followed by mRNA deadenylation and decay. Science *336*, 237-240.

Doench, J.G., Petersen, C.P., and Sharp, P.A. (2003). siRNAs can function as miRNAs. Genes Dev *17*, 438-442.

Doudna, J.A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science *346*, 1258096.

Duchaine, T.F., Wohlschlegel, J.A., Kennedy, S., Bei, Y., Conte, D., Jr., Pang, K., Brownell, D.R., Harding, S., Mitani, S., Ruvkun, G.*, et al.* (2006). Functional proteomics reveals the biochemical niche of C. elegans DCR-1 in multiple small-RNA-mediated pathways. Cell *124*, 343-354.

Ebert, M.S., and Sharp, P.A. (2012). Roles for microRNAs in conferring robustness to biological processes. Cell *149*, 515-524.

Elbashir, S.M., Lendeckel, W., and Tuschl, T. (2001). RNA interference is mediated by 21- and 22-nucleotide RNAs. Genes Dev *15*, 188-200.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154-156.

Fattahi, F., Steinbeck, J.A., Kriks, S., Tchieu, J., Zimmer, B., Kishinevsky, S., Zeltner, N., Mica, Y., El-Nachef, W., Zhao, H.*, et al.* (2016). Deriving human ENS lineages for cell therapy and drug discovery in Hirschsprung disease. Nature *531*, 105-109.

Finger, S., Heavens, R.P., Sirinathsinghji, D.J., Kuehn, M.R., and Dunnett, S.B. (1988). Behavioral and neurochemical evaluation of a transgenic mouse model of Lesch-Nyhan syndrome. J Neurol Sci *86*, 203-213.

Flamand, M., and Duchaine, T.F. (2012). SnapShot: endogenous RNAi machinery and mechanisms. Cell *150*, 662-662 e662.

Fukagawa, T., Nogami, M., Yoshikawa, M., Ikeno, M., Okazaki, T., Takami, Y., Nakayama, T., and Oshimura, M. (2004). Dicer is essential for formation of the heterochromatin structure in vertebrate cells. Nat Cell Biol *6*, 784-791.

Gafni, O., Weinberger, L., Mansour, A.A., Manor, Y.S., Chomsky, E., Ben-Yosef, D., Kalma, Y., Viukov, S., Maza, I., Zviran, A.*, et al.* (2013). Derivation of novel human ground state naive pluripotent stem cells. Nature *504*, 282-286.

Gill, J.G., Langer, E.M., Lindsley, R.C., Cai, M., Murphy, T.L., Kyba, M., and Murphy, K.M. (2011). Snail and the microRNA-200 family act in opposition to regulate epithelial-to-mesenchymal transition and germ layer fate restriction in differentiating ESCs. Stem Cells *29*, 764-776.

Gonzalez, F., Zhu, Z., Shi, Z.D., Lelli, K., Verma, N., Li, Q.V., and Huangfu, D. (2014). An iCRISPR platform for rapid, multiplexable, and inducible genome editing in human pluripotent stem cells. Cell Stem Cell *15*, 215-226.

Gradwohl, G., Dierich, A., LeMeur, M., and Guillemot, F. (2000). neurogenin3 is required for the development of the four endocrine cell lineages of the pancreas. Proc Natl Acad Sci U S A *97*, 1607-1611.

Graham, B., Marcais, A., Dharmalingam, G., Carroll, T., Kanellopoulou, C., Graumann, J., Nesterova, T.B., Bermange, A., Brazauskas, P., Xella, B.*, et al.* (2016). MicroRNAs of the miR-290-295 Family Maintain Bivalency in Mouse Embryonic Stem Cells. Stem Cell Reports *6*, 635-642.

Gregory, R.I., Yan, K.P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. Nature *432*, 235-240.

Greve, T.S., Judson, R.L., and Blelloch, R. (2013). microRNA control of mouse and human pluripotent stem cell behavior. Annu Rev Cell Dev Biol *29*, 213-239.

Grishok, A., Pasquinelli, A.E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D.L., Fire, A., Ruvkun, G., and Mello, C.C. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. Cell *106*, 23-34.

Guo, H., Ingolia, N.T., Weissman, J.S., and Bartel, D.P. (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature *466*, 835-840.

Guo, Y., Liu, J., Elfenbein, S.J., Ma, Y., Zhong, M., Qiu, C., Ding, Y., and Lu, J. (2015). Characterization of the mammalian miRNA turnover landscape. Nucleic Acids Res *43*, 2326-2341.

Gurtan, A.M., Lu, V., Bhutkar, A., and Sharp, P.A. (2012). In vivo structure-function analysis of human Dicer reveals directional processing of precursor miRNAs. RNA *18*, 1116-1122.

Hammond, S.M., Bernstein, E., Beach, D., and Hannon, G.J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature *404*, 293-296.

Han, J., Lee, Y., Yeom, K.H., Kim, Y.K., Jin, H., and Kim, V.N. (2004). The Drosha-DGCR8 complex in primary microRNA processing. Genes Dev *18*, 3016-3027.

Harfe, B.D., McManus, M.T., Mansfield, J.H., Hornstein, E., and Tabin, C.J. (2005). The RNaseIII enzyme Dicer is required for morphogenesis but not patterning of the vertebrate limb. Proc Natl Acad Sci U S A *102*, 10898-10903.

Hockemeyer, D., and Jaenisch, R. (2010). Gene targeting in human pluripotent cells. Cold Spring Harb Symp Quant Biol *75*, 201-209.

Hockemeyer, D., and Jaenisch, R. (2016). Induced Pluripotent Stem Cells Meet Genome Editing. Cell Stem Cell *18*, 573-586.

Houbaviy, H.B., Murray, M.F., and Sharp, P.A. (2003). Embryonic stem cell-specific MicroRNAs. Dev Cell *5*, 351-358.

Hu, S., Wilson, K.D., Ghosh, Z., Han, L., Wang, Y., Lan, F., Ransohoff, K.J., Burridge, P., and Wu, J.C. (2013). MicroRNA-302 increases reprogramming efficiency via repression of NR2F2. Stem Cells *31*, 259-268.

Hutvagner, G., McLachlan, J., Pasquinelli, A.E., Balint, E., Tuschl, T., and Zamore, P.D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science *293*, 834-838.

James, D., Levine, A.J., Besser, D., and Hemmati-Brivanlou, A. (2005). TGFbeta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. Development *132*, 1273-1282.

Johnson, C.D., Esquela-Kerscher, A., Stefani, G., Byrom, M., Kelnar, K., Ovcharenko, D., Wilson, M., Wang, X., Shelton, J., Shingara, J.*, et al.* (2007). The let-7 microRNA represses cell proliferation pathways in human cells. Cancer Res *67*, 7713-7722.

Jouneau, A., Ciaudo, C., Sismeiro, O., Brochard, V., Jouneau, L., Vandormael-Pournin, S., Coppee, J.Y., Zhou, Q., Heard, E., Antoniewski, C.*, et al.* (2012). Naive and primed murine pluripotent stem cells have distinct miRNA expression profiles. RNA *18*, 253-264.

Judson, R.L., Babiarz, J.E., Venere, M., and Blelloch, R. (2009). Embryonic stem cell-specific microRNAs promote induced pluripotency. Nat Biotechnol *27*, 459-461.

Kanellopoulou, C., Gilpatrick, T., Kilaru, G., Burr, P., Nguyen, C.K., Morawski, A., Lenardo, M.J., and Muljo, S.A. (2015). Reprogramming of Polycomb-Mediated Gene Silencing in Embryonic Stem Cells by the miR-290 Family and the Methyltransferase Ash1l. Stem Cell Reports *5*, 971-978.

Kanellopoulou, C., Muljo, S.A., Kung, A.L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D.M., and Rajewsky, K. (2005). Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev *19*, 489-501.

Ketting, R.F., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J., and Plasterk, R.H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes Dev *15*, 2654-2659.

Khvorova, A., Reynolds, A., and Jayasena, S.D. (2003). Functional siRNAs and miRNAs exhibit strand bias. Cell *115*, 209-216.

Kim, Y.K., Kim, B., and Kim, V.N. (2016). Re-evaluation of the roles of DROSHA, Export in 5, and DICER in microRNA biogenesis. Proc Natl Acad Sci U S A *113*, E1881-1889.

Kim, Y.S., Yi, B.R., Kim, N.H., and Choi, K.C. (2014). Role of the epithelial-mesenchymal transition and its effects on embryonic stem cells. Exp Mol Med *46*, e108.

Knight, S.W., and Bass, B.L. (2001). A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in Caenorhabditis elegans. Science *293*, 2269-2271.

Krill, K.T., Gurdziel, K., Heaton, J.H., Simon, D.P., and Hammer, G.D. (2013). Dicer deficiency reveals microRNAs predicted to control gene expression in the developing adrenal cortex. Mol Endocrinol *27*, 754-768.

Kuhnert, F., Mancuso, M.R., Hampton, J., Stankunas, K., Asano, T., Chen, C.Z., and Kuo, C.J. (2008). Attribution of vascular phenotypes of the murine Egfl7 locus to the microRNA miR-126. Development *135*, 3989-3993.

Kurzynska-Kokorniak, A., Koralewska, N., Pokornowska, M., Urbanowicz, A., Tworak, A., Mickiewicz, A., and Figlerowicz, M. (2015). The many faces of Dicer: the complexity of the mechanisms regulating Dicer gene expression and enzyme activities. Nucleic Acids Res *43*, 4365-4380.

Landthaler, M., Yalcin, A., and Tuschl, T. (2004). The human DiGeorge syndrome critical region gene 8 and Its D. melanogaster homolog are required for miRNA biogenesis. Curr Biol *14*, 2162-2167.

Lango Allen, H., Flanagan, S.E., Shaw-Smith, C., De Franco, E., Akerman, I., Caswell, R., International Pancreatic Agenesis, C., Ferrer, J., Hattersley, A.T., and Ellard, S. (2011). GATA6 haploinsufficiency causes pancreatic agenesis in humans. Nat Genet *44*, 20-22.

Lee, M.R., Prasain, N., Chae, H.D., Kim, Y.J., Mantel, C., Yoder, M.C., and Broxmeyer, H.E. (2013). Epigenetic regulation of NANOG by miR-302 cluster-MBD2 completes induced pluripotent stem cell reprogramming. Stem Cells *31*, 666-681.

Lee, N.S., Kim, J.S., Cho, W.J., Lee, M.R., Steiner, R., Gompers, A., Ling, D., Zhang, J., Strom, P., Behlke, M.*, et al.* (2008). miR-302b maintains "stemness" of human embryonal carcinoma cells by post-transcriptional regulation of Cyclin D2 expression. Biochem Biophys Res Commun *377*, 434-440.

Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*, 843-854.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S.*, et al.* (2003). The nuclear RNase III Drosha initiates microRNA processing. Nature *425*, 415-419.

Lee, Y., Kim, M., Han, J., Yeom, K.H., Lee, S., Baek, S.H., and Kim, V.N. (2004a). MicroRNA genes are transcribed by RNA polymerase II. EMBO J *23*, 4051-4060.

Lee, Y.S., Nakahara, K., Pham, J.W., Kim, K., He, Z., Sontheimer, E.J., and Carthew, R.W. (2004b). Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell *117*, 69-81.

Levy, C., Khaled, M., Robinson, K.C., Veguilla, R.A., Chen, P.H., Yokoyama, S., Makino, E., Lu, J., Larue, L., Beermann, F.*, et al.* (2010). Lineage-specific transcriptional regulation of DICER by MITF in melanocytes. Cell *141*, 994-1005.

Lewis, B.P., Burge, C.B., and Bartel, D.P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell *120*, 15-20.

Li, Z., Yang, C.S., Nakashima, K., and Rana, T.M. (2011). Small RNA-mediated regulation of iPS cell generation. EMBO J *30*, 823-834.

Liao, B., Bao, X., Liu, L., Feng, S., Zovoilis, A., Liu, W., Xue, Y., Cai, J., Guo, X., Qin, B.*, et al.* (2011). MicroRNA cluster 302-367 enhances somatic cell reprogramming by accelerating a mesenchymal-to-epithelial transition. J Biol Chem *286*, 17359-17364.

Liao, J., Karnik, R., Gu, H., Ziller, M.J., Clement, K., Tsankov, A.M., Akopian, V., Gifford, C.A., Donaghey, J., Galonska, C.*, et al.* (2015). Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. Nat Genet *47*, 469-478.

Lin, S.L., Chang, D.C., Lin, C.H., Ying, S.Y., Leu, D., and Wu, D.T. (2011). Regulation of somatic cell reprogramming through inducible mir-302 expression. Nucleic Acids Res *39*, 1054-1065.

Lipchina, I., Elkabetz, Y., Hafner, M., Sheridan, R., Mihailovic, A., Tuschl, T., Sander, C., Studer, L., and Betel, D. (2011). Genome-wide identification of microRNA targets in human ES cells reveals a role for miR-302 in modulating BMP response. Genes Dev *25*, 2173-2186.

Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.J., Hammond, S.M., Joshua-Tor, L., and Hannon, G.J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. Science *305*, 1437-1441.

Liu, N., Bezprozvannaya, S., Williams, A.H., Qi, X., Richardson, J.A., Bassel-Duby, R., and Olson, E.N. (2008). microRNA-133a regulates cardiomyocyte proliferation and suppresses smooth muscle gene expression in the heart. Genes Dev *22*, 3242-3254.

Lund, E., Guttinger, S., Calado, A., Dahlberg, J.E., and Kutay, U. (2004). Nuclear export of microRNA precursors. Science *303*, 95-98.

Macrae, I.J., Li, F., Zhou, K., Cande, W.Z., and Doudna, J.A. (2006a). Structure of Dicer and mechanistic implications for RNAi. Cold Spring Harb Symp Quant Biol *71*, 73-80.

Macrae, I.J., Zhou, K., Li, F., Repic, A., Brooks, A.N., Cande, W.Z., Adams, P.D., and Doudna, J.A. (2006b). Structural basis for double-stranded RNA processing by Dicer. Science *311*, 195-198.

Mallanna, S.K., and Rizzino, A. (2010). Emerging roles of microRNAs in the control of embryonic stem cells and the generation of induced pluripotent stem cells. Dev Biol *344*, 16-25.

Mallo, M., Wellik, D.M., and Deschamps, J. (2010). Hox genes and regional patterning of the vertebrate body plan. Dev Biol *344*, 7-15.

Maniataki, E., and Mourelatos, Z. (2005). A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. Genes Dev *19*, 2979-2990.

Marson, A., Levine, S.S., Cole, M.F., Frampton, G.M., Brambrink, T., Johnstone, S., Guenther, M.G., Johnston, W.K., Wernig, M., Newman, J.*, et al.* (2008). Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell *134*, 521-533.

Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci U S A *78*, 7634-7638.

Martin, M. (2011). Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads. EMBnet journal *17.1*, 10-12.

Mashal, R.D., Koontz, J., and Sklar, J. (1995). Detection of mutations by cleavage of DNA heteroduplexes with bacteriophage resolvases. Nat Genet *9*, 177-183.

Mathieu, J., and Ruohola-Baker, H. (2013). Regulation of stem cell populations by microRNAs. Adv Exp Med Biol *786*, 329-351.

Medeiros, L.A., Dennis, L.M., Gill, M.E., Houbaviy, H., Markoulaki, S., Fu, D., White, A.C., Kirak, O., Sharp, P.A., Page, D.C.*, et al.* (2011). Mir-290-295 deficiency in mice results in partially penetrant embryonic lethality and germ cell defects. Proc Natl Acad Sci U S A *108*, 14163-14168.

Meerbrey, K.L., Hu, G., Kessler, J.D., Roarty, K., Li, M.Z., Fang, J.E., Herschkowitz, J.I., Burrows, A.E., Ciccia, A., Sun, T.*, et al.* (2011). The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. Proc Natl Acad Sci U S A *108*, 3665-3670.

Melton, C., Judson, R.L., and Blelloch, R. (2010). Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. Nature *463*, 621-626.

Mendell, J.T. (2008). miRiad roles for the miR-17-92 cluster in development and disease. Cell *133*, 217-222.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell *113*, 631-642.

Miyazaki, H., Higashimoto, K., Yada, Y., Endo, T.A., Sharif, J., Komori, T., Matsuda, M., Koseki, Y., Nakayama, M., Soejima, H.*, et al.* (2013). Ash1l methylates Lys36 of histone H3 independently of transcriptional elongation to counteract polycomb silencing. PLoS Genet *9*, e1003897.

Mogilyansky, E., and Rigoutsos, I. (2013). The miR-17/92 cluster: a comprehensive update on its genomics, genetics, functions and increasingly important and numerous roles in health and disease. Cell Death Differ *20*, 1603-1614.

Morin, R.D., O'Connor, M.D., Griffith, M., Kuchenbauer, F., Delaney, A., Prabhu, A.L., Zhao, Y., McDonald, H., Zeng, T., Hirst, M.*, et al.* (2008). Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. Genome Res *18*, 610-621.

Morrisey, E.E., Tang, Z., Sigrist, K., Lu, M.M., Jiang, F., Ip, H.S., and Parmacek, M.S. (1998). GATA6 regulates HNF4 and is required for differentiation of visceral endoderm in the mouse embryo. Genes Dev *12*, 3579-3590.

Moss, E.G., and Tang, L. (2003). Conservation of the heterochronic regulator Lin-28, its developmental expression and microRNA complementary sites. Dev Biol *258*, 432-442.

Mukherjee, K., Campos, H., and Kolaczkowski, B. (2013). Evolution of animal and plant dicers: early parallel duplications and recurrent adaptation of antiviral RNA binding in plants. Mol Biol Evol *30*, 627-641.

Munoz-Sanjuan, I., and Brivanlou, A.H. (2002). Neural induction, the default model and embryonic stem cells. Nat Rev Neurosci *3*, 271-280.

Murchison, E.P., Partridge, J.F., Tam, O.H., Cheloufi, S., and Hannon, G.J. (2005). Characterization of Dicer-deficient murine embryonic stem cells. Proc Natl Acad Sci U S A *102*, 12135-12140.

Nichols, J., and Smith, A. (2009). Naive and primed pluripotent states. Cell Stem Cell *4*, 487-492.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell *95*, 379-391.

Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet *24*, 372-376.

Okamura, K., Hagen, J.W., Duan, H., Tyler, D.M., and Lai, E.C. (2007). The mirtron pathway generates microRNA-class regulatory RNAs in Drosophila. Cell *130*, 89-100.

Okamura, K., and Lai, E.C. (2008). Endogenous small interfering RNAs in animals. Nat Rev Mol Cell Biol *9*, 673-678.

Oshlack, A., Robinson, M.D., and Young, M.D. (2010). From RNA-seq reads to differential expression results. Genome biology *11*, 220.

Parchem, R.J., Moore, N., Fish, J.L., Parchem, J.G., Braga, T.T., Shenoy, A., Oldham, M.C., Rubenstein, J.L., Schneider, R.A., and Blelloch, R. (2015). miR-302 Is Required for Timing of Neural Differentiation, Neural Tube Closure, and Embryonic Viability. Cell Rep *12*, 760-773.

Parchem, R.J., Ye, J., Judson, R.L., LaRussa, M.F., Krishnakumar, R., Blelloch, A., Oldham, M.C., and Blelloch, R. (2014). Two miRNA clusters reveal alternative paths in late-stage reprogramming. Cell Stem Cell *14*, 617-631.

Park, C.Y., Choi, Y.S., and McManus, M.T. (2010). Analysis of microRNA knockouts in mice. Hum Mol Genet *19*, R169-175.

Pera, M.F., and Tam, P.P. (2010). Extrinsic regulation of pluripotent stem cells. Nature *465*, 713-720.

Pernaute, B., Spruce, T., Smith, K.M., Sanchez-Nieto, J.M., Manzanares, M., Cobb, B., and Rodriguez, T.A. (2014). MicroRNAs control the apoptotic threshold in primed pluripotent stem cells through regulation of BIM. Genes Dev *28*, 1873-1878.

Pham, J.W., and Sontheimer, E.J. (2005). Molecular requirements for RNA-induced silencing complex assembly in the Drosophila RNA interference pathway. J Biol Chem *280*, 39278-39283.

Polo, J.M., Anderssen, E., Walsh, R.M., Schwarz, B.A., Nefzger, C.M., Lim, S.M., Borkent, M., Apostolou, E., Alaei, S., Cloutier, J.*, et al.* (2012). A molecular roadmap of reprogramming somatic cells into iPS cells. Cell *151*, 1617-1632.

Qi, J., Yu, J.Y., Shcherbata, H.R., Mathieu, J., Wang, A.J., Seal, S., Zhou, W., Stadler, B.M., Bourgin, D., Wang, L.*, et al.* (2009). microRNAs regulate human embryonic stem cell division. Cell Cycle *8*, 3729-3741.

Qi, Y., Zhang, X.J., Renier, N., Wu, Z., Atkin, T., Sun, Z., Ozair, M.Z., Tchieu, J., Zimmer, B., Fattahi, F.*, et al.* (2017). Combined small-molecule inhibition accelerates the derivation of functional cortical neurons from human pluripotent stem cells. Nat Biotechnol *35*, 154-163.

Reinhart, B.J., and Bartel, D.P. (2002). Small RNAs correspond to centromere heterochromatic repeats. Science *297*, 1831.

Rezania, A., Bruin, J.E., Arora, P., Rubin, A., Batushansky, I., Asadi, A., O'Dwyer, S., Quiskamp, N., Mojibian, M., Albrecht, T.*, et al.* (2014). Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. Nat Biotechnol *32*, 1121-1133.

Rodriguez, A., Griffiths-Jones, S., Ashurst, J.L., and Bradley, A. (2004). Identification of mammalian microRNA host genes and transcription units. Genome Res *14*, 1902-1910.

Rosa, A., Spagnoli, F.M., and Brivanlou, A.H. (2009). The miR-430/427/302 family controls mesendodermal fate specification via species-specific target selection. Dev Cell *16*, 517-527.

Rouet, P., Smih, F., and Jasin, M. (1994). Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. Proc Natl Acad Sci U S A *91*, 6064-6068.

Ruby, J.G., Jan, C.H., and Bartel, D.P. (2007). Intronic microRNA precursors that bypass Drosha processing. Nature *448*, 83-86.

Samavarchi-Tehrani, P., Golipour, A., David, L., Sung, H.K., Beyer, T.A., Datti, A., Woltjen, K., Nagy, A., and Wrana, J.L. (2010). Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. Cell Stem Cell *7*, 64-77.

Schwarz, D.S., Hutvagner, G., Du, T., Xu, Z., Aronin, N., and Zamore, P.D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. Cell *115*, 199-208.

Shen, X., Liu, Y., Hsu, Y.J., Fujiwara, Y., Kim, J., Mao, X., Yuan, G.C., and Orkin, S.H. (2008). EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. Mol Cell *32*, 491-502.

Shi, Z.D., Lee, K., Yang, D., Amin, S., Verma, N., Li, Q.V., Zhu, Z., Soh, C.L., Kumar, R., Evans, T.*, et al.* (2017). Genome Editing in hPSCs Reveals GATA6 Haploinsufficiency and a Genetic Interaction with GATA4 in Human Pancreatic Development. Cell Stem Cell.

Sinkkonen, L., Hugenschmidt, T., Berninger, P., Gaidatzis, D., Mohn, F., Artus-Revel, C.G., Zavolan, M., Svoboda, P., and Filipowicz, W. (2008). MicroRNAs control de novo DNA methylation through regulation of transcriptional repressors in mouse embryonic stem cells. Nat Struct Mol Biol *15*, 259-267.

Smibert, P., Yang, J.S., Azzam, G., Liu, J.L., and Lai, E.C. (2013). Homeostatic control of Argonaute stability by microRNA availability. Nat Struct Mol Biol *20*, 789-795.

Smith, J.R., Maguire, S., Davis, L.A., Alexander, M., Yang, F., Chandran, S., ffrench-Constant, C., and Pedersen, R.A. (2008). Robust, persistent transgene expression in human embryonic stem cells is achieved with AAVS1-targeted integration. Stem Cells *26*, 496-504.

Smyth, G.K. (2005). Limma: linear models for microarray data. In Bioinformatics and Computational Biology Solutions using R and Bioconductor, V.C. R. Gentleman, S. Dudoit, R. Irizarry, W. Huber (eds), ed. (New York: Springer), pp. 397-420.

Spruce, T., Pernaute, B., Di-Gregorio, A., Cobb, B.S., Merkenschlager, M., Manzanares, M., and Rodriguez, T.A. (2010). An early developmental role for miRNAs in the maintenance of extraembryonic stem cells in the mouse embryo. Dev Cell *19*, 207-219.

Stadler, B., Ivanovska, I., Mehta, K., Song, S., Nelson, A., Tan, Y., Mathieu, J., Darby, C., Blau, C.A., Ware, C.*, et al.* (2010). Characterization of microRNAs involved in embryonic stem cell states. Stem Cells Dev *19*, 935-950.

Subramanyam, D., Lamouille, S., Judson, R.L., Liu, J.Y., Bucay, N., Derynck, R., and Blelloch, R. (2011). Multiple targets of miR-302 and miR-372 promote reprogramming of human fibroblasts to induced pluripotent stem cells. Nat Biotechnol *29*, 443-448.

Suh, M.R., Lee, Y., Kim, J.Y., Kim, S.K., Moon, S.H., Lee, J.Y., Cha, K.Y., Chung, H.M., Yoon, H.S., Moon, S.Y.*, et al.* (2004). Human embryonic stem cells express a unique set of microRNAs. Dev Biol *270*, 488-498.

Tabibzadeh, S., and Hemmati-Brivanlou, A. (2006). Lefty at the crossroads of "stemness" and differentiative events. Stem Cells *24*, 1998-2006.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Tanaka, Y., Kawahashi, K., Katagiri, Z., Nakayama, Y., Mahajan, M., and Kioussis, D. (2011). Dual function of histone H3 lysine 36 methyltransferase ASH1 in regulation of Hox gene expression. PLoS One *6*, e28171.

Tang, F., Kaneda, M., O'Carroll, D., Hajkova, P., Barton, S.C., Sun, Y.A., Lee, C., Tarakhovsky, A., Lao, K., and Surani, M.A. (2007). Maternal microRNAs are essential for mouse zygotic development. Genes Dev *21*, 644-648.

Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature *448*, 196-199.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145-1147.

Thomson, J.M., Newman, M., Parker, J.S., Morin-Kensicki, E.M., Wright, T., and Hammond, S.M. (2006). Extensive post-transcriptional regulation of microRNAs and its implications for cancer. Genes Dev *20*, 2202-2207.

Thornton, J.E., and Gregory, R.I. (2012). How does Lin28 let-7 control development and disease? Trends Cell Biol *22*, 474-482.

Tsumura, A., Hayakawa, T., Kumaki, Y., Takebayashi, S., Sakaue, M., Matsuoka, C., Shimotohno, K., Ishikawa, F., Li, E., Ueda, H.R.*, et al.* (2006). Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. Genes Cells *11*, 805-814.

Vallier, L., Alexander, M., and Pedersen, R.A. (2005). Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. J Cell Sci *118*, 4495-4509.

Ventura, A., Young, A.G., Winslow, M.M., Lintault, L., Meissner, A., Erkeland, S.J., Newman, J., Bronson, R.T., Crowley, D., Stone, J.R.*, et al.* (2008). Targeted deletion reveals essential and overlapping functions of the miR-17 through 92 family of miRNA clusters. Cell *132*, 875-886.

Wang, G., Guo, X., Hong, W., Liu, Q., Wei, T., Lu, C., Gao, L., Ye, D., Zhou, Y., Chen, J.*, et al.* (2013a). Critical regulation of miR-200/ZEB2 pathway in Oct4/Sox2-induced mesenchymal-to-epithelial transition and induced pluripotent stem cell generation. Proc Natl Acad Sci U S A *110*, 2858-2863.

Wang, T., Chen, K., Zeng, X., Yang, J., Wu, Y., Shi, X., Qin, B., Zeng, L., Esteban, M.A., Pan, G.*, et al.* (2011). The histone demethylases Jhdm1a/1b enhance somatic cell reprogramming in a vitamin-C-dependent manner. Cell Stem Cell *9*, 575-587.

Wang, Y., Baskerville, S., Shenoy, A., Babiarz, J.E., Baehner, L., and Blelloch, R. (2008). Embryonic stem cell-specific microRNAs regulate the G1-S transition and promote rapid proliferation. Nat Genet *40*, 1478-1483.

Wang, Y., Medvid, R., Melton, C., Jaenisch, R., and Blelloch, R. (2007). DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. Nat Genet *39*, 380-385.

Wang, Y., Melton, C., Li, Y.P., Shenoy, A., Zhang, X.X., Subramanyam, D., and Blelloch, R. (2013b). miR-294/miR-302 promotes proliferation, suppresses G1-S restriction point, and inhibits ESC differentiation through separable mechanisms. Cell Rep *4*, 99-109.

Weinberger, L., Ayyash, M., Novershtern, N., and Hanna, J.H. (2016). Dynamic stem cell states: naive to primed pluripotency in rodents and humans. Nat Rev Mol Cell Biol *17*, 155-169.

White, J., and Dalton, S. (2005). Cell cycle control of embryonic stem cells. Stem Cell Rev *1*, 131-138.

White, S.A., and Allshire, R.C. (2004). Loss of Dicer fowls up centromeres. Nat Cell Biol *6*, 696-697.

Williams, L.A., Davis-Dusenbery, B.N., and Eggan, K.C. (2012). SnapShot: directed differentiation of pluripotent stem cells. Cell *149*, 1174-1174 e1171.

Xu, R.H., Chen, X., Li, D.S., Li, R., Addicks, G.C., Glennon, C., Zwaka, T.P., and Thomson, J.A. (2002). BMP4 initiates human embryonic stem cell differentiation to trophoblast. Nat Biotechnol *20*, 1261-1264.

Yekta, S., Shih, I.H., and Bartel, D.P. (2004). MicroRNA-directed cleavage of HOXB8 mRNA. Science *304*, 594-596.

Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev *17*, 3011-3016.

Zamore, P.D., Tuschl, T., Sharp, P.A., and Bartel, D.P. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell *101*, 25-33.

Zhang, H., Kolb, F.A., Brondani, V., Billy, E., and Filipowicz, W. (2002). Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. EMBO J *21*, 5875-5885.

Zhang, H., Kolb, F.A., Jaskiewicz, L., Westhof, E., and Filipowicz, W. (2004). Single processing center models for human Dicer and bacterial RNase III. Cell *118*, 57-68.

Zhang, J., and Ma, L. (2012). MicroRNA control of epithelial-mesenchymal transition and metastasis. Cancer Metastasis Rev *31*, 653-662.

Zhang, P., Li, J., Tan, Z., Wang, C., Liu, T., Chen, L., Yong, J., Jiang, W., Sun, X., Du, L.*, et al.* (2008). Short-term BMP-4 treatment initiates mesoderm induction in human embryonic stem cells. Blood *111*, 1933-1941.

Zhang, X., Huang, C.T., Chen, J., Pankratz, M.T., Xi, J., Li, J., Yang, Y., Lavaute, T.M., Li, X.J., Ayala, M.*, et al.* (2010). Pax6 is a human neuroectoderm cell fate determinant. Cell Stem Cell *7*, 90-100.

Zhao, Y., Ransom, J.F., Li, A., Vedantham, V., von Drehle, M., Muth, A.N., Tsuchihashi, T., McManus, M.T., Schwartz, R.J., and Srivastava, D. (2007). Dysregulation of cardiogenesis, cardiac conduction, and cell cycle in mice lacking miRNA-1-2. Cell *129*, 303-317.

Zheng, G.X., Ravi, A., Calabrese, J.M., Medeiros, L.A., Kirak, O., Dennis, L.M., Jaenisch, R., Burge, C.B., and Sharp, P.A. (2011). A latent pro-survival function for the mir-290-295 cluster in mouse embryonic stem cells. PLoS Genet *7*, e1002054.

Zhu, Z., Gonzalez, F., and Huangfu, D. (2014). The iCRISPR platform for rapid genome editing in human pluripotent stem cells. Methods Enzymol *546*, 215-250.

Zhu, Z., and Huangfu, D. (2013). Human pluripotent stem cells: an emerging model in developmental biology. Development *140*, 705-717.

Zhu, Z., Li, Q.V., Lee, K., Rosen, B.P., Gonzalez, F., Soh, C.L., and Huangfu, D. (2016). Genome Editing of Lineage Determinants in Human Pluripotent Stem Cells Reveals Mechanisms of Pancreatic Development and Diabetes. Cell Stem Cell *18*, 755-768.

Zovoilis, A., Smorag, L., Pantazi, A., and Engel, W. (2009). Members of the miR-290 cluster modulate in vitro differentiation of mouse embryonic stem cells. Differentiation *78*, 69-78.

Zwaka, T.P., and Thomson, J.A. (2003). Homologous recombination in human embryonic stem cells. Nat Biotechnol *21*, 319-321.