THE UNIVERSITY *of* EDINBURGH

INVESTIGATING THE ROLE OF MICRORNAS IN MAMMALIAN DEVELOPMENTAL TRANSITIONS

LAURA BAILEY

Thesis presented for the degree of Doctor of Philosophy

The University of Edinburgh

December 2011

I declare that this thesis and the work declared herein is my own, unless otherwise stated. The work described in this thesis has not been submitted for any other degree or professional qualification.

Laura Bailey

Abstract

miRNAs are short, non-coding RNA molecules that regulate gene expression post-transcriptionally through inhibition of translation and/or mRNA degradation. Mammalian development is a complex series of developmental transitions, which relies on accurate spatial and temporal regulation of gene expression and we are interested in the role that miRNAs may play in these developmental transitions. An initial objective was to establish which, if any, miRNAs were dynamically regulated in a cell model of an early developmental transition, and to establish whether differential expression of any particular miRNA played a functional role in this developmental process. Having established a role for specific miRNAs, further objectives were to assess the reliability of current miRNA-mRNA target identification procedures and to assess the general role of miRNAs in cellular differentiation.

In order to explore the roles of miRNAs during an early developmental transition, an embryonic stem (ES) cell model of trophectoderm differentiation was used. In this model system the expression of the key ES cell regulatory gene, *Oct4*, can be conditionally repressed, which induces the ES cells to differentiate down the trophectoderm lineage. The expression of microRNAs was profiled in this model system by cloning and sequencing of small RNAs. This approach identified miRNAs that were dynamically regulated during differentiation. The expression patterns of differentially regulated miRNAs were confirmed by miRNA northern analysis. The miRNA profiling data showed that *mmu-miR-294* and *mmu-mir-295* are expressed at similar levels in ES cells and differentiated cells, which disagrees with previous reports that these miRNAs are ES cell specific. Several of the miRNAs with higher expression levels in differentiated cells are encoded within a placental-enriched polycomb group gene, *Sfmbt2*, suggesting an important role for these miRNAs in extraembryonic development. One of the miRNAs that was expressed at higher levels in ES cells than in differentiated cells, *mmu-miR-92a*, was shown to play a role in regulation of cell proliferation.

Three current methods of identifying miRNA targets were assessed. A sequence-based method using the web-based utility miRecords, which amalgamates results from numerous target prediction databases, was used to generate lists of potential targets of the *Sfmbt2* miRNA cluster and of *mmu-miR-92a*. Amalgamating

results from multiple target prediction programs may improve the likelihood that the predicted targets are real. Exemplifying this, the single *mmu-miR-92a* target that was predicted by six different target prediction programs had been previously experimentally verified. An experimental method of identifying direct miRNA targets, PAR-CLIP, was investigated but proved technically limiting for routine use in the laboratory. A proteome-based experimental method for identifying potential miRNA targets, called SILAC, was successfully used to identify proteins that were differentially expressed in the cell model of trophectoderm differentiation. Differential expression of two of these proteins, CTBP2 and CKB, was confirmed by western analysis. miRecords was then used to assess whether the differentially expressed proteins were likely to be targets of the differentially expressed miRNAs that had been identified in the miRNA profiling analysis.

The general role of miRNAs in cell differentiation was investigated using a cell line that does not express miRNAs. This ES cell line is deficient for the miRNA-processing enzyme DGCR8, which results in loss of expression of mature miRNAs in these cells. Compared to wild type ES cells, miRNA-deficient ES cells expressed normal levels of the ES cell marker genes *Oct4* and *Sox2* but elevated levels of *Nanog*. In contrast to wild type ES cells, miRNA-deficient ES cells did not upregulate the mesoderm marker gene *Brachyury* during embryoid body differentiation and showed reduced upregulation of the endoderm marker gene *Gata6*. These findings suggest that miRNAs are not required for maintenance of pluripotency, but are essential for proper ES cell differentiation.

The results presented in this thesis show that miRNAs are dynamically expressed during a mammalian developmental transition and are involved in regulating early developmental processes. We believe that miRNAs act as an additional level of genetic regulation to ensure canalisation during embryonic development.

For Mum and Dad.

"I can no other answer make but thanks,
And thanks, and ever thanks ″
Twelfth Night, 3.3.15-16

Acknowledgements

There are many people that have helped me to reach this point: from my late Grandad Mower who captured the imagination of a little girl with his ideas of perpetual motion scribbled on the back of an envelope, to my excellent school biology teacher Mr Strang, who first introduced me to the field of genetics. To all those who have contributed to my fascination and love for science, thank you.

During my PhD I have greatly enjoyed working alongside the members of the Clinton, Burdon, Sang and McGrew groups, past and present. In particular, Catriona and Lorna kept me laughing, Sunil was a constant voice of calm and Joni was there beside me from the beginning, right to the very end.

I am indebted to the members of staff who provided me with support in the laboratory. In particular, Derek gave me superb guidance through the practical world of molecular biology and was always on hand to answer my many questions, and Linda introduced me to embryonic stem cell culture and provided constant practical (and emotional) support throughout my PhD. I also greatly appreciate the technical advice given to me by Debiao, Stephen, Alison, Richard, Wilfrid and Helen.

I have been fortunate to enjoy a good working relationship with both of my supervisors. My primary supervisor Mike ensured that my work retained direction, and always gave up his time to offer help and advice when I needed it. Tom, my second supervisor, provided guidance on all things embryonic stem cell and his enthusiam for science was a great motivator. I thank you both.

My parent's encouragement and belief in my ability gave me the confidence to study for a PhD and made it easier to continue when things were not going so well, and the dedication that my brother shows his subject has been a constant source of inspiration to me.

To Andy, simply, thank you.

Contents

List of Figures

List of Tables

List of Abbreviations

2i	Two inhibitors
3' UTR	3' untranslated region
ATP	Adenosine trophosphate
BMP	Bone morphogenetic protein
BSA	Bovine serum albumin
ChIP	Chromatin immunoprecipitation
CHIR	Chiron
cDNA	Complimentary DNA
DAPI	4', 6-diamidino-2-phenylindole
DMEM	Dulbecco's modified eagle medium
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxyribonucleotide triphosphate
Dox	Doxycycline
DTT	Dithiothreitol
E	Embryonic day
EB	Embryoid body
EDTA	Ethylenediaminetetraacetic acid
EG	Embryonic germ (cells)
EGFP	Enhanced green fluorescent protein
EpiSC	Epiblast stem cell
ES	Embryonic stem (cells)
FACS	Fluorescent activated cell sorting
FITC	Fluorescin isothiocyanate
FCS	Foetal calf serum
GMEM	Glasgow minimum essential medium
GSK3	Glycogen synthase kinase 3
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRP	Horseradish peroxidase

ICM	Inner cell mass
IgG	Immunoglobulin G
iPS	Induced pluripotent stem (cells)
JAK	Janus activated kinase
lacZ	Beta-galactosidase
LIF	Leukaemia inhibitory factor
LNA	Locked nucleic acid
mRNA	Messenger RNA
miRNA	MicroRNA
MEFs	Mouse embryonic fibroblast (cells)
MOPS	3-(N-morpholino)propanesulfonic acid
MTT	(3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffer saline
PBST	PBS with Triton
PFA	Paraformaldehyde
PCR	Polymerase chain reaction
pri-miRNA	Primary microRNA
PVDF	Polyvinylidene fluoride
QPCR	Quantitative PCR
RBP	RNA binding protein
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNase	Ribonuclease
SD	Standard deviation
SDS	Sodium Dodecyl Sulphate
SEM	Standard error of the mean
SILAC	Stable isotope labelling with amino acids in cell culture
siRNA	Short interfering RNA
SSC	Saline-sodium citrate
TBE	Tris-borate-EDTA
TBS	Tris-buffered saline
TBST	TBS with Tween 20
TE	Trophectoderm
TF	Transcription factor
Tris	Tris (hydroxymethyl) aminomethane
TS	Trophoblast stem
TVP	Trypsin versene phosphate
UV	Ultra-violet
XEN	Extraembryonic endoderm

Chapter 1

Introduction

1.1 General Introduction

The process of embryonic development, whereby an organism develops from a single cell to a complex form consisting of highly specialised structures, tissues and cell types, is truly remarkable and far from being understood. In order to study mammalian development, *in vivo* and *in vitro* methods can be used, including the derivation and culture of embryonic cell types such as embryonic stem (ES) cells. ES cells are unique in their ability to self-renew indefinitely and retain pluripotency: the ability to differentiate into cell types of any germ layer. Transfer of ES cells to appropriate conditions can induce their differentiation into more specialised cell types, allowing the factors controlling these transitions to be investigated. Understanding the factors that control developmental transitions will not only provide insight into the regulation of normal and abnormal embryonic development, but may also lead to the establishment of an unlimited source of specialised cell types and organs for therapeutic benefit.

In order to generate the vast array of cell types observed during the development process and in the adult organism, accurate spatial and temporal regulation of gene expression is essential. Some means of regulating gene expression, such as control of transcription and protein modification, have long been established. Recently, an entirely novel level of regulation has been described – post-transcriptional regulation of gene expression by microRNAs (miRNAs). miRNAs

are short, non-coding RNA molecules of approximately 22 nucleotides in length that have been found in all eukaryotic organisms investigated to date. Some miRNAs are highly conserved across species, and research suggests that miRNAs have existed since early animal evolution (Grimson *et al.*, 2008). miRNAs were initially discovered as regulators of development in nematode worms and have since been shown to have roles in the developmental of numerous animals and plants. It is possible that miRNAs function as master regulators of developmental processes or to ensure canalisation during development. A combination of both these functions is also possible. We believe that miRNAs act primarily as an additional level of genetic regulation to ensure canalisation during embryonic development.

This thesis describes the results of investigating the role of miRNAs in early mammalian developmental transitions. In order to investigate the role of miRNAs in regulating developmental processes, ES cells were used to model differentiation events *in vitro*. To explore the specific roles of miRNAs during an early developmental transition, an embryonic stem (ES) cell model of trophectoderm differentiation was used, and to explore the general roles of miRNAs in developmental transitions, the differentiation capabilities of wild type and miRNA-deficient ES cells were compared. This introduction will: (1) describe what is currently known about the biogenesis, regulation and function of miRNAs; (2) detail early mouse development; (3) discuss what is known about ES cells and why they are suitable as models of developmental processes; and (4) describe the known roles of miRNAs in development, with particular attention to the roles of miRNAs in early mouse development.

1.2 The biology of miRNAs

MicroRNAs (miRNAs) are short, single-stranded RNA molecules that represent an entirely novel mechanism of genetic regulation, whereby gene activity is attenuated post-transcriptionally through translational inhibition and/or destabilisation of messenger-RNA. miRNAs were first discovered in the nematode worm, *C. elegans*, but are now thought to function in most, if not all, multicellular organisms as well as some viruses.

1.2.1 The history of miRNA

1.2.1.1 The discovery of lin-4

The first description of an endogenous small RNA with translational suppression capabilities was published in Cell in 1993 (Lee *et al.*, 1993). This report showed that in *C. elegans* a non-protein coding transcript of 22nt in length, known as *lin-4*, was complementary to a repeated sequence found in the 3' UTR of *lin-14* mRNA. A loss of function of *lin-4* resulted in early developmental fates re-occurring at late developmental stages whereas the opposite was true for loss of function of LIN-14. It was therefore thought that *lin-4* could function to negatively regulate the level of LIN-14 protein. Lee *et al.* cloned the *lin-4* locus and showed complementarity between *lin-4* and *lin-14*. Because *lin-4* did not produce a protein-coding transcript, it was suggested that the regulation represented inhibition of translation via an antisense RNA-RNA interaction.

1.2.1.2 miRNAs as general regulators of gene expression

It wasn't until 2001 that the term 'microRNA' (abbreviated to miRNA) was first coined in a series of three papers presented in Science (Lee & Ambros, 2001, Lau *et al.*, 2001, Lagos-Quintana *et al.*, 2001). Lee & Ambros used bioinformatics approaches and cDNA cloning to discover 15 new miRNA genes while Lau *et al.* discovered 55 new miRNAs by isolating and cloning endogenous RNAs. Both of these studies were performed in *C. elegans*. By isolating and cloning miRNAs from HeLa cells and *D. melanogaster*, Lagos-Quintana *et al.* showed that miRNAs are expressed in both vertebrates and invertebrates. This series of articles led to speculation that miRNAs were numerous, widespread mediators of post-transcriptional genetic regulation. Since these papers were published the field of miRNA biology has expanded rapidly and miRNAs are now thought to exist in most multicellular organisms.

1.2.1.3 The extent of miRNA regulation

The extent to which miRNA regulation operates is gradually being realised. The latest release of the miRBase database (http://www.mirbase.org/, Griffiths-Jones, 2004), which lists published microRNA sequences, contains '18226 entries

representing hairpin precursor miRNAs, expressing 21643 mature miRNA products, in 168 species'. This includes over 700 different mouse miRNAs. In 2005 it was reported that human miRNAs might regulate up to 30% of human protein coding genes (Lewis *et al.*, 2005). A more recent estimate suggests that the human genome may encode over 1000 different miRNAs, which are predicted to regulate >60% of human protein coding genes (Friedman *et al.*, 2009). However, only 2286 miRNA:mRNA interactions appear in the miRecords database (http://mirecords.biolead.org, Xiao *et al.*, 2009), which is the largest collection of 'validated' miRNA targets available. Bearing in mind that each miRNA is predicted to target hundreds of mRNAs (Lim *et al.*, 2005), and the number of confirmed interactions makes up a very small proportion of the interactions that are likely to occur.

Additionally, only a small proportion of known miRNAs have been functionally characterised to date. Mice deficient for specific miRNAs or miRNA clusters have been generated and show roles for *miR-1* and *miR-208* in cardiogenesis (van Rooij *et al.*, 2007, Zhao *et al.*, 2007), *miR-155* and *miR-146a* in the immune system (Rodriguez *et al.*, 2007, Thai *et al.*, 2007, Boldin *et al.*, 2011), and the *miR-17-92a* cluster in development of the heart, lung and immune system (Ventura *et al.*, 2008). Recently, a library of germline-transmissible mouse ES cell lines that harbour deletions for 392 miRNA genes has been generated (Prosser *et al.*, 2011). This resource will provide material for the functional characterisation of many more mouse miRNAs. One area of miRNA biology that is reasonably well understood is miRNA biogenesis.

1.2.2 miRNA biogenesis

The stages of miRNA biogenesis from transcription to incorporation into the RISC complex are shown in *Figure 1.1.* The following is an account of the current understanding regarding miRNA biogenesis.

1.2.2.1 Genomic location of miRNAs

MicroRNAs are transcribed from the genome as long precursor molecules called pri-miRNAs. miRNAs may be located in intergenic regions of the genome or

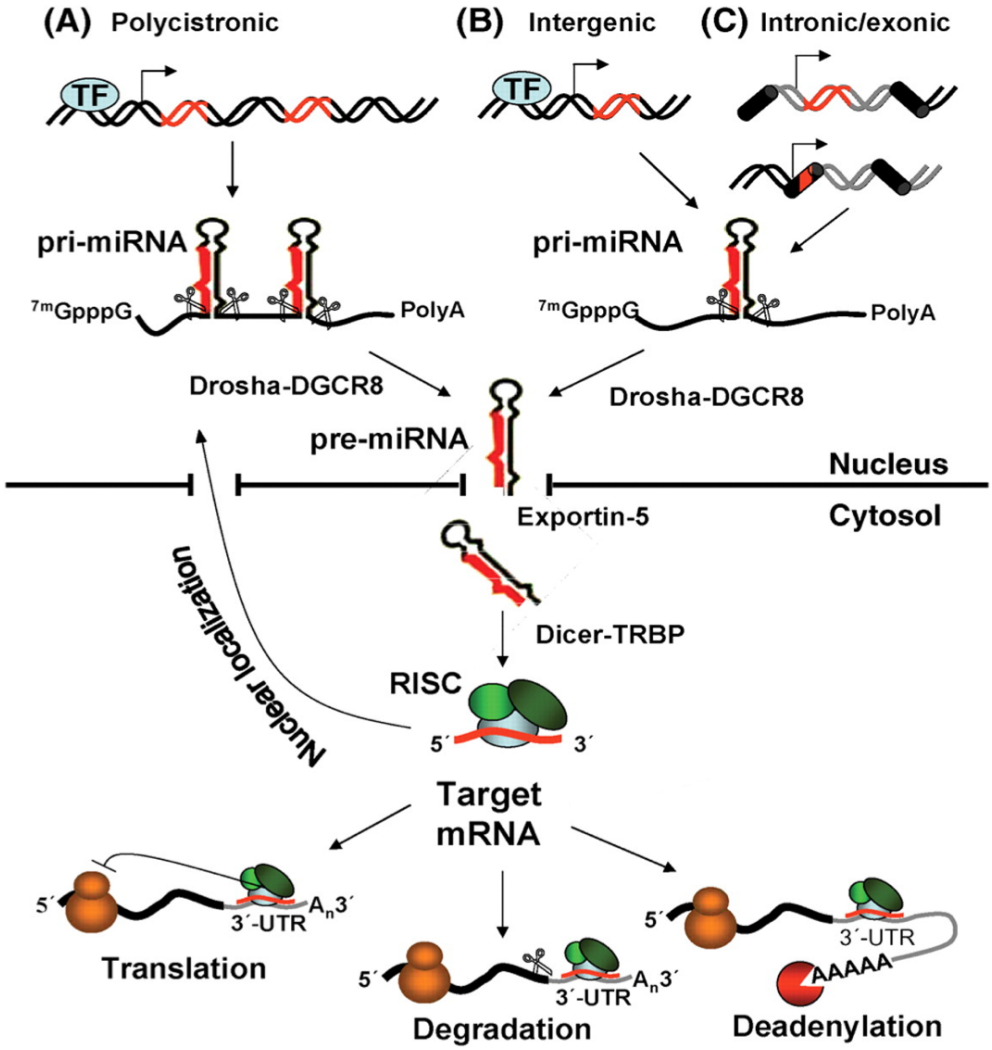


Figure 1.1 – *miRNA genomic organisation, biogenesis and function.* miRNAs are localised in the genome as polycistrons that are subsequently cleaved into multiple pre-miRNAs (A), in intergenic regions (B) or in intronic/exonic regions of protein coding genes (C). Primary miRNAs (Pri-miRNA) are transcribed from the genome by RNA polymerase II and have a 7-methylguanosine (7mGpppG) Cap and a poly(A) tail. Pri-miRNA molecules are processed into ~60nt hairpin pre-miRNAs by the microprocessor complex composed of the RNAse III enzyme DROSHA and the ds RNA binding protein DGCR8. Exportin 5 (XPO5) actively transports the pre-miRNA into the cytosol, where it is processed by the RNAse III enzyme DICER with its partner, the RNA binding protein, TRBP into mature ~22nt double stranded miRNA. TRBP facilitates recruitment of AGO2, a key component of the RNA-induced silencing complex (RISC) and a single strand of the miRNA (the 'guide' strand) is incorporated into the RISC complex. The mature miRNA sequence then guides the RISC complex to regions of partial complementarity in the target mRNA sequence and represses target expression by one of a number of mechanisms: repression of mRNA translation, miRNA cleavage and subsequent degradation or deadenylation of the mRNA and subsequent degradation. Figure adapted from Fazi & Nervi, 2008.

within the introns of protein-coding genes (Ying & Lin, 2005). It was initially assumed that all intronic miRNAs are transcribed as part of the host protein-coding transcript. However, it is now known that some intronic miRNAs have their own promoter and enhancer, which means the miRNAs will be synthesised as a separate transcriptional unit from the host gene (Ozsolak *et al.*, 2008, Monteys *et al.*, 2010). This may explain why the expression profiles of some intronic miRNAs do not correlate with the expression pattern of their host gene (Ozsolak *et al.*, 2008, Sikand *et al.*, 2009). Some miRNAs are organised in clusters in the genome. By analysis of nucleosome positioning patterns and ChIP-chip screens. Ozsolak *et al.* found either one promoter or no promoter upstream of miRNA clusters, which suggested either a shared transcriptional start site or the presence of additional promoters not identified in their analysis. Together with the additional observation that polycistronic transcripts for several miRNA clusters exist (Lee *et al.*, 2002), this suggests that miRNA clusters can be transcribed as a single pri-miRNA.

1.2.2.2 miRNA transcription

While the majority of small RNAs (e.g. tRNAs and U6 snRNAs) are transcribed using an RNA polymerase III dependent mechanism, most microRNAs are transcribed by RNA polymerase II (Lee *et al.*, 2004). pri-miRNAs have characteristics of protein-coding transcripts exhibiting 5' caps and 3' poly(A) tails in their transcripts. Lee *et al.* showed that transcription of miRNA genes was dependent on RNA polymerase II, and therefore miRNA genes could be classified as being of class II. However, there is evidence to suggest that at least some miRNAs are transcribed by RNA polymerase III if their genomic location is within an Alu repeat region as Alu transcription occurs through RNA polymerase III (Borchert *et al.*, 2006). Transcription results in the formation of long, precursor molecules called pri-miRNAs, which enter the miRNA biogenesis pathway.

1.2.2.3 Processing of Pri-miRNAs by the microprocessor complex

Pri-miRNAs may range in size from a few hundred nucleotides up to several kilobases in length. The pri-miRNAs form hairpin structures, which are cleaved in

the nucleus to form precursor miRNAs (pre-miRNAs). This cleavage is dependent on the RNase III nuclease, DROSHA (Lee *et al.*, 2003). RNase III proteins are endonucleases specific to double stranded RNA. They cleave the RNA helix producing characteristic 2nt 3' overhangs. DROSHA's status as an RNase III nuclease, its involvement in pre-rRNA processing, and the fact it localises predominantly to the nucleus made it a good candidate as a nuclear processor of pri-miRNAs. Lee *et al.* (2003) immunoprecipitated DROSHA and found it had the capacity to process pre-miRNA from pri-miRNA *in vitro*. Similarly, depletion of *Drosha* by siRNA resulted in a build-up of pri-miRNAs and a reduction in pre-miRNAs.

Characterisation of native complexes containing DROSHA (Denli *et al.*, 2004) showed it to be a component of a 500 kDa complex termed the microprocessor complex. The complex was shown to include another protein, which was dubbed PASHA (partner of Drosha). PASHA is equivalent to vertebrate DGCR8 (DiGeorge syndrome critical region gene 8), and functions within the microprocessor complex to stably bind the pri-miRNA and act as a molecular ruler to determine the precise cleavage site for generation of pre-miRNAs (Han *et al.*, 2006) (*Figure 1.2*).

1.2.2.4 Nuclear export of pre-miRNAs by Exportin-5

The pre-miRNAs that are generated by the microprocessor complex form short hairpins of approximately 70 nucleotides in length that are transported from the nucleus to the cytoplasm. Export of pre-miRNAs is dependent on Exportin-5 (XPO5), a RanGTP-dependent dsRNA-binding protein. Competition experiments showed that export of labelled pre-miRNAs could be inhibited by the presence of unlabelled competitor pre-miRNAs, but that export can be rescued by the addition of exogenous XPO5. Conversely, a reduction in cellular levels of XPO5 by RNAi resulted in a reduction in mature miRNA levels (Lund *et al.*, 2004). It is thought that the spatial separation of the nucleases DROSHA and DICER in the nucleus and cytoplasm respectively, allows their sequential action; XPO5 may play a role between these two functions to make sure pre-miRNAs exported from the nucleus are of high integrity (Wang *et al.*, 2011). This is because XPO5 recognises the pre-miRNA independently of its sequence or loop structure but requires a defined length of double-stranded stem and 2nt 3' overhangs for

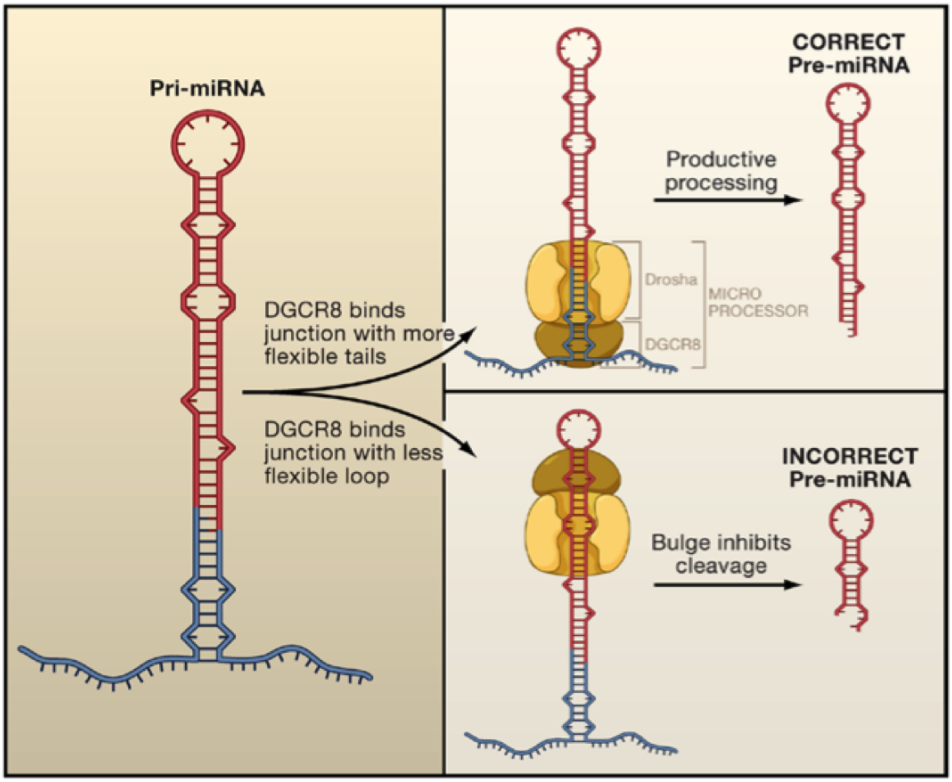


Figure 1.2 – *The role of DGCR8 in the microprocessor.* DGCR8 is thought to bind more favourably to the junction between the rigid stem structure of the pri-miRNA and the flexible 5' and 3' single stranded segments than between regions of the stem and the more constrained loop structure. Correct binding of DGCR8 is thought to result in correct positioning of DROSHA so that the cleavage site of the RNA stem loop will occur ~11 bp up the stem, resulting in an ~70 nucleotide long pre-miRNA hairpin. Thus, DGCR8 acts with structural features of pri-miRNAs to promote their accurate processing into pre-miRNAs by the microprocessor complex. From Seitz & Zamore, 2006.

successful binding (Lund *et al.*, 2004, Zeng & Cullen, 2004). Thus, only correctly processed miRNAs will be exported into the cytoplasm where they are processed to double-stranded miRNA molecules by a second endonuclease called DICER.

1.2.2.5 Processing of pre-miRNAs by *Dicer* and TRBP

In the cytoplasm another RNase III enzyme, DICER, acts with its partner TRBP (TAR RNA-binding protein) to process pre-miRNA into mature double stranded miRNA with characteristic 2nt 3' overhangs (Hutvagner *et al.*, 2001, Chendrimada *et al.*, 2005). DICER processing is required for production of mature siRNAs as well as miRNAs (Bernstein *et al.*, 2001). Grishok *et al.* first showed the involvement of *Dicer* in the miRNA pathway in *C. elegans*; the progeny of animals treated with *Dicer* siRNA exhibited accumulation of the long forms (~70nt) of the miRNAs lin-4 and let-7, while the levels of short forms (~22nt) were reduced (Grishok *et al.*, 2001). DICER contains an NH2-terminal DEXH-box ATP-dependent RNA helicase domain, a PAZ domain, tandem RNAse III motifs, and a COOH-terminal dsRNA-binding domain (Hutvagner *et al.*, 2001). TRBP has three double-stranded RNA-binding domains and is required for the recruitment of AGO2 to the small interfering RNA (siRNA) bound by DICER (Chendrimada *et al.*, 2005). DICER processing results in formation of mature miRNAs, but DICER processing is not always required because recently non-canonical pathways of miRNA biogenesis have been described.

1.2.2.6 Non-canonical pathways of miRNA processing

Some miRNAs are generated by non-canonical biogenesis pathways (*Figure 1.3*). For example, intronic miRNAs (called mirtrons) may form a stem-loop structure where the 3' end of the stem-loop precursor is coincident with the 3' splice site. This structure is cleaved by nuclear pre-mRNA splicing rather than DROSHA/DGCR8 processing (Okamura *et al.*, 2007, Ruby *et al.*, 2007). Additionally, there is evidence of miRNA processing in the absence of *Dicer*. The level of *miR-451* expression is not affected by loss of *Dicer*, but is reduced in AGO2 mutants. Investigation of this revealed that *pre-miR-451* is not processed by *Dicer*, but is loaded into AGO and is cleaved by the AGO2 catalytic centre to generate an in-

termediate 3' end, which is then further trimmed (Cheloufi *et al.*, 2010, Cifuentes *et al.*, 2010).

1.2.2.7 Regulation of miRNA biogenesis

miRNAs are post-transcriptional regulators of gene expression, but they are also the subject of regulation themselves, both directly and through regulation of their associated proteins. miRNAs are regulated at the transcriptional level, during their biogenesis, through RNA editing of their nucleotide sequence, at the functional level, through the cellular localisation of the miRNA and through regulation of their decay (reviewed in Krol *et al.*, 2010b). As an example, the miRNA *let-7* is regulated by LIN28 protein at multiple stages of its biogenesis. *Let-7* is regulated during processing by the microprocessor complex, during DICER processing and through polyuridylation and subsequent degradation (Viswananthan *et al.*, 2008, Newman *et al.*, 2008, Rybak *et al.*, 2008, Heo *et al.*, 2008). Other examples of regulation of miRNA expression are described below.

Regulation of miRNA expression can occur at the transcriptional level. Most miRNAs are transcribed by RNA polymerase II although there is evidence for transcription of some human miRNAs by RNA polymerase III (Borchert *et al.*, 2006). Both RNA polymerases recognised different sequence elements resulting in a large variety of regulatory options. Expression of selected miRNAs is under the control of transcription factors such as C-MYC (O'Donnell *et al.*, 2005) and dependent on the methylation status of the miRNA promoter region (Saito *et al.*, 2006). Differential expression of proteins involved in miRNA biogenesis also contributes to regulation of miRNA expression. This includes regulation of stability of the key RISC component AGO. For example, AGO is hydroxylated at proline 700 by the type I collagen prolyl-4-hydroxylase (C-P4H(I)), which increases protein stability and facilitates AGO localisation to P-bodies (Qi *et al.*, 2008). As can be seen the generation of mature miRNAs is subject to numerous regulatory mechanisms.

1.2.2.8 Mature miRNAs

The mature miRNA is a short double-stranded RNA molecule of between 20 and 24 nucleotides in length. Processed miRNAs generally have a longer half-

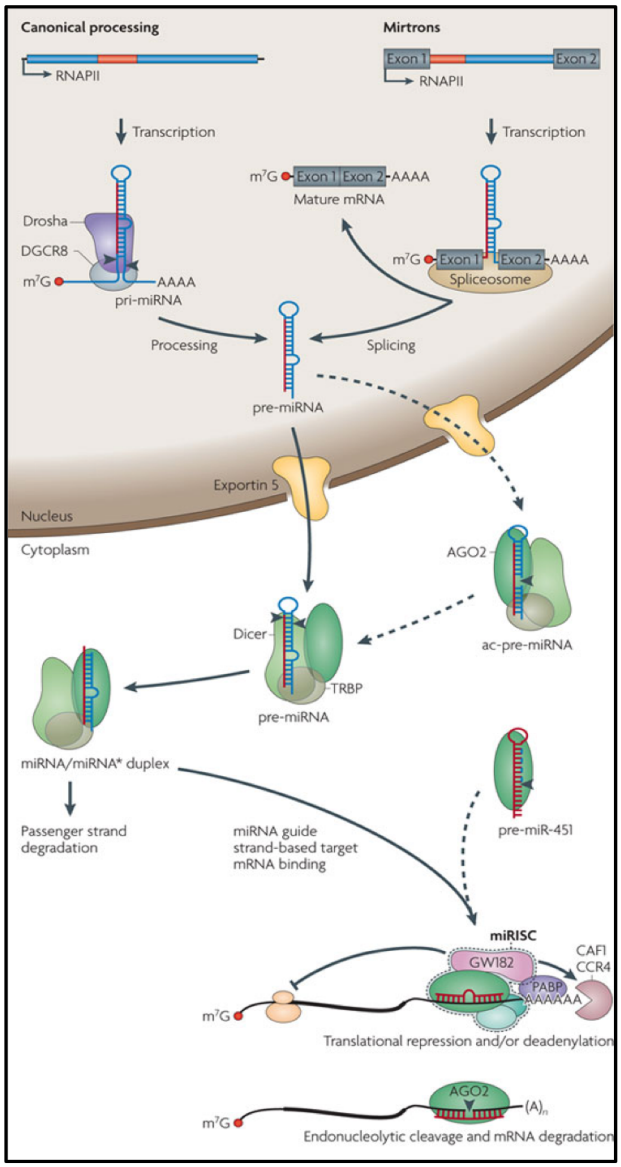


Figure 1.3 – *Non-canonical miRNA biogenesis.* During canonical miRNA processing, miRNAs reach their mature form after two RNAse III-mediated cleavage events involving the DROSHA-DGCR8 complex in the nucleus and the DICER-TRBP complex in the cytoplasm. In some cases, miRNAs are encoded within short introns (miRtrons) and are processed into pre-miRNA form by the splicing apparatus, thereby bypassing the DROSHA-DGCR8 processing step (Ruby *et al.*, 2007). In other cases, AGO2, which has endonuclease activity, can support DICER processing by cleaving the 3' arm of the pre-miRNA hairpin (Diederichs & Haber, 2007). This generates an additional processing intermediate called AGO2-cleaved precursor miRNA (ac-pre-miRNA). In addition, processing of *miR-451* requires AGO2 cleavage but not DICER cleavage (Cheloufi *et al.*, 2010, Cifuentes *et al.*, 2010). Thus, miRNAs can be generated in a *Drosha*-independent and *Dicer*-independent manner. Figure from Krol *et al.* (2010b).

life than the average mRNA molecule. For example, the half-life of *miR-223* was determined to be approximately 25 hours whereas the concentration of an mRNA derived from the same precursor transcript, NGFR, decreased 10-fold in 24 hours (Baccarini *et al.*, 2011). Despite this there are probably mechanisms, as yet unknown, which regulate turnover of miRNAs because some show rapid decreases in expression upon certain stimuli (Pedersen *et al.*, 2007, Krol *et al.*, 2010a). For example, expression of interferon beta, a protein produced in host cells in response to the presence of pathogens, leads to rapid reduction in expression of *miR-122*, a miRNA that has been shown to be essential for hepatitis C virus replication (Pedersen *et al.*, 2007).

Once mature miRNAs have been synthesised they are generally localised to the cytoplasm but a distinct hexanucleotide terminal motif on *hsa-miR-29b* results in its localisation to the nucleus (Hwang *et al.*, 2007). This suggests that miRNAs contain cis-regulatory motifs and is interesting because there is evidence that miRNAs could have a function in the nucleus as regulators of gene transcription (Place *et al.*, 2008). Kinetic analysis of miRNA expression showed that miRNAs regulate multiple mRNAs during their lifetime; an individual miRNA is thought to regulate at least two target transcripts before it is degraded (Baccarini *et al.*, 2011). Target regulation limits miRNA recycling because it promotes posttranscriptional modifications to the 3' end of the miRNA, which results in modulation of miRNA stability (Baccarini *et al.*, 2011).

1.2.3 Interaction of miRNAs with target transcripts

1.2.3.1 miRNAs and the RISC complex

The complex responsible for targeting miRNAs (and also other small RNA molecules such as siRNAs) is known as the RNA-induced silencing complex (RISC). The term RISC was first coined by Hammond *et al.* (2000) in a paper showing degradation of specific mRNAs in *Drosophila* cells after transfection of double stranded RNAs (dsRNAs). The researchers reasoned that this degradation was due to sequence specific nuclease activity, as non-specific mRNAs were not degraded.

1.2.3.2 Components of the RISC complex

ARGONAUTE proteins All RISC complexes characterised so far contain at least one ARGONAUTE (AGO) protein, which are the core effectors of the miRNA pathway. The AGO protein family has multiple members, all of which are involved in dsRNA gene silencing (Czech & Hannon, 2011). The number of AGO paralogs present varies from only one in *S. pombe* (Verdel *et al.*, 2004), to over twenty in *C. elegans* (Carmell *et al.*, 2002). Mouse and human genomes encode four AGO proteins and mouse ES cells that are deficient for AGO1-4 are completely defective in miRNA silencing (Su *et al.*, 2009). Proteins within the AGO family have a PAZ domain and a PIWI domain (Rand *et al.*, 2004). The PAZ domain binds RNA with the specific 2-nucleotide 3' overhang that is characteristic of short RNA duplexes. The PIWI domain has been shown to mediate protein-protein interactions between AGO and *Dicer* (Tahbaz *et al.*, 2004) and constitutes the catalytic domain of the enzyme (Liu *et al.*, 2004). The structure of AGO is shown in *Figure 1.4*. The four mouse AGO proteins are ubiquitously expressed and share significant sequence homology. However, only AGO2 is capable of 'slicer' activity, catalysing the cleavage of the miRNA target (Liu *et al.*, 2004). Purified AGO proteins precipitate with single stranded miRNAs (Meister *et al*, 2004), which suggests that the double-stranded mature miRNA is either unwound by an as yet unidentified helicase or that one strand may be cleaved by the 'Slicer' activity of AGO2 (Matranga *et al.*, 2005). The guide strand is retained by AGO because of a lower thermodynamic stability at the 5'end, while the other strand (referred to as miRNA*) is degraded (Schwarz *et al.*, 2003).

Other components of the RISC complex Other proteins that associate with the RISC complex in *Drosophila* include a protein with five staphylococcal/micrococcal nuclease domains called Tudor-SN (Caudy *et al.*, 2003), Fragile X-related protein and Vasa intronic gene (VIG) (Caudy *et al.*, 2002). The RISC complex allows the recruitment of the miRNA to mRNA targets through target recognition by perfect or imperfect complementarity between RNA sequences, and facilitates processing of the mRNA.

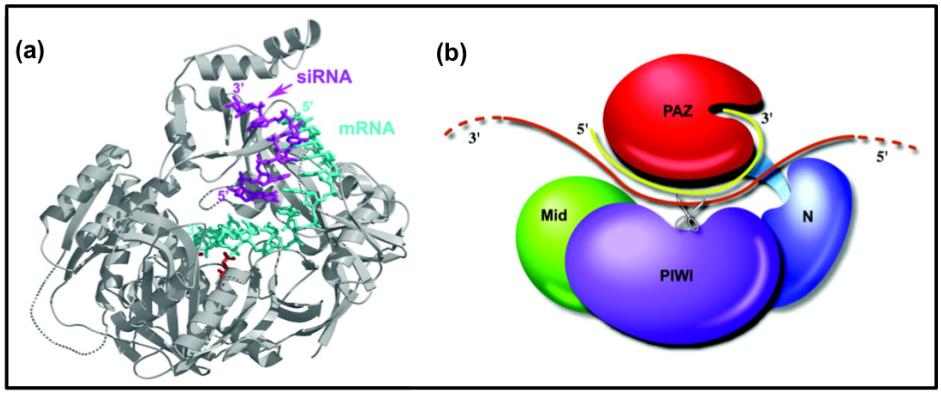


Figure 1.4 – *The structure of Argonaute proteins.* Song *et al.* (2004) identified the crystal structure of AGO protein from the archaebacterium *Pyrococcus furiosis* and presented a model for siRNA-guided mRNA cleavage by AGO. (a) shows the 3' potion of an siRNA (purple) and the 5' end of the mRNA stand (green), which were placed by superimposing the PAZ domains from PfAGO and hAGO1. The view is similar to that seen in (b), which is a schematic depiction of the model proposed by Song *et al.* for siRNA-guided mRNA cleavage. The siRNA (pale green) is localised with its 3' end in the PAZ cleft. The rest of the siRNA lies in the AGO groove. The mRNA comes in between the N-terminal and PAZ domains and out between the PAZ and MID domains. The active site is in the PIWI domain and cleaves the mRNA opposite the middle of the siRNA guide. The MID-PIWI domains have been crystallised from eukaryotic AGO1 protein (from *Neurospora crassa*) and show similar domain orientation to those in prokaryotic AGO proteins suggesting a conserved mode of binding with the RNA substrate. (Boland *et al.*, 2011). Figure adapted from Song *et al.* (2004).

1.2.3.3 Factors important for target recognition

The most important region of the miRNA for target recognition is nucleotides 2-7 from the 5' end, termed the 'seed region'. In most cases this region pairs perfectly with the mRNA target (Bartel, 2009). Different types of seeds have been characterised depending on the extent of the binding within this region and whether there is binding at position 8 or an adenine complementary to position 1 of the miRNA. Although supplementary binding at the 3' end of the miRNA has been assumed, there is little evidence to directly support its frequent function (Doench and Sharp, 2004). However, 3' compensatory binding is thought to play a role in situations where seed pairing is compromised, for example by a bulge or mismatch in the seed region (Yekta *et al.*, 2004).

1.2.3.4 The miRNA 'seed' region

The seed region was first identified as being important in miRNA targeting through computational prediction of miRNA targets followed by experimental verification of some predicted targets (Stark *et al.*, 2003). Stark *et al.* observed that the few animal miRNA-target duplexes that had been verified previously contained mismatches, gaps and G:U basepairs at different positions. Even allowing for G:U basepairs, the longest continuous sequence alignments were between 8 and 10 nucleotides in length. They developed a computer based screening strategy for predicting miRNA targets, which predicted all known miRNA-mRNA target duplexes. Six novel predicted miRNA targets were verified experimentally. Comparison of the features of the five previously identified miRNA targets, and the six newly validated targets revealed that all target sites shared perfect sequence identity with nucleotides 2-8 of the miRNA. Soon after, Lewis *et al.* made a similar discovery and called the region most important for targeting the 'miRNA seed' (Lewis *et al.*, 2003).

1.2.4 Mechanisms of miRNA function

miRNA regulation of target mRNA is thought to occur through translation repression and/or mRNA destabilisation. Translational repression of target mRNAs

has been reported to occur in a number of ways: through inhibition of translational initiation or elongation, co-translational protein degradation and also premature termination of translation. Methods of mRNA regulation by miRNAs are summarised in *Figure 1.5.*

1.2.4.1 The mechanism of miRNA function in plants and animals

It was thought that, in animals, miRNAs functioned primarily through translational repression whereas in plants miRNAs acted almost entirely through mRNA degradation. These differences arose because plant miRNAs show near perfect complementarity with their mRNA targets, while in animals the base pairing is generally much less extensive. More recent data has shown that this functional distinction is not correct and miRNAs can induce mRNA degradation in animals and translation repression in plants (Behm-Ansmant *et al.*, 2006, Brodersen *et al.*, 2008). The question of which one of these mechanisms acts predominantly is still under debate (reviewed in Huntzinger & Isaurralde, 2011). However, the consensus still stands that the method of regulation is dependent on the degree of mismatch between the guide miRNA and its target mRNA; miRNAs with near perfect complementarity to target mRNAs guide mRNA cleavage, while those with a large number of mismatches inhibit translation.

1.2.4.2 Translational inhibition

Translational inhibition was first recognised as a potential route for miRNA function due to findings in *C. elegans*; miRNA-regulated mRNA levels were not substantially changed, but levels of the encoded protein were reduced (Olsen *et al.*, 1999). In addition, the mRNAs were found to be associated with ribosomes that could go on to perform translational elongation *in vitro.* Since then, further evidence has emerged that miRNAs function through translational inhibition after protein synthesis has been initiated (Petersen *et al.*, 2006). Studies have also shown evidence of inhibition of protein synthesis at the initiation stage through impaired cap recognition (Pillai *et al.*, 2005, Humphreys *et al.*, 2005).

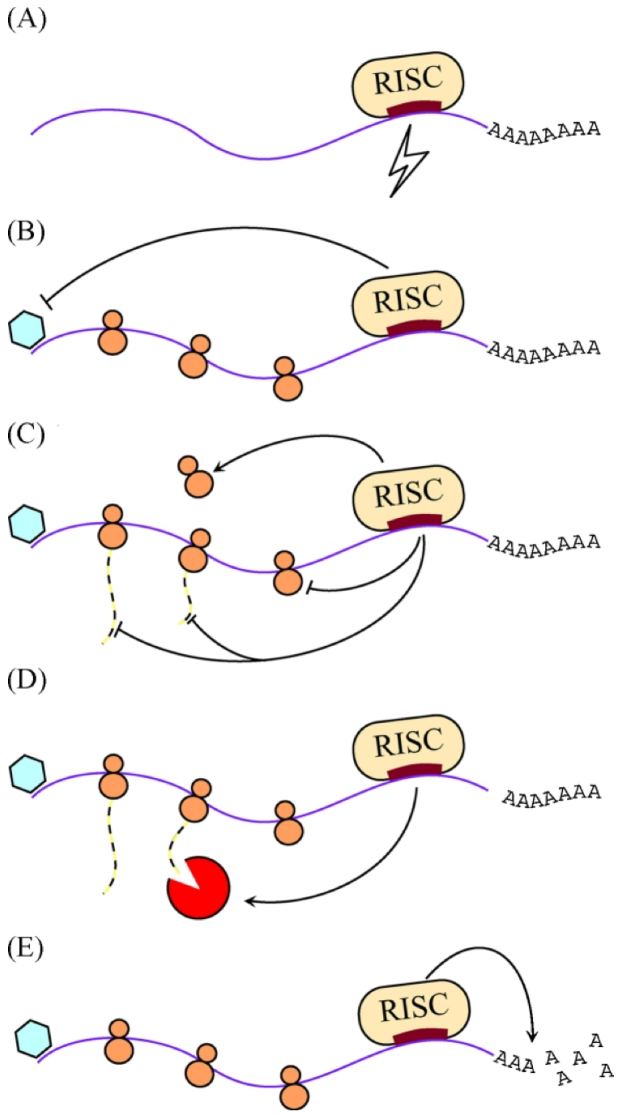


Figure 1.5 – *Mechanisms of miRNA function.* Schematic showing various proposed mechanisms of miRNA function. Target mRNA is shown as a violet line, miRNA is shown as thick red line associated with RISC complex, the blue hexagon represents the translation initiation complex and the orange circles represent ribosomes. (a) Perfect complementarity between miRNA and target mRNA leads to cleavage of the mRNA by the catalytic site within the PIWI domain of AGO proteins followed by mRNA degradation. Imperfect complementarity between miRNA and mRNA leads to inhibition of translation initiation (b) or inhibition of translation elongation/termination and promotion of ribosome drop-off (c). There is also evidence that regulation can occur through recruitment of a proteolytic enzyme to degrade emerging polypeptides (d) and deadenylation followed by degradation of the mRNA target (e). Figure from Lewis & Steel (2010).

1.2.4.3 Localisation of miRNA targets to P-bodies

A relationship between miRNA-mediated regulation and cytoplasmic P-bodies has also been proposed. AGO proteins localise to P-bodies along with repressed mRNAs, which localise in a miRNA dependant manner (Pillai *et al.*, 2005). The P-bodies contain a range of enzymes involved in mRNA turnover and so a model was suggested in which RISC binding to mRNA results in delivery to a P-body where the mRNA comes into contact with enzymes responsible for deadenylation and decapping or is held in stasis. Stasis could potentially be achieved due to the absence of ribosomes within P-bodies. However, it is now considered that P-bodies are a consequence, rather than a cause of miRNA-mediate gene silencing because silencing can occur in cells devoid of functional P-bodies (Eulalio *et al.*, 2007).

1.2.4.4 miRNA-mediated destabilisation of target transcripts

While it was originally thought that animal miRNAs functioned mainly through translational inhibition, it is now known that destabilisation of target mRNAs is a common result of miRNA target recognition. The first evidence of this was through transfection of a miRNA into cells that did not normally express the specific miRNA. Twelve hours after transfection a significant decrease (between 1.5 and 3.6-fold) was observed in the levels of many mRNAs, most of which had potential target sites for the miRNA (Lim *et al.*, 2005). Multiple further studies have made the same observation (including Sebach *et al.*, 2008, Baek *et al.*, 2008, Guo *et al.*, 2010). The majority of this mRNA degradation is thought to occur through miRNAs directing their targets to the 5' to 3' mRNA decay pathway where mRNAs are deadenylated by the CAF1-CCR4-NOT deadenylase complex and then decapped by DCP2 (Rehwinkel *et al.*, 2005, Behm-Ansmant *et al.*, 2006). Although miRNAs can direct endonucleolytic cleavage of fully complementary targets, this seems to be a rare occurrence in animals. Only two examples of such 'Slicing' have been published: *miR-196* mediated cleavage of HOXB8 mRNA and *miR-127* and *miR-136* mediated cleavage of retrotransposon-like 1 (rtl1) mRNA (Yekta *et al.*, 2004, Davis *et al.*, 2005).

1.2.4.5 Promotion of transcription

There is evidence to suggest that miRNAs may also act to promote gene transcription. Transfection of *miR-373* or *pre-miR-373* into mammalian cells was shown to increase the levels of E-cadherin expression, which contained a *miR-373* target site in its promoter region (Place *et al.*, 2008). Additionally, design and introduction of synthetic dsRNAs complementary to regions within gene promoters was shown to result in sequence-specific and long term induction of the targeted gene (Li *et al.*, 2006, Janowski *et al.*, 2007). However, although it has been shown that miRNA-mediated activation of gene transcription is sequence dependent and requires functional *Dicer* and AGO2 there is, as yet, no suggested mechanism for this novel miRNA function.

1.2.5 miRNAs and disease

miRNAs have been shown to have roles in many, if not all, biological processes including apotosis, cell signalling, organogenesis and development. Soon after their discovery, miRNAs were linked with human health and shown to have roles in neurological diseases, immunity, viral and endocrine diseases and in cancer.

1.2.5.1 miRNAs in neurological disorders

The neurological disorder most frequently associated with miRNAs is fragile X syndrome, which is the most common inherited mental retardation disease. This syndrome is caused by mutations that prevent production of fragile X mental retardation protein (FMRP), an RNA-binding protein that is thought to regulate the expression of numerous genes at the protein synthesis level. The mechanism by which FMRP does this is not yet known. FMRP normally forms ribonucleoprotein complexes and can interact with the RISC complex as well as directly with miRNAs (Caudy *et al.*, 2002, Jin *et al.*, 2004). In *Drosophila*, AGO1 is essential for FMRP function (Jin *et al.*, 2004). These findings suggest that FMRP may regulate gene expression through an RNAi-related mechanism.

1.2.5.2 miRNAs and the immune system

Many roles for miRNAs in the immune system have been documented predominantly involving the miRNAs *miR-146a* and *miR-155*. These miRNAs are strongly upregulated in response to lipopolysaccharide (LPS) stimulation of the immune system (Tili *et al.*, 2007, Nahid *et al.*, 2009). *miR-146a* is induced by the immune system regulator NF-KB and directly regulates expression of TRAF6 and IRAK1, which are upstream of NF-KB in this immune system cascade of gene expression (Taganov *et al.*, 2006). Therefore *miR-146a* creates a negative-feedback loop that controls the NF-KB mediated response to LPS stimulation.

1.2.5.3 miRNAs and cancer

Most research of miRNAs in disease has been in the field of cancer. The first evidence for a correlation between miRNAs and cancer was the observation that miRNAs are frequently located in Cancer-Associated Genomic Regions (CAGRs) (Calin *et al.*, 2004). Work following on from this has revealed that miRNA expression profiles in tumours are significantly different from those in healthy tissue (Lu *et al.*, 2005). Multiple studies have made similar findings and have led to lists of miRNAs that are thought to be mainly oncogenic and those that are thought to act as tumour suppressors. Among the miRNAs associated with cancer are the *miR-15a-16* cluster, which is deleted or downregulated in chronic lymphocytic leukaemia (Calin *et al.*, 2002), *miR-17* and *miR-21*, which show upregulation in colon, lung, stomach, prostrate and pancreatic tumours (Volinia *et al.*, 2005) and the *miR-17-92a* cluster of miRNAs. The overexpression of the *miR-17-92a* cluster has been frequently associated with solid tumours or in haematological malignancies (reviewed in Olive *et al.*, 2010). Conditional knock out of *miR-17-92a* showed that expression of *miR-17-92a* is required to suppress apoptosis in MYC-driven B-cell lymphomas (Mu *et al.*, 2009). Interestingly, *miR-17-92a* is a direct transcriptional target of C-MYC and both C-MYC and the *miR-17-92a* cluster are expressed at high levels in ES cells, which are thought to share similar proliferative properties with cancer cells (Thomson *et al.*, 2004, Kim *et al.*, 2010a).

1.2.5.4 miRNAs in viruses

miRNAs are encoded within the genomes of some viruses. Viruses were first shown to express miRNAs in 2004, when five miRNAs were produced in human B cells after infection with the Epstein-Barr virus (EBV), a γ-herpesvirus (Pfeffer *et al.*, 2004). Since then, miRNAs have been discovered in all herpesviruses examined. However, other unrelated DNA viruses seem to contain only one or two miRNAs, or none at all, and viruses with an RNA genome have been reported to lack miRNAs altogether (Cullen, 2009). Interestingly, there is evidence of viral miRNAs that are functional orthologs of eukaryotic miRNAs. For example, functional orthologs of *miR-155* exists in Karposi's sarcoma-associated herpesvirus (KSHV) (Gottwein *et al.*, 2007, Skalsky *et al.*, 2007) and in Marek's Disease Virus (MDV) (Zhao *et al.*, 2009) and are capable of regulating common targets with *miR-155*.

Currently, it is thought that viral miRNAs may serve two major functions. One function is the inhibition of cellular factors that are produced as a result of viral infection. For example, EBV miR-BART5 inhibits expression of PUMA, an antiviral factor that promotes cell apoptosis (Choy *et al.*, 2008). A second function is the inhibition of viral proteins, including key early regulatory proteins. This may be required for latency of viral infections. During latency, herpes simplex virus (HSV-1) produces high levels of miRNAs that down regulate expression of ICP0 and ICP4, two immediate-early transactivators that are thought to play key roles in productive HSV-1 infection (Umbach *et al.*, 2008). Viral miRNAs may be key for the oncogenic capabilities of some herpesviruses such as KSHV and MDV. Indeed, deletion of the six-miRNA cluster 1 from the genome of MDV abolished the oncogenicity of the virus (Zhao *et al.*, 2011), which appeared to be primarily due to loss of the MDV functional ortholog of *miR-155*, miR-M4.

1.2.5.5 miRNAs as therapeutics

Expression patterns of miRNAs associated with specific diseases, particularly human cancers, have led to the theory that miRNAs may be used as biomarkers for disease (Trang *et al.*, 2009). Additionally, manipulation of miRNA expression may present a useful therapeutic approach. The future of miRNA research into

human disease will undoubtedly reveal more conditions with a miRNA component and may lead to novel biomarkers and therapeutics for disease.

1.2.6 Other small RNA molecules

1.2.6.1 siRNAs

As well as miRNAs, other classes of small RNA molecules are also found in mammalian cells. These are the endogenous small interfering RNAs (siRNAs) and Piwi-interacting RNAs (piRNAs) (Watanabe *et al.*, 2006). siRNAs are derived from long double stranded RNA molecules or long hairpin molecules, which are processed by *Dicer* to generate short RNA duplexes of ~21nt in length. As for miRNAs, one strand of the siRNA associates with the RISC complex targeting it to mRNAs with sequence complementarity to the siRNA. However, in contrast to miRNAs, siRNAs bind with perfect sequence complementarity to their target mRNAs resulting in cleavage of the target transcript (Zamore *et al.*, 2000). The finding that introduction of double-stranded (dsRNA) complementary to endogenous mRNA transcripts resulted in specific and potent suppression of gene expression has resulted in its used as a valuable research tool (Fire *et al.*, 1998). Thus far, endogenous siRNAs have only been detected in abundance in mouse oocytes and ES cells (Tam *et al.*, 2008, Watanabe *et al.*, 2008, Babiarz *et al.*, 2008). In mammals, these small RNAs are often encoded within pseudogenes and proposed functions are the regulation of the source-genes of the pseudogenes as well as regulation of retrotransposon expression (Tam *et al.*, 2008, Watanabe *et al.*, 2008).

1.2.6.2 piRNAs

The piRNAs differ from both miRNAs and siRNAs in that they do not require *Dicer* for their processing. These short RNA molecules are between 25 and 32 nt in length and are expressed predominantly in the germline in mammals (Girard *et al.*, 2006). piRNAs are characterised by their interaction with PIWI proteins, which are a distinct family of AGO proteins, although the mechanisms of piRNA biogenesis and function are poorly understood (Klattenhoff & Theurkauf, 2008).

Studies in mice implicate these small RNA molecules in germline development, silencing of transposons and maintenance of germline DNA integrity in mammals (Kuramochi-Miyagawa *et al.*, 2004, Carmell *et al.*, 2007).

1.2.7 Summary

While much is now known about the biogenesis, regulation and function of miRNAs there are still key questions to be answered. There is still disagreement, for example, about whether miRNA destabilisation or translational suppression is the major mode of miRNA function. The significance and functional details of non-canonical pathways of miRNA biogenesis need to be determined, as does the emerging concept of regulation of expression of miRNAs themselves. Additionally, we are far from elucidating the biological functions of the large number of individual miRNAs that have been discovered in the last decade.

1.3 Early mouse development

1.3.1 Fertilisation and onset of cleavage

Mammalian development begins when the meiotic cycle of the oocyte, which was suspended in metaphase of the second meiotic division, is reactivated by calcium release into the oocyte as a result of sperm entry (Swann, 1990). This occurs in the oviduct as the ovulated oocyte travels from the ovary to the uterus. Approximately 24 hours after completing meiosis the first cleavage event occurs. At the 2-cell stage the mouse embryo begins zygotic transcription, and this produces proteins that drive future cleavage events and further development (Piko & Clegg, 1982). The first cleavage event is meridional, but then rotational cleavage occurs; this means that during the second cleavage one blastomere divides meridionally, and the other divides equatorially (Gulyas, 1975). In contrast to some other embryos, the cells in mammalian embryos do not all divide synchronously so embryos frequently contain odd numbers of cells.

1.3.2 The morula stage embryo

After the third cleavage event the blastomeres are loosely packed, but during the 8-cell stage compaction occurs. Changes in the distribution of cell surface proteins such as E-cadherin cause the cells to form a tight ball, which is stabilised by tight junctions between the outside cells of the embryo (Fleming *et al.*, 2001). Cells within the embryo contain gap junctions allowing the passage of small molecules. Formation of tight junctions results in deformation of the round shape of the blastomeres and formation of the morula embryo. At this stage the first obvious differentiation event begins to occur. Most of the external cells of the morula will produce trophoblast lineages, while the majority of the internal cells will form the inner cell mass (ICM). It is the inner cell mass from which ES cells are derived. By the 64-cell stage the trophectoderm and ICM are completely distinct, and neither are capable of contributing cells to the alternative layer (Fleming, 1987).

1.3.3 Trophectoderm differentiation

1.3.3.1 Reciprocal Inhibition of *Oct4* and *Cdx2*

Cells of the inner cell mass express the transcription factor octamer-binding protein 4 (*Oct4*) and cells of the trophectoderm express the homeobox transcription factor caudal type homeobox 2 (*Cdx2*) (*Figure 1.6*). *Cdx2* is thought to be a key regulator of trophectoderm differentiation, as a loss-of-function *Cdx2* mutation results in defects in trophoblast differentiation (Strumpf *et al.*, 2005). In studies using mouse ES cells it was found that the TE is specified as a result of reciprocal inhibition between *Oct4* and *Cdx2* (Niwa *et al.*, 2005). This produces an *Oct4*-positive, *Cdx2*-negative compartment and an *Oct4*-negative, *Cdx2*-positive compartment in the early embryo that will become ICM and TE respectively.

1.3.3.2 *Tead4* in trophectoderm specification

The *Oct4*-*Cdx2* reciprocal-inhibition model requires a slight asymmetry to be introduced to the system. Recently, it has been shown that this asymmetry is introduced to the system via activation of *Tead4*, a transcriptional effector of

the Hippo-YAP signalling pathway, in the outer cells of the embryo. *Tead4*-null mutants arrest prior to expression of *Cdx2* and do not form a blastocoel (Nishioka *et al.*, 2008). This suggested that *Tead4* is upstream of *Cdx2* in the hierarchy of TE specifying genes. However, TEAD4 is ubiquitously expressed in the early embryo. Although TEAD4 is expressed in the inner and outer cells of the early embryo, its coactivator, YAP, is restricted to the nuclei of outer cells. In inner cells YAP is phosphorylated by LATS, a component of the Hippo signalling pathway, which results in the cytoplasmic localisation of YAP (Nishioka *et al.*, 2009). YAP is found to be cytoplasmic only in cells with a high degree of cell-cell contact that are internal to cell aggregates. This model provides evidence for position-dependent TEAD4 activity and cell fate specification. *in vitro*, ES cells induced to missexpress TEAD4 differentiate into TE lineages and can be induced to form TS-like cells when grown in TS cell culture conditions (Nishioka *et al.*, 2009). After the TE has been specified, the TE cells secrete fluid into the internal portion of the embryo. This creates the blastocoel cavity, and the embryo is now described as a blastocyst.

1.3.4 Differentiation of the primitive endoderm

The second lineage specification event, which occurs in the pre-implantation blastocyst embryo, is the segregation of the epiblast and the primitive endoderm. The epiblast gives rise to the embryo itself whereas the primitive endoderm gives rise to extra-embryonic structures, which support growth of the embryo and are involved in embryonic patterning. While specification of the trophectoderm and ICM is driven by cell position, the opposite seems to be true for specification of the primitive endoderm: cell fate predetermines cell position. Specification of the primitive endoderm is thought to begin at E3.5 with a 'salt and pepper' mosaic pattern of expression of the epiblast-specific transcription factor *Nanog*, and the essential master-regulator of primitive endoderm, *Gata6* (*Figure 1.6*). This means that cells within the late ICM express *Nanog* or *Gata6* exclusively (Chazaud *et al.*, 2006). Lineage tracing revealed that the descendants of individual E3.5 ICM cells contribute to epiblast or primitive endoderm lineages, but never both (Chazaud *et al.*, 2006). This showed that at E3.5 the fate of ICM cells has already been determined and led to a model of epiblast/primitive endoderm formation in which an initial mosaic of lineage progenitors at E3.5 are sorted

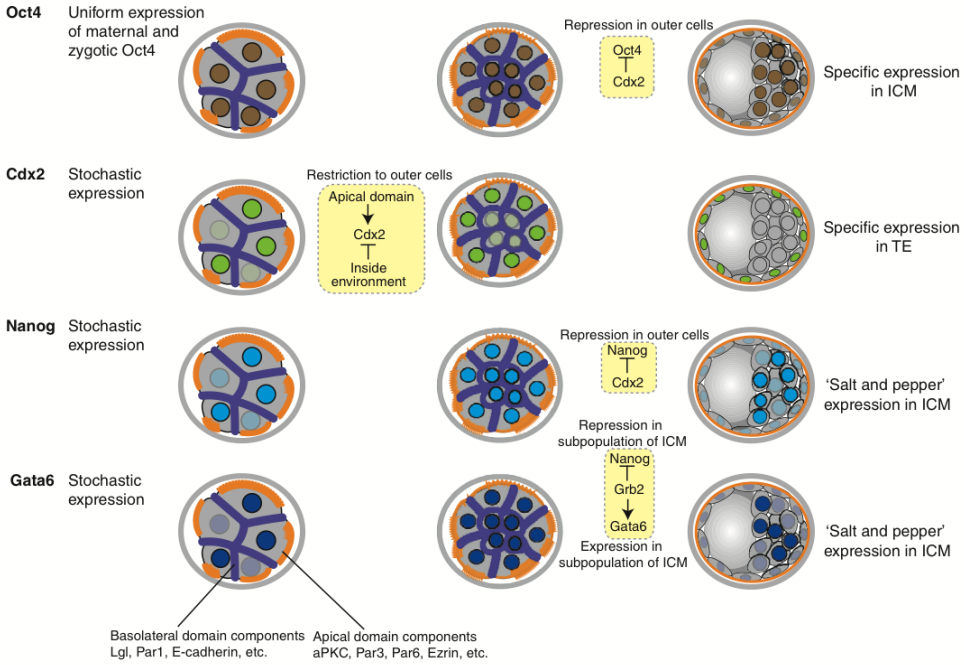


Figure 1.6 – *Molecular players in the formation of the first lineages in the mouse embryo.* Four lineage-specific transcription factors, *Oct4*, *Cdx2*, *Nanog* and *Gata6*, are important for specification of the first three lineages in the mouse embryo: the trophectoderm (TE), the primitive endoderm and the epiblast. *Oct4* is expressed in all blastomeres throughout the early cleavage stages. At the blastocyst stage, *Oct4* is downregulated in the outer TE cells through inhibition by *Cdx2*. The expression of *Cdx2* is stochastic during cleavage and becomes restricted to outer cells by the blastocyst stage through a position-dependent mechanism. *Nanog* and *Gata6* are expressed in all cells until the early blastocyst stage. *Nanog* expression is then restricted to a subpopulation of ICM cells through inhibition by *Cdx2* and Grb2-dependent signaling, while *Gata6* expression is maintained by Grb2-dependent signaling. By the late blastocyst stage cells of the ICM will express either *Nanog* or *Gata6*. Through a process of cell sorting and relocation, cells positive for *Gata6* or *Nanog* will be located to appropriate locations within the ICM and will generate primitive endoderm and epiblast, respectively. Figure from Rossant & Tam (2009).

and relocated to appropriate locations within the ICM by E4.5 (Rossant and Tam, 2009). Active signalling through a GRB2 (growth receptor bound protein 2)-dependent mechanism is important for primitive endoderm formation because in the absence of GRB2, no primitive endoderm forms and all cells of the ICM express *Nanog* (Chazaud *et al.*, 2006).

1.3.5 Implantation of the embryo

As the embryo moves through the oviduct the surrounding zona pellucida prevents it from adhering to the oviduct walls; however, on reaching the uterus, the embryo must 'hatch' from the zona pellucida to allow implantation in the uterus wall. A small hole is lysed in the zona and the mouse embryo is extruded and is able to implant into the uterine wall. This occurs at approximately E4.5. Implantation occurs as a result of apposition of the mural trophectoderm to the uterine epithelium. The decidual response occurs in the area immediately surrounding the point of embryo attachment and leads to formation of the maternal component of the placenta, which is composed of maternal vasculature and uterine decidual cells. The embryonic component of the placenta is comprised of layers of differentiated trophoblast cells (extraembryonic ectoderm of the chorion) and fetal vascular system of mesoderm (allantois) origin (Rinkenberger *et al.*, 1997, Rossant and Cross, 2001).

1.3.6 Gastrulation

The post-implantation blastocyst is composed of trophectoderm, primitive endoderm and epiblast lineages. The epiblast contains all the cells that will go on to form the components of the embryo proper. The three primary germlayers –endoderm, ectoderm and mesoderm – are produced as a result of gastrulation. Gastrulation begins at the embryonic/extraembryonic junction with formation of the primitive streak, which progresses distally through the embryo. Although the signals that initiate gastrulation at a specific point are not known, the secreted factors, *Nodal* and *Wnt3* are essential for this process (Conlon *et al.*, 1994, Liu *et al.*, 1999). The most anterior portion of the primitive streak becomes the node. Cells migrate individually through the primitive streak and are patterned according to differential gene expression along the length of the streak. Mesoderm

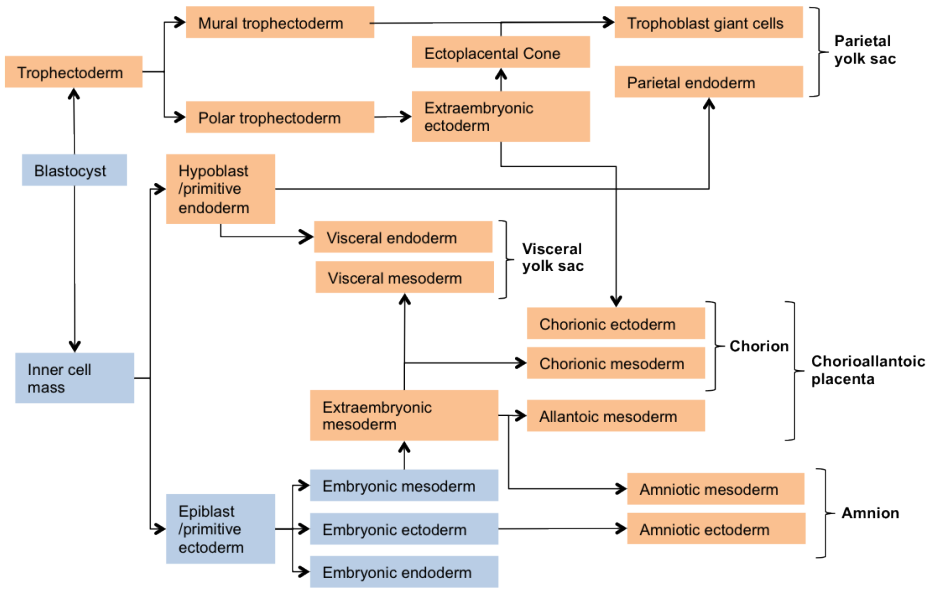


Figure 1.7 – *Lineages of tissues constituting the mouse embryo.* Blue boxes – all tissues that will give rise to the embryo proper and extraembryonic cells. Orange boxes – extraembryonic tissues.

and endoderm cells migrate through the primitive streak whereas cells migrating through the node will become notochord. Expression of the T-box transcription factor, *Brachyury*, is essential for normal morphogenetic mesoderm movements during gastrulation (Wilson *et al.*, 1995). Fate maps have been generated for the gastrulating embryo by labelling individual cells of the epiblast (Lawson *et al.*, 1991). *Figure 1.7* shows the origins of the early tissue lineages of the mouse.

1.4 Mouse embryonic stem cells

Embryonic stem (ES) cells are pluripotent stem cell lines that are derived from the inner cell mass (ICM) of blastocyst stage embryos. They have two remarkable properties: the ability to self-renew indefinitely, and the ability to produce cells belonging to the three major tissue germlayers: endoderm, mesoderm and ectoderm (*Figure 1.8*). This is demonstrated by the generation of chimeric mice; ES cells will integrate fully into the developing embryo, and can contribute to all tissues including the production of functional gametes. ES cells are also amenable to genome engineering; this property can be used in combination with their contribution to the germline to introduce mutations into lines of mice.

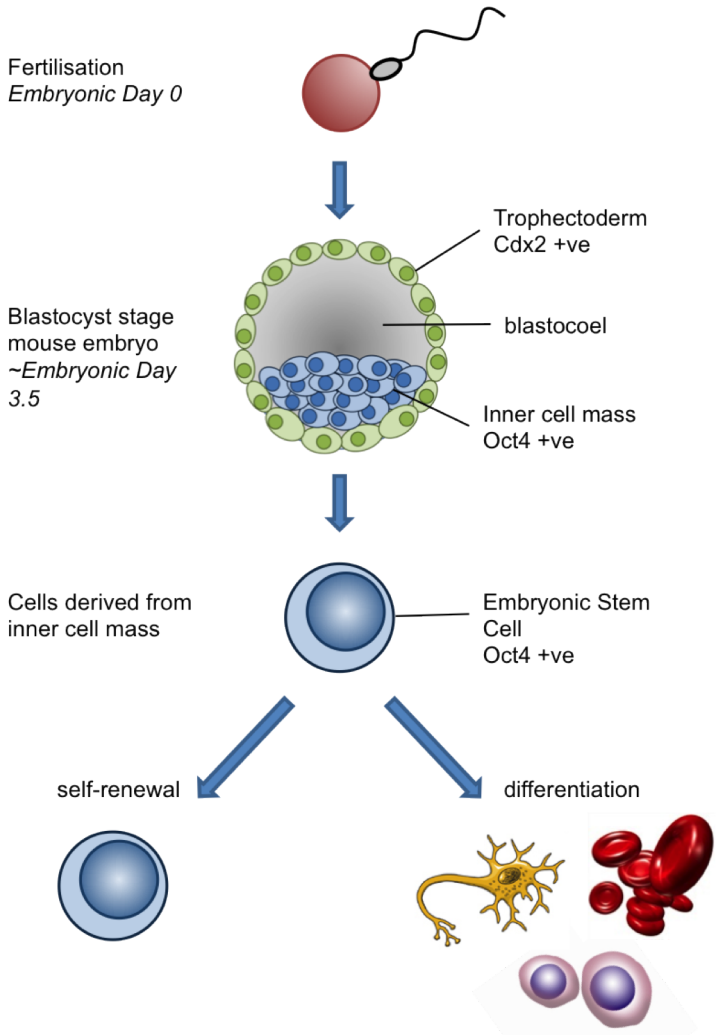


Figure 1.8 – *Mouse embryonic stem cells.* Mouse ES cells are derived from the inner cell mass of blastocyst stage embryos. If grown in the appropriate conditions ES cells will self-renew indefinitely. They also have the ability to differentiate into cells belonging to the three germ lineages *in vivo* and *in vitro*.

The incredible properties of ES cells have resulted in a large research effort to try to understand the mechanisms governing the pluripotency and differentiation of these cells. ES cells not only have the potential to provide study systems for development and disease, but may one day lead to therapeutic treatments for degenerative diseases such as Alzheimer's and Parkinson's disease. ES cells can be induced to differentiate *in vitro*, mimicking early developmental transitions such as trophectoderm differentiation, primitive endoderm differentiation and germ layer differentiation. In this way ES cells provide excellent model systems for the study of early developmental events *in vitro*.

1.4.1 Origins and culture of mouse ES cells

1.4.1.1 Embryonal carcinoma cells

The discovery and maintenance of ES cells was made possible by prior work on teratocarcinomas and the embryonal carcinoma (EC) cells derived from teratomas. In 1970 it was found that by grafting mouse epiblast cells onto a permissive host site a tumour consisting of many different cell types could be created (Solter *et al.*, 1970). These teratocarcinomas contained cells belonging to each tissue germlayer, and an undifferentiated component that could be maintained by serial transplantation. Cells derived from this undifferentiated component were termed EC cells (Evans, 1972). While these cells could contribute to embryogenesis and form chimaeric embryos, their genetic constitution was compromised and most EC cells show restricted developmental potential (Martin, 1980). However, study of EC cells laid the foundations for the study of ES cells, which were first reported in 1981 (Evans & Kaufman, 1981; Martin, 1981). A key finding was that co-culture of EC cells with mitotically-inactivated fibroblast cells led to efficient establishment of EC cultures with a high differentiative capacity (Martin & Evans, 1975).

1.4.1.2 Feeder-dependent ES cell culture

The basic protocol for ES cell derivation is now relatively straightforward (Robertson, 1987): Blastocyst stage embryos (either intact or after immunosurgical isolation of the ICM) are placed onto a layer of mitotically inactive feeder cells in

culture. After several days of culture, epiblast outgrowths become apparent and can be disaggregated and transferred to fresh cultures containing feeder cells. Colonies develop and are assessed, and undifferentiated ones are dissociated and replated. If secondary undifferentiated colonies arise, then these can generally be expanded and cultured as ES cells. It was reasoned that the fibroblast cells were producing some critical factor for ES cell maintenance, hence the term 'feeder cells' (*Figure 1.9a*). ES cells can be neither derived nor maintained in the presence of media supplemented with serum alone.

1.4.1.3 Culture of ES cells in serum and Lif

The requirement for a feeder layer was abrogated through culture of ES cells in conditioned media, confirming that the fibroblast cells were producing a trophic factor that supported self-renewal (Smith & Hooper, 1983; Smith & Hooper, 1987). Further investigation showed that addition of the cytokine, leukaemia inhibitory factor (LIF), was sufficient to sustain ES cell self-renewal in the absence of a feeder later (Smith *et al.*, 1988, Williams *et al.*, 1988) (*Figure 1.9b*). Binding of Lif to its receptor initiates heterodimerisation of the Lif receptor with the gp130 receptor (Yoshida *et al.*, 1994). Downstream of this receptor two signalling cascades are initiated; activation of STAT3 (signal transducer and activator of transcription 3) by JAK (janus kinase) (Lutticken *et al.*, 1994), and stimulation of the RAS-ERK mitogen-activated protein kinase cascade (Takahashi-Tezuka *et al.*, 1998).

The role of STAT3 STAT3 activation is thought to be essential for the self-renewal of ES cells grown in serum + Lif conditions because ES cells expressing a dominant interfering mutant of STAT3 could not be isolated (Niwa *et al.*, 1998). Furthermore, inducing expression of the dominant interfering form of STAT3 in ES cells abrogated self-renewal and promoted differentiation. Lif/STAT3 are thought to control ES cell self-renewal and pluripotency in part through a MYC-dependent mechanism. Mouse ES cells express high levels of the MYC transcription factor and constitutive expression of stable MYC results in Lif independent self-renewal and pluripotency (Cartwright *et al.*, 2005).

1.4.1.4 Serum-free culture of ES cells

Lif and BMP A conclusion that STAT3 activation is the only requirement for ES cell propagation is premature, because the addition of serum to culture medium may be fulfilling other requirements. Therefore a key goal was to remove serum from ES cell culture in order to grow the cells in completely defined conditions. ES cells can be grown in medium containing serum substitute supplemented with Lif but only at high densities. Colony formation from single cells is not possible in these conditions (Ying *et al.*, 2003). In the absence of serum, neural differentiation is only partly inhibited by the presence of Lif and self-renewal is attenuated. A breakthrough came when it was shown that addition of the anti-neural factor BMP4 in combination with Lif was sufficient to support ES cell self-renewal in the serum-free media N2B27 (Ying *et al.*, 2003) (*Figure 1.9c*). The BMP4 pathway signals through SMAD 1, 5 and 8 induce expression of ID proteins, which suppress differentiation. Constitutive expression of ID1, 2 or 3 is sufficient to replace the requirement for BMP or serum (Ying *et al.*, 2003).

Culture of ES cells in the presence of small molecule inhibitors More recently, it has been shown that ES cells do not require extrinsic factors for self-renewal or maintenance of their pluripotent state (Ying *et al.*, 2008) (*Figure 1.9d*). Small molecule inhibition of the ERK signalling cascade and GSK3B allowed propagation of mouse ES cells with a similar proliferation rate to cells grown in serum with Lif. Expression of pluripotency markers was maintained and colonies could be grown from single cells. ES cells could be derived in the inhibitor medium and, crucially, contributed to tissues from all three germ layers in chimaeric mice as well as showing germline transmission of their genome.

The inhibitor medium works through addition of small molecule inhibitors of Fgf receptor tyrosine kinases (SU5402) and mitogen activated protein/extracellular signal-regulated kinase (ERK) kinase (MEK) (PD184352), which together inhibit pro-differentiation signalling via mitogen-activated protein kinase. These two inhibitors can be replaced by a single, more potent, MEK inhibitor (PD0325901). Additional inhibition of GSK3B by CHIR99021 is thought to improve growth and viability of the ES cell cultures. The fact that ES cells can maintain their pluripotency and ability to self-renew in the absence of extrinsic stimuli suggests

that they may represent a primitive 'ground state', which is self-maintaining if shielded from differentiation stimuli.

1.4.2 Pluripotency of mouse ES cells

1.4.2.1 Intrinsic determinants

Oct4 The embryonic stem cell state is maintained by expression of a network of key transcription factors with *Oct4*, *Sox2* and *Nanog* at its core. *Oct4* (octamer-binding transcription factor 4), also known as POU5f1 (POU domain, class 5, transcription factor 1), is a homeodomain transcription factor, which is expressed in pluripotent cells of the early embryo, germ cells and embryonic stem cells (Okamoto *et al.*, 1990, Rosner *et al.*, 1990, Scholer *et al.*, 1990). *Oct4* is required for maintenance of the pluripotent cell population in the early embryo and its absence is thought to cause differentiation along the trophoblast lineage (Nichols *et al.*, 1998).

Sox2 The SRY-related HMG-box (SOX) family of transcription factors are involved in the regulation of embryonic development and cell fate and have a single HMG DNA-binding domain. *Sox2* is developmentally regulated (Wiebe *et al.*, 2000), expressed in the pluripotent lineage of the early mouse embryo and its down-regulation correlates with a commitment to differentiate (Avilion *et al.*, 2003). If *Sox2* is removed from ES cells, as in the *Sox2*-inducible null ES cell line generated by Masui *et al.*, the cells differentiate, and the majority become trophectoderm-like cells as assessed by expression of trophectoderm marker genes (Masui *et al.*, 2007).

Nanog *Nanog* was added to the core transcriptional network of ES cells in 2003 following its discovery by Chambers *et al.* (2003) and Mitsui *et al.* (2003). The homeobox transcription factor is expressed in the inner cells of a compacted morula and blastocyst as well as in early germ cells and ES cells. *Nanog* is key for early cell fate specification because its ablation causes early pluripotent cells to adopt a differentiated parietal/visceral endoderm fate (Mitsui *et al.*, 2003). Additionally, *Nanog* overexpression allows self-renewal of ES cells independent

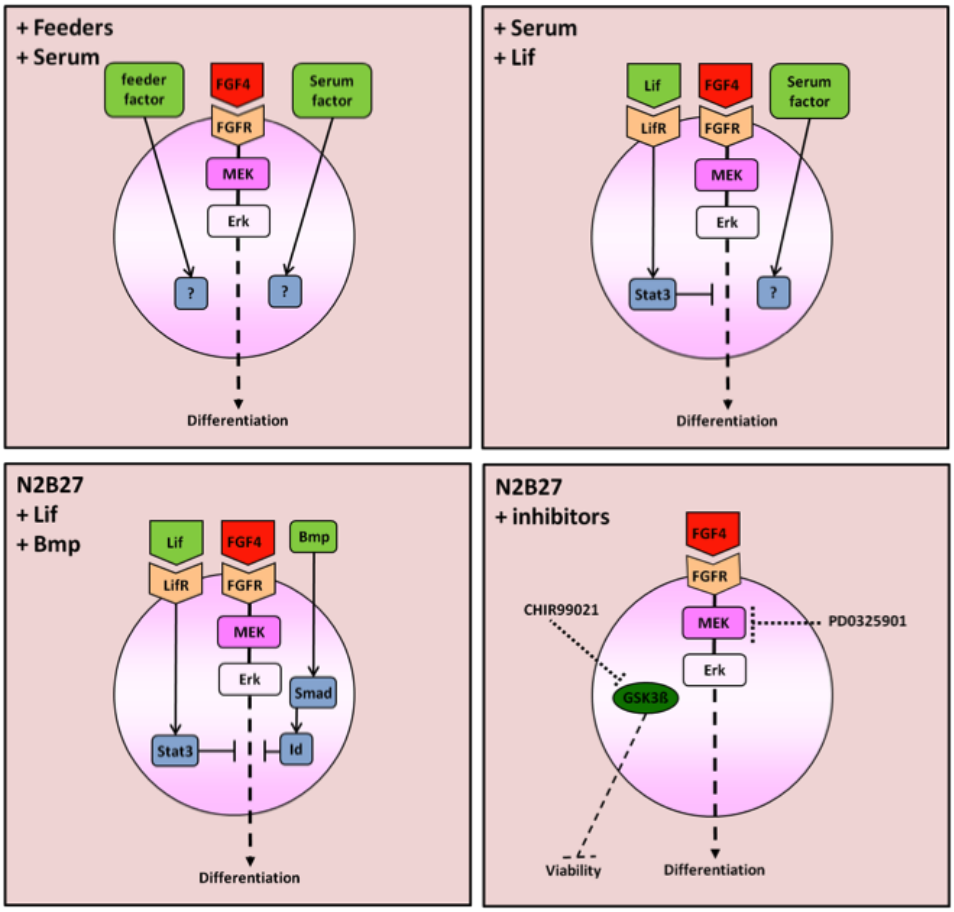


Figure 1.9 – *Summary of ES cell culture conditions.* Mouse ES cells were initially grown on a layer of mouse embryonic fibroblast cells in the presence of serum (+ Feeders + Serum) and maintained ES cells in a undifferentiated state through the presence of undefined factors from serum and produced by feeder cells. The cytokine Lif can substitute for the presence of feeder layer through activation of STAT3, which is thought to inhibit pro-differentiation signalling downstream of Erk (+ Serum + Lif). ES cells can be grown in serum-free conditions by addition of the Lif cytokine and bone morphogenic proteins (Bmps), which activate the Smad pathway and induce expression of Inhibitor of Differentiation (Id) proteins (N2B27 +Lif + Bmp). ES cells can be grown in the absence of extrinsic factors through growth in serum-free conditions with the addition of small molecule inhibitors of the Erk signalling pathway and of GSK3B (N2B27 + inhibitors).

of Lif stimulation of STAT3 (Mitsui *et al.*, 2003 & Chambers *et al.*, 2003) and reduces the differentiation potential of the ES cells (Chambers *et al.*, 2003).

1.4.2.2 Transcriptional regulation by pluripotency factors

ChIp analysis has shown that *Oct4*, *Sox2* and *Nanog* show extensive co-binding of the same regions of the mouse genome (Chen *et al.*, 2008, Marson *et al.*, 2008). In addition to these three core pluripotency factors, are multiple other transcription factors that are important in maintaining ES cell identity, including SMAD1, STAT3 and TCF3, the downstream effectors of signalling pathways triggered by BMP, Lif and Wnt, respectively. SMAD1, STAT3 and TCF3 were found to co-occupy promoter regions with *Oct4* and *Nanog*, linking extracellular signalling with the core transcriptional network of ES cells (Chen *et al.*, 2008, Cole *et al.*, 2008).

The key pluripotency factors bind to the promoters of a large number of genes, many of which are transcription factors (Loh *et al.*, 2006). In this way they may initiate cascades of downstream regulatory events from upstream events. *Oct4* can also regulate expression of other target genes through its positive regulation of histone modifiers, including *Jmjd1a* and *Jmjd2c*, which prevent the accumulation of repressive methylation at the promoters of genes maintaining pluripotency (Loh *et al.*, 2007). Indeed the chromatin of mouse ES cells has hallmarks of highly active chromatin including a higher-order structure (Meshorer & Misteli, 2006), which may prove essential for maintenance of pluripotency.

1.4.3 Other pluripotent cell types

1.4.3.1 EpiSCs

The epiblast will go on to form all the cells of the embryo proper, and is a distinct entity from the inner cell mass, which forms the epiblast and primitive endoderm lineages. Cells derived from the epiblast are termed epiblast stem cells (EpiSCs) and require different culture conditions from ES cells, which are derived from the ICM (Brons *et al.*, 2007, Tesar *et al.*, 2007). These cells are 'primed' pluripotent stem cells because they retain the ability to form teratomas in nude mice but not

the ability to form chimaeric embryos. EpiSCs also differ from ES cells in that they express some genes normally expressed by the pre-gastrulation epiblast that are expressed at low levels in ES cells such as *Fgf5*, and lack expression of some ES cell associated genes such as *Rex1* (Brons *et al.*, 2007).

1.4.3.2 iPS cells

ES cells retain the potential to differentiate into any tissue of the three germ lineages if grown in the appropriate conditions. This property is lost as cells differentiate down a particular lineage. However, seminal work by Yamanaka's laboratory showed that it is possible to take differentiated cells and convert them back into a pluripotent state (Takahashi & Yamanaka, 2006). The so-called induced pluripotent stem (iPS) cells show similar characteristics to ES cells and have the ability to contribute to chimaeric animals and transmit their genome through the germline. Since the first iPS cells were generated from mice in 2006 many laboratories have used the technology to attempt generation of iPS cells from different species (Martins-Taylor & Xu, 2010). Different technical approaches are now available but the basic principle remains the same: A combination of defined factors, which may include *Oct4*, *Sox2*, C-MYC, *Klf4*, *Nanog* and *Lin28*, are introduced into somatic cells by retroviral or adenoviral transduction or transfection of expression vectors. Outgrowing colonies are picked, expanded and assessed for expression of pluripotency genes before being subjected to tests of their developmental potential.

The observation that differentiated cells can be reprogrammed to a pluripotent state is perhaps not surprising, given that a nucleus from a somatic cell can be reprogrammed by its nuclear transfer into an enucleated oocyte (Wilmut *et al.*, 1997) and that cell fusion of embryonic germ cells with somatic cells results in pluripotent cells that can contribute to the three germ layers in chimaeric embryos (Tada *et al.*, 1997). Nevertheless, transcription factor induced reprogramming is the best evidence so far of the plasticity of cells in response to changes in their transcriptional regulators. Furthermore, this method can be used to generate pluripotent cells from human somatic cells, which could be used to generate patient-specific cells for tissue repair or replacement thus avoiding immunological rejection and ethical issues associated with alternate sources of pluripotent cells.

1.4.3.3 Similarities between iPS cells and ES cells

There are numerous studies that claim that iPS cells and ES cells are morphologically, functionally and molecularly equivalent but also a substantial number of studies claiming molecular differences between both mouse and human iPS cells and ES cells (reviewed in Plath & Lowry, 2011). Notably, two groups have shown that the DNA methylation pattern of the original cell persists in iPS cells and that this affects their ability to differentiate down particular lineages (Polo *et al.*, 2010, Kim *et al.*, 2010c). Therefore, although the potential of iPS cells as sources of clinically relevant populations of differentiated cells is unparalleled, current problems with reprogramming mean that ES cell research is equally as relevant with the advent of iPS cells as it was before.

1.4.4 Differentiation of mouse ES cells

1.4.4.1 *in vitro* differentiation

The potential of ES cells to form any tissue from the three germ lineages has tremendous potential for the study of development and disease, testing of therapeutics and generation of clinically relevant populations of cells. ES cell differentiation into multiple cell types has been demonstrated *in vitro* through the formation of three-dimensional aggregates called embryoid bodies, through monolayer culture on extracellular matrix proteins and through culture of ES cells on supportive stromal layers (Murry & Keller, 2008). However, although multiple cell types may terminally differentiate from ES cells, it is believed that ES cells only have the capacity to directly differentiate into three cell types: the primitive ectoderm, the primitive endoderm and the trophectoderm (Niwa H., 2010).

Many methods of differentiating ES cells involve the supplementation of culture media with fetal calf serum, which is a poorly defined combination of factors with batch-to-batch variation. This had led to problems with reproduction of differentiation protocols but advances including the use of serum-free media with specific inducers are beginning to address this issue (Murry & Keller, 2008).

1.4.4.2 Heterogeneity of ES cell cultures

Even in defined differentiation conditions, however, not all ES cells will differentiate down the lineage of choice (Lowell *et al.*, 2006). While the reason behind this is not completely clear, it is known that individual ES cells within a culture show distinct expression patterns of a small number of genes. *Nanog* expression is heterogeneous in ES cells (Singh *et al.*, 2007) as is expression of *Gata6*, *Rex1* and *Stella* (Singh *et al.*, 2007, Toyooka *et al.*, 2008, Hayashi *et al.*, 2008b). This heterogeneity does not represent distinct stable cell populations because isolated positive and negative populations will regain heterogeneous expression after a period of culture. However, positive or negative populations show different differentiation potential, suggesting that ES cells can oscillate between states predisposed to self-renewal (naïve) or differentiation into a particular lineage (primed).

1.4.4.3 Signalling pathways controlling ES cell differentiation

It has already been mentioned that suppression of the ERK signalling pathway (in combination with GSK3B inhibition) permits self-renewal of ES cells in the absence of extrinsic stimuli. This is thought to be because signalling through FGF/ERK initiates differentiation of ES cells (Kunath *et al.*, 2007, Stavridis *et al.*, 2007). In this way ES cells can be said to inherently drive differentiation, because they produce FGF4, which activates autocrine ERK signalling. Traditional culture of ES cells in Lif and serum/BMP4 is thought to counterbalance the pro-differentiative autocrine FGF4/ERK signal by inhibiting differentiation downstream of ERK signalling (Ying *et al.*, 2008).

1.4.4.4 A new theory of differentiation

Traditionally, it has been thought that ES cell pluripotency genes must be down regulated in order for an ES cell to differentiate. Recently, a new theory of pluripotency has been proposed. A review by Loh and Lim (2011) introduced the concept that rather than existing as a stable 'ground' state, the pluripotent state is a result of competition between different pluripotency factors that each have lineage-specifying actions. This theory is supported by a paper from Ramanathan's laboratory, which claims that *Oct4* and *Sox2* orchestrate germ layer fate selection

with *Oct4* supressing neuroectodermal differentiation and promoting mesendodermal differentiation, and *Sox2* inhibiting mesendodermal differentiation while promoting neuroectodermal differentiation (Thomson *et al.*, 2011). The concept of pluripotency as an unstable battlefield rather than a stable self-maintaining state needs further investigation but may help to explain how pluripotency is maintained in the early embryo in the presence of pro-differentiation signalling.

1.4.4.5 The relationship between mouse ES cells and the early embryo

Cells of the early embryo express similar pluripotency markers to ES cells in culture and show heterogeneous expression of *Nanog* and *Gata6*, which is also seen in ES cells (Singh *et al.*, 2007). One major anomaly between traditional mouse ES cell culture and the early embryo, is that while activation of STAT3 (induced through Lif signalling) is required for maintenance of ES cells in culture, lif null mice develop normally (Stewart *et al.*, 1992) and STAT3 null embryos do not arrest until after implantation having formed normal preimplantation blastocysts (Takeda *et al.*, 1997). However, although gp130 null embryos can proceed through early embryogenesis, they cannot maintain a pluripotent cell population during diapause (Nichols *et al.*, 2001). This suggests that activation of STAT3 is required for *in vivo* maintenance of the pluripotent cell population during developmental arrest and provides a physiological rationale for the responsiveness of mouse ES cells to signalling via gp130 cytokines.

Early mouse embryos can be cultured in the presence of the MEK inhibitor, PD0325901, and behave in a similar manner to cultured cells grown in these conditions, with homogeneous expression of *Nanog* and loss of paternal X chromosome inactivation in XX embryos (Nichols *et al.*, 2009). Cells of the embryo grown in the presence of the MEK inhibitor maintain pluripotency as assessed by formation of germline-competent chimaeras and are amenable to efficient derivation of ES cell clones from single cells. The authors concluded that mouse ES cells are in the same state as naïve epiblast and intrinsic self-replication is not an adaptation to culture.

1.4.5 Mouse ES cells as models of early differentiation and development

Mouse ES cells are amenable to genetic manipulation and can therefore be used as a tool to study gene function through generation of knockout mouse lines. In addition to this, their similarity with cells of the early mouse embryo and ability to differentiate in culture mean that ES cells are an excellent model system for the study of molecular mechanisms underpinning cell fate transitions. Below is an account of *in vitro* differentiation systems for ES cells that model early developmental transitions.

1.4.5.1 Trophectoderm differentiation from ES cells

As ES cells are derived from the ICM and do not contribute efficiently to trophectodermal lineages following injection into blastocysts (Beddington & Robertson, 1989), it was originally considered that ES cells could not generate trophectoderm directly. However, it is now known that ES cells can be induced to form trophoblast stem (TS) cells in culture through manipulation of specific transcription factors (Niwa, 2010) including the ES cell master regulator *Oct4* (Niwa *et al.*, 2000) or the homeobox gene *Cdx2* (Niwa *et al.*, 2005, Hay *et al.*, 2004).

Trophectoderm differentiation can be induced in the ZHBTc4 ES cell system (Niwa *et al.*, 2000) by the rapid downregulation of *Oct4* expression, described in section 3.1.1. Upon rapid *Oct4* downregulation, several trophoblast-associated genes are upregulated, including the transcription factors *Cdx2* and *Eomes* (Niwa *et al.*, 2000). Missexpression in ES cells of both *Cdx2* and *Eomes* induces the expression of other trophoblast-associated genes such as placental cadherin (*Cdh3*) (Niwa *et al.*, 2005). In particular, missexpression of *Cdx2* leads to TE differentiation, where TS cells can be derived that will contribute to the placental lineage *in vivo*. In summary, although the ICM and TE differentiate prior to the point of ES cells isolation from the ICM, ES cells retain the ability to differentiate into trophoblast lineages and can be induced to do so by manipulation of key transcription factors.

1.4.5.2 Primitive endoderm differentiation

Differentiation into primitive endoderm can be initiated in 3D aggregates of ES cells called embryoid bodies (EBs) where it forms as a layer on the EB surface. EBs loosely mimic early embryonic development and the similarities are clear when comparing *Gata6* knockout embryos, which do not form primitive endoderm, and EBs generated from *Gata6* knockout ES cells, which mimic this phenotype (Capo-Chichi *et al.*, 2005, Cai *et al.*, 2008). Primitive endoderm cells from embryos can be maintained *in vitro* as extraembryonic endoderm (XEN) cells (Kunath *et al.*, 2005) and XEN-like cells can be induced directly from ES cells by forced expression of *Gata4* or *Gata6* (Fujikura *et al.*, 2002).

1.4.5.3 Primitive ectoderm differentiation

Primitive ectoderm can be formed from ES cells in EB culture, in which cells in the interior of the structure form pluripotent epithelium equivalent to primitive ectoderm and express the primitive ectoderm marker *Fgf5* (Haub & Goldfarb, 1991, Hebert *et al.*, 1991). Monolayer culture of ES cells in the absence of Lif also leads to *Fgf5* upregulation, suggesting that a proportion of these cells acquire a primitive ectoderm like state. Crucially, upregulation of *Fgf5* expression always occurs prior to differentiation of various embryonic lineages but not extraembryonic lineages. This means than in order for ES cells to differentiate into embryonic lineages it is thought that they must first go through a primitive ectoderm-like state, recapitulating what is seen in the early embryo (Niwa, 2010, Shimozaki *et al.*, 2003).

1.4.5.4 Germ layer differentiation

In the early embryo, the three germ layers are produced from primitive ectoderm during the process of gastrulation. Although it has been known for a long time that cells belonging to the three germ lineages can be formed during EB differentiation (Doetschman *et al.*, 1985), it was only shown recently that a gastrulation-like process can occur in EBs (ten Berge *et al.*, 2008). Embryoid bodies were

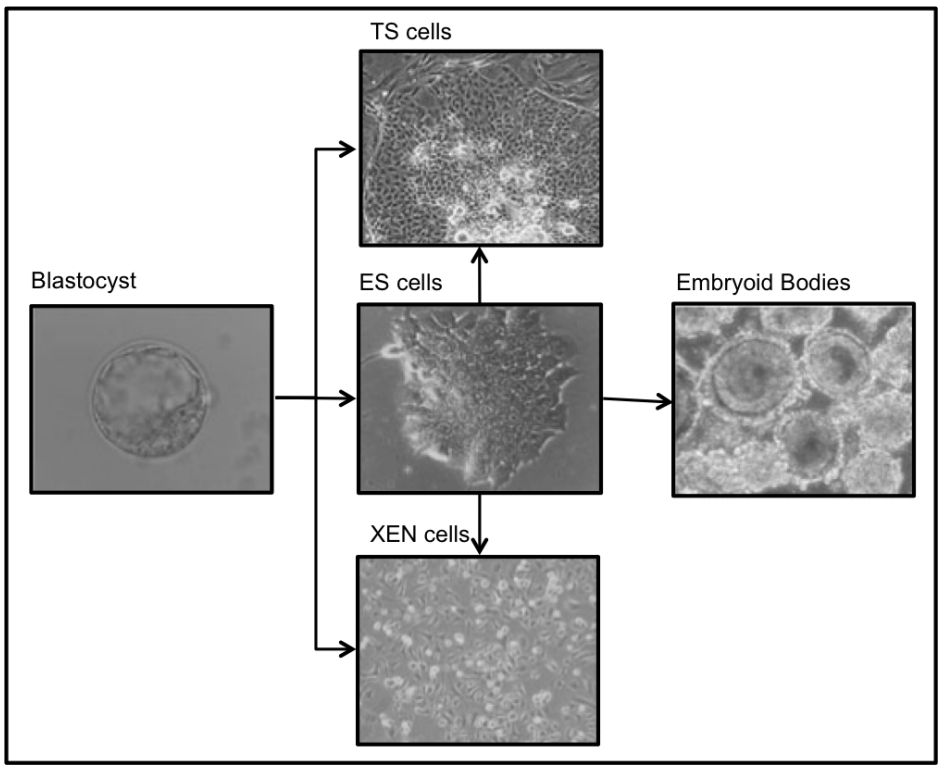


Figure 1.10 – *ES cells as models of developmental transitions.* Three cell types can be derived from late blastocyst embryos and propagated *in vitro*: trophoblast stem (TS) cells, embryonic stem (ES) cells and extra-embryonic endoderm (XEN) cells. ES cells can be induced to differentiate into TS cells or XEN cells *in vitro* by genetic manipulation. They can also form embryoid bodies, which generate cells belonging to the three germ lineages. Figure adapted from Niwa (2010).

shown to exhibit anterior-posterior polarity and formation of a primitive streak-like region, which was dependent on local activation of the Wnt signalling pathway (ten Berge *et al.*, 2008).

1.4.5.5 ES cells provide an ideal system to model early developmental transitions

The study of early mammalian developmental transitions is hampered by the fact that development occurs *in utero*. However, the derivation and culture of embryonic cells *in vitro* provides an ideal opportunity to study the factors involved in developmental processes. As detailed above, cells representing the three earliest mouse lineages, the TE, the ICM and the extraembryonic endoderm, can be

maintained in culture, retain properties of the tissue of origin, and can be reintroduced to an early embryo and contribute solely to the tissue of origin. ES cells are similar to cells of the ICM, retain pluripotency in culture and can be induced to differentiate, mimicking early developmental processes *in vitro*. ES cells can be induced to differentiate into cell types representative of the TE and extraembryonic endoderm (*Figure 1.10*), and differentiated cells will contribute exclusively to these lineages upon introduction to an embryo (Shimosato *et al.*, 2007, Niwa *et al.*, 2005). Additionally, ES cells can be induced to form embryoid bodies (*Figure 1.10*), which recapitulate aspects of embryonic development *in vitro*. Together, these findings show that ES cells provide an ideal model system for the study of factors involved in the regulation of early developmental decisions.

1.4.6 Summary

While much is now understood about the transcription factor network and signalling pathways underpinning the pluripotent state, the details of how this state is maintained are still to be confirmed. However, it has been shown that ES cells share similar properties with pluripotent cells of the early embryo and can be used as excellent model systems of early developmental events.

1.5 miRNAs in developmental transitions

miRNAs were first discovered in *C. elegans*, as regulators of developmental transitions. Since then these small molecules have been shown to have roles in many physiological processes, including essential functions for the correct development of multiple species (Wienholds *et al.*, 2003, Lee *et al.*, 1993, Reinhart *et al.*, 2000, Brennecke *et al.*, 2003, Bernstein *et al.*, 2003, Wang *et al.*, 2007). This section will introduce the modes of regulation by which miRNAs may operate during development, as well as what is currently known about the roles of miRNAs in early mammalian development and ES cells.

1.5.1 miRNAs regulate developmental processes

miRNAs act through several modes of regulation during development including developmental switches, adding robustness to developmental programs and regulating proliferation and apoptosis. Some examples of the role of miRNAs during development are mentioned below.

1.5.1.1 miRNAs as developmental switches

In the nematode worm *C. elegans* the first discovered miRNA, *lin-4*, is an example of a miRNA acting as a developmental switch. If *lin-4* is lost *C. elegans* cannot undergo the transition from the first to the second larval stage (Lee *et al.*, 1993). Cell lineage patterns normally specific for the first larval stage are reiterated at later larval stages, with the animals going through additional larval moults. This phenotype is due to the loss of repression of *lin-14*, a direct target of *lin-4*, which is normally down regulated during the first larval stage.

miRNAs are also required for the developmental switch from neurogenesis to gliogenesis in the developing spinal cord of the mouse (Zheng *et al.*, 2010). Neural and glial cells are produced sequentially from ventricular neural progenitor cells in the developing CNS and a key question is how the switch from neurogenesis to gliogenesis is controlled. Zheng *et al.* generated Olig1Cre-mediated conditional *Dicer*$^{-/-}$ mice, which were deficient for miRNA processing in the ventral neuroepithelium. Astrogliogenesis and oligodendrogenesis was severely disrupted in the conditional *Dicer*$^{-/-}$ mice but neural patterning and motor neuron development was not affected. This suggests that miRNAs are required for the developmental switch from neurogenesis to gliogenesis.

1.5.1.2 miRNAs add robustness to developmental programs

It is thought that miRNAs are predominantly involved in increasing the robustness of developmental decisions, rather than directly regulating them. For example the *miR-430* family are required in the zebrafish embryo during the process of activation of zygotic transcription. In these embryos the *miR-430* family are expressed at high levels coincident with the start of zygotic gene expression. If

these embryos are depleted for miRNAs, brain defects result, but these defects can be rescued by the injection of *miR-430* (Giraldez *et al.*, 2005). Several hundred transcripts are regulated by *miR-430*, a large proportion of which are maternal transcripts (Giraldez *et al.*, 2006). This suggests that *miR-430* is required to accelerate the clearance of maternal transcripts at the start of zygotic transcription adding robustness to the material to zygotic transition.

1.5.1.3 miRNAs regulate proliferation and apoptosis

Other processes essential for correct development are also regulated by miRNAs including proliferation and apoptosis. An example of miRNA-regulation of apoptosis is seen in *Drosophila* embryos that lack the miRNA *bantam* are small and die as early pupae (Brennecke *et al.*, 2003). Normally, *bantam* is thought to promote tissue growth by regulating the proapoptotic gene *hid*, which contains five *bantam* target sites in its 3' UTR. *Hid* expression induced apoptosis, but *Hid*-induced apoptosis could be suppressed by introduction of *bantam* showing that at least part of the phenotype of the *bantam* mutant is due to loss of regulation of *Hid* expression. miRNAs are known to play roles in regulating proliferation through involvement in cell cycle regulation (Wang *et al.*, 2008) and through regulating the switch between proliferation and differentiation (Yi *et al.*, 2008).

1.5.1.4 miRNAs are essential for embryonic development

The requirement of miRNAs for correct development is widely conserved and is exemplified by the creation of mutants that are negative for components of the miRNA-processing pathway. *Dicer*1-null zebrafish arrest at day 10 of development (Wienholds *et al.*, 2003), correct development timing in the nematode worm *C. elegans* depends upon the expression of miRNAs *lin-4* and *let-7* (Lee *et al.*, 1993, Reinhart *et al.*, 2000) and the *Drosophila* miRNA bantam controls cell proliferation during fruit fly development (Brennecke *et al.*, 2003).

miRNAs are also essential for early mammalian development. *Dicer*$^{-/-}$ murine embryos are reported to arrest around E7.5, when the embryo is undergoing gastrulation (Bernstein *et al.*, 2003) (*Figure 1.11*) and DGCR8$^{-/-}$ embryos are reported to arrest prior to E6.5, the point at which gastrulation commences (Wang

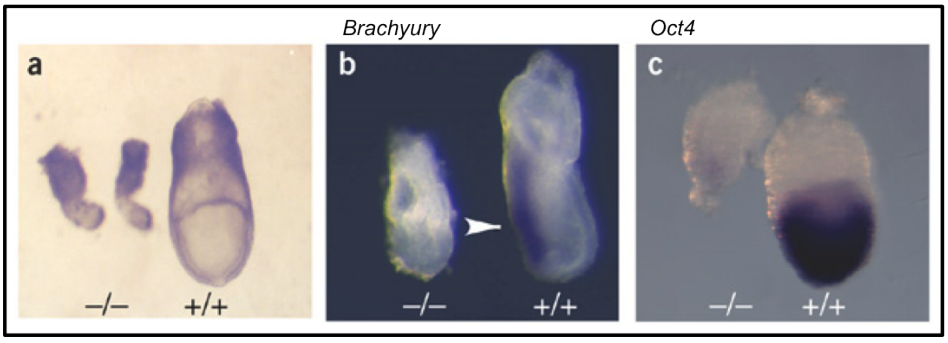


Figure 1.11 – *Dicer*$^{-/-}$ *and wt embryos.* Bernstein *et al.* (2003) disrupted the *Dicer* gene in mice. Homozygous null embryos arrested at approximately E7.5. (a) Typical E7.5 wt and mutant embryos. Mutant embryos appear small and morphologically abnormal although embryonic-extraembryonic regions are distinguishable. (b) *In situ* hybridisation for brachyury, a marker of the primitive streak, showed that mutant embryos did not express brachyury. (c) *In situ* hybridisation for *Oct4* showed that staining was much reduced in mutant embryos. Figure from Bernstein *et al.* (2003).

et al., 2007). In situ hybridisation showed reduced *Oct4* expression in *Dicer*$^{-/-}$ embryos compared to wt embryos and no detectable expression of *Brachyury*. The authors concluded that *Dicer*$^{-/-}$ embryos exhibited a loss of stem cells and a failure to specify the primitive streak. The exact reason for developmental arrest of miRNA-deficient embryos is not known but both *Dicer*$^{-/-}$ and DGCR8$^{-/-}$ embryos are fully resorbed by E10 with only evidence of implantation remaining.

1.5.2 miRNAs in early mouse development

miRNAs are expressed in mouse embryos in a spatial and temporal fashion (Wheeler *et al.*, 2006, Mineno *et al.*, 2006, Takada *et al.*, 2006). Loss of the miRNA processing enzymes *Dicer*1 and DGCR8 demonstrate that the global miRNA population is essential for mammalian development (Bernstein *et al.*, 2003, Wang *et al.*, 2007) and more recent work has shown that they are involved in multiple processes during early mouse development.

1.5.2.1 miRNAs in early lineage specification and gastrulation

miRNAs are involved in developmental processes as early as gametogenesis. Mice deficient for *Dicer*1 have primordial germ cells (PGCs) with reduced proliferation

rates and have retarded spermatogenesis (Hayashi *et al.*, 2008a). In contrast to their essential role in murine spermatogenesis, miRNAs do not appear to function during oogenesis. Indeed, oocytes deficient for DGCR8 have no phenotype and show no change in mRNA levels relative to wild type oocytes (Suh *et al.*, 2010).

Spruce *et al.* (2010) performed a comprehensive analysis of early mouse embryos with zygotic loss of *Dicer*. Such embryos showed initial specification of the epiblast with strong expression of *Oct4* and *Cripto*. Expression of markers of the primitive streak, *Eomes* (which is also a trophectoderm marker), *Brachyury* and *Nodal*, showed that gastrulation was initiated but delayed in mutant embryos. Additionally, markers of definitive endoderm were either drastically reduced or lost in the majority of mutant embryos. Overall, the results showed that *Dicer*$^{-/-}$ embryos could form an epiblast and initiate gastrulation but that progression of patterning was abnormal (Spruce *et al.*, 2010). A large caveat remains with this data however, because the embryos were deficient for zygotic *Dicer*, but contribution of miRNAs generated by maternal *Dicer* could not be excluded. In fact, the authors found similar levels of *miR-291-3p*, *miR-295* and *miR-291-5p* in mutant and wild type blastocysts. Therefore the authors could not determine if miRNAs have a role in specification of the first lineages of the blastocyst. Additionally, the findings of Spruce *et al.* disagreed with those of Bernstein *et al.* (mentioned above) who found that *Dicer*$^{-/-}$ embryos exhibited loss of the stem cell population and an inability to specify the primitive streak (Bernstein *et al.*, 2003).

1.5.2.2 miRNA knockout mice

The generation of mice deficient for miRNA clusters has provided insight into the role of specific miRNAs during mouse development. Deficiency for the *miR-290-295* cluster of miRNAs results in high incidences of embryonic lethality, with homozygous null embryos showing general developmental delays or partial or complete localisation outwith the yolk sac (Medeiros *et al.*, 2011). However, a proportion of homozygous null mice reach birth and are viable, although female mice are infertile. This shows that while this miRNA cluster is important for correct development, it is not essential. Mice homozygous null for the *miR-17-92* cluster of miRNAs die shortly after birth with lung hypoplasia, ventricular septal defects and a reduction in the number of pre-B cells in fetal livers (Ventura *et al.*, 2008). Interestingly, while ablation of the *miR-106b-25* or *miR-106a-363*

clusters had no obvious phenotypic effect, embryos lacking both the *miR-17-92* and *miR-106b-25* clusters or all three miRNA clusters had a much more severe phenotype than the *miR-17-92* knockout alone and died midgestation. Together, these results show that miRNAs play essential roles during mouse embryonic development.

1.5.3 miRNAs in ES cells

1.5.3.1 Expression of miRNAs in ES cells

miRNA expression was first assessed in mouse ES cells in 2003 when Phillip Sharp's laboratory cloned short 20-26nt RNAs from undifferentiated and differentiated ES cells (Houbaviy *et al.*, 2003). They found miRNAs associated with the undifferentiated and differentiated states including the *miR-290-295* family, which were highly expressed in undifferentiated mouse ES cells. Since then, another study has shown that the *miR-290-295* cluster made up between 23% and 29% of miRNA sequences from mouse ES cells, the largest proportion attributed to a single miRNA cluster (Calabrese *et al.*, 2007). Although the *miR-290-295* cluster are expressed at very high levels in mouse ES cells, they are not found in human ES cells (Suh *et al.*, 2004). However, human ES cells express high levels of the *miR-371-373* and *miR-302-367* cluster miRNAs, which are sequence related to *miR-290-295* (Suh *et al.*, 2004). Despite the early arrest of miRNA-deficient embryos, *Dicer*$^{-/-}$ and DGCR8$^{-/-}$ ES cells can be maintained *in vitro* (Murchison *et al.*, 2005, Kanellopoulou *et al.*, 2005, Wang *et al.*, 2007). These cells express markers of pluripotency and replicate indefinitely but show defects in rate of growth and differentiation capacity.

1.5.3.2 The phenotype of *Dicer*$^{-/-}$ ES cells

Dicer is an RNAse III enzyme that is involved in miRNA and siRNA biogenesis. *Dicer*$^{-/-}$ ES cells have been generated by two independent laboratories (Murchison *et al.*, 2005, Kanellopoulou *et al.*, 2007) and are deficient for miRNAs and siRNAs. Murchison *et al.* generated conditional *Dicer*$^{-/-}$ ES cells and made the observation that their proliferation rate was slower than that of WT ES cells.

This was attributed to an alteration of the cell cycle, which showed an increase in cells in G1 and G0 phases and a corresponding decrease of cells in G2 and M phase. The authors reported that the *Dicer*$^{-/-}$ ES cells expressed ES cell marker genes including *Oct4*, but showed an increase in transcripts derived from heterochromatic repeats. Kanellopoulou *et al.* also generated *Dicer*$^{-/-}$ ES cells and reported that they expressed the ES cell marker *Oct4* at comparable levels to WT ES cells, but proliferated more slowly and were defective in epigenetic silencing of centromeric repeat sequences similar to the findings of Murchison *et al.* Kanellopoulou *et al.* also investigated the differentiation potential of the *Dicer*$^{-/-}$ ES cells. They found that embryoid body induced differentiation of *Dicer*$^{-/-}$ ES cells showed severe defects. There was little morphological evidence of differentiation and no expression of differentiation markers of the three germ lineages was detectable.

1.5.3.3 The phenotype of DGCR8$^{-/-}$ ES cells

DGCR8 is a component of the microprocessor complex, which is involved in miRNA biogenesis (*Figure 1.2*). Removal of DGCR8 from mouse ES cells resulted in viable cells that did not express canonical miRNAs (Wang *et al.*, 2007). DGCR8$^{-/-}$ ES cells have provided an excellent model for elucidating the function of miRNAs in ES cells because individual miRNAs can be reintroduced to these cells and the effect, if any, on the phenotype of the cells can be assessed. In this way miRNAs controlling proliferation have been determined. DGCR8$^{-/-}$ ES cells show a reduction in proliferation rate, which is associated with accumulation of the cells in the G1 phase of the cell cycle (Wang *et al.*, 2008). Wang and others used a screen to test the effects of re-introducing 461 individual miRNAs on proliferation of the DGCR8$^{-/-}$ ES cells. They identified 14 different miRNAs that rescued the proliferation defect (Wang *et al.*, 2008). These miRNAs belonged to three families: the *miR-290*, the *miR-302* and the *miR-17-92a* clusters of miRNAs, which are all highly expressed in ES cells and down-regulated upon differentiation. Collectively, these miRNAs were termed ESCC (for ESC-specific cell cycle regulating) miRNAs. These findings suggest that a major role for miRNAs in ES cells is to ensure rapid cell cycle progression. The miRNAs achieve this, at least in part, by inhibiting inhibitors of the CyclinE/CDK2 pathway (*p21cip*, *Rbl2* and *Lats2*), which regulates the G1/S transition (Wang *et al.*, 2008). The phenotype of miRNA-deficient ES cells is discussed in more detail in section 5.1.1.

1.5.4 The role of miRNAs in ES cell differentiation

1.5.4.1 The let-7 family of miRNAs

Profiling of miRNAs during ES cell differentiation has revealed specific miRNAs that increase in expression during differentiation (Krichevsky *et al.*, 2006, Ivey *et al.*, 2008). One family of miRNAs that is strongly associated with ES cell differentiation is the *let-7* family (Viswanathan *et al.*, 2008). DGCR8$^{-/-}$ ES cells are reported to show an inability to silence self-renewal when induced to differentiate, but introduction of *let-7* miRNAs is sufficient to rescue this defect (Melton *et al.*, 2010). Interestingly, while mature *let-7g* is not found in undifferentiated ES cells, the *pri-let7g* transcript is readily detectable (Viswanathan *et al.*, 2008). Viswananthan *et al.* (2008) determined that a protein highly expressed in ES cells, LIN28, interacted with *pri-miR-let7g* and inhibited its cleavage by the microprocessor complex. Additionally, LIN28 also regulates expression of *let-7* miRNAs by binding to *pre-let7* and promoting its polyuridylation targeting it for degradation (Heo *et al.*, 2008). LIN28 is expressed highly in undifferentiated ES cells and its expression decreases during differentiation. This allows expression of mature *let-7* miRNAs to increase and inhibit self-renewal genes adding a layer of robustness to the differentiation switch.

1.5.4.2 miRNA regulation of pluripotency genes

There are examples of miRNAs that directly regulate key ES cell pluripotency genes such as *Oct4*, *Sox2* and *Nanog* during differentiation. Upon retinoic acid induced differentiation of mouse ES cells, *miR-134*, *miR-296* and *miR-470* are up-regulated. *miR-470* directly regulates expression of *Nanog* and *Oct4*, *miR-296* directly regulates expression of *Nanog* and *miR-134* directly regulates *Sox2* expression (Tay *et al.*, 2008a). Another study showed that expression of *Sox2* and *Klf4* is down regulated by *miR-200c* in mouse ES cells, but did not confirm a direct interaction (Wellner *et al.*, 2009). In human ES cells *miR-145* has been shown to target *Oct4*, *Sox2* and *Klf4* during ES cell differentiation and this interaction is part of a negative feedback loop because the *miR-145* promoter is bound and repressed by *Oct4* in human ES cells (Xu *et al.*, 2009).

1.5.4.3 miRNA regulation of epigenetics in ES cells

Characterisation of *Dicer*$^{-/-}$ ES cells also showed that expression of the de novo DNA methytransferase genes *Dnmt3a*, *Dnmt3b* and *Dnmt3l* were significantly down regulated in the *Dicer*$^{-/-}$ cells (Sinkkonen *et al.*, 2008), but that this could be rescued by expression of *miR-290* family miRNAs. *Oct4* silencing in differentiating *Dicer*$^{-/-}$ ES cells was accompanied by the accumulation of repressive histone marks but not by DNA methylation, which indicated an inability to stably repress *Oct4*. Introduction of *Dnmts* or the *miR-290* cluster rescued the defective *Oct4* promoter methylation indicating a requirement for these miRNAs for stable repression of pluripotency genes during differentiation.

1.5.5 The role of miRNAs in reprogramming

As well as roles in differentiation, miRNA have also been shown to have roles in the process of de-differentiation, or reprogramming. Introduction of the embryonic stem cell specific miRNAs *miR-291-3p*, *miR-294* and *miR-295* along with retroviruses expressing *Oct4*, *Sox2* and *Klf4* enhanced reprogramming in the absence of C-MYC (Judson *et al.*, 2009). C-MYC binds the promoter of these miRNAs, which may suggest that the requirement for C-MYC to increase the efficiency of reprogramming is mediated through these miRNAs. Reports have also shown that overexpression of *miR-302* can reprogram human cancer cells (Li *et al.*, 2008) and human hair follicle cells (Li *et al.*, 2010) into an iPS cell-like state.

Recently, it was reported that it is possible to reprogram mouse and human somatic cells to pluripotency using miRNAs alone (Anokye-Danso *et al.*, 2011). Anokye-Danso and others expressed the *miR-302-367* miRNA cluster in mouse embryonic fibroblasts and human foreskin and dermal fibroblasts and observed that reprogrammed iPS cell colonies were generated. These clones showed characteristics of iPS cells reprogrammed by the original 4-factor approach including teratoma formation and generation of chimaeric mice. If reproducible, this has immense potential for iPS cell research because it abrogates the need for introduction of pluripotency transcription factors and because the miRNAs show high species conservation, which may make this approach readily usable in other species.

1.5.6 Summary

Proper embryonic development is dependent on accurate spatial and temporal regulation of gene expression. miRNAs regulate gene expression post-transcriptionally, and are essential for proper embryonic development. miRNA deficient ES cells have been generated and shown to be viable although the cells show defects in proliferation and differentiation. The miRNA-deficient ES cells provide an excellent system for the investigation of miRNA regulation of developmental processes.

1.6 Aims and Objectives

Mammalian development is a complex series of developmental transitions, which relies on accurate spatial and temporal regulation of gene expression. miRNAs regulate gene expression post-transcriptionally through inhibition of translation and/or mRNA degradation. The overall aim of this project was to investigate the role of miRNAs in regulating developmental processes. An initial objective was to establish whether miRNAs were dynamically regulated in a cell based model of an early developmental transition and whether differentially expressed miRNAs, if any, could be functionally associated with the developmental process. In order to explore the roles of miRNAs during an early developmental transition, an embryonic stem (ES) cell model of trophectoderm differentiation was used. In this model system the expression of the key ES cell regulatory gene, *Oct4*, can be conditionally repressed, which induces the ES cells to differentiate down the trophectoderm lineage. Identification of miRNA target transcripts remains a key goal of miRNA research. Therefore, having established a role for specific miRNAs in developmental transitions, a further objective was to assess the reliability of current sequence-based and experimental methods of miRNA-mRNA target identification. In order to investigate the general role of miRNAs in regulating developmental processes, the final objective was to compare the differentiation capabilities of wt ES cells with those of an ES cell line that does not express miRNAs.

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 Chemicals and Equipment

All chemicals were of molecular biology or analytical grade and were obtained from a variety of manufacturers: Sigma, Fluka, Fisher Scientific, BDH, Fisons and Electran. All commercial reagents and chemicals were stored according to the manufacturer's instructions. All water used for the preparation of solutions was double distilled, deionised (Purite Select Analyst) and sterilised by autoclaving (121°C, 15-20 minutes) except when working with RNA and during the PAR-CLIP method. In these cases, deionised sterile nuclease-free water from Ambion was used. Radioisotopes (EasyTides Adenoside 5'-triphosphate-[γ-32P]-) were obtained from Perkin-Elmer. Work with radioisotopes was carried out within a designated area of the laboratory. All solutions and microcentrifuge tubes used for experiments involving RNA and/or PCR were autoclaved twice (121°C, 15-20 minutes). Filter tips (Art or Greiner) were used for all cell and molecular biology work and designated tips and pipettes kept for RNA and/or PCR work. Pipettes and working area were treated with RNAse-Zap (Ambion) prior to work involving RNA.

2.1.2 Reagents

2.1.2.1 LNA oligonucleotides

miRNA inhibition and detection was achieved using RNA oligonucleotides with approximately every third RNA nucleotide replaced by an LNA-modified DNA nucleotide (capitals). Sequences below:

Small RNA	Manufacturer	Sequence of LNA
mmu-miR-92a	Sigma-proligo	cAggCcgGgaCaaGtgCaaTa
mmu-miR-92a	Exiqon	caGgcmCggGacAagTgcAatA
mmu-miR-294	Exiqon	acAcaCaaAagGgaAgcActTt
mmu-miR-199a	Sigma-proligo	gaAcaGgtAgtCtgAacActGgg
mmu-miR-155	Exiqon	acmCccTatmCacAatTagmCatTaa
mmu-miR-363	Exiqon	taCagAtgGatAccGtgCaaTt
mmu-miR-25	Exiqon	tcAgamCcgAgamCaaGtgmCaaTg
mES1	Sigma-proligo	acAtaTacAtgCacGcaCttAt
mES2	Sigma-proligo	ccAagTacTgaAatTaaAggCt
mmu-miR-30C	Sigma-proligo	gcTgaGagTgtAggAtgTttAca
*mmu-miR-467a**	Exiqon	gtGtaGgtGtgTgtAtgTatAt
mmu-miR-669b	Sigma-proligo	acAtgCacAtgCacAcaAaaCt

Table 2.1 – *Sequences of LNA oligonucleotides.*

2.1.2.2 miRNA precursors

Exogenous miRNA expression was achieved using Pre-miR miRNA Precursors from Ambion:

*mmu-miR-467a** stem-loop sequence: CCUGUGUGCAUAAGUGCGCGCAUGUAUAUG CGUGUAUAUUUUAUGCAUAUACAUACACA CACCUACACACACAUGCACACAGA CA

mmu-miR-210 stem-loop sequence: ACCCGGCAGUGCCUCCAGGCGCAGGGCAGCC CCUGCCCACCGCACACUGCGCUGCCCCAGACCCACUGUGCGUGUGACAGCGGCU GAUCUGUGCCUGGGCAGCGCGACCC

2.1.2.3 siRNA oligonucleotides

Specific mRNA knockdown was achieved using siRNA oligonucleotides (Dharmacon). Sequences below:

siRNA targetting *Oct4*: AAGGAUGUGGUUCGAGUAUGG

siRNA targetting *EGFP*: AAGAACGGCAUCAA GGUGAAC

2.1.2.4 QPCR Primers

Gene	Forward primer 5'-3' sequence	Reverse primer 5'-3' sequence
Oct4	ATCACTCACATCGCCAATCA	AAGGTGTCCCTGTAGCCTCA
Sox2	CGGCTCTGTTATTGGAATCAG	TCTCAAACTGTGCATAATGGAGT
Nanog	TGCAATAAGTTCAAGGCCAAC	GACTCCAAGGACAAGCAAGC
Esrrß	CAGTCCTTCGTGCTGTCTCA	TGGGACTGGATGGGAGATAA
Rex1	ACAAAGGGGACGAAGCAAGAGA	CCAGCATCGATAAGACACCACA
Stella	TTGTTGTCGGTGCTGAAAGA	CATCTGAATGGCTCACTGTC
Cdx2	GCGACAAGGGCTTGTTTAGA	GAGGGAAGGGACAGGAAGTC
Eomes	CTTGGTCATCCCCACTTCC	CTGGTCCCTTATTGAACCACAT
Hand1	GTTCCCATTCGTTGCTGAAT	CTGCGAGTGGTCACACTGAT
Mash2	GATGACCTCTGTCCCTCACC	GCCAAACATCAGCGTCAGTA
Id2	AGGTGGAGCGTGAATACCAG	CAGCATTCAGTAGGCTCGTG
Sfmbt2	GCTGATAAACCCACGATGCT	TGACGCTGCTTGTGGTAATC
Mylip	GAAGGAAGCCATGCTGTGTA	GTAGACGTGCTGGACATGCT
Trim33	TTCCCAGCATTACCAAATCC	TTGTGTGTCTGCATAAACTTGAA
Dnmt3a	AAGGAGGACATCCTGTGGTG	CTTCAGCGGAGCAAGAG
Gata4	CGAGGGTGAGCCTGTATGTAA	ATTCAGGTTCTTGGGCTTCC
Gata6	CTACACAAGCGACCACCTCA	TGTAGAGGCCGTCTTGACCT
Fgf5	GGGATTGTAGGAATACGAGGAG	CGCGGACGCATAGGTATTAT
Brachyury	AACTGCGAGTGGGTCTGGAAG	TGGGTCTCGGGAAAGCAGTG
Snail	CTGCACGACCTGTGGAAAG	AGTGGGAGCAGGAGAATGG
Goosecoid	GAAGCCCTGGAGAACCTCTT	AGGATCGCTTCTGTCGTCTC
Nestin	AGGCTGAGAACTCTCGCTTG	AGAGAAGGATGTTGGGCTGA
Odz4	AGGAATGGGACAACAGCAAG	AGCCGTAGAGCTGGTCAAAG
Nfyc	GATGAACTGAAACCTCCAAAGC	GCTGAGCCAGCGTGAAGTA
C77370	GGCCTGCGAGATGAGAAAT	CCCAAATGCTGTTTCTGCTT
ActB	TGACAGGATGCAGAAGGAGA	GTACTTGCGCTCAGGAGGAG
DGCR8	AAGGTCTCTGTGCTCCCAAG	ATCTTGGTCATCATTGGCTGT

Table 2.2 – *QPCR primer sequences.*

2.1.2.5 Antibodies for western analysis

Protein levels were assessed by western blot analysis using the following antibodies:

Oct4 - Santa Cruz sc5279 (1:1000); ActB – Sigma Aldrich A5441 (1:5000); DGCR8 – ProteinTech 10996-1-AP (1:1000); Ago2 – WAKO BioProducts 014-2202 (1:1000); pSTAT3 – Cell Signalling Technologies 9134S (1:1000); STAT3 – Transduction Laboratories 610190 (1:1000); pAKT - Cell Signalling Technologies 4060S (1:1000); AKT - Cell Signalling Technologies 4685 (1:1000); pERK - Cell Signalling Technologies 4377S (1:1000); ERK – Transduction Laboratories 61014 (1:1000); pGSK3B - Cell Signalling Technologies 9336S (1:1000); GSK3B - Cell Signalling Technologies 9315S (1:1000).

Secondary antibodies were used at 1:2000 dilutions: Horse-anti-mouse IgG (HRP-linked) Cell Signalling Technologies 7076S Goat-anti-rabbit IgG (HRP-linked) Cell Signalling Technologies 7074S

2.1.2.6 Antibodies for immunohistochemistry analysis

Protein expression was visualised by immunohistochemistry using the following antibodies: Oct4 - Santa Cruz sc5279 (1:200); Nanog - Abcam ab80892 (1:200); Nestin - DSHB Rat-401 (1:20); Tuj1 - Covance mms-435P (1:500); E-cadherin – Takara ECCD-2 (1:200). Secondary antibodies were Invitrogen Alexa Fluor dye conjugated antibodies raised in donkey or goat corresponding to appropriate antibody isotypes. Secondary antibodies were used at a 1:2000 dilution.

2.2 General Methodology

2.2.1 Cell Culture Techniques

2.2.1.1 Embryonic Stem Cell Culture

ES cells were cultured at 37°C (5% CO2) in GMEM supplemented with 10% fetal calf serum, 2mM L-Glutamine, 1x non-essential amino acids, 1mM sodium pyruvate, 100uM ß-Mercaptoethanol and Lif. Lif was synthesised in the laboratory

and its efficacy was determined by titration against commercial Lif before use (performed by Linda Sutherland). ES cells were cultured on gelatin-coated tissue culture treated plastic, media was replaced every 24 hours and cells were passaged every 48 hours or when 80% confluency was reached. TVP was used to dissociate ES cell cultures for passaging. For growth in serum-free conditions, ES cells were first grown in standard ES cell media, cells were washed twice with 1 x PBS and then N2B27 (Ying & Smith, 2003) or 2i media was added to cultures. 2i media was made up fresh for each usage by addition of PD0325901 and CHIR99021 to N2B27 media to final concentrations of 1uM and 3uM, respectively. When necessary, cells were counted using a haemocytometer before being diluted to an appropriate concentration for plating. ES cells were passaged a maximum of 50 times. Vials containing ES cells were frozen in 10% DMSO, 50% FCS in ES cell media at -80°C for 1 week, and then transferred to -150°C for long-term storage.

2.2.1.2 Differentiation of ZHBTc4 ES cells

$5x10^5$ ZHBTc4 ES cells were transferred to gelatin-coated T25 tissue culture flasks in standard ES cell culture media. 24 hours later the media was removed and replaced with standard ES cell media or ES cell media containing 1µg/ml doxycycline. ZHBTc4 ES cells were grown in the presence or absence of doxycycline for up to 96 hours with the media being replaced with fresh media (with or without 1µg/ml doxycycline) every 24 hours. Protein and RNA samples were taken every 24 hours.

2.2.1.3 Derivation of mouse embryonic fibroblasts

MEFs were derived from E12.5 and E13.5 C57/Bl6 x CBA/Ca F1 hybrid mouse embryos. Embryos were washed in PBS with 2x Penicillin:streptomycin and internal organs were removed using forceps. Embryos were washed twice in PBS with 2x Pencillin:Streptomycin. Single embryos were placed in 7ml bijous containing 2 mls of TVP and were cut into small pieces using scissors. Bijous containing embryos were incubated for 5 minutes at 37°C and then vortexed thoroughly. This process was repeated to obtain a cloudy suspension. 3 mls of MEF media (GMEM supplemented with 10% fetal calf serum, 2mM L-Glutamine, 1x non-essential amino acids, 1mM sodium pyruvate and 100uM ß-Mercaptoethanol)

was added to the bijou to stop the trypsinisation reaction and bijous were vortexed. Bijous were left to stand until any remaining pieces of tissue had sunk to the bottom of the bijous. The top 3 mls of media was then removed from the bijou and transferred to a gelatinised T75 tissue culture flask containing 12 mls of MEF media. Cultures were incubated overnight at 37°C in 5% CO2. After 24 hours dead cells were removed from the flask by replacing the media with fresh MEF media. When the cultures were 70-80% confluent, cells were trypsinised and resuspended in 10% DMSO, 50% FCS in MEF media. 2 vials of cells per T75 flask were frozen at -80°C and transferred to -150°C one week later for long-term storage.

2.2.1.4 Fibroblast Cell Culture

3T3 cells and mouse embryonic fibroblasts (MEFs) were grown at 37°C (5% CO2) in GMEM supplemented with 10% fetal calf serum, 2mM L-Glutamine, 1x non-essential amino acids, 1mM sodium pyruvate and 100uM ß-Mercaptoethanol. Cells were grown on gelatin-coated tissue culture-treated plastic. Media was replaced every 24 (3T3 cells) or 48 (MEFs) hours and cultures were passaged when 80% confluency was reached (approximately every 48 hours for 3T3 cells and every 72-96 hours for MEFs). When necessary, cells were counted using a haemocytometer before being diluted to an appropriate concentration for plating. 3T3 cell stocks were frozen in 10% DMSO, 50% FCS in MEF media at -80°C for 1 week, and then transferred to -150°C for long-term storage. For MEFs to be used as feeder cells, cells were grown to 70-80% confluency, trypsinised, resuspended in MEF media in a 50ml centrifuge tube and irradiated (using a GammaCell 1000 Elite) with a 50 Gray dose. Cells were then counted and plated at a density of $1.5x10^5$ per T25.

2.2.1.5 Culture of Embryoid Bodies

For formation of embryoid bodies (EBs) $1x10^6$ ES cells were transferred to a single well of a 6-well Ultra-Low Attachment® Plate (Corning®) in GMEM supplemented with 10% fetal calf serum, 2mM L-Glutamine, 1x non-essential amino acids, 1mM sodium pyruvate and 100uM ß-Mercaptoethanol. Media was removed and replaced with fresh EB media every 48 hours. After 8 days in

suspension culture 10 EBs were transferred to a single well of a 6-well gelatin-coated tissue culture-treated plastic plate. EBs adhered to the surface of the plate and were grown for a further 8 days. Media was removed and replaced with fresh EB media every 48 hours. Samples were taken at the start of the experiment (time 0), day 4, day 8, day 12 and day 16 of culture. For RNA analysis, EBs media was replaced with 1xPBS, and then 1xPBS was removed and replaced with 1ml RNA-Bee (AMS Biotechnology), which lysed the EBs.

2.2.1.6 Transfection of ES cells and 3T3 cells

ES cells and 3T3s were transfected using the transfection reagents Lipofectamine™ 2000 (Invitrogen) or Lipofectamine™ LTX (Invitrogen). Both transfection reagents were used according to the manufacturer's instructions. Briefly, cells were cultured in 6-well plates or 96-well plates for 24 hours and the media was removed approximately 1 hour prior to transfection and replaced with 1ml of media in 6-well plates or 100µl of media in 96-well plates. DNA/RNA for transfection was diluted in Opti-MEM® (Invitrogen), followed by dilution of Lipofectamine 2000 or Lipofectamine LTX in Opti-MEM and incubation at room temperature for 5 minutes. An equal volume of Optimem plus Lipofectamine was added to each tube of diluted DNA/RNA and incubated at room temperature for at least 30 minutes. An appropriate volume of transfection mix was added per well of cells. Plates of transfected cells were incubated at 37^{o}C (5% CO_2) overnight for a maximum of 18 hours. Reactions were stopped by removal of media and replacement with fresh media.

MiRNA knockdown miRNA expression was inhibited using between 37.5nM and 150nM of complementary RNA oligonucleotide with approximately every third RNA nucleotide replaced by an LNA-modified DNA nucleotide (from Exiqon or Sigma-Proligo). Transfections were carried out as described above.

MiRNA Overexpression miRNAs were overexpressed in cells by transfection of between 30nM and 200nM pre-miR miRNA precursors (Ambion). Transfections were carried out as described above.

SiRNA Knockdown For siRNA knockdown 1x10^5 cells were transferred to each well of a 6-well plate. After 24 hours, media was removed and replaced with 1ml of fresh media. Cells were transfected using the Lipofectamine™ 2000 (Invitrogen) reagent with 100ng/ml siRNA targeting Oct4 or EGFP (control). Transfections were carried out as described above. Transfections were repeated 24 hours later using the same protocol. Samples were lysed for RNA (in 1ml RNA-Bee) 72 hours after the first transfection.

2.2.1.7 Cell proliferation assays

The reduction of MTT (3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) to formazan was monitored using the Vybrant® MTT Cell Proliferation Assay Kit (Molecular Probes) as described in the manufacturer's instructions. For analysis of cell proliferation after miRNA knockdown, 1x10^3 ES cells or 0.5x10^3 3T3 cells were transferred to each well of a gelatin-coated tissue culture-treated 96-well plate. After 24 hours, media was replaced with 100µl of fresh media and cells were transfected overnight with LNA oligonucleotides complementary to various miRNAs using 1µl per well of Lipofectamine LTX. Transfection reactions were stopped after a maximum of 18 hours by removal of the media and replacement with fresh media. The following afternoon, or 24 or 48 hours later, the media was replaced with 100µl phenol-red-free media and 10µl of 12mM MTT stock solution per well. Reactions were incubated at 37°C for either 1 hour (ES cells) or 2 hours (3T3 cells). Then 100 µl of SDS::HCL solution was added to each well and the contents mixed thoroughly by pipette. Reactions were incubated overnight at 37°C in a humidified chamber. After a maximum time of 18 hours the contents of the wells were mixed by pipette and the absorbance of each well at 570nm was recorded by a 96-well Multiscan Ascent plate reader (Thermo Electron Corporation, Thermo Scientific). Background readings were subtracted and results were normalised to control values.

The CyQUANT Direct Cell Proliferation Assay Kit (Invitrogen) was used to assay cell proliferation according to the manufacturer's instructions. For analysis of cell proliferation after miRNA knockdown, 1x10^3 ES cells were transferred to each well of a gelatin-coated tissue culture-treated 96-well plate. After 24 hours, media was replaced with 100µl of fresh media and cells were transfected overnight with LNA oligonucleotides complementary to various miRNAs using 1µl per well

of Lipofectamine LTX. Transfection reactions were stopped after a maximum of 18 hours by removal of the media and replacement with fresh media. The following afternoon, or 24 or 48 hours later, the media was replaced with 100µl media and 100µl of 2x detection reagent per well. Reactions were incubated at 37°C for 1 hour and the absorbance at 485nm of the contents of each well was recorded by a 96-well Multiscan Ascent plate reader (Thermo Electron Corporation, Thermo Scientific). Background readings were subtracted and results were normalised to control values.

2.2.1.8 Cell Counting by Haemocytometer

Haemocytometers were cleaned with 70% ethanol before coverslips were moistened and affixed to the haemocytometer using gentle pressure until Newton's rings could be observed (this ensures that the chamber is of the correct depth). The suspension of cells for counting was mixed thoroughly and 10µl was pipetted into the haemocytometer chamber. Using a hand tally counter, cells within the 3 corner squares (each made up of 16 squares) were counted. Cells within the square and on the right hand and bottom boundary line were counted. The average number of cells in the 3 corner squares is equivalent to the number of cells in the original suspension $x10^4$per ml.

2.2.1.9 Cell Counting by Nucleocounter

The Nucleocounter NC-100 (Chemometec) was used according to the manufacturer's instructions. A representative cell sample was mixed with an equal volume of reagent A-100 (lysis/disaggregation buffer) followed by regent B-100 (stabilising buffer). Samples were loaded onto a Nucleocassette, which was placed in the Nucleocounter for cell number analysis. For assessing the effect on cell number of miRNA inhibition, $2x10^5$ cells were transferred to each well of a 6-well gelatin-coated tissue culture-treated plate. After 24 hours, cells were transfected overnight with 1µg/well LNA oligonucleotides using 5µl per well of Lipofectamine LTX. After a further 24 hours, cells were removed from the plate by trypsinisation, counted by nucleocounter and transferred to fresh 6-well plates (800000 ZHBTc4 ES cells or 700000 E14 ES cells/well) or to 96-well plates for MTT analysis and Cyquant Analysis (1000 cells/well for ZHBTc4 ES cells only). After a further 48

hours of culture, cells were removed from the plate and counted and MTT and CYQUANT assays were performed. The remaining cells were lysed for RNA (in 1ml RNA-Bee) in order to confirm inhibition of miRNA expression.

2.2.1.10 Cell Imaging

Photographs of cells were taken using the Nikon Eclipse TE2000-U microscope. Bright field photographs were captured using white light at an exposure of 10ms and adjusted for white background. Photos for immunohistochemistry were captured using FITC (green, 515-555nm), Texas Red (red, 600-660nm) or UV (blue, 420nm) filters at an exposure of 100-800ms.

2.2.1.11 Neural Differentiation of mES cells

For neural differentiation $5x10^3$ ES cells were transferred to each well of 20µg/ml laminin-coated 4-well plates (plates were coated with laminin at 37ºC for at least 2 hours) in N2B27 media (Ying & Smith, 2003) supplemented with 0.25µM PD0325901 and 1.5µM CHIR99021 (custom synthesised by the Division of Signal Transduction Therapy, University of Dundee). The media was changed every 48 hours. After 7 days the cells were fixed in 4% paraformaldehyde and immunohistochemistry was used to visualise expression of proteins associated with neural cells (protocol below).

2.2.1.12 Subcloning of DGCR8$^{-/-}$ ES cells

For subcloning 250 DGCR8$^{-/-}$ ES cells were transferred to 10cm^2 gelatin-coated tissue culture-treated plastic dishes and cultured in ES cell media. The media was changed every 72-96 hours. After >30 days of culture colonies were scraped and sucked into a p20 pipette and transferred to single wells of a 96-well gelatin-coated tissue culture-treated plate. Clones were cultured in ES cell media and cell numbers were expanded until there were sufficient cells for testing of miRNA-deficiency.

2.2.1.13 Genetic modification of mES cells

For generation of genetically modified ES cell lines, $5x10^6$ ES cells were transferred to each well of a gelatin-coated tissue culture-treated 6-well plate and cultured in ES cell media. 24 hours later cells were transfected overnight with 4µg/well of expression vector using Lipofectamine 2000 (Invitrogen). 48 hours later transfected cells were trypsinised and transferred to 2 x gelatin-coated tissue culture treated 10cm plates at different densities (3/4 and 1/4 of total cells) in ES cell media. 48 hours later 0.5µg/ml puromycin was added to ES cell cultures in ES cell media. Media plus 0.5µg/ml puromycin was replaced every 3-4 days. 12 days after the initial addition of puromycin, the puromycin concentration was increased to 0.75µg/ml in ES cell media. 3 days later clones stably expressing the EGFP expression vector were carefully transferred to single wells of 96-well plates using a p20 pipette. 2 days later clones stably expressing the PCAGIP empty vector were transferred to 96-well plates in the same manner. These clones were cultured and cell numbers expanded until there were enough cells to make frozen stocks.

2.2.2 RNA Techniques

2.2.2.1 RNA extraction

Cells were lysed in RNA-Bee (AMS Biotechnology) for RNA extraction. Samples were vortexed thoroughly and frozen in RNA-Bee at -80°C for future processing, or processed directly as follows: 200µl of chloroform was added per 1ml of RNA-Bee, samples were shaken vigorously for 30 seconds and then centrifuged at 4°C for at least 15 minutes at 12,000xg. The aqueous phase was transferred to a fresh 1.5ml microcentrifuge tube and 0.7 volumes of isopropanol was added. Samples were vortexed briefly then centrifuged at 4°C for at least 30 minutes at 12,000xg. The supernatant was removed by pipette and 0.5ml of ice-cold 85% ethanol was added to the microcentrifuge containing the RNA pellet. The sample was centrifuged at 4°C for at least 10 minutes at 12,000xg. The supernatant was removed by pipette and the microcentrifuge tube containing the RNA pellet was put in ice with the lid open for 10 minutes in order for any residual ethanol to evaporate. RNA pellets were resuspended in nuclease-free water by thorough vortexing and

heating for 2 minutes at 65°C. For long-term storage RNA was precipitated with 0.1 volumes of sodium acetate pH5.2 and 2.5 volumes of 100% ethanol and stored at -20°C.

2.2.2.2 Recovery of precipitated RNA

Samples containing precipitated RNA were vortexed thoroughly and a volume of sample containing the desired quantity of RNA was removed to a fresh 0.5ml microcentrifuge tube. The remaining stocks containing precipitated RNA were returned to -20^{o}C storage and the decanted samples were centrifuged at 4°C for at least 30 minutes at 12,000xg. The supernatant was removed by pipette and 500µl of ice-cold 85% ethanol was added to the microcentrifuge tube containing the RNA pellet. The sample was centrifuged at 4°C for at least 10 minutes at 12,000xg. The supernatant was removed and by pipette and the microcentrifuge tube containing the RNA pellet was put at 45°C with the lid open for 2 minutes in order for any residual ethanol to evaporate. RNA pellets were resuspended in the desired volume of nuclease-free water by thorough vortexing and heating for 2 minutes at 65°C.

2.2.2.3 Quantitative and qualitative analysis of RNA

The concentration of RNA in solution was quantitated using the Nanodrop spectrophotometer (ThemoScientific) according to the manufacturer's instructions. Sample pedestals were cleaned before use and nuclease-free water was used to set the reference spectrum of the spectrophotometer. A 1µl volume of RNA sample was pipetted onto the sample pedestal and the absorbance was read at the optimal path length (0.05mm – 1mm). A full spectral output was given showing the concentration of RNA in solution and the presence of contaminants. The quality of RNA within a sample was assessed using the Agilent Bioanalyser 2100 according to the manufacturer's instructions. The Agilent RNA 6000 Nano Kit was used to assess sample quality except in cases where low RNA concentrations were expected, in which case the Agilent RNA 6000 Pico Kit was used.

2.2.2.4 Northern analysis of miRNA levels

5µg of total RNA in nuclease-free water was mixed with an equal volume of 2x TBE-Urea loading buffer (Invitrogen) up to a maximum sample volume of 30µl. Samples were heated to 70^{o}C for 3 minutes in order to minimise the presence of RNA secondary structures. Samples were pipetted into wells of a 15% TBE-Urea Criterion polyacrylamide gel (Bio-Rad) and separated by gel electrophoresis in 1x TBE buffer at 125V for approximately 1.5 hours. The resulting gel and a piece of Hybond™-N membrane (pre-wetted in 0.5 x TBE buffer) were placed between pieces of filter paper soaked in 0.5 x TBE buffer. RNA was transferred from the gel onto Hybond™-N membrane (Amersham) using the Bio-Rad semi-dry transfer system (220mA for 1 hour). RNA was cross-linked to the membrane (120mJ/cm^2) using the UV Stratalinker 2400 (Stratagene).

For detection of specific miRNA molecules, membranes were pre-hybridised in ULTRAhyb Oligo buffer (Ambion) at 42^{o}C for a least one hour and incubated overnight at 42^{o}C with complementary radiolabelled LNA-modified oligonucleotides in ULTRAhyb Oligo buffer (Ambion). Oligonucleotides were radiolabelled at the 5' end with [P32]γ-ATP using the mirVana Probe and Marker Kit (Ambion). After removal of unbound oligonucleotide (2 x 15 minute washes in 2xSSC, 0.1%SDS at 42°C), membranes were wrapped in saran wrap, placed in an exposure cassette and exposed to photographic film (Kodak, Biomax MS film) at -80°C in the presence of a MS intensifying screen (Kodak) for between 1 hour and 24 hours. Photographic film was developed (Konica SRX-101A X-ograph machine) and the resulting image was assessed for the presence of non-specific signals. Membranes went through cycles of washes in 0.1xSSC, 0.1% SDS at increasing temperatures followed by exposure to photographic film until non-specific signals were removed. Different exposure times were used to get an optimum autoradiograph showing expression of the specific miRNA. Membranes were washed in boiling 0.1% SDS in order to remove any traces of radiolabelled oligonucleotide and could then be pre-hybridised and incubated with a radiolabelled oligonucleotide complementary to a different miRNA. Alternatively, membranes could be stored in saran wrap at -20^{o}C indefinitely.

2.2.2.5 Phosphorimager quantitation

To quantitate hybridised radiolabelled nucleic acid after northern analysis of miRNA levels, membranes were exposed to phosphorimager screens at room temperature for the required length of time. Phosphorimager screens were first cleared by exposure on a light box for 5 minutes. To quantify the stored signal, screens were scanned (BioRad Molecular Imager FX) and resulting data was analysed using Quantity One software. Values for miRNA expression were normalised for variation in loading by correcting for control values (typically the spliceosome RNA U6).

2.2.2.6 cDNA synthesis

Prior to cDNA synthesis, contaminating DNA was removed from RNA samples using the DNA-free kit from Ambion according to the manufacturer's instructions. 0.1 volumes of NaAc pH5.2 and 2.5 volumes of 100% ethanol were then added to the sample and samples were transferred to -20^{o}C for at least 1 hour in order to precipitate the RNA. Precipitated RNA was recovered (see above for protocol), resuspended in nuclease-free water and quantitated by Nanodrop. Up to 52µg of RNA was used for cDNA synthesis. cDNA was synthesised using the First-Strand cDNA synthesis kit (GE Healthcare) according to the manufacturer's instructions. cDNA was aliquotted to avoid repeated freezing and thawing and stored at -20°C.

2.2.2.7 QPCR primer design and optimisation

Sequences for QPCR primers were either taken from publications or designed using Primer3 software (http://frodo.wi.mit.edu/primer3/) with the following criteria: ~ 100 base pair product, ~20 nucleotide primer sequence, TM (melting temperature) of $\geq$60°C, low primer complementarity, Max Poly-X (The maximum allowable length of a mononucleotide repeat, for example AAAA) = 3. Primers were optimised by performing QPCR on a pool of cDNA samples from undifferentiated and differentiated ES cell cultures over a 1x10^{6} range of cDNA concentrations. Primers were suitable for use if they amplified a single product over >1000-fold range with 95-105% efficiency and a correlation of >0.95.

2.2.2.8 QPCR analysis of gene expression

20-35ng of cDNA in 9.5µl nuclease-free H_2O was amplified over 40 QPCR cycles using the SYBR Green PCR Master Mix (Invitrogen) and 1.25µl of each 10uM forward and reverse primer using a Stratagene lightcycler 3000P QPCR machine. The PCR temperature cycles were set-up according to the manufacturer's protocol. Samples containing no cDNA, H_2O-only and RNA that had not been reverse-transcribed were used as controls. MxPro software (Statagene) was used for analysis of QPCR data. The fluorescence threshold was set at 0.01 (so that amplification curves crossed the threshold during exponential PCR amplification of product) and Ct values were normalised to β-Actin expression. miRNA qPCR was performed using Taqman MicroRNA Assays (Applied Biosystems) according to the manufacturers protocol using 10ng of total RNA in 5µl of nuclease-free H_2O for miRNA-specific cDNA synthesis and QPCR (using a Stratagene lightcycler 3000P QPCR machine). Duplicate cDNA-synthesis reactions were carried out for each sample because normalisation to a housekeeping gene was not possible due to the fact that the cDNA synthesis step, as well as the QPCR, were specific to the miRNA being assayed.

2.2.2.9 MiRNA profiling

In order to profile miRNAs during the ES cell to trophoblast transition, ZHBTc4 ES cells were induced to differentiate by addition of doxycycline. ZHBTc4 ES cells were also grown in the absence of doxycycline and therefore remained undifferentiated. Samples were taken every 24 hours from the undifferentiated and differentiated ES cells. miRNA libraries were generated from undifferentiated ES cells at time 0, 24 hours and 48 hours, and from differentiated ES cells 24, 48 and 72 hours after addition of doxycycline. For miRNA profiling $5x10^6$ ES cells were transferred to T150 tissue culture flasks in 36mls of ES cell culture media. Cells were grown for 24 hours in doxycycline-free media before doxycycline was added to half of the flasks (10ng/ml). Cells were harvested at the point of doxycycline addition (Time 0) and after 24, 48, 72 and 96 hours of growth in media with or without doxycycline. RNA was extracted from ES cells using RNABee (AMS Biotechnology) according to the manufacturer's instructions. The quantity and quality of RNA was assessed using the Labtech Nanodrop ND-1000 and the Agilent 2100 Bioanalyser RNA 6000 Nano Assay. RNA was stored

at -20oC as ethanol precipitate. miRNA was isolated from total RNA by electrophoresis using the FlashPAGE™ Fractionator from Ambion according to the manufacturer's instructions. Eluted miRNA was stored at -20oC as ethanol precipitate. MicroRNA libraries were generated essentially as described by Lau et al. (2001). Briefly, RNA substrate and decade marker RNA were end labelled with γ-32P-dATP. An miRNA cloning linker (17.91x) was ligated to the 3' ends of the miRNA and the ligated products were gel purified. The 5' end adaptor (17.92R) was then ligated to the recovered material and the final ligation products were gel purified and recovered. Ligated miRNAs were converted to cDNA using reverse transcriptase and a primer designed to recognize the 3' adaptor oligonucleotide ligated to the miRNA (15.22). cDNA was amplified by PCR using primers (17.92 and 17.93D) designed to recognise the 5' and 3' adaptors ligated onto the ends of the miRNA. The 5' and 3' PCR primers had BanI restriction sites at their 5' and 3' ends respectively. After PCR amplification, cDNA was digested with BanI restriction enzyme and the products were ligated using T4 DNA ligase. An adenine nucleotide was added on to the 3' ends of resulting concatemers using Taq polymerase and dATP. Products were ligated into a TOPO cloning vector (Invitrogen) and transformed into chemically competent cells. Colonies were picked and sent for sequencing by DNA sequencing services. miRNA profiling was carried out by Tom Burdon and Derek McBride. Primer sequences are below:

17.91x pCTGTAGGCACCATCAAx (x:DMT-O-C3-CPG)

17.93R ATCGTaggcacctgaaa (RNA/DNA version, lowercase RNA)

15.22 ATTGATGGTGCCTAC

17.93D ATCGTAGGCACCTGAAA (DNA version, calcTm = 42°C)

17.92 ATTGATGGTGCCTACAG (calcTm = 40°C)

2.2.3 Protein Techniques

2.2.3.1 Protein extraction and quantitation

Cell pellets were lysed by addition of an equivalent volume of lysis buffer containing EDTA-free protease inhibitors (Roche). Lysates were incubated on ice for 10

minutes then centrifuged at 13,000rpm for 15 minutes. Supernatant was retained and protein quantity was determined using the Bio-Rad Protein Assay according to the manufacturer's instructions. Protein was stored at -80^{o}C in lysis buffer.

2.2.3.2 Western analysis for protein expression

A volume of protein sample containing 50µg of protein was combined with an equivalent volume of 2x Laemmli Sample Buffer (Bio-Rad) and boiling the sample for 10 minutes denatured proteins. Samples were pipetted into the wells of an 8-20% denaturing polyacrylamide gel (Bio-Rad) and separated by gel electrophoresis at 115V for 1.5-2 hours. Filter papers were soaked in one of three different buffers: Anode 1 buffer (18.1g Tris, 100ml MeOH, 400ml H_2O), Anode 2 buffer (1.5g Tris, 100ml MeOH, 400ml H_2O) or Cathode buffer (2.6g 6-amino-n-hexanoic acid, 100ml MeOH, 400ml H_20). A gel 'sandwich' was assembled on the platinum anode surface of the Trans-Blot SD Semi-Dry Electrophorectic Transfer Cell (Bio-Rad) comprising of: 2 x filter papers soaked in Anode 1 buffer, 2 x filter papers soaked in Anode 2 buffer, PVDF membrane (GE Healthcare) pre-wetted in methanol and soaked in Anode 2 buffer, the polyacrylamide gel soaked in Anode 2 buffer, and finally 4 x filter papers soaked in cathode buffer. Proteins were transferred from the gel onto a PVDF membrane at 90mA per gel for 90 minutes. Membranes were pre-incubated in 0.1% Tween (in 1 x PBS) containing 5% Marvel Milk powder, 3% BSA, and 1% serum corresponding to secondary antibody host, and incubated overnight at 4^{o}C with primary antibody at the recommended concentration. Membranes were washed in 0.1% Tween (in 1 x PBS) and incubated with HRP-conjugated secondary antibody at the recommended concentration (generally 1:2000) for 2 hours. Membranes were again washed in 0.1 % Tween (in 1 x PBS). ECL reagents (GE Healthcare) and exposure to photographic film (Kodak, X-OMAT LS film, Cat: 8646770) were used to visualise HRP activity in accordance with manufacturer's instructions. Photographic films were developed using the Konica SRX-101A X-ograph machine.

For Chapter 5 western analysis Cells were lysed in 1 x SDS loading buffer (10% glycerol, 3% SDS, 62.5mM Tris HCL pH6.8, 0.005% Bromophenol Blue, 3% Beta-mercaptoethanol) and exposed to sonication (Sonicator Ultrasonic Processor XL, Misonix) for 8 minutes. Equivalent sample volumes were pipetted into

sample wells and separated by electrophoresis on a Novex NuPage 4-20% Bis-Tris gel (Invitrogen) in 1 x MOPS SDS running buffer (Invitrogen). Hybond-ECL membrane (GE Healthcare) was pre-wetted with dH_2O for 10 min before being soaked in cold transfer buffer (100ml 10X transfer buffer (29.3g Glycine, 58g Tris and 18.8ml 20% SDS made up to 1l dH_20), 200ml EtOH, 700ml dH_20). Filter papers and fiber pads were also soaked in cold transfer buffer and used to assemble a gel 'sandwich' within a transfer cassette: 1 x fibre pad, 2 x filter paper, Bis-Tris gel, Hybond-ECL membrane, 2 x filter paper, 1 x fibre pad. Using a wet transfer system (Criterion blotter, Bio-Rad) proteins were transferred from the gel onto a Hybond-ECL membrane (GE Healthcare) at 90V for 75 minutes. Membranes were pre-incubated in 5% Marvel Milk powder in TBST (80ml 25x TBS (125g Tris, 400g NaCl, made up to 2L with H_20, pH7.6), 2ml Tween 20 (Sigma), made up to 2L with H_20), and incubated overnight at 4^oC with the recommended primary antibody concentration in 5% BSA in TBST. Membranes were washed with TBST (4 x 15 min) and then incubated with HRP-conjugated secondary antibody at the recommended concentration (generally 1:2000) for 2 hours in 0.5% Marvel Milk in TBST. Membranes were washed with TBST (4 x 14 min). ECL reagents (Invitrogen) and exposure to photographic film (Amersham Hyperfilm ECL) were used to visualise HRP activity in accordance with manufacturer's instructions. Photographic films were developed using the Konica SRX-101A X-ograph machine.

2.2.3.3 Immunohistochemistry

$5x10^4$ cells were transferred to each well of a 4-well plate, grown for 24 hours and then fixed in 4% paraformaldehyde (2 x PBS washes, then 15 minutes in 4% PFA at room temperature, then 2 x PBS washes) and stored at 4°C in PBS. Plates were washed (4 x 5 minute washes) with PBST (1 x PBS + 3% Triton), pre-hybridised in 1% BSA, 10% serum (from secondary antibody host) in PBST, and incubated with the suggested concentration of primary antibody in pre-hybridisation solution at 4°C overnight. Plates were washed (4 x 5 minute washes) with PBST and incubated with fluorophore-conjugated secondary antibody (1:1000 dilution) for at least 2 hours at room temperature in the dark. Plates were washed with PBST (4 x 5 minute washes) and stained with DAPI (Invitrogen 1:5000 in PBST) for 5 minutes at room temperature in the dark. Plates were washed with PBS

(2 x 5 minute washes) and stored in PBS at 4°C in the dark. The presence of fluorophore-conjugated antibodies was visualised by exposure to the appropriate wavelength of light.

2.2.4 Bioinformatics

2.2.4.1 Bioinformatics analysis of miRNA profiling data

For bioinformatics analysis the UNIX operating system was used as a platform to run programs written by Wilfrid Carre in Perl. The programs were used to carry out the following processes: RNA sequences were 'trimmed' by removing the 5' and 3' sequences that corresponded to the ligated adaptor sequences. The remaining short-RNA sequences were catalogued and the frequency of occurrence of individual sequences was calculated. Sequences from the profiling data were compared with sequences present within miRBase to identify previously reported miRNAs. Sequences were then matched against the mouse genome to identify their genomic locations. Genomic sequences 10nt 5' and 70nt 3' or 70nt 5' and 10nt 3' of the identified short-RNA sequence were extracted and assessed for the potential to form a hairpin structure by RNAfold.

2.2.4.2 Data analysis of miRNA profiling data

Using Excel, the number of times a miRNA was sequenced in each miRNA library was normalised to the size of the library. Sequences that did not map to the mouse genome were removed, as were sequences that were not predicted to form a hairpin secondary structure by RNAfold. Sequences present in undifferentiated and differentiating cells were compared in order to identify differentially expressed miRNAs. Differentially expressed miRNAs were defined as miRNAs recorded >10 instances that differed by >2-fold between amalgamated undifferentiated and differentiated miRNA libraries.

2.3 Detailed Methods

2.3.1 SILAC

The SILAC (stable isotope labelling of amino acids in cell culture) method relies on metabolic incorporation of 'heavy' or 'light' isotopes of amino acids into newly synthesised proteins in cell culture. The heavy amino acids contain substituted stable isotopic nuclei (in this case 13C) while the light amino acids contain standard isotopic nuclei (12C). Two cell populations are grown in culture media that is identical except that one of them contains light forms of a particular amino acid and the other contains heavy forms. After a number of cell divisions, each instance of the amino acid will be replaced by its isotope labelled analog. Cell populations are then combined and proteins are extracted and prepared for mass spectrometry analysis. Differential expression of proteins can be determined because the heavy labelled proteins have an m/z ratio that is slightly higher than that of light labelled proteins.

In this experiment ZHBTc4 ES cells were grown in the presence of R0K0 (arginine and lysine with 12C nuclei) or R6K6 (arginine and lysine with 13C nuclei). After 96 hours doxycycline was added to the R6K6 (or heavy) labelled population, which made this population differentiate into trophectoderm. 48 hours later equal numbers of undifferentiated R0K0 cells and differentiated R6K6 cells were combined. Proteins were extracted and prepared for mass spectrometry analysis, which allowed comparison of protein expression levels in the two cell populations. A detailed SILAC method follows:

The necessary steps were taken to avoid keratin contamination during this protocol. Peptides tend to stick to plastic surfaces on storage so 1.5 ml Eppendorf LoBind tubes (cat no. 2243108-1) were used throughout.

2.3.1.1 Cell Culture

SILAC ES cell media was prepared using pre-made DMEM media with standard amino acids (R0K0) or heavy labelled amino acids (R6K6) from Dundee cell products and supplemented as for standard ES cell media (10% dialysed fetal calf

serum, 2mM L-Glutamine, 1x non-essential amino acids, 1mM sodium pyruvate, 100uM ß-Mercaptoethanol and Lif).

$2x10^6$ ES cells were transferred to a T25 flask containing either R0K0 or R6K6 ES cell media and the media was removed and replaced with fresh R0K0 or R6K6 media after 24 hours. After 48 hours the cultures were 80% confluent. TVP treatment (1 x PBS wash followed by 3 minutes in 0.8mls TVP at 37^{o}C) was used to detach cells from the surface of the T25. 5ml of R0K0 or R6K6 media was added and cells were transferred to a 15 ml falcon tube, pelleted and resuspended in 10ml of media for cell counting. $2x10^6$ ES cells from each flask were transferred to a fresh T25 flask containing R0K0 or R6K6 ES cell media.

ES cells were grown in R0K0 or R6K6 media for a further 48 hours and the media was removed and replaced with fresh R0K0 or R6K6 media after 24 hours. After 48 hours the cultures were 80% confluent and $2x10^6$ ES cells were transferred to fresh T25 flasks as above. The R0K0 cultures continued to be grown in R0K0 and 1µg/ml doxycycline was added to the R6K6 cultures in R6K6 media. The media was removed and replaced with fresh R0K0 media or R6K6 media with 1µg/ml doxycycline after 24 hours.

TVP treatment (1 x PBS wash followed by 3 minutes in 0.8mls TVP at 37^{o}C) was used to detach cells from the surface of the T25. 5ml of R0K0 or R6K6 media was added and cells were transferred to a 15 ml falcon tube, pelleted and resuspended in 10 ml of 1 x PBS for cell counting by nucleocounter. An equal number ($5x10^6$) of cells from the two cell populations were added to a 15 ml falcon tube, pelleted, washed with PBS, pelleted and resuspended in 100µl of T-Per buffer (Pierce) and protease inhibitors (Roche). Samples were homogenised using a glass homogenisor, sonicated (Clifton Ultrasonic Cleaner) and frozen at -80^{o}C. Volumes of cell suspension from the two separate cell populations containing $1x10^6$ cells were transferred to 1.5ml microcentrifuge tubes, centrifuged, resuspended in 100µl T-PER buffer with protease inhibitors and frozen at -80^{o}C. These samples were later thawed and the protein was quantitated by Bio-Rad protein assay according to the manufacturer's instructions. This gave an indication of the quantity of protein in the sample for SILAC analysis (approximately 0.5µg/µl).

The protein sample for SILAC analysis was thawed, homogenised with a glass homogenisor, sonicated (Clifton Ultrasonic Cleaner) and centrifuged at 10,000xg for 5 min. The supernatant was transferred to a fresh 1.5ml microcentrifuge tube

and 25µl was taken for SILAC processing (this corresponded to approximately 125µg of protein). The remaining sample was frozen at -80^{o}C.

2.3.1.2 Reduction and Alkylation

The protein sample for SILAC analysis was reduced and alkylated by DTT/Iodoacetamide: Disulphide linkages between cysteine residues were broken by addition of DTT to a final concentration of 10 mM (BDH 443553B) and boiling of the sample for 1-2 min. Alkylation then prevented the disulphide bonds from re-forming. The sample was alkylated by addition of iodoacetamide (Sigma I1149) to a final concentration of 50 mM followed by incubation at room temperature in the dark for 30 min. 2x Laemmli loading buffer (BD Biosciences) was added to the samples and the samples were boiled at 85°C for 2 minutes.

2.3.1.3 Seperating proteins on gels and excising bands

Protein samples were separated by electrophoresis on a Novex 4-12% polyacrylamide gel (Invitrogen) at 125V for 90 minutes. The resulting gel was cut into 10 slices in a sterile 14cm tissue culture dish using a sterile scalpel. 5µl of the sample was run on a second gel, which was stained using Blum's silver stain (Blum et al., 1987) and scanned between two pieces of acetate.

Working as quickly as possible, the gel slices were cut into smaller pieces using sterile scapels and placed in 1.5 ml Lobind Eppendorf tubes (one slice per tube). The fine end of a sterile P1000 pipette tip was fused using heat and used to mash up the gel pieces within each microcentrifuge tube.

2.3.1.4 Digestion of band pieces

300 µl of dH_2O was added to gel pieces for 15 min then 300 µl of CH_3CN (Acetonitrite; Sigma A3396) was added for a further 15 min. The supernatant was removed by pipette and 300 µl of 20 mM NH_4HCO_3 (Sigma A6141) was added to gel pieces for 15 min. The supernatant was removed and 300 µl of 20 mM NH_4HCO_3/CH_3CN (50:50 v/v) was added to the gel pieces for 15 min at which

point the gel pieces shrank and appeared opaque. The supernatant was removed. 100 µl of CH_3CN was added to dehydrate the gel pieces for 5 min. The gel pieces shrank and appeared completely white. The supernatant was removed and the band pieces were dried in a Speedvac (miVac DNA concentrator, Genevac) for 5 minutes.

50 µl/gel slice of 25 µg/ml modified trypsin (ABI) in 20 mM NH_4HCO_3 was added to each sample. Gel pieces were rehydrated in trypsin digestion buffer for 30 min. If required, more NH_4HCO_3 was added in order to cover the gel pieces. Samples were mixed gently, centrifuged to bring gel pieces to the bottom of the microcentrifuge tubes and incubated at 37°C overnight (>16h).

2.3.1.5 Extraction of peptides

All gel extraction steps were performed at 30°C on a shaking platform to ensure complete extraction of peptides. An equal volume of CH3CN was added to the digested peptides and the sample was incubated at 30°C for 30 minutes. The supernatant (containing the peptides for analysis) was transferred to a fresh Eppendorf LoBind tube. A volume of 1% formic acid (BDH cat No. 101155F) was added to cover the gel pieces and samples were incubated for 20 minutes. The supernatant (containing more peptides) was transferred to the tube containing the peptides for analysis. A second wash with 1% formic acid was performed and the supernatant (containing more peptides) was transferred to the tube containing the peptides for analysis. 150 µl of CH3CN was added to the gel pieces and incubated for 10 minutes. The gel pieces shrank and turned white. The supernatant (containing more peptides) was transferred to the tube containing the peptides for analysis. The peptides were stored at -80°C and sent to Dundee Cell Products (http://www.dundeecellproducts.com/) for MS analysis.

2.3.2 PAR-CLIP

PAR-CLIP is a novel method that can be used to identify the binding sites of cellular RNA-binding proteins (RBPs) including Ago2 (Hafner et al., 2010a & 2010b). This method relies on the incorporation of photoreactive ribonucleoside analogs into RNA transcripts during cell culture. Irradiation of the cells by UV

light at 365nm induces crosslinking of the photoreactive-nucleoside-labelled RNA molecules to their interacting RBPs. The RBP (in this case Ago2) is immunoprecipitated and co-immunoprecipited RNAs are isolated. The isolated RNA is converted to cDNA, which is used to make DNA libraries for sequencing. A detailed PAR-CLIP method follows:

Solutions can be made up in advance. Add protease inhibitors and DTT directly before use.

2.3.2.1 Expanding Cells

ES cells were expanded in ES cell growth medium to obtain 100-400 x 10^6 cells (approx. 10 x 10cm cell culture plates) at approximately 80% confluency. 14 hours before crosslinking 4-thiouridine (Sigma) was added to a final concentration of 100 μM (1:10000 v/v of a 1 M 4-thiouridine stock solution - 260.27 mg 4-thiouridine dissolved in 1 ml DMSO) directly to the cell culture medium. For the miR-467a* expression experiment, 30nM pre-miR miRNA Precursor for mmu-miR-467a* was transfected into 10 x 10cm plates of DGCR8$^{-/-}$ ES cells with 12.5μl Lipofectamine 2000 (Invitrogen) per plate according to the manufacturer's instructions. Lipofectamine 2000 alone was added to another 10 x 10cm plates of DGCR8$^{-/-}$ ES cells.

2.3.2.2 UV-Crosslinking

Cells were washed once with 10 ml ice-cold PBS per plate and PBS was removed completely. Plates of cells were placed on a tray with ice and irradiated uncovered with 0.15 J/cm2 of 365 nm UV light in a Stratalinker 2400 (Stratagene). Cells were scraped off plates with a cell scraper in 1 ml ice-cold PBS per plate, transferred to 15 ml centrifugation tubes (cells from replicate plates were added to 1 x 15ml tube) and collected by centrifugation at 1000 x g for 5 min at 4°C. The supernatant was discarded. 100 x 10^6 ES cells (10 x 10 cm plates) yielded approx. 1 ml of wet cell pellet. Cells pellets were shock-frozen in liquid nitrogen and stored at -80°C. Cell pellets can be stored like this for at least 12 months.

2.3.2.3 Cell lysis and RNaseT1 digest

The cell pellet was resuspended in 3 volumes of 1x NP40 lysis buffer (50 mM HEPES, pH 7.5, 150 mM KCl, 2 mM EDTA, 1 mM NaF, 0.5% (v/v) NP40, 0.5 mM DTT, Roche complete EDTA-free protease inhibitor cocktail) and incubated on ice for 10 min. Cell lysates were cleared of unbroken cells and debris firstly by centrifugation at 13,000 x g for 15 min at 4°C, and then by filtering through a 0.2 µm membrane syringe filter (Pall Acrodisc or equivalent). RNase T1 (Fermentas, 1000 U/µl) was added to a final concentration of 1 U/µl and samples were incubated in a water bath for 15 min at 22°C. Reactions were subsequently cooled for 5 min on ice before proceeding.

2.3.2.4 Immunoprecipitation and recovery of crosslinked target RNA fragments

For immunoprecipitation, 10 µl of Dynabeads Protein G magnetic particles (Invitrogen) were transferred per ml of cell lysate (for a typical experiment it should be approx. 40-50 µl of beads) to a 1.5 ml microfuge tube. Beads were washed twice with 1 ml of citrate-phosphate buffer (4.7 g/l citric acid, 9.2 g/l Na2HPO4, pH 5.0) and resuspended in twice the volume of citrate-phosphate buffer relative to the original volume of bead suspension. 12.5 µg (12.5µl) of mouse anti-Ago2 monoclonal antibody (Wako) was added per ml of bead suspension (for a final mass of 1µg per pulldown assuming 80µl of beads/pulldown) and incubated on a rotating wheel for 40 min at room temperature. Beads were washed twice in 1 ml of citrate-phosphate buffer to remove unbound antibody before being resuspended in twice the volume of citrate-phosphate buffer relative to the original volume of bead suspension.

20 µl of freshly prepared antibody-conjugated magnetic beads were added per ml of partial RNase T1 treated cell lysate and samples were incubated in 15 ml centrifugation tubes on a rotating wheel for 1 h at 4°C. Magnetic beads were collected on a magnetic particle collector for 15 ml centrifugation tubes (Invitrogen) and transferred to 1.5 ml microfuge tubes, then washed 3 times in 1 ml of IP wash buffer (50 mM HEPES-KOH, pH 7.5, 300 mM KCl, 0.05% (v/v) NP40, 0.5 mM DTT, Roche complete EDTA-free protease inhibitor cocktail). A magnetic particle collector for 1.5 ml tubes (Invitrogen) was used to separate the beads

from the supernatant according to the manufacturer's instructions. RNaseT1 (Fermentas, 1000 U/µl) was added to a final concentration of 100 U/µl and samples were incubated in a water bath for 15 min at 22 °C and cooled subsequently on ice for 5 min. Beads were washed 3 times in 1 ml of high-salt wash buffer (50 mM HEPES-KOH, pH 7.5, 500 mM KCl, 0.05% (v/v) NP40, 0.5 mM DTT, Roche complete EDTA-free protease inhibitor cocktail) then resuspended in 1 volume of dephosphorylation buffer (50 mM Tris-HCl, pH 7.9, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT). Calf intestinal alkaline phosphatase (CIAP, NEB) was added to a final concentration of 0.5 U/µl, and the suspension was incubated for 10 min at 37°C. Beads were washed twice in 1 ml of phosphatase wash buffer (50 mM Tris-HCl, pH 7.5, 20 mM EGTA, 0.5% (v/v) NP40) then twice in polynucleotide kinase (PNK) buffer without DTT (50 mM Tris-HCl, pH 7.5, 50 mM NaCl, 10 mM $MgCl_2$). Beads were resuspended in one original bead volume of PNK buffer (50 mM Tris-HCl, pH 7.5, 50 mM NaCl, 10 mM $MgCl_2$, 5 mM DTT).

ϒ-32P-ATP was added to the bead suspension to a final concentration of 0.5 µCi/µl and T4 PNK (NEB) was added to a final concentration of 1 U/µl. Samples were incubated for 30 min at 37°C. Non-radioactive ATP (Sigma) was added to obtain a final concentration of 100 µM and samples were incubated for another 5 min at 37°C. Beads were washed 5 times with 800 µl of PNK buffer without DTT and resuspended in 70 µl of SDS-PAGE loading buffer (10% glycerol (v/v), 50 mM Tris-HCl, pH 6.8, 2 mM EDTA, 2% SDS (w/v), 100 mM DTT, 0.1% bromophenol blue).

The radiolabeled suspension was incubated for 5 min in a heat block at 95°C to denature and release the immunoprecipitated RBP with crosslinked RNA and vortexed. Magnetic beads were removed on the separator and the supernatant was transferred to a clean 1.5 ml microfuge tube. 40ul of the supernatant was pipetted into wells of a Novex Bis-Tris 4-12% (Invitrogen) precast polyacrylamide gel and run at 200 V for 55 min. The gel was removed from the plastic casing and laid on a pre-developed photographic film. Glo-gos (Agilent Technologies) were used to facilitate the alignment of the gel to the film. The gel and photographic film were wrapped in plastic film (e.g. Saran wrap) and exposed to Kodak MS film overnight with an intensifying screen at -80°C in an exposure cassette. The gel was aligned with the film and slices corresponding to the expected size of RBP (Ago2, approx 100 kDa) were cut out using sterile scalpels. Gel slices were transferred to D-Tube Dialyzer Midi Tubes with 800 µl 1x SDS running buffer.

The crosslinked RNA-RBP complex was electroeluted in 1x SDS running buffer at 100 V for 2 h according to the instructions of the manufacturer.

The electroeluate (800µl) was pipetted from the Dialyzer Midi Tubes into 2.0 ml microcentrifuge tube tubes. An equal volume of 2x Proteinase K Buffer (100 mM Tris-HCl, pH 7.5, 150 mM NaCl, 12.5 mM EDTA, 2% (w/v) SDS) was added to the electroeluate followed by the addition of Proteinase K (Roche) to a final concentration of 1.2 mg/ml. Samples were incubated for 30 min at 55°C. RNA was recovered by acidic phenol/chloroform (125:24, pH 4.7) followed by a chloroform extraction. At this stage it was necessary to split the sample between multiple 1.5ml microcentrifuge tubes because sample volumes became too large for single tubes to be used. RNA was precipitated overnight with 2.5 volumes of 100% ethanol, 0.1 volume of Sodium Acetate and 1 µl of glycogen (10mg/ml stock). RNA was recovered and resuspended in 6µl dH_2O for downstream analysis.

2.3.2.5 Small RNA Library Preparation

Small RNA libraries were prepared by following the 'Small RNA v1.5 Sample Preparation Guide' (Illumina). Briefly, 5'and 3' adaptor molecules were ligated onto the PAR-CLIP recovered RNA. RNA was precipitated, recovered, reverse transcribed and PCR amplified using primers designed to amplify across the ligated adaptors with Ban1 sites at their termini. The resulting DNA was concatemerised, separated on a 10% TBE PAGE gel (Novex, Invitrogen) and visualised by SYBR Gold (Invitrogen) staining of the gel. Bands corresponding to the insert sequence and ligated adaptors (93bp and above) were cut out from the gel and DNA was eluted overnight then precipitated. DNA was recovered from precipitate and ligated into a pGEM-T Easy Vector (Promega) according to the manufacturer's instructions. JM109 Competent Cells (Promega) were transformed according to the manufacturer's instructions and cells were plated on ampicillin/XGAL/IPTG LB Agar plates and grown overnight at 37°C. Colonies were picked and grown overnight at 37°C (225-250rpm) in 2mls of LB Broth with 100µg/ml ampicillin. Plasmid DNA was recovered by Miniprep (QIAPrep Spin Miniprep Kit, QIAGEN) and sent for sequencing by Dundee Sequencing Services (http://www.dnaseq.co.uk/).

Chapter 3

miRNAs in trophoblast differentiation

3.1 Introduction

miRNAs have been shown to be very important in early mammalian development. Deleting the canonical miRNA-processing pathway in murine embryos leads to early embryonic lethality: Dicer1 null murine embryos arrest development prior to E7.5 (Bernstein *et al.*, 2003) and DGCR8 knockout embryos arrest development prior to E6.5 (Wang *et al.*, 2007). It is therefore likely that miRNAs are involved in regulating individual developmental transitions. A developmental transition is defined here as the transformation of a cell into a different cell state with a concomitant effect on the developmental potential of the cell. In order to investigate what role, if any, that miRNAs play in individual developmental transitions, we wished to: (i) determine if miRNAs are differentially represented during developmental transitions, and (ii) determine if any differentially represented miRNAs play functional roles in the developmental process. In order to answer this question miRNA expression was evaluated through generation and sequencing of miRNA libraries (miRNA profiling) in a cell-based model of an early developmental transition. An *in vitro* model of a developmental transition was used rather than *in vivo* miRNA profiling in order to reduce the complexity of the system being examined and to identify miRNAs that were directly associated with a single developmental process. The ZHBTc4 ES cell system was

chosen not only because it models the earliest obvious differentiation event in the early embryo – that of trophoblast differentiation, but also because this is a regulated system that allows reproducible, homogeneous differentiation of the cultured cells.

3.1.1 The ZHBTc4 model system

Oct4 is a transcription factor that is a known master-regulator of ES cell differentiation; it is expressed at high levels in embryonic stem cells and is downregulated upon ES cell differentiation (Okamoto *et al.*, 1990, Rosner *et al.*, 1990, Scholer *et al.*, 1990). Hitoshi Niwa and others have investigated the effect of missexpressing *Oct4* in mouse ES cells by generating genetically modified ES cell lines. In one of these lines, the ZHBTc4 ES cell line, both endogenous *Oct4* alleles have been inactivated and the cells carry a conditional *Oct4* transgene, which is regulated by tetracycline (Niwa *et al.*, 2000). In standard culture conditions, the ZHBTc4 ES cells express *Oct4* at approximately 60% of normal levels, however, in the presence of tetracycline, the *Oct4* transgene is completely repressed and the cells rapidly lose all *Oct4* expression. Loss of *Oct4* expression leads to rapid differentiation of the ZHBTc4 ES cells. On the basis of the morphological changes and because two trophoblast-associated genes, *Hand1* and *Cdx2*, are upregulated, Niwa *et al.* reported that the cells differentiated into trophectoderm on downregulation of *Oct4*. Furthermore, when the resulting differentiated cells are grown in trophoblast stem cell conditions (Tanaka *et al.*, 1998), trophoblast stem cells are generated that can be passaged repeatedly and are capable of differentiating into trophoblast subtypes upon removal of FGF4.

Since the original publication, ZHBTc4 ES cells have been used to study *Oct4* involvement in ES cell differentiation (Hay *et al.*, 2004), ES cell cycle control (Lee *et al.*, 2010), ES cell survival under stress conditions (Guo *et al.*, 2008) and germ cell specification (Okamura *et al.*, 2008). This cell line has also been used to investigate the capacity of the two human *Oct4* isoforms to confer self-renewal in mouse ES cells (Lee *et al.*, 2006), and the tumorigenicity of the translocation products EWS-Oct4 and EWS-Oct4B (Lee *et al.*, 2007, Kim *et al.*, 2010d).

3.1.2 The role of miRNAs during ES cell differentiation

Maintenance of ES cells relies on suppression of the MAPK signalling pathway or pro-differentiation signals downstream of it, while differentiation of ES cells requires activation of ERK1/2 by FGF4 (Ying *et al.*, 2008). Differentiation of ES cells is associated with a loss of pluripotency characterised by down-regulation of the ES cell master regulators *Oct4* and *Sox2* (Nichols *et al.*, 1998, Avilion *et al.*, 2003), with obvious changes in morphology and proliferation rate, and with alteration to signalling status, fate potential, dependence on growth factors and cell cycle regulation. Differentiation associated changes to the cell cycle include the establishment of a G1 phase, which is largely absent in the rapidly proliferating ES cell (Savatier *et al.*, 1994), and a decrease in the rate of cell cycle progression. It is possible that the cell-cycle properties of ES cells may be functionally important for pluripotency and that acquisition of G1/S regulation may be necessary for differentiation to occur (Burdon *et al.*, 2002). It is hypothesised that miRNAs could be regulating any aspect, or multiple aspects of the differentiation process.

3.2 Aims of Chapter 3

Using the ZHBTc4 ES cell model system, the primary aims of this chapter were to determine if miRNAs were differentially represented in a model of trophectoderm differentiation and if differentially represented miRNAs, if any, were functionally associated with elements of the developmental process.

3.3 Results

Prior to the project start, the ZHBTc4 ES cell system was initially characterised and culture samples for miRNA profiling prepared by T. Burdon. RNA extraction and miRNA library preparation was carried out by D. McBride. DNA sequencing was performed commercially by DNA Sequencing Services (Dundee). Programmes used for analysis of miRNA sequence data were written and tested by W. Carre.

3.3.1 Characterisation of the ZHBTc4 model system

To ensure that our *in vitro* system represented an appropriate model of trophoblast differentiation, the original characterisation of the ZHBTc4 system was repeated and extended. In the ZHBTc4 ES cell culture system (Niwa *et al.*, 2000) *Oct4* is expressed from a doxycycline-regulated transgene as illustrated in *Figure 3.1a*. In this system, both endogenous copies of *Oct4* have been disrupted and addition of doxycycline switches off the *Oct4* transgene, which results in rapid differentiation of the cells. Using protein samples previously prepared by T. Burdon, the expression of OCT4 in undifferentiated and differentiated ZHBTc4 ES cells was investigated. Protein samples from ZHBTc4 ES cells that had been cultured without doxycycline or in the presence of doxycycline for 12, 24, 48, 72 and 96 hours were used for western analysis of OCT4 expression. As can be seen in *Figure 3.1b*, in untreated cells OCT4 protein levels increase slightly between plating (T0) and 12 hours in culture, and thereafter similar OCT4 protein levels are maintained for up to 96 hours. In contrast, in doxycycline treated cells, OCT4 protein levels decrease dramatically to undetectable levels in as little as 12 hours following addition of doxycycline. This means that both suppression of the transgene by doxycycline and the turnover rate of the transgenic protein is extremely rapid in these cells. However, it must be noted that the changes observed in OCT4 protein expression arise from analysis of a single experiment and are therefore not definitive.

ZHBTc4 ES cells were grown without doxycycline or with the addition of doxycycline for up to 96 hours in order to assess their morphology (*Figure 3.2*). Cells grown without doxycycline remained ES cell-like in their morphology, growing very tightly with large nuclei. The cells grown in the presence of doxycycline for 24 hours appeared ES cell-like. After a further 24 hours with doxycycline, cells within colonies became more tightly associated before spreading out and taking on a very different appearance. The boundaries between cells were difficult to distinguish. At the edges of colonies cells appeared very large with flattened morphology and clear to granular cytoplasm. With increasing time, in doxycycline treated cultures the proportion of these large granular cells increased although it is worth noting that even 96 hours after addition of doxycycline there were still colonies of cells with an ES-like morphology. A greater number of floating, dead cells were observed in the cultures grown without doxycycline. This

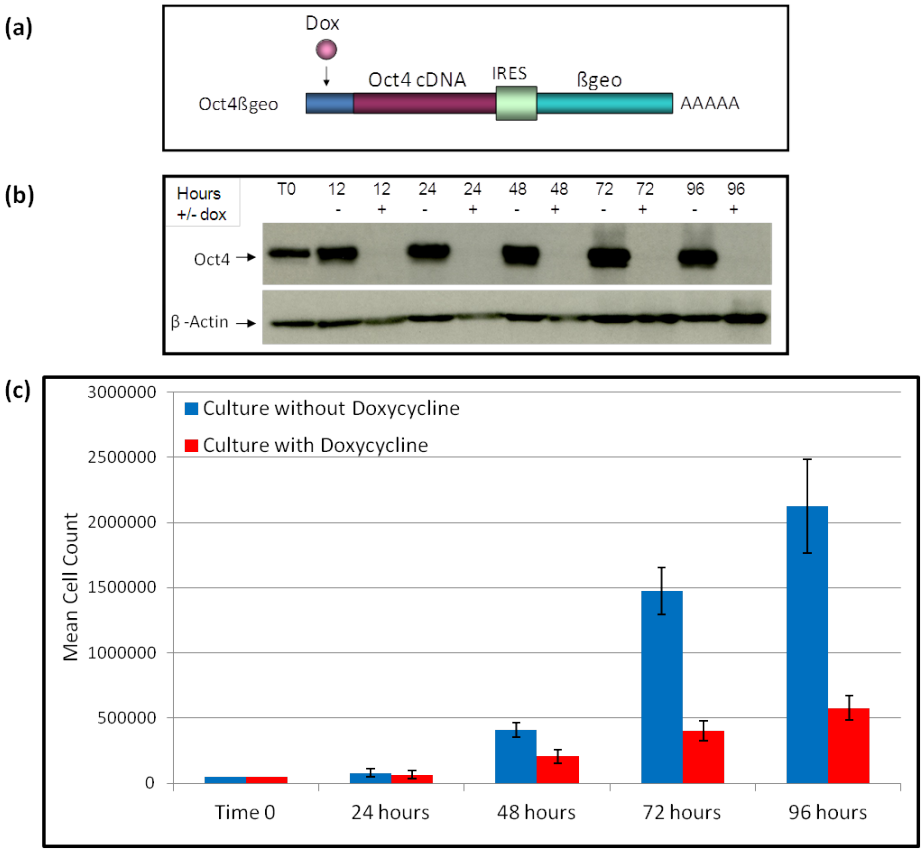


Figure 3.1 – *Characterisation of ZHBTc4 ES cell system.* (a) Schematic diagram illustrating the doxycycline-regulated *Oct4* transgene present in ZHBTc4 ES cells. (b) Western analysis of OCT4 expression in ZHBTc4 ES cells at time of plating (Time 0) and ZHBTc4 ES cells grown without doxycycline (- dox) or in the presence of 1µg/ml doxycycline (+dox) for 12, 24, 48, 72 or 96 hours. Western membrane was re-probed for β-ACTIN expression to estimate protein-loading levels in individual lanes (loading control), n=1. (c) ZHBTc4 ES cells were grown without doxycycline or with doxycycline for 24, 48, 72 and 96 hours and cells were harvested, dissociated and counted by nucleocounter at each timepoint. Time 0 shows the number of cells initially plated (50,000). N=3, error bars show standard deviation.

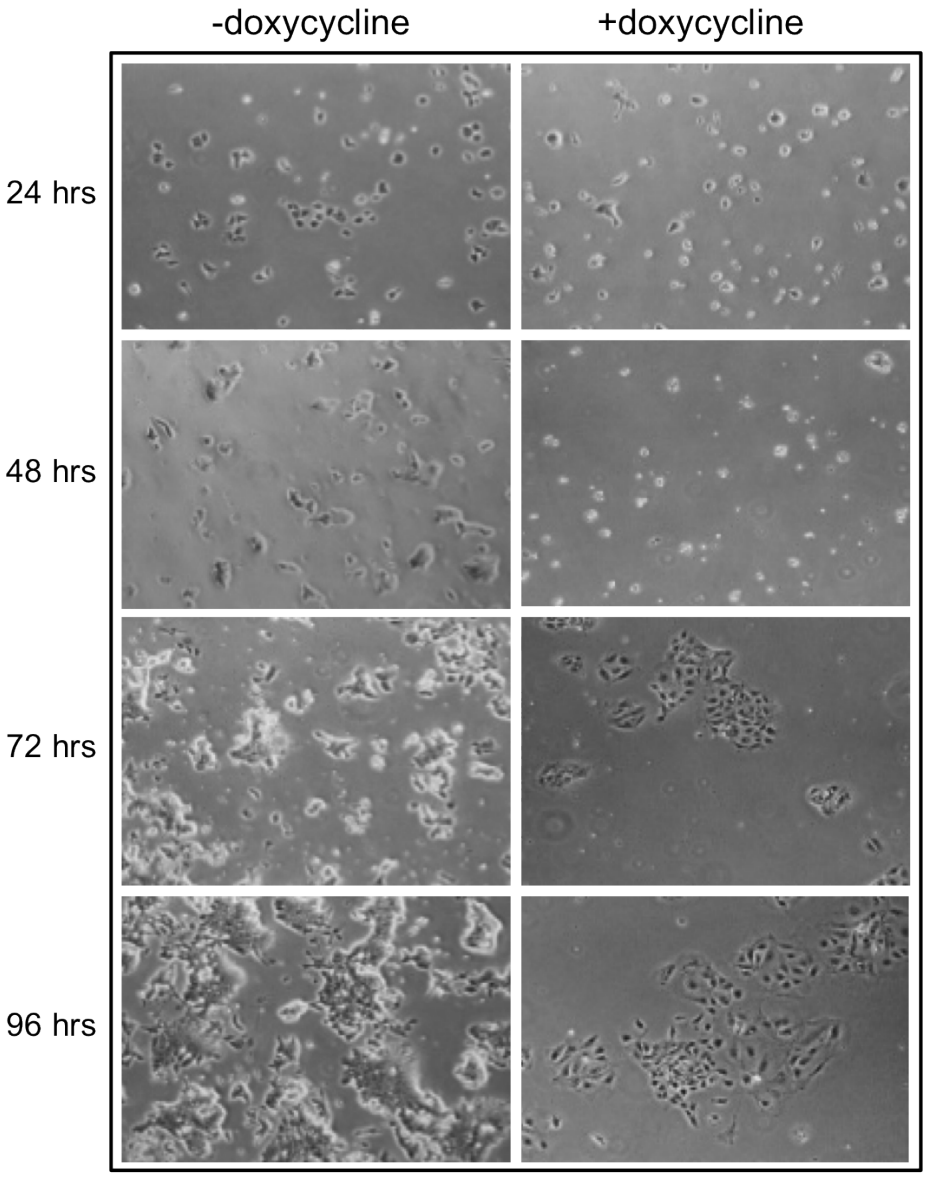


Figure 3.2 – *Morphology of untreated and doxycycline-treated ZHBTc4 ES cells.* ZHBTc4 ES cells were grown without doxycycline (- doxycycline) or in the presence of 1µg/ml doxycycline (+ doxycycline) for up to 96 hours and bright-field photographs (x100 magnification) were taken every 24 hours.

Gene	Full Gene Name	Description
Oct4 (Pou5f1)	POU domain, class 5, transcription factor 1	ES cell maintenance of pluripotency (Nichols et al., 1998)
Sox2	SRY-box containing gene 2	ES cell maintainance of pluripotency (Ivanova et al., 2006)
Nanog	Nanog homeobox	ES cell maintainance of pluripotency (Mitsui et al., 2003)
Esrrb	estrogen related receptor beta	Blocks ES cell differentiation (Loh et al.,2006), Trophoblast differentiation and proliferation (Luo et al., 1997)
Cdx2	Caudal type homeobox 2	Trophectoderm differentiation (Strumpf et al., 2005)
Eomes	Eomesodermin homolog	Trophectoderm differentiation (Strumpf et al., 2005)
Hand1	heart and neural crest derivatives expressed transcript 1	Trophoblast Giant Cell differentiation (Riley et al., 1998)
Mash2 (Ascl2)	achaete-scute complex homolog 2	Spongiotrophoblast and precursors (Guillemot et al., 1994)
Id2	Inhibitor of DNA binding 2	Marker of trophoblast stem cells (Erlebacher et al., 2004)

Table 3.1 – *Panel of genes used as markers of ES cells and trophoblast lineages*

may be explained by the fact that the undifferentiated cultures proliferated more rapidly and became over-confluent. To quantitate the difference in growth rates between undifferentiated cells and those induced to differentiate by the addition of doxycycline, cell counts were performed on untreated and doxycycline treated cultures at various timepoints after plating (T0). *Figure 3.1c* shows that cells grown in the absence of doxycycline proliferate more quickly than cells grown with doxycycline. This is to be expected as ES cells proliferate more rapidly than differentiated cells due to differences in the ES cell cycle, in which G1 control is reduced or absent (Savatier *et al.*, 1994).

It has been reported that addition of doxycycline to ZHBTc4 ES cells results in upregulation of TE marker genes, including *Cdx2* and *Hand1*. To get a more complete picture of the differentiation of these cells, QPCR assays were designed and optimised for a panel of ES cell and TE marker genes (*Table 3.1*). Primer optimisation was carried out on pooled cDNA derived from undifferentiated and differentiated ES cell cultures. Optimisation QPCR was performed on a series of 10-fold cDNA dilutions over a $1x10^6$ fold range. Primers were considered suitable for use if they amplified a single product over >1000-fold range with an efficiency between 95-105% and a 'goodness of fit' (RSq) of >0.95. Primer sequences can be found in Materials and Methods, *Table 3.2* shows the amplification efficiencies, 'goodness of fit' (RSq) and amplification range for each primer pair. Appendix A

Gene	Single Product?	Effective amplification range and Ct spread	RSq	Amplification efficiency (%)
Oct4	Yes	10,000 fold (Ct 22-35.5)	0.994	99.9
Sox2	Yes	100,000 fold (Ct 16-32)	0.995	103.9
Nanog	Yes	100,000 fold (Ct 16-33)	0.995	96.5
EsrrB	Yes	10,000 fold (Ct 22-36)	0.992	100.1
Cdx2	Yes	100,000 fold (Ct 22-38)	0.968	98.5
Eomes	Yes	10,000 fold (Ct 23.5-36)	0.928	106.8
Hand1	Yes	10,000 fold (Ct 22-35.5)	0.988	103.2
Mash2	Yes	1000 fold (Ct 24-33.5)	0.996	98.8
Id2	Yes	1000 fold (Ct 23-33.5)	0.999	94.7
Sfmbt2	Yes	100,000 fold (Ct 20-36)	0.998	101.1
Nfyc	Yes	100,000 fold (Ct 21-38)	0.997	96.8
Odz4	Yes	1000 fold (Ct 27-37)	0.991	96.5
C77370	Yes	1000 fold (Ct 27-37)	0.995	102.3
ActB	Yes	100,000 fold (Ct 16-33)	0.999	101.1

Table 3.2 – *Optimisation details for QPCR primers used in Chapter 3*

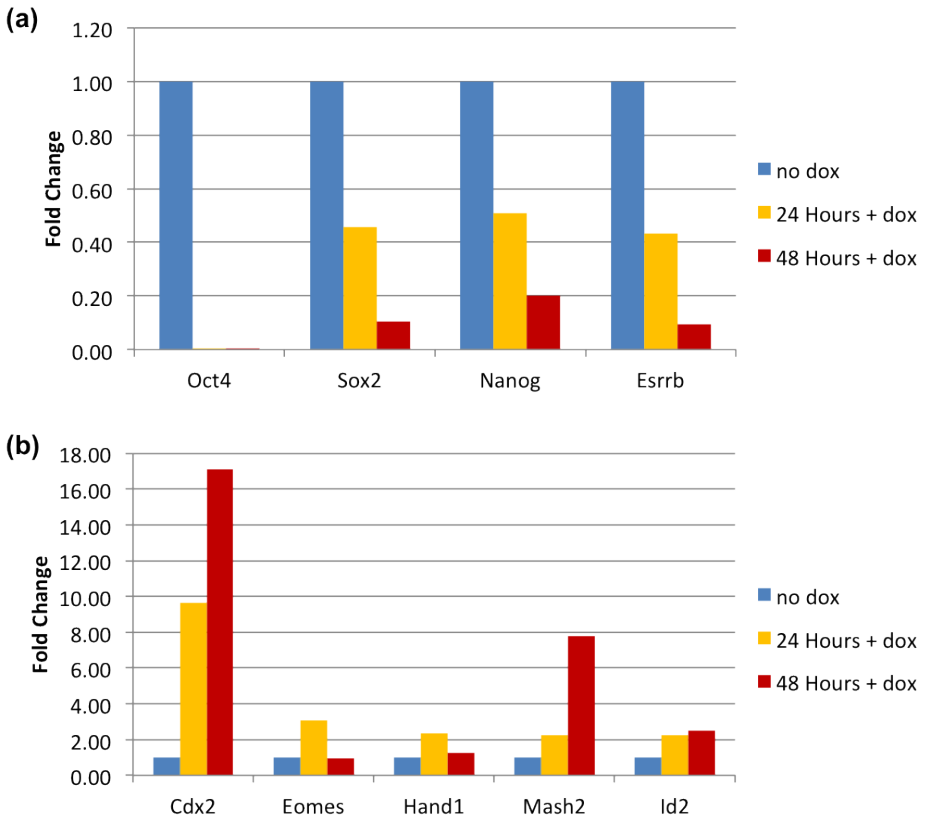


Figure 3.3 – *Analysis of gene expression in undifferentiated and differentiated ZHBTc4 ES cells.* ZHBTc4 ES cells were grown without doxycycline or in the presence of 1μg/ml doxycyline for 24 or 48 hours. RNA was extracted from untreated and doxycycline-treated cell cultures, DNAse treated and used to generate cDNA for QPCR analysis of expression of ES cell marker genes (a) and trophoblast marker genes (b). QPCR Ct values were normalised to *β-Actin* expression. Untreated cultures are represented by a value of 1.0 and results show gene expression in doxycycline treated cultures relative to non-treated cultures at each timepoint. N=1.

contains details of all primer pair optimisations. Gene expression was examined by QPCR 24 and 48 hours after addition of doxycycline in treated cultures as well as untreated cultures at identical timepoints (*Figure 3.3a & b*). Addition of doxycycline resulted in decreased expression of markers of pluripotency (*Oct4*, *Sox2*, *Nanog*) and increased expression of known markers of trophectoderm lineages (*Cdx2*, *Eomes*, *Hand1*, *Mash2*, *Id2*). Interestingly *Esrrß* was designed as a marker of trophectoderm differentiation because mice deficient for this protein exhibit placental defects (Luo *et al.*, 1997), but its expression pattern suggested that higher levels were associated with undifferentiated ES cells. Conclusions about expression of pluripotency and trophectoderm marker genes in the ZHBTc4 ES cell system are preliminary because they arise from a single experimental analysis.

3.3.2 miRNA profiling during the ES:trophoblast transition

The ZHBTc4 system was used to profile miRNAs present in ES cells and TE cells in order to identify miRNAs present in this cell type and those that were differentially regulated as a result of the transition of ES cells into trophoblast. For miRNA profiling analysis, T. Burdon carried out the tissue culture, D. McBride prepared the miRNA libraries and W. Carré wrote Bioinformatics programs for analysis of the miRNA profiling output. RNA was extracted from cultures grown with or without doxycycline for 0, 12, 24, 48, 72 and 96 hours. For the generation of miRNA libraries, the miRNA fraction was purified from total-RNA collected from the 0, 24 and 48 hour timepoints of undifferentiated culture samples, and from total-RNA collected from the 24, 48 and 72 hour timepoints of differentiating culture samples (*Figure 3.4a*). The resulting miRNA populations were cloned and sequenced as illustrated in *Figure 3.4a*. The bioinformatics programs performed the following functions: individual short RNA sequences were extracted from vector sequence and linker sequences; extracted sequences were searched in miRBase to identify homology with known miRNAs; results were combined and the data was sorted to show the number of times each miRNA was sequenced from the different libraries; the mouse genome was searched for homology to short RNA sequences and if a match was found, a portion of the upstream and downstream DNA sequence was extracted and analysed for the ability to form a hairpin structure.

200 clones were sequenced from each of the six different libraries, which resulted

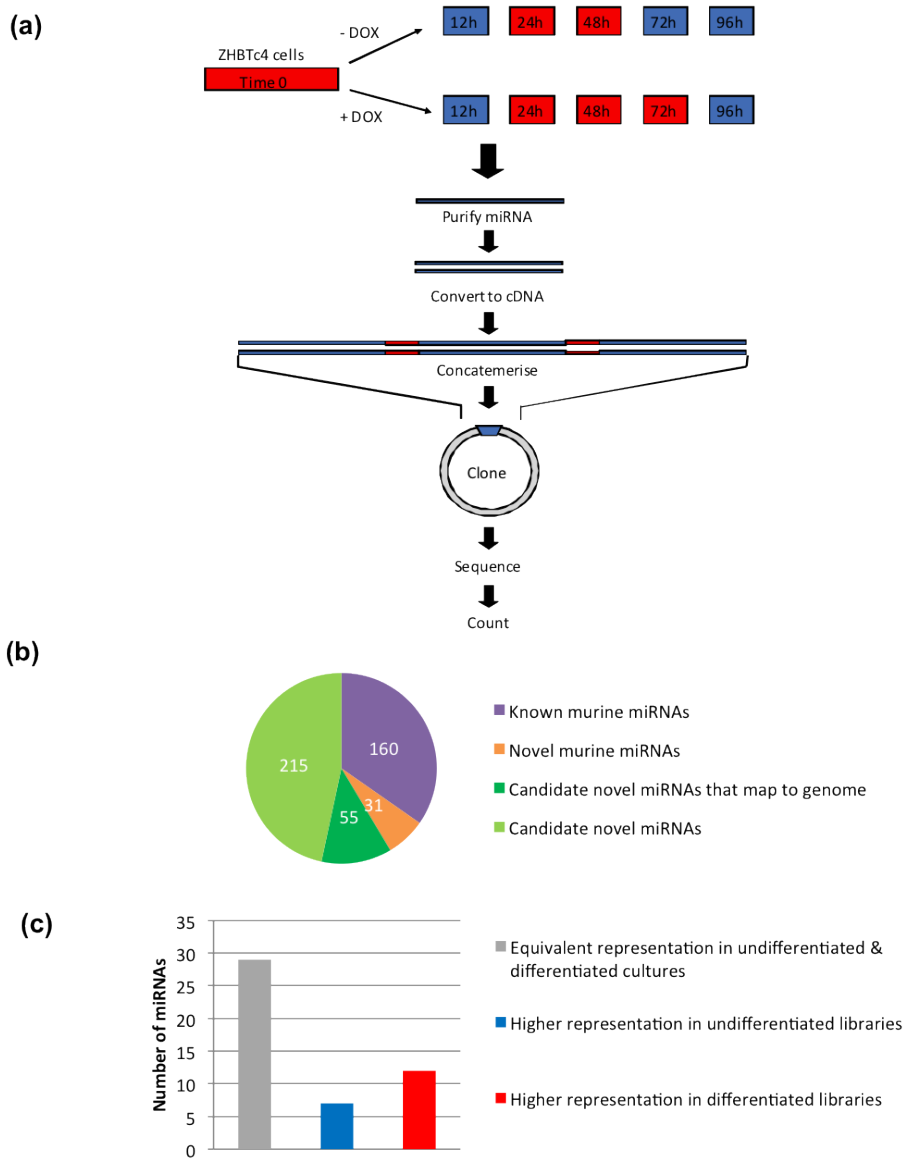


Figure 3.4 – *miRNA profiling in undifferentiated and differentiated ZHBTc4 ES cells.* (a) Schematic illustration of miRNA profiling method. ZHBTc4 ES cells were grown either without doxycycline (- DOX) or in the presence of 1µg/ml doxycycline (+ DOX) for 12, 24, 48, 72 or 96 hours. RNA was extracted from the starting population of undifferentiated ZHBTc4 ES cells (Time 0) and the timepoints indicated. MiRNAs were purified from the timepoints shown in red boxes. Following miRNA purification, RNA was converted to cDNA, concatemerised, cloned and sequenced. (b) Pie chart showing the proportions of total miRNA sequences that represent previously known mouse miRNAs, novel miRNAs and candidate novel miRNAs. (c) Bar chart showing the number of miRNAs with equivalent representation in undifferentiated and differentiated miRNA libraries and the number of miRNAs represented at higher levels in undifferentiated ZHBTc4 ES cells or differentiated cells.

in a total of 5683 individual short-RNA sequences. Individual library clones encoded concatemerised sequences representing an average of 4/5 miRNA sequences per clone. Bioinformatics analysis revealed that 461 different individual sequences had been identified, of which 160 had been previously been recorded as mouse miRNAs. Of the remaining 301 sequences, which are potentially novel miRNAs, 86 map to the published mouse genome and of these, 31 map in a region that folds appropriately into a hairpin structure as defined by RNA-fold (*Figure 3.4b*). These 31 sequences are likely to represent novel miRNAs. The remaining 270 sequences that do not match a region in the mouse genome or fold into a hairpin structure are 'candidate' miRNA sequences. Full miRNA profiling data is available in Appendix B.

3.3.3 Analysis of miRNA profiling results

Differentially represented miRNAs were defined as miRNAs recorded >10 instances and differing by >2-fold between undifferentiated and differentiated cultures. By the criteria adopted, 19 miRNAs were differentially regulated as a result of trophoblast induction (*Figure 3.4c* & *Table 3.3*). Of these, the levels of 12 miRNAs increased and the levels of 7 miRNAs decreased as a result of differentiation. Increases in representation ranged from 2-fold to 11-fold and decreases in representation ranged from 2-fold to 8-fold. Most of the differentially represented miRNAs were previously known, but one miRNA that was observed to be represented more highly in differentiated libraries than undifferentiated libraries, *mES2*, was novel. *mES2* was represented at 5-fold higher levels in differentiated libraries compared to undifferentiated libraries and was encoded within intron 10 of *Sfmbt2*. This suggests that *mES2* is a novel miRNA that is part of the *Sfmbt2* miRNA cluster.

To confirm that the profiling data reflected gene expression, a subset of the differentially represented miRNAs was selected for northern analysis. RNA from ZHBTc4 ES cells grown without doxycycline or in the presence of doxycycline for various lengths of time was used for miRNA northern analysis of *miR-30C*, *miR-669b*, *mES2* and *miR-467a** expression (*Figure 3.5a*). Phosphoimager analysis of miRNA northern membranes was used to quantitate miRNA expression in undifferentiated and differentiated ZHBTc4 ES cells (*Figure 3.5b*).

miRNA name	# of sequences in undifferentiated libraries	# of sequences in differentiated libraries	Approximate Fold Change	Genomic Location
'UP' miRNAs (representation increases on differentiation)				
mmu-miR-467a-1*	2	21	+11	Sfmbt2
mmu-miR-23a	2	14	+7	Intergenic
mmu-miR-467c	2	12	+6	Sfmbt2
mES1 (miRNA_plate7_ES-T24_71-1)	2	9	+5	Sfmbt2
mmu-miR-467e	3	14	+5	Sfmbt2
mmu-miR-669d	2	9	+4	Sfmbt2
mmu-miR-297a	3	9	+3	Sfmbt2, Intergenic, Sobp201
mmu-miR-191	4	11	+3	intergenic
mmu-miR-669a	4	9	+3	Sfmbt2
mmu-miR-99b	6	15	+2	intergenic
mmu-miR-210	4	9	+2	intergenic
mmu-miR-669b	4	8	+2	Sfmbt2
'DOWN' miRNAs (representation decreases on differentiation)				
mmu-miR-672	8	1	-8	C77370
mES2 (miRNA_plate7_ES-T24_85-1)	15	3	-5	intergenic
mmu-miR-708	19	5	-4	Odz4
mmu-miR-30c	11	3	-3	Intergenic, Nfyc
mmu-miR-124	16	6	-3	Intergenic
mmu-miR-92a	16	7	-2	Intergenic, Mirhg1
mmu-miR-106a	10	5	-2	Kis2
mmu-miR-19a	12	6	-2	Mirhg1

Table 3.3 – *miRNAs identified as being differentially represented during trophectoderm differentiation of ZHBTc4 ES cells.* The total number of times a miRNA was sequenced is shown for differentially represented miRNAs in ZHBTc4 ES cell cultures grown without doxycycline ('Undifferentiated' column) or with doxycycline ('Differentiated' column). The total number of sequencing events is pooled from 3 undifferentiated libraries and 3 differentiated libraries after normalisation for library size. 'Fold Change' is the ratio (to the nearest whole number) of the number of sequencing events in differentiated libraries compared to undifferentiated libraries. The genomic location column shows if the miRNA is intergenic or encoded within a protein-coding gene.

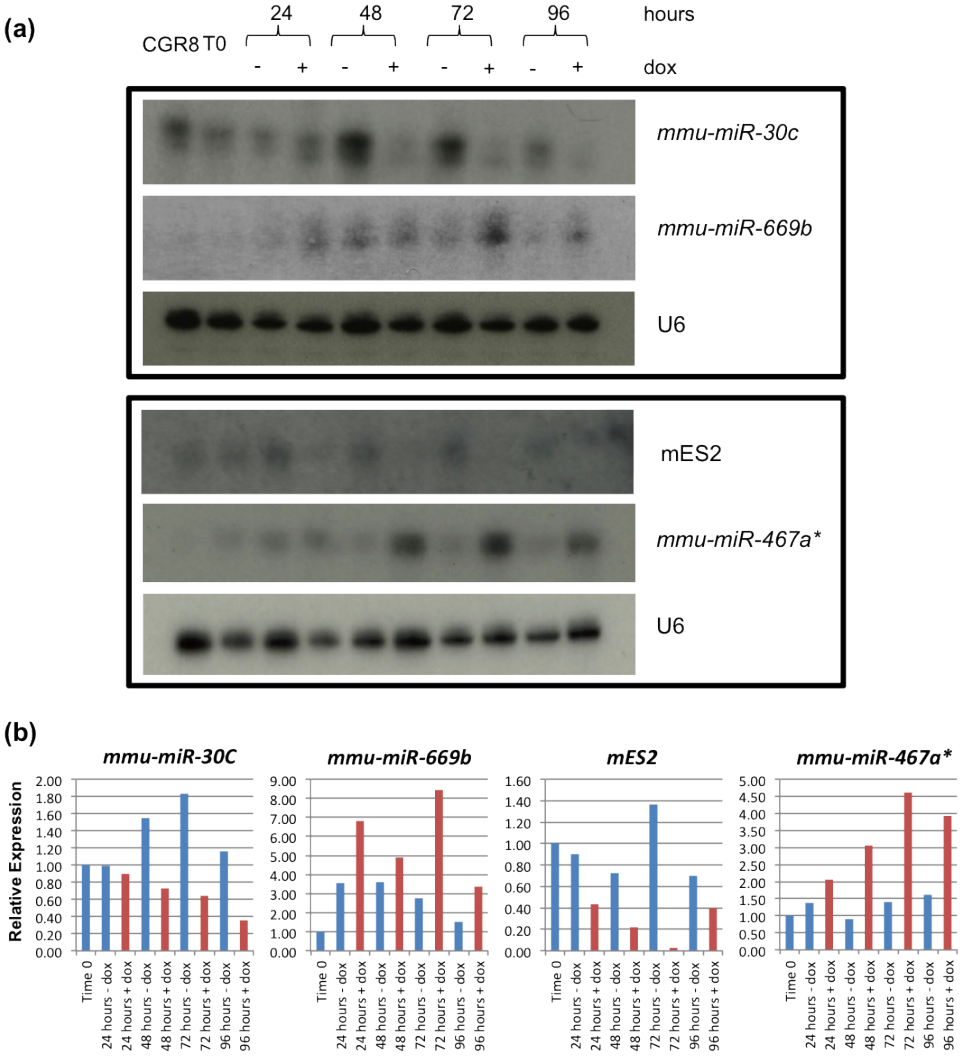


Figure 3.5 – *Expression analysis of miRNAs in undifferentiated and differentiated ZHBTc4 ES cell cultures.* (a) Total RNA (5µg) from ZHBTc4 ES cells at the time of plating (T0) or ZHBTc4 ES cells grown without doxycycline (-dox) or in the presence of 1µg/ml doxycycline (+dox) for 24, 48, 72 or 96 hours was used for miRNA northern analysis. Expression of *miR-30C*, *miR-669b*, *miR-467a**, the novel miRNA, *mES2*, and the spliceosomal RNA, U6, is shown. (b) miRNA northern membranes were exposed to a phosphoimager screen for quantitation of expression of each miRNA. Values for each sample were normalised against values obtained for U6 and then calculated relative to the T0 sample. N=1.

Expression of *miR-30C* was observed to increase at 48 and 72 hours in undifferentiated ZHBTc4 ES cells. In differentiated cultures however, *miR-30C* expression decreased gradually from 24 hours to 96 hours after addition of doxycycline. The overall miRNA profiling showed a 3-fold decrease in representation of *miR-30C* in differentiated ZHBTc4 libraries compared to undifferentiated libraries. Phosphoimager quantitation showed an approximately 2-fold decrease in expression at 48 hours and a 3-fold decrease in expression at 72 and 96 hours.

Expression of *miR-669b* was observed to increase at 24, 48 and 72 hours in undifferentiated ZHBTc4 ES cells relative to expression at T0. In differentiated cultures, *miR-669b* expression was higher than in undifferentiated cultures at all timepoints. Differences in expression levels between undifferentiated and differentiated samples at different timepoints varied between approximately 1.4 and 2.6-fold. For comparison, overall miRNA-profiling analysis showed 2-fold higher representation of *miR-669b* in differentiated libraries compared to undifferentiated libraries.

Expression of the novel miRNA, *mES2* was seen to stay at approximately constant levels in undifferentiated ZHBTc4 ES cells. However, in differentiated cultures *mES2* expression was lower in differentiated samples compared to undifferentiated samples at all timepoints with a dramatic difference in expression levels observed 72 hours after doxycycline addition. The miRNA profiling data showed a 5-fold decrease in representation of *mES2* in differentiated ZHBTc4 libraries compared to undifferentiated libraries. Phosphoimager quantitation showed a decrease in expression of between approximately 45-fold (at 72 hours) and 1.75-fold (at 96 hours).

Expression of *miR-467a** was observed to stay at an approximately constant level in undifferentiated ZHBTc4 ES cells relative to expression at T0. In differentiated cultures, *miR-467a** expression was higher than in undifferentiated cultures at all timepoints and showed an increase in expression from 24 hours to 72 hours after addition of doxycycline. Differences in expression levels between undifferentiated and differentiated samples at each timepoint varied between approximately 1.4-fold at 24 hours and 3-fold at 72 hours. miRNA-profiling analysis showed 11-fold higher representation of *miR-467a** in differentiated libraries compared to undifferentiated libraries.

Analysis of expression of several miRNAs predicted to be differentially represented

by the miRNA-profiling data confirmed differential expression as predicted by the profiling data. Furthermore, for three out of four miRNAs the differences in expression levels observed between undifferentiated and differentiated samples were broadly similar to the differences in representation levels from the miRNA profiling data. However, while the expression analysis described here appears to support the miRNA profiling data, it must be noted that these data are from a single experimental analysis and are therefore not definitive.

3.3.4 Analysis of binding of miRNA promoters by ES cell transcription factors

The possibility that differentially represented miRNAs identified in this study may be known to be regulated by factors important in the self-renewal of ES cells was investigated. In a study investigating how the key transcriptional regulators of ES cells regulate miRNA gene expression, Marson *et al.* identified miRNA promoters that contained binding sites for *Oct4*, *Sox2*, *Nanog* and *Tcf3*, as well as miRNA promoters that bind SUZ12 and carry the H3K27me3 chromatin marks associated with repressed genes (Marson *et al.*, 2008). The Marson *et al.* dataset was examined to determine whether the promoters of differentially represented miRNAs identified in this study were bound by these key transcription factors (*Table 3.4*).

Of the 7 miRNAs that decreased in representation on differentiation, 6 have binding sites for *Oct4* within their promoter regions and of these 4 are bound by all 4 transcription factors analysed. This suggests that expression of these miRNAs may be directly regulated by key ES cell transcription factors. In addition, the *Sfmbt2* gene promoter region contained predicted binding sites for repressor proteins, consistent with the observation that the representation of 8 of the miRNAs encoded within this cluster is lower in ES cells than in differentiated cultures. *Oct4* and *Sox2* bound two of the miRNAs that were upregulated during differentiation. These transcription factors can act as repressors as well as activators of transcription depending on context so perhaps these miRNAs are being repressed by *Oct4* and *Sox2* in ES cells.

miRNA name	Scoring	Binding					
	Conserved	Oct4	Sox2	Nanog	Tcf3	Suz12	H3K27me3
***Sfmbt2* miRNA cluster**							
*mmu-miR-467a-1**		-	-	-	-	Suz12	K27me3
mmu-miR-467c		-	-	-	-	Suz12	K27me3
mmu-miR-467e		-	-	-	-	Suz12	K27me3
mmu-miR-669d		-	-	-	-	Suz12	K27me3
mmu-miR-297a		-	-	-	-	Suz12	K27me3
mmu-miR-669a		-	-	-	-	Suz12	K27me3
mmu-miR-669b		-	-	-	-	Suz12	K27me3
UP' miRNAs (representation increases on differentiation)							
mmu-miR-23a		Oct4	Sox2	-	-	-	-
mmu-miR-191		Oct4	Sox2	-	-	-	-
mmu-miR-99b		-	-	-	-	-	-
mmu-miR-210		-	-	-	Tcf3	-	-
DOWN' miRNAs (representation decreases on differentiation)							
mmu-miR-672		Oct4	Sox2	Nanog	Tcf3	-	-
mmu-miR-708	Cons.	Oct4	Sox2	Nanog	-	Suz12	K27me3
mmu-miR-30c	Cons.	-	-	-	-	-	-
mmu-miR-124	Cons.	Oct4	Sox2	Nanog	Tcf3	Suz12	K27me3
mmu-miR-92a		Oct4	Sox2	Nanog	Tcf3	-	-
mmu-miR-106a		Oct4	Sox2	Nanog	Tcf3	-	-
mmu-miR-19a		Oct4	-	-	-	-	-

Table 3.4 – *Binding of miRNA promoters by ES cell factors.* Table shows if the promoters of differentially represented miRNAs are bound by ES cell factors (*Oct4*, *Sox2*, *Nanog* and *Tcf3*) or the polycomb group protein (SUZ12) or modified by the repressive chromatin mark H2K27H3. Data from the miRNA profiling analysis was matched to promoter binding data from Marson *et al.* (2008). MiRNAs encoded in more than one genomic location (eg. *miR-30C*) are listed using miRBase designation and promoter-binding data for each location is shown.

3.3.5 Expression analysis of genes containing differentially represented miRNAs

Several of the differentially represented miRNAs were encoded within introns of others genes (see *Table 3.3*). In the majority of such cases it is generally thought that the miRNA is transcribed as part of the mRNA transcript (Baskerville & Bartel, 2005) although there are cases where separate miRNA transcripts are encoded within introns. It might be expected that the genes containing differentially represented miRNAs would be differentially expressed during the ES-trophoblast transition. In order to examine this possibility ZHBTc4 ES cells were grown in the presence of doxycycline for various lengths of time and RNA was extracted, DNAse treated and converted to cDNA. QPCR analysis was performed on cDNA from differentiating ZHBTc4 ES cells in order to analyse the expression of genes containing differentially represented miRNAs. QPCR analysis was also performed on undifferentiated cells from corresponding timepoints to the differentiating cells. This analysis was performed as a single experiment and therefore conclusions drawn from the results are preliminary.

Expression of *Sfmbt2* (which contains a large cluster of miRNAs including *miR-467a** and *miR-669b*) gradually increased throughout the 96-hour period of trophectoderm differentiation, reaching levels that were 10-fold higher in differentiated ZHBTc4 ES cultures than in undifferentiated cultures (*Figure 3.6a*). miRNA-profiling analysis showed differential representation of miRNAs from the *Sfmbt2* cluster, which varied between, 2-fold and 11-fold.

The expression level of *C77370* (which encodes *miR-672*) decreased rapidly on differentiation, falling approximately 30-fold over the 48 hours following addition of doxycycline to ZHBTc4 ES cells (*Figure 3.6b*). This is a more dramatic decrease in expression than was observed for *miR-672*, which was 8-fold lower in differentiated cells compared to undifferentiated cells according to the miRNA-profiling data. *C77370* is an expressed sequence, which is believed to be protein coding, but the function of which is not known.

Expression of *Odz4* (which encodes *miR-708*) also decreased during trophectoderm differentiation, falling approximately 10-fold during the first 72 hours of differentiation (*Figure 3.6c*). The level of *miR-708* was 4-fold lower in differentiated libraries compared to undifferentiated libraries according to the miRNA-profiling

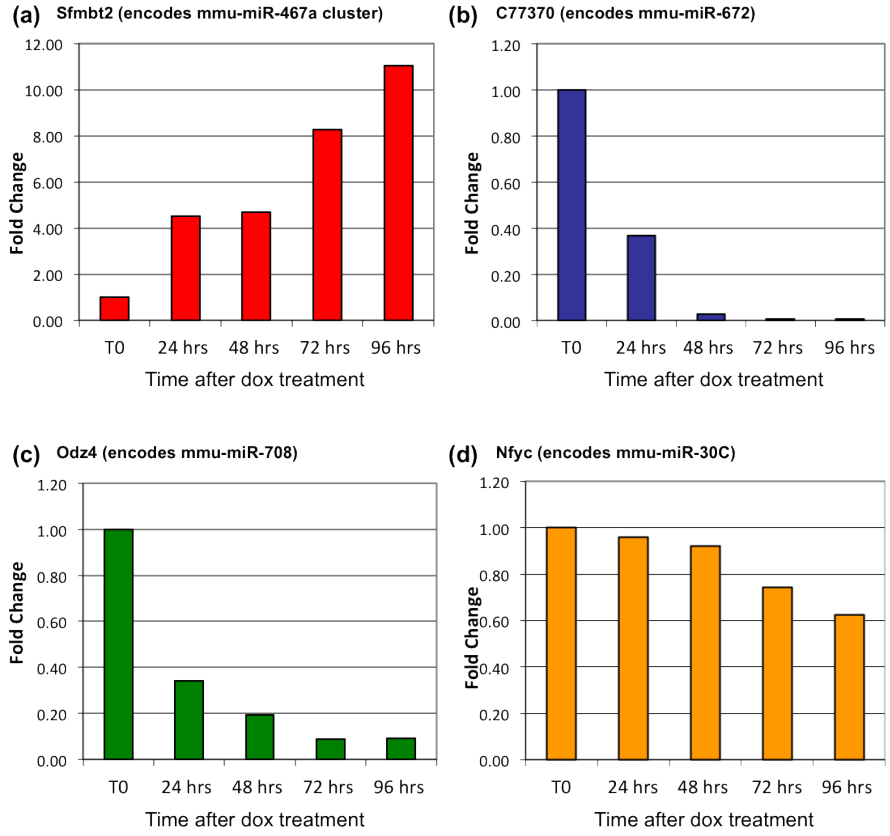


Figure 3.6 – *Expression analysis of genes containing miRNAs.* RNA was harvested from ZHBTc4 ES cells (T0) and ZHBTc4 ES cells that were grown in the presence of 1μg/ml doxycycline for 24, 48, 72 or 96 hours. RNA was DNAse treated, converted to cDNA and used for QPCR analysis of expression of *Sfmbt2* (a), *C77370* (b), *Odz4* (c) and *Nfyc* (d) in differentiating ZHBTc4 ES cells. QPCR Ct values were normalised to *β-Actin* expression and expression is shown relative to the expression in ZHBTc4 ES cells (T0). N=1.

data. *Odz4* (also known as *Ten4*) is a member of a transmembrane protein family that have the ability to form homo and heterodimers and may be involved in intercellular signalling during development (Feng *et al.*, 2002, Tucker & Chiquet-Ehrismann, 2006). *Odz4* is expressed during development and is essential for mouse gastrulation (Lossie *et al.*, 2005).

The patterns of expression of *Sfmbt2*, *C77370* and *Odz4* mirrored the representation patterns seen for the miRNAs encoded within them although the extent of the differential expression varied. In contrast, *Nfyc* (which is thought to contain *miR-30C*) was not differentially expressed (*Figure 3.6d*). *miR-30C* was confirmed as being differentially expressed by miRNA northern analysis (see *Figure 3.5a* & *b*) so the results from the expression analyses do not agree. The disagreement between the expression patterns of *Nfyc* and *miR-30C* may be explained by the fact that *miR-30C* is found at two different locations within the mouse genome, one within *Nfyc* and the other in an intergenic region. The differential expression seen here suggests that the *miR-30C* that is detected is expressed from the location on Chromosome 1 rather than within the *Nfyc* gene on Chromosome 4. Alternatively, *miR-30C* may have its own promoter and be processed as a separate transcript to *Nfyc*. *Nfyc* encodes one subunit of a trimeric complex forming a highly conserved transcription factor and has been shown to suppress expression of the developmentally important hox gene *egl-5* in *C. elegans* (Sinha *et al.*, 1996, Deng *et al.*, 2007). The finding that genes encoding miRNAs follow the same pattern as the miRNA means that more information can be extracted from miRNA profiling than the representation of the miRNAs themselves. These types of analyses may not only identify differentially represented miRNAs, but differentially expressed protein-coding genes as well.

3.3.6 Expression analysis of *Sfmbt2*

Of the 12 miRNAs upregulated on differentiation, 8 are contained within a single multi-miRNA cluster located in intron 10 of the Polycomb group gene, *Sfmbt2* (Scm (sex comb on midleg)-like with four mbt (malignant brain tumour repeat) domains 2) (*Figure 3.7a*).

Figure 3.6a shows that *Sfmbt2* expression mirrors that of the miRNA cluster contained within it. *Sfmbt2* expression increased during trophoblast differentia-

tion of ZHBTc4 ES cells. In order to confirm that upregulation of *Sfmbt2* was associated with a reduction in *Oct4* expression and also to assess the effect of an alternate method of *Oct4* depletion on expression of ES cell and TE marker genes, gene expression was examined after siRNA knockdown of *Oct4* expression in CGR8 ES cells (the parental line of ZHBTc4 ES cells). As a control for the transfection procedure, cells were treated with an siRNA against EGFP and expression of marker genes was analysed. RNA available in the laboratory (prepared by T. Burdon) was DNAse treated and used to generate cDNA for QPCR analysis of gene expression.

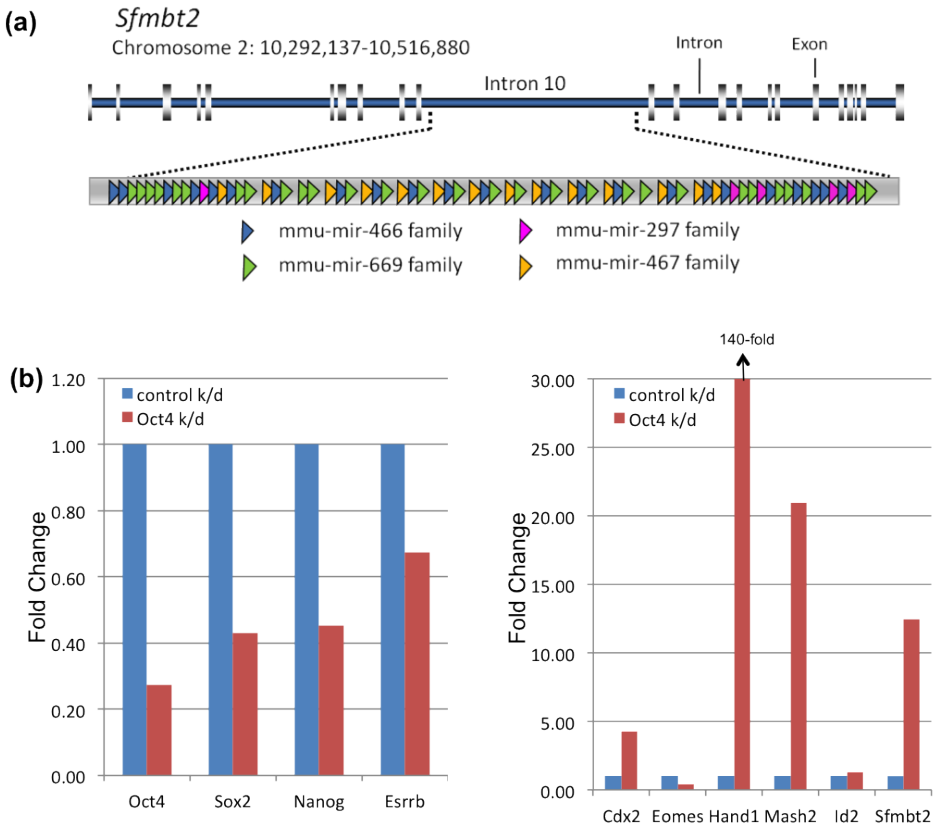


Figure 3.7 - *Schematic diagram of Sfmbt2 and encoded miRNA cluster and expression analysis of Sfmbt2* (continued on next page).

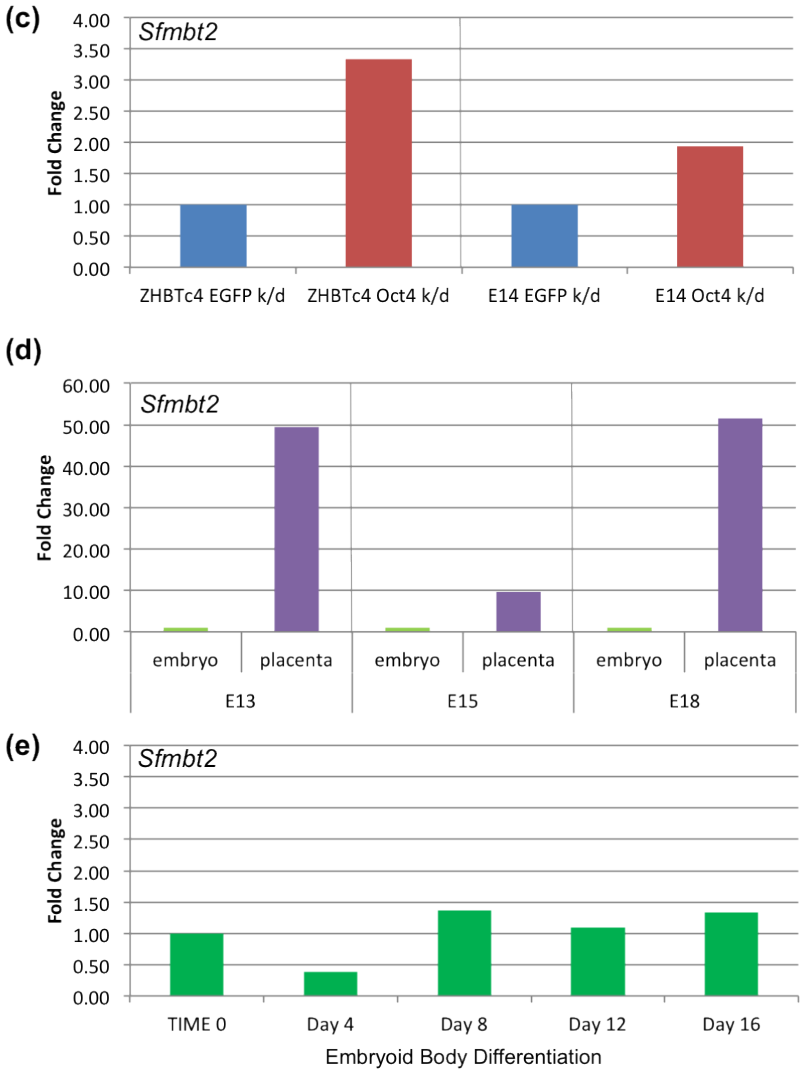


Figure 3.7 – *Schematic diagram of Sfmbt2 and encoded miRNA cluster and expression analysis of Sfmbt2.* (a) Schematic diagram of *Sfmbt2* gene showing introns as blue bars and exons as grey bars. The *Sfmbt2* miRNA cluster is encoded in intron 10 and is composed of 71 miRNAs belonging to 4 miRNA families. (b-c) CGR8 ES cell cultures (b) or ZHBTc4 and E14 ES cell cultures (c) were transfected with siRNAs designed to target *Oct4* or *EGFP* (control). 72 hours after transfection RNA was extracted, DNAse treated, converted to cDNA and used for QPCR analysis of gene expression. QPCR Ct values were normalised to *β-Actin* levels and expression is shown relative to expression of cultures transfected with siRNA targeting EGFP. N=1. (d) RNA from the embryonic or extra-embryonic (placenta) compartments of E13, E15 and E18 mouse embryos was DNAse treated, converted to cDNA and used for QPCR analysis of *Sfmbt2* expression. QPCR Ct values were normalised to *β-Actin* levels and expression in each extra-embryonic sample is shown relative to expression in the corresponding embryonic sample. N=1. (e) E14 ES cells (T0) were used to form embryoid bodies and RNA was extracted after 4, 8, 12 or 16 days of differentiation (Day 4, Day 8, Day 12, Day 16). RNA was DNAse treated, converted to cDNA and used for QPCR analysis of *Sfmbt2* expression. QPCR Ct values were normalised to *β-Actin* and expression at each timepoint of EB differentiation is shown relative to expression at T0. N=1.

Oct4 levels were reduced by >50% following siRNA knockdown of *Oct4* (*Figure 3.7b*). The decrease in *Oct4* expression after siRNA knockdown of *Oct4* is not as dramatic as that seen in the ZHBTc4 ES cell system. This is to be expected because not every cell will be transfected with siRNA targeting *Oct4* and also because the transcription of *Oct4* has not been prevented in these cells, as it is in the doxycycline-treated ZHBTc4 ES cell cultures. Following inhibition of *Oct4* expression by siRNA, ES cell marker genes decreased in expression and, *Cdx2*, *Hand1* and *Mash2* marker genes increased in expression (*Figure 3.7b*) similar to gene expression in doxycycline-treated ZHBTc4 ES cell cultures. The downregulation of ES cell markers is not as dramatic as in the ZHBTc4 ES cell system and neither is the upregulation of *Cdx2*, *Eomes* and *Id2*. Interestingly the increase in the level of *Hand1*, a marker of trophoblast giant cells, is dramatically greater after siRNA knockdown of *Oct4* than in the ZHBTc4 system. While this is difficult to rationalise in the context of trophectoderm differentiation, it should be noted that *Hand1* is not only a factor that promotes trophoblast giant cell differentiation, but is also involved in the differentiation of extra-embryonic mesoderm and during heart development (Firulli *et al.*, 1998). Additionally, the siRNA knockdown samples were analysed 72 hours after *Oct4* knockdown but the ZHBTc4 ES cell samples were analysed 24 and 48 hours after knockdown. This may also account for a higher level of *Hand1*, which is a marker of later trophoblast differentiation. Overall, these data show that changes in gene expression after knockdown of *Oct4* expression in ES cells are broadly similar following siRNA knockdown in wt ES cells or doxycycline addition to the ZHBTc4 ES cell system. However, it must be pointed out that these data arise from a single experimental replicate and experimental repeats would be required to render conclusions about gene expression definitive.

The expression of *Sfmbt2* increased after siRNA knockdown of *Oct4* expression compared to knockdown of EGFP expression. In order to confirm that upregulation of *Sfmbt2* was associated with a reduction in *Oct4* expression, *Sfmbt2* expression was examined in two different ES cell lines after siRNA knockdown of *Oct4* expression compared to siRNA knockdown of EGFP expression. ES cells were plated and grown for 24 hours followed by two overnight transfections (24 hours apart) with siRNAs predicted to target *Oct4* or EGFP. 72 hours after the initial transfection, RNA samples were harvested, DNAse treated and used to generate cDNA for QPCR analysis of *Sfmbt2* expression. It was observed that

Sfmbt2 expression increased following *Oct4* knockdown by siRNA relative to a control knockdown. This effect was observed in both ES cell lines (*Figure 3.7c*) although data was from a single experiment in each cell line so conclusions are not definitive. These data suggest that *Sfmbt2* expression increases as a result of *Oct4* downregulation in the ZHBTc4 ES cell system and after siRNA knockdown of *Oct4* expression. This effect could be due to direct regulation of *Sfmbt2* by *Oct4* as there is a putative *Oct4* binding site within the promoter of *Sfmbt2* (Loh *et al.*, 2006).

To assess if *Sfmbt2* (and by association its encoded miRNA cluster) is associated with the TE *in vivo* its expression was assessed in mouse embryos. TE contributes to the placenta so, using material available within the laboratory, *Sfmbt2* expression was assessed in placentas and associated embryos (*Figure 3.7d*). Embryos and placentas were collected from pregnant mice at E13, E15 and E18 and RNA was extracted by S. Butterwith and stored as an ethanol precipitate. RNA was recovered from precipitate, DNAse treated and used to generate cDNA. QPCR analysis of *Sfmbt2* expression was performed on cDNA generated from embryos and associated placentas. As can be seen from *Figure 3.7d*, at E13, E15 and E18 *Sfmbt2* is expressed between 10 and 50 fold higher in placental tissue than in embryonic tissue. It is likely then that *Sfmbt2* and its encoded miRNA cluster are functionally associated with TE in early development although because these data arise from a single experiment this conclusion is not definitive.

To investigate if *Sfmbt2* upregulation is a general feature of differentiation or if it is specific to TE differentiation the expression of *Sfmbt2* was investigated during embryoid body differentiation of E14 ES cells. E14 ES cells were plated in serum media without the addition of Lif in non-adherent culture conditions to stimulate EB formation and were grown in non-adherent culture conditions for 8 days. On day 8 of differentiation EBs were transferred to adherent culture conditions and cultured for a further 8 days. RNA was extracted from undifferentiated ES cells (Time 0) and on days 4, 8, 12 and 16 of differentiation. RNA was DNAse treated and used to generate cDNA for QPCR analysis of *Sfmbt2* expression. Expression of *Sfmbt2* remained broadly similar during embryoid body differentiation of E14 ES cells (*Figure 3.7e*). Because embryoid bodies do not readily differentiate into TE lineages (Weitzer, 2008) this finding suggests that the upregulation of *Sfmbt2* is specific to TE differentiation. Further experimental repeats would be required to confirm this preliminary conclusion.

3.3.7 Expression analysis of *miR-92a* and related miRNAs

Members of the *miR-17-92a* cluster and the sequence-related *miR-106-363* cluster of miRNAs were more highly represented in ES cells than in trophoblast cells. *miR-92a* is encoded within both clusters (*Figure 3.8a*) and is related in sequence to *miR-25* and *miR-363* (*Figure 3.8b*), sharing an identical seed sequence with these miRNAs. Of the 3 differentially represented miRNAs encoded within these clusters, *miR-92a* was the one most frequently recorded in our profiling analysis.

To confirm that *miR-92a* was differentially expressed during trophectoderm differentiation of ZHBTc4 ES cells, miRNA northern analysis was performed. RNA from ZHBTc4 ES cells grown without doxycycline or in the presence of doxycycline for various lengths of time was used for miRNA northern analysis of *miR-92a* expression and analysis of expression of the related miRNAs, *miR-25* and *miR-363* (*Figure 3.9a*). Phosphoimager analysis of miRNA northern membranes was used to quantitate miRNA expression in undifferentiated and differentiated ZHBTc4 ES cells (*Figure 3.9b*). Expression of *miR-92a* remained broadly similar in undifferentiated ZHBTc4 ES cells. In differentiated cultures however, *miR-92a* expression was consistently lower in differentiated cultures compared to undifferentiated cultures. The miRNA profiling showed a 2-fold decrease in expression of *miR-92a* in differentiated ZHBTc4 libraries compared to undifferentiated libraries. Phosphoimager quantitation showed an approximately 2-fold decrease in expression at 48 hours and 96 hours after addition of doxycycline, which is consistent with the fold change in representation calculated from the profiling data. Analysis of *miR-25* expression revealed that it is also differentially expressed in the ZHBTc4 ES cell system although it did not meet the arbitrary criteria that were defined for differentially represented miRNAs (*Figure 3.9a & b*). *miR-25* expression was similar in undifferentiated ZHBTc4 ES cells at T0, 24 and 48 hours, but showed an increase in expression at 72 and 96 hours. In differentiated cultures, expression of *miR-25* was lower than expression in undifferentiated cultures at each timepoint. Variation in expression levels varied between approximately 1.3-fold and 2-fold. *miR-363* was expressed at very low levels consistent with the fact that it was not detected in the miRNA profiling analysis (*Figure 3.9a*). Although these observations are supported by the findings of the miRNA profiling analysis, it must be noted that they arise from a single experimental replicate and have the potential to be misleading.

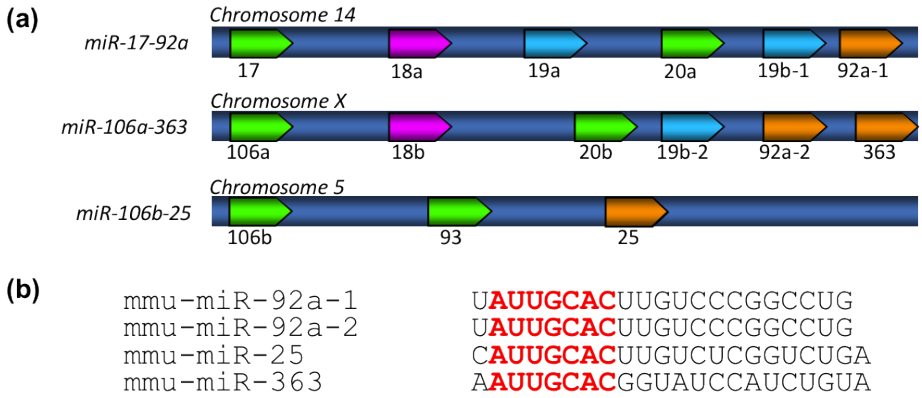


Figure 3.8 – *Diagram of the miR-17-92a cluster and related miRNA clusters.* (a) Diagrams of the *miR-17-92a* cluster on chromosome 14 and related clusters *miR-106a-363* and *miR-106b-25* on chromosomes X and 5, respectively. Each miRNA sequence is represented by an arrow. miRNAs with identical seed sequences (nucleotides 2-8 from the 5' end of the miRNA) are shown as the same colour. (b) Mature miRNA sequences of *miR-92a*, *miR-25* and *miR-363*. The seed sequence is highlighted in red.

In an attempt to gain some insight into the function of these miRNAs, expression was analysed in a variety of adult tissues. The expression of *miR-92a*, *miR-25* and *miR-363* was assessed in adult mouse tissues using RNA available in the laboratory (prepared by D. McBride). RNA was recovered from precipitate and used for miRNA northern analysis of expression of *miR-92a*, *miR-25* and *miR-363*. miRNA northern analysis showed that *miR-92a* has variable expression in the tissues of the mouse (*Figure 3.10a & b*). *miR-92a* is expressed in all tissues examined but has particularly high levels of expression in thymus, spleen, bone marrow and lung. *miR-25* has a more restricted pattern of expression and is expressed most highly in heart, spleen, lung and bone marrow. The differences in expression patterns can most readily be explained by the fact that *miR-25* and *miR-92a* are expressed from different genomic locations. In contrast to *miR-92a* and *miR-25*, expression of *miR-363* was only detected in thymus by miRNA northern analysis under the conditions employed here. Experimental repeats would be required in order to make conclusions about the expression patterns of these miRNAs definitive.

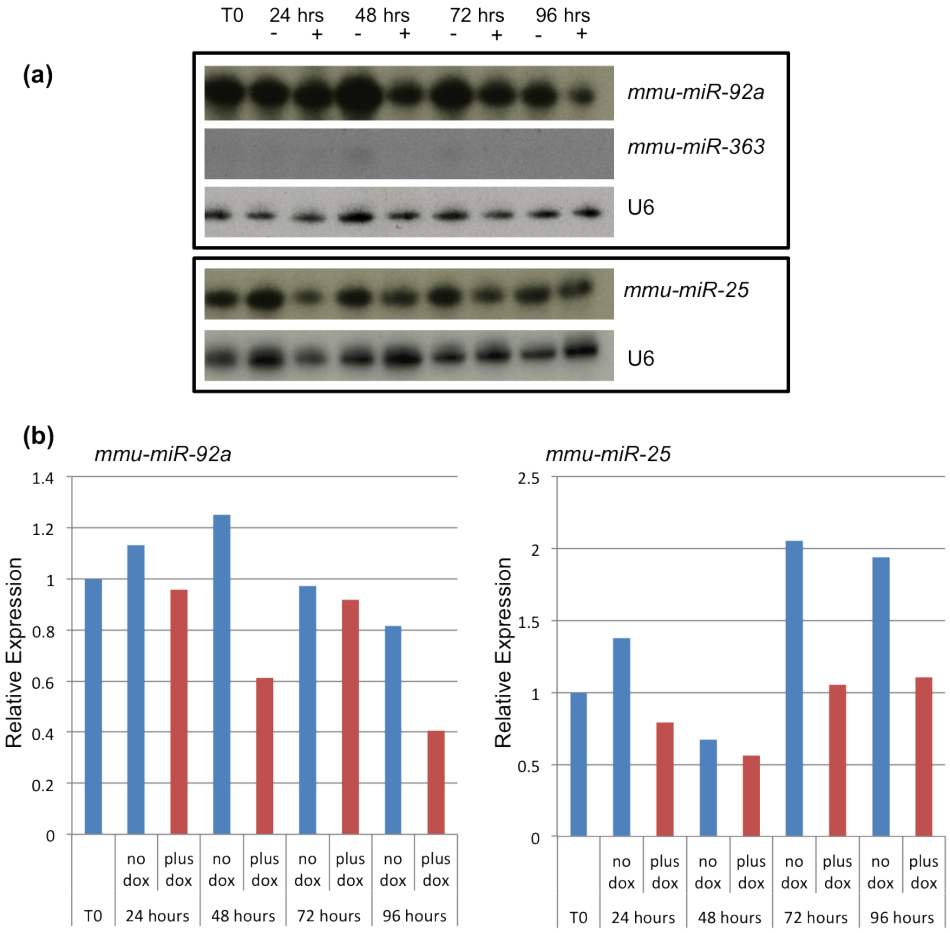


Figure 3.9 – *Expression analysis of miR-92a and related miRNAs in undifferentiated and differentiated ZHBTc4 ES cells.* (a) Total RNA (5µg) from ZHBTc4 ES cells at the time of plating (T0) or ZHBTc4 ES cells grown without doxycycline (-dox) or in the presence of 1µg/ml doxycycline (+dox) for 24, 48, 72 or 96 hours was used for miRNA northern analysis. Expression of *miR-92a*, *miR-25*, *miR-363* and the spliceosomal RNA, U6, is shown. (b) miRNA northern membranes were exposed to a phosphoimager screen for quantitation of expression of *miR-92a* and *miR-25*. Expression of *miR-363* was too low for quantitation by phosphoimager. Expression of each miRNA was normalised to U6 expression and is shown relative to expression of the T0 sample. N=1.

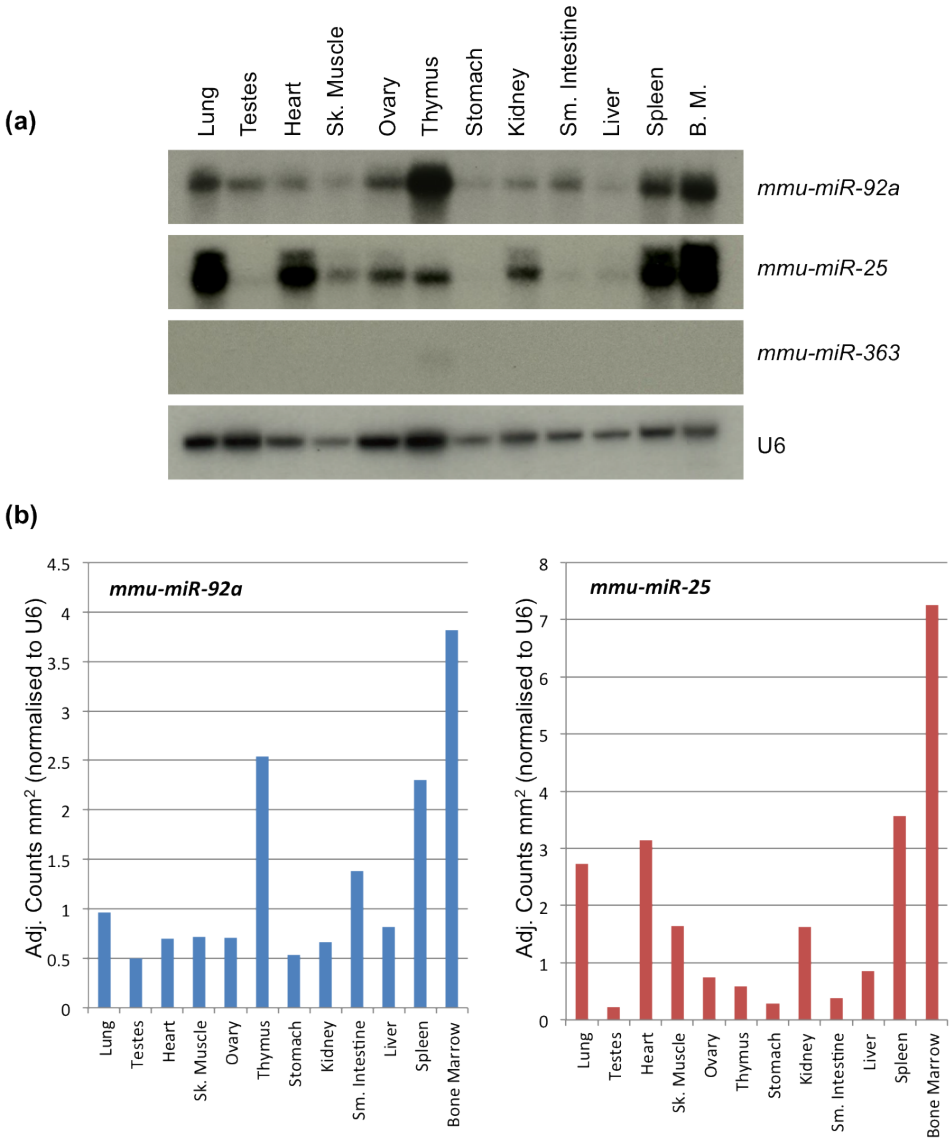


Figure 3.10 – *Expression of miR-92a and related miRNAs in various adult mouse tissues.* (a) Total RNA (5µg) from adult mouse tissues was used for miRNA northern analysis. Expression of *miR-92a*, *miR-25*, *miR-363* and the spliceosomal RNA, U6, is shown. (b) miRNA northern membranes were exposed to a phosphoimager screen for quantitation of expression of *miR-92a* and *miR-25*. Expression of *miR-363* was too low for quantitation by phosphoimager. Expression of each miRNA is shown normalised to U6 expression. N=1.

3.3.8 Optimisation of systems for miRNA functional analysis

In order to investigate the role of *miR-92a* during the ES cell to trophoblast transition, systems for functional analysis were optimised. The possibility of inhibiting miRNA expression in ES cells by addition of LNA molecules complementary to the mature miRNAs was examined. LNAs are RNA nucleotides where the O2' and C4' atoms are linked by a methylene group, which locks the molecule into the ideal conformation for Watson-Crick binding. The 'LNAs' used in this thesis are hybrid oligonucleotides comprised of DNA bases with LNA nucleotides inserted at every third position. This results in an oligonucleotide that pairs with complementary sequence more rapidly and increases the stability of the resulting complex.

To determine the appropriate concentration of LNA required to inhibit miRNA expression in undifferentiated ES cells, ZHBTc4 ES cells were transfected with a range of concentrations of an LNA designed to target *miR-30C*. RNA was collected from cultures after 24 hours of transfection and used for miRNA northern analysis of expression of *miR-30C*. *Figure 3.11a* shows that addition of a complementary LNA inhibited *miR-30C* expression in a concentration-dependent manner. No *miR-30C* expression was detected after transfection of ES cells with >37.5nM complementary LNA. miRNA northern membranes were exposed to a phosphoimager screen for quantitation of *miR-30C* expression. This analysis showed that >85% knockdown could be achieved at 24 hours using a concentration of 9.4-150nM (62.5-250ng/ml) of LNA (*Figure 3.11b*). Cell numbers were assessed after transfection of various concentrations of LNA to ensure that transfection did not cause cell death. The number of cells present after transfection with various concentrations of LNA was assessed by dissociation of the cells followed by cell counting using a nucleocounter (*Figure 3.11c*). No effect on cell numbers was observed after transfection of even the highest concentration of LNA.

To ensure that the above findings were not specific to *miR-30C*, and to optimise transfection of differentiated ZHBTc4 ES cells, the optimisation procedure was repeated in differentiating ZHBTc4 ES cells with transfection of an LNA complementary to *miR-467a**. ZHBTc4 ES cells were grown in the presence of doxycycline for 72 hours to ensure robust expression of *miR-467a** (see *Figure 3.5a*). Cells were transfected with a range of concentrations of an LNA designed to tar-

get *miR-467a**. RNA was collected from cultures after 24 hours of transfection and used for miRNA northern analysis of expression of *miR-467a**. As observed for *miR-30C* inhibition, *Figure 3.12a* shows that addition of a complementary LNA inhibited *miR-467a** expression in a concentration-dependent manner. No *miR-467a** expression was visible after transfection of ES cells with >37.5nM complementary LNA. miRNA northern membranes were exposed to a phosphoimager screen for quantitation of *miR-467a** expression. This analysis showed that >90% knockdown could be achieved at 24 hours using a concentration of 9.4-150nM (62.5-250ng/ml) of LNA (*Figure 3.12b*). As for *miR-30C*, no effect on cell numbers was observed after transfection of even the highest concentration of LNA complementary to *miR-467a** (*Figure 3.12c*).

In order to determine how long the miRNA inhibition lasted, undifferentiated ZHBTc4 ES cells were exposed to transfection reagent alone or transfected overnight with 22.5nM or 45nM of an LNA complementary to *miR-30C* and subsequently left to grow for various lengths of time. RNA was harvested from cells 24, 48 and 72 hours after transfection and used for miRNA northern analysis of *miR-30C* expression. *Figure 3.13a* shows that the miRNA knockdown is transient and the time taken for expression to reappear is dependent on the concentration of LNA added. miRNA northern membranes were exposed to a phosphoimager screen for quantitation of *miR-30C* expression (*Figure 3.13b*). These data show that a 22.5 or 45nM LNA concentration achieved approximately 80% miRNA knockdown for 48 hours. At 72 hours post transfection 50% knockdown was observed for a 22.5nM concentration of LNA and >60% inhibition was observed for a 45nM concentration of LNA. It is likely that significant suppression of miRNA levels continues beyond 72 hours after transfection.

Optimisation of LNA-mediated knockdown of miRNA expression led to the following conclusions: LNA oligonucleotides complementary to mature miRNA sequences can be used to inhibit miRNA expression in undifferentiated and differentiated ZHBTc4 ES cells. Transfection of up to 150nM of LNA into undifferentiated or differentiated ZHBTc4 ES cells does not result in increased cell death. Inhibition of miRNA expression by LNAs is concentration and time dependent. However, it must be noted that while these data provide indications about the conditions required for miRNA inhibition by LNA, because they come from single experiments these conclusions are not definitive. As a result of these experiments the concentration of LNA used for future miRNA knockdown experiments was

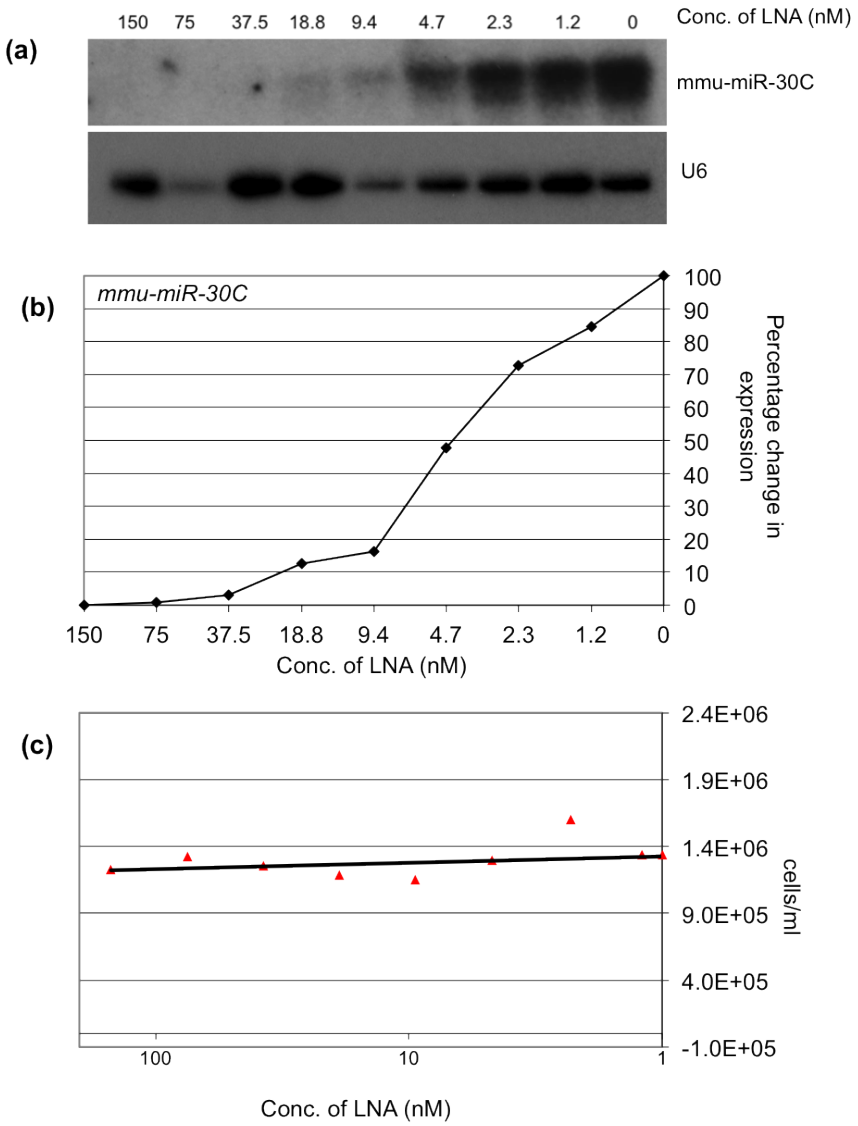


Figure 3.11 – *Optimisation of miRNA knockdown by LNA in undifferentiated ZHBTc4 ES cells.* 50,000 ZHBTc4 ES cells were grown in serum + Lif media without doxycycline for 72 hours then grown overnight in the presence of transfection reagent (Lipofectamine 2000) or transfected overnight with between 1.8nM and 150nM of LNA complementary to *miR-30C*. RNA was extracted 24 hours after the start of transfection. (a) Total RNA (5ug) from untransfected ZHBTc4 ES cells and from ZHBTc4 ES cells 24 hours after transfection with different concentrations of LNA targeting *miR-30C* was used for miRNA northern analysis. Expression of *miR-30C* and the spliceosomal RNA, U6, is shown. N=1. (b) miRNA northern membranes were exposed to a phosphoimager screen for quantitation of expression of *miR-30C* and U6. Expression of *miR-30C* was normalised to U6 expression and is shown as a percentage of expression in untransfected ZHBTc4 ES cells. N=1. (c) Cells were dissociated 24 hours after transfection and counted in duplicate by nucleocounter. Scatter chart shows the mean number of cells per ml for untransfected ZHBTc4 ES cells (shown on the log scale as 1nM concentration of LNA) and for ZHBTc4 ES cells transfected with various concentrations of LNA. A linear trendline is shown. N=1.

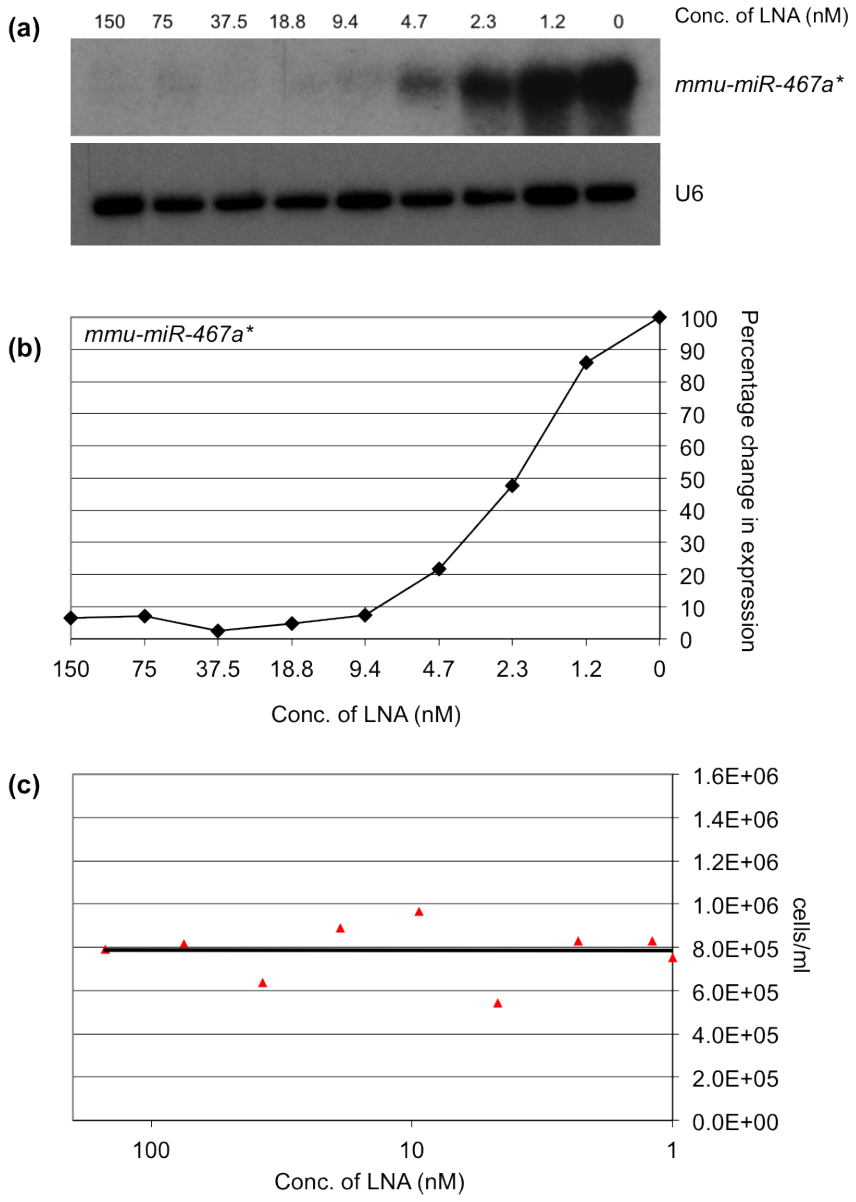


Figure 3.12 – *Optimisation of miRNA knockdown by LNA in differentiated ZHBTc4 ES cells.* 50,000 ZHBTc4 ES cells were grown in serum + Lif media with 1µg/ ml doxycycline for 72 hours then grown overnight in the presence of transfection reagent (Lipofectamine 2000) or transfected overnight with between 1.8nM and 150nM of LNA complementary to *miR-467a**. RNA was extracted 24 hours after the start of transfection. (a) Total RNA (5ug) from untransfected ZHBTc4 ES cells and from ZHBTc4 ES cells 24 hours after transfection with different concentrations of LNA targeting *miR-467a** was used for miRNA northern analysis. Expression of *miR-467a** and the spliceosomal RNA, U6, is shown. N=1. (b) miRNA northern membranes were exposed to a phosphoimager screen for quantitation of expression of *miR-467a** and U6. Expression of *miR-467a** was normalised to U6 expression and is shown as a percentage of expression in untransfected ZHBTc4 ES cells. N=1. (c) Cells were dissociated 24 hours after transfection and counted in duplicate by nucleocounter. Scatter chart shows the mean number of cells per ml for untransfected ZHBTc4 ES cells (shown on the log scale as 1nM concentration of LNA) and for ZHBTc4 ES cells transfected with various concentrations of LNA. A linear trendline is shown. N=1.

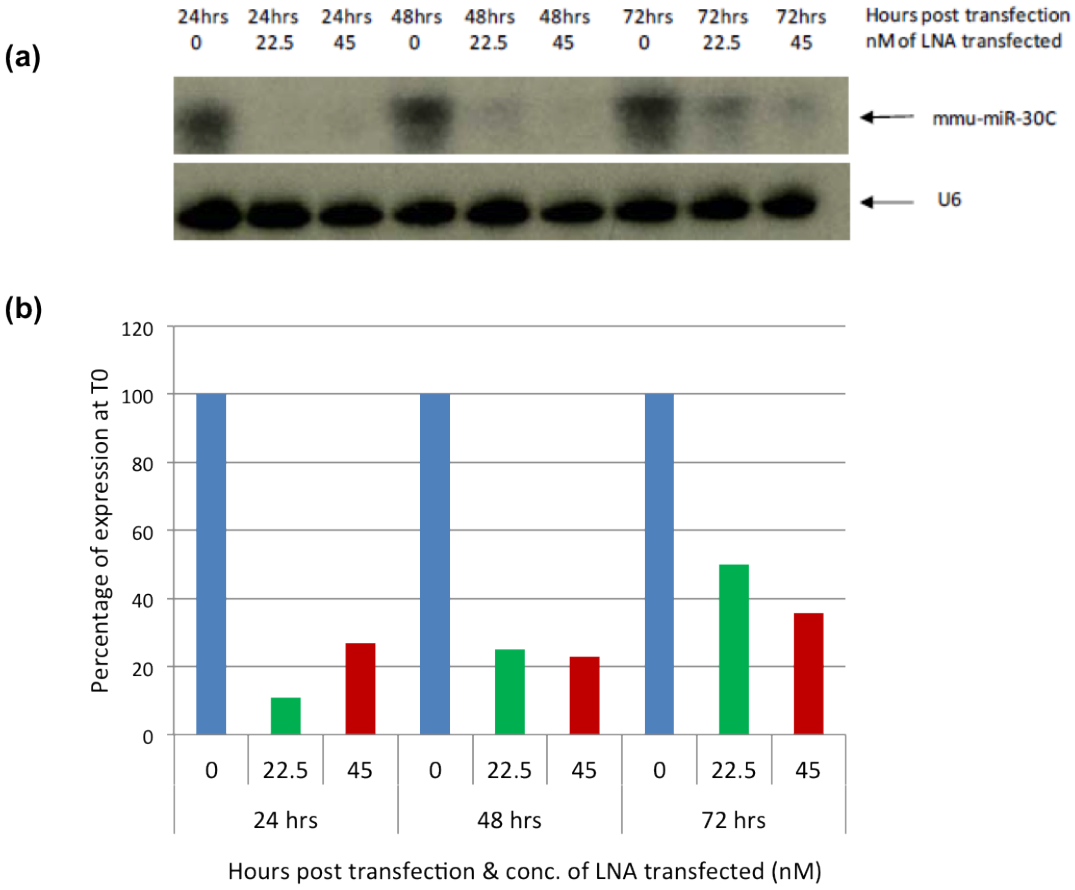


Figure 3.13 – *Analysis of length of time miRNA knockdown persists for after transfection of different concentrations of LNA.* 50,000 ZHBTc4 ES cells were grown in serum + Lif media without doxycycline for 48 hours then grown overnight in the presence of transfection reagent (Lipofectamine 2000) or transfected overnight with 22.5nM or 45nM of LNA complementary to *miR-30C*. RNA was extracted 24, 48 and 72 hours after the start of transfection. (a) Total RNA (15ug) from untransfected ZHBTc4 ES cells (0nM) and from ZHBTc4 ES cells 24, 48 and 72 hours after transfection of 22.5nM or 45nM of LNA targeting *miR-30C* was used for miRNA northern analysis. Expression of *miR-30C* and the spliceosomal RNA, U6, is shown. N=1. (b) miRNA northern membranes were exposed to a phosphoimager screen for quantitation of expression of *miR-30C* and U6. Expression of *miR-30C* was normalised to U6 expression and is shown as a percentage of expression in untransfected ZHBTc4 ES cells at each timepoint. N=1.

between 37.5 and 150 nM.

3.3.9 The effect of *miR-92a* suppression on gene expression

According to our analysis, *miR-92a* is differentially expressed during trophoblast differentiation and shows reduced levels as a result of the transition from ES cell to trophoblast lineage. In order to assess if *miR-92a* and related miRNAs had any direct effect on differentiation, QPCR analysis of ES cells and TE markers was carried out on ES cells after combined inhibition of *miR-92a*, *miR-25* and *miR-363* using 100nM of LNA oligonucleotides complementary to the three miRNAs. Control cultures were transfected with an LNA designed to inhibit a novel chicken miRNA (DM1) that is not found in mice (*Figure 3.14a*). QPCR assays for the ES cell and TE markers detailed in *Table 3.1* were used in this analysis. *Figure 3.11a* shows that there is no significant difference in the expression of either ES cell or trophoblast markers following miRNA knockdown in ZHBTc4 ES cells. Identical treatments and analysis were also carried out in two wild type ES cell lines (E14 and CGR8 ES cells) and although there was slightly more variation seen than with ZHBTc4 cells, there was no obvious effect of miRNA knockdown (*Figure 3.14b* & *c*). In each case the specificity of the knockdown of *miR-92a* was confirmed by miRNA northern analysis (*Figure 3.11a-c*). Although these data come from a single experimental replicate, the fact that the same observation is made in each ES cell line strengthens the preliminary conclusions made from these results.

From this experiment it can be seen that knockdown of *miR-92a* and related miRNAs had no effect on the expression of ES cell markers or TE markers. Therefore it is unlikely that the function of these miRNAs in this system relates to the differentiation process *per se*. However, *miR-92a* and related miRNAs are likely to have a role in other aspects of cell function that are associated with this developmental transition. For example, there is evidence that some miRNAs encoded by these clusters have roles in cell proliferation (Mendell, 2008) and *Figure 3.1* shows that the proliferation rate of ES cells decreases during TE differentiation. In light of this the possibility that *miR-92a* and related miRNAs were involved in regulating cell proliferation in ES cells was investigated.

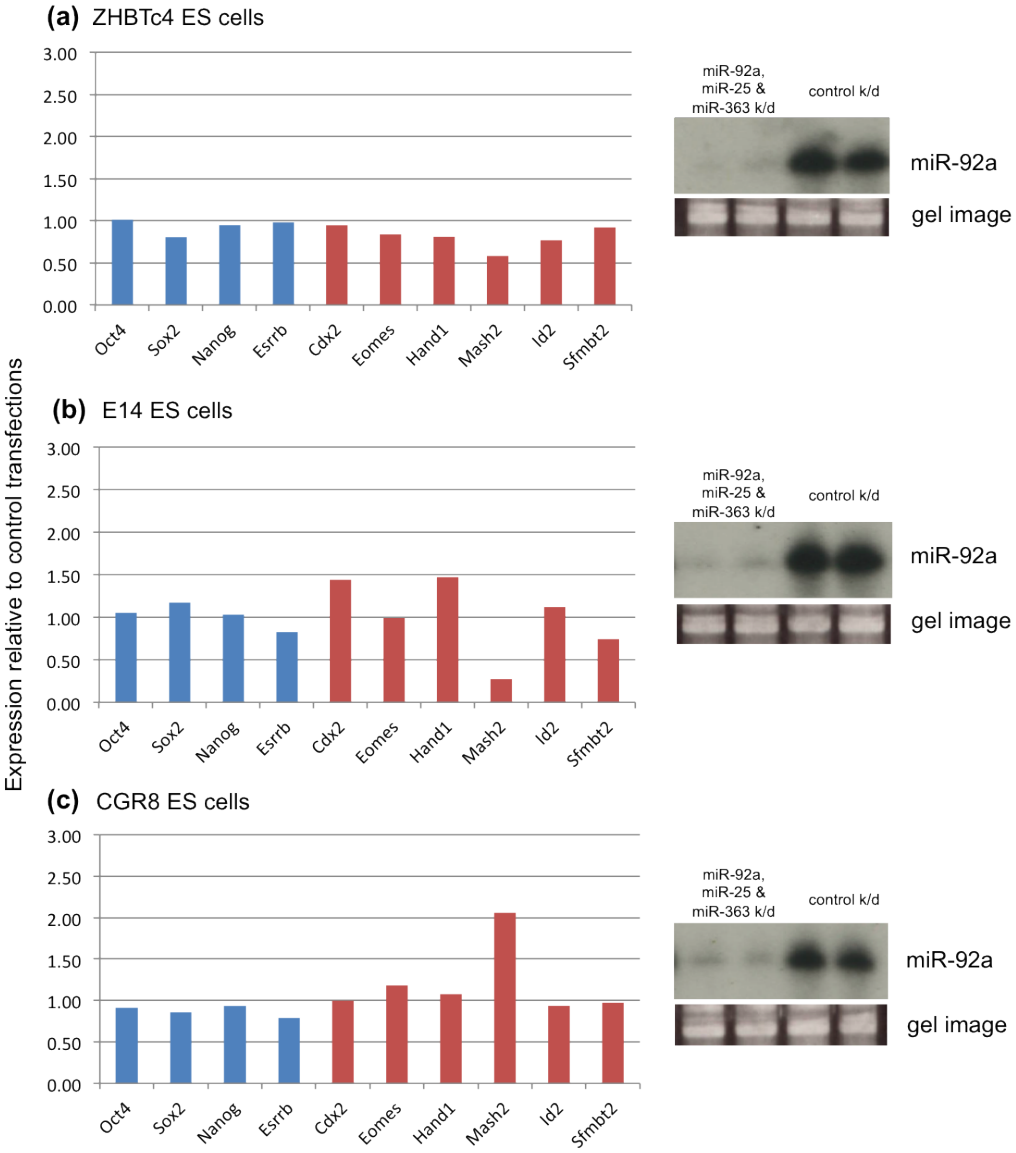


Figure 3.14 – *Expression of ES cell and trophoblast marker genes after inhibition of miRNA expression.* ES cells were transfected with a combination of LNAs targeting *miR-92a*, *miR-25* and *miR-363* or an LNA targeting DM1 (*gga-miR-2954* - control transfection). RNA was extracted 96 hours after transfection. RNA was DNAse treated, converted to cDNA and used for QPCR analysis of expression of ES cell marker genes (Oct3/4, *Sox2*, *Nanog* and *Esrrß*) and trophoblast marker genes (*Cdx2*, *Eomes*, *Hand1*, *Mash2*, *Id2* and *Sfmbt2*). QPCR Ct values were normalised to *β-Actin* and expression is shown for the *miR-92a*, *miR-25* and *miR-363* knockdown relative to expression in the control transfection. Results are shown for transfection of ZHBTc4 ES cells (a), E14 ES cells (b) and CGR8 ES cells (c). Total RNA (5ug) from control transfections and transfection of LNAs targeting *miR-92a*, *miR-25* and *miR-363* was used for miRNA northern analysis. Expression of *miR-92a* is shown. The gel image shows the presence of RNA in each lane. N=1.

3.3.10 Optimisation of cell proliferation assays

Optimisation of MTT and Cyquant assays identified the linear range of these assays (described in Appendix C) and as a result 1000-1500 ZHBTc4 or 3T3 cells were plated per well of 96-well plates for future MTT and Cyquant assays. These numbers were chosen because cells would be grown for at least 48 hours in the plates and so the final number of cells to be assayed should fall within the linear range. Additional optimisation resulted in use of Lipofectamine LTX as the transfection reagent for proliferation experiments, and knockdown of *miR-155* as the transfection control (Appendix C). *miR-155* was chosen for use as a control because in addition to it having little effect on cell numbers, this miRNA was not sequenced in the miRNA profiling experiment and was not detected in mouse ES cells by miRNA northern analysis (data not shown).

3.3.11 Effect of *miR-92a* suppression on cell number

In order to investigate the effect of *miR-92a* inhibition on ZHBTc4 ES cell proliferation, ZHBTc4 ES cells were plated per well of a 96-well tissue-culture plate, miRNA expression was inhibited by transfection of a complementary LNA and the cultures were subjected to MTT analysis 24, 48 or 72 hours after transfection. As well as inhibition of *miR-92a*, the effect of inhibiting *miR-25* or *miR-363* expression was also investigated. *Figure 3.15a* shows that knock down of *miR-25* and *miR-363* had no effect on cell numbers in relation to *miR-155* (control) knockdown. In contrast, knockdown of *miR-92a* significantly reduced cell numbers as early as 24 hours after treatment and this effect was maintained for up to at least 72 hours.

To test whether there were any additive effects of knocking down *miR-92a* and related miRNAs, ZHBTc4 ES cells were transfected with an LNA targeting *miR-92a* alone, or in combination with LNAs targeting *miR-25* or *miR-25* and *miR-363* and cell numbers were assessed by MTT assay. Cell numbers for each combination of miRNA knockdowns was normalised to cell numbers in control transfection of an equivalent concentration of LNA targeting *miR-155*. The reduction in cell numbers seen with *miR-92a* knockdown alone was similar to that observed when knockdown was combined with knockdown of *miR-25* and *miR-363* suggesting

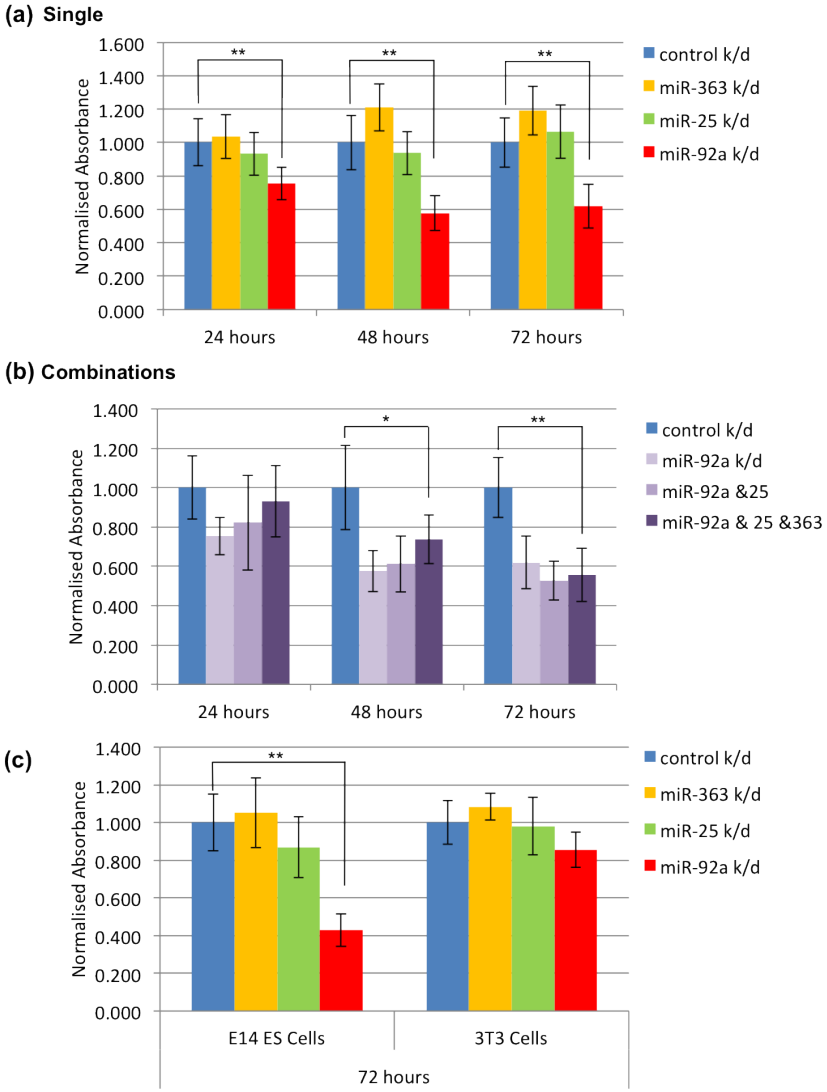


Figure 3.15 – *Effect of miRNA knockdown on cell numbers.* (a) 1000 ZHBTc4 ES cells were plated per well of 96-well plates and grown for 24 hours before being transfected overnight (using Lipofectamine LTX) with 75nM of LNA targeting various miRNAs. Control knockdown is transfection of LNA targeting *miR-155*. Cells were grown for a further 24, 48 or 72 hours and subjected to MTT analysis. Bar chart shows mean absorbance relative to control knockdown from 2 independent experiments with 8 replicate wells per experiment. Error bars show standard error of the mean. **P<0.01. (b) 1000 ZHBTc4 ES cells were plated per well of 96-well plates and grown for 24 hours before being transfected overnight (using Lipofectamine LTX) with LNAs targeting various combinations of miRNAs (75nM per LNA). Cells were grown for a further 24, 48 or 72 hours and subjected to MTT analysis. Bar chart shows mean absorbance relative to an equivalent concentration of control LNA (*miR-155*) from 2 independent experiments with 8 replicate wells per experiment. Error bars show standard error of the mean. (c) 1000 E14 ES cells or 3T3 cells were plated per well of 96-well plates and grown for 24 hours before being transfected overnight with 75nM of LNA targeting various miRNAs. Control knockdown is transfection of LNA targeting *miR-155*. Cells were grown for a further 72 hours before being subjected to MTT analysis. Bar chart shows mean absorbance relative to control knockdown from 2 independent experiments with 8 replicate wells per experiment. Error bars show standard error of the mean. **P<0.01.

that the effect is highly sequence specific (*Figure 3.15*b). To determine whether this effect was restricted to ZHBTc4 ES cells, the effects of 72-hour *miR-92a* knockdown were also assessed in E14 ES cells and 3T3 cells. Knockdown of *miR-92a* in E14 ES cells resulted in a significant decrease in cell numbers compared to knockdown of *miR-155* as measured by MTT assay. Surprisingly, knockdown of *miR-92a* in 3T3 cells had no effect on cell numbers compared to knockdown of *miR-155* (*Figure 3.15c*). Reducing levels of *miR-92a* appears to affect cell proliferation in ES cells, but not in fibroblast cells.

The MTT assay is a colorimetric assay for measuring the activity of enzymes that reduce MTT ((3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide)) to formazan dyes and although changes in MTT assay absorbance are considered to reflect changes in cell number, it is possible that the changes that have been recorded actually reflect changes to cell metabolism. Given this possibility, the effect of *miR-92a* knockdown was also assessed using two alternative means of estimating cell numbers. The effect of *miR-92a* knockdown in ZHBTc4 ES cells was assessed by both CYQUANT analysis, and by estimation of absolute cell numbers. The effects of *miR-92a* knockdown on a wt ES cell line was also assessed by estimating absolute cell numbers. For this analysis, ZHBTc4 were plated in 6-well plates, grown for 24 hours and transfected overnight with LNAs targeting *miR-155* or *miR-92a*. Approximately 6 hours after transfection, cells were dissociated, counted and replated in 6-well plates (for cell counting) or 96-well plates (for Cyquant analysis). After 48 hours of growth Cyquant assays were performed on 96-well plates and cells in 6-well plates were dissociated and counted by nucleocounter. RNA from cultures was extracted and used for miRNA northern analysis of *miR-92a* expression. This confirmed knockdown of *miR-92a* in the ZHBTc4 ES cells and E14 ES cells transfected with an LNA complementary to *miR-92a* (*Figure 3.16a*). Cyquant analysis showed a reduction of cell numbers after *miR-92a* knockdown compared to *miR-155* knockdown (*Figure 3.17b*). Quantitation of absolute cell numbers also confirmed that ZHBTc4 cell numbers decrease upon *miR-92a* knockdown (*Figure 3.16c*). To ensure that the effect on cell numbers was not restricted to ZHBTC4 ES cells, a similar analysis was performed on a second ES cell line - E14 ES cells. Knockdown of *miR-92a* in E14 ES cells also caused a reduction in cell numbers confirming that this effect occurs in other ES cells lines (*Figure 3.16d*).

The decrease in cell numbers induced by depletion of *miR-92a* by LNA anti-miR

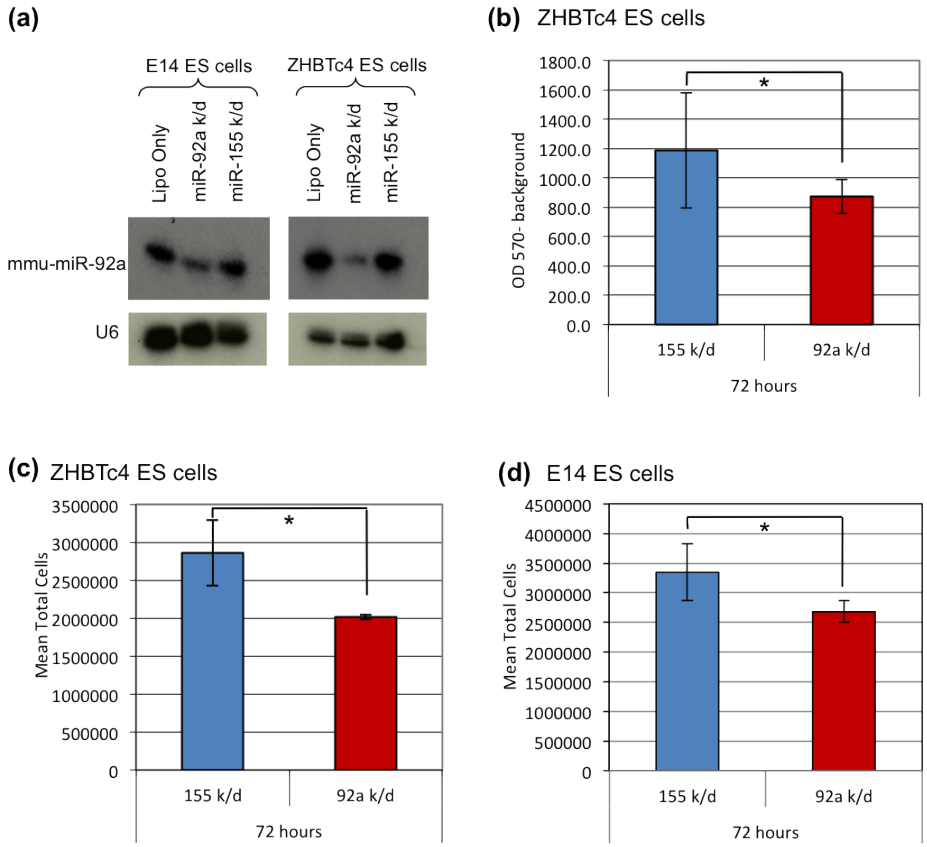


Figure 3.16 – *Effect of miRNA knockdown on cell numbers assessed by CYQUANT assay and cell counting.*(a) Total RNA (5µg) from ZHBTc4 ES cells grown overnight in the presence of Lipofectamine only or transfected overnight with control LNA (targeting miR-155) or an LNA targeting *miR-92a* and then grown for a further 72 hours. RNA was used for miRNA northern analysis of *miR-92a* expression and expression of the spliceosomal RNA, U6. RNA was from cells used for cell count experiments (this figure c & d). (b) 200,000 ZHBTc4 ES cells were plated per well of 6-well plates and grown for 24 hours before being transfected (using Lipofectamine LTX) overnight with 150nM of LNA targeting *miR-92a* or miR-155 (control). After the transfection had been stopped the cells were left to grow for ~6 hours before being dissociated and replated at 1000 cells per well of 96-well plates. Cells were grown for a further 48 hours (total - 72 hours post transfection) before being subjected to CYQUANT assay. Bar chart shows mean absorbance relative to control knockdown from 3 experimental replicates with 8 replicate wells per experiment (after background absorbance was deducted). Error bars show standard error of the mean. *$P<0.05$. (c & d) 200,000 ZHBTc4 ES cells (c) or E14 ES cells (d) were plated per well of 6-well plates and grown for 24 hours before being transfected overnight with 150nM of LNA targeting *miR-92a* or miR-155 (control). After the transfection had been stopped the cells were left to grow for ~6 hours before being dissociated and replated at 800,000 (ZHBTc4) or 700,000 (E14) cells per well of 6-well plates. Cells were grown for a further 48 hours (total - 72 hours post transfection) before being dissociated and counted by nucleocounter. Bar chart shows the mean of the total number of cells from duplicate counts of three experimental replicates. Error bars show standard error of the mean. *$P<0.05$.

mirrors the reduction in cell numbers seen during differentiation of ZHBTc4 ES cells (*Figure 3.1c*). It seems likely that, having previously shown that *miR-92a* is unlikely to be associated with differentiation, expression of *miR-92a* is related to proliferation within this system, and the decrease in cell numbers observed during the developmental transition may be partly due to the decrease in *miR-92a* expression.

3.3.12 The effect of perturbing BMP levels on *miR-92a* expression

In 2011 Wang *et al.* showed that the *miR-17-92a* cluster was regulated by BMP signalling in P19 (mouse embryonal carcinoma) cells and 293T (human embryonic kidney stem) cells. ES cells cultured in classic conditions are dependant on BMP signalling to induce inhibitor-of-differentiation (ID) proteins (Ying *et al.*, 2003). BMP signalling is therefore thought to promote self-renewal of ES cells. It was of interest to investigate whether BMP signalling regulated the expression of *miR-92a*, which is highly expressed in ES cells, because this may suggest that BMP signalling promotes proliferation of ES cells as well as suppressing their differentiation.

In order to assess the effect of BMP signalling on *miR-92a* expression, E14 and CGR8 ES cells were grown in a variety of culture conditions where levels of BMP signalling were affected by serum supplementation, the addition of purified BMP4 or the addition of the BMP inhibitor, dorsomorphin (Yu *et al.*, 2008). *Figures 3.17a & b* shows expression of *miR-92a* in wt ES cells treated to either induce or inhibit BMP signalling. This shows that inhibition of BMP by the addition of dorsomorphin to ES cells leads to a decrease in *miR-92a* expression. In contrast, *miR-92a* expression was induced after addition of BMP4 to cells cultured in N2B27 + Lif. This effect was observed only in CGR8 ES cells and not in E14 ES cells. One possible explanation for this would arise if the ES cells themselves are producing BMP. The BMP produced by the ES cells during the 48 hour culture in N2B27 & Lif may be sufficient to induce *miR-92a* expression to maximal levels, which would then mean that further addition of BMP4 did not induce greater expression. This is supported by the observation that E14 ES cells grown in the absence of BMP with addition of DMSO for 24 hours showed a similar level of *miR-92a* expression to those cells cultured in the presence of BMP4.

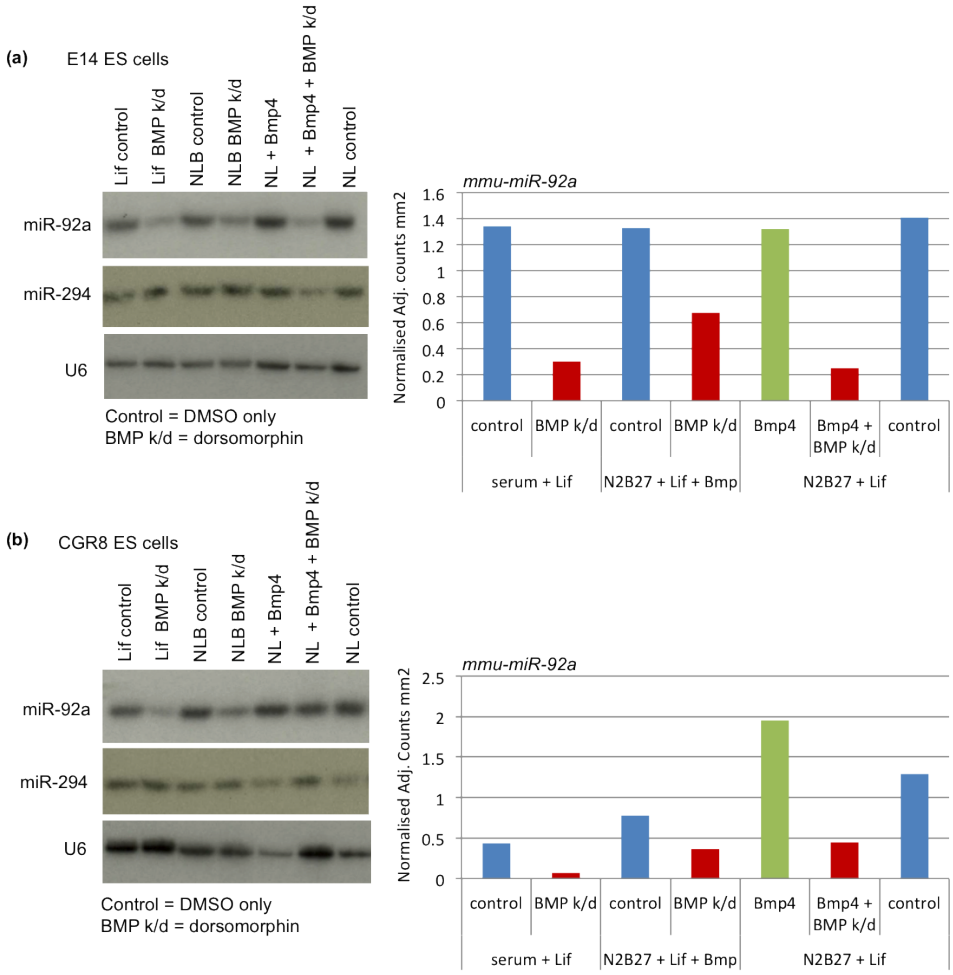


Figure 3.17 – *Expression of miR-92a after inhibition of Bmp in ES cells.* 500,000 E14 ES cells (a) or CGR8 ES cells (b) were plated per well of 6-well plates and grown in the following conditions: serum+Lif control - cells cultured in serum+Lif media for 24 hours then 10µl DMSO added for 48 hours; serum+Lif dorso - cells cultured in serum+Lif media for 24 hours then dorsomorphin (final concentration 5uM) added for 48 hours; N2B27+Lif+BMP control - cells cultured in N2B27 media with 1000U Lif and 10ng/ml BMP4 for 24 hours then 10ul DMSO added for 48 hours; N2B27+Lif+BMP dorso - cells cultured in N2B27 media with 1000 U Lif and 10ng/ml BMP4 for 24 hours then dorsomorphin (final concentration 5uM) added for 48 hours; N2B27+Lif BMP4 - cells cultured in N2B27 media with 1000U Lif for 48 hours then 10ng/ml BMP4 added for 24 hours; N2B27+Lif BMP4 dorso - cells cultured in N2B27 media with 1000 U Lif for 48 hours then 10ng/ml BMP4 and dorsomorphin (final concentration 5uM) added for 24 hours; N2B27+Lif control – cells cultured in N2B27 media with 1000U Lif for 48 hours then 10ul DMSO added for 24 hours. RNA was extracted from each culture and total RNA (5ug) was used for miRNA northern analysis of expression of *miR-92a*, *miR-294* and the spliceosomal RNA, U6. miRNA northern membranes were exposed to a phosphoimager screen for quantitation of expression of *miR-92a* and U6. Bar charts show expression of *miR-92a* after normalisation to U6 expression. N=1.

In order to confirm that the reduction in the expression of *miR-92a* seen with dorsomorphin treatment was specific, and not a general effect on the entire miRNA population, the expression of *miR-294* in these samples was also analysed by miRNA northern (*Figure 3.17a & b*). In contrast to expression of *miR-92a*, *miR-294* expression levels were not affected by the addition of dorsomorphin to the culture media. It must be noted that conclusions from these data are preliminary because the data arises from a single experimental replicate in each cell line.

3.4 Discussion

The primary aims of this chapter were to determine if miRNAs were differentially represented in a model of trophectoderm differentiation and if differentially represented miRNAs could be functionally associated with elements of the differentiation process. miRNA profiling by cloning and sequencing of small RNAs identified miRNAs that were present in undifferentiated mouse ES cells and in cells that had been induced to differentiate into trophectoderm. This approach identified miRNAs that were differentially expressed during the differentiation process. Several of the miRNAs with higher representation levels in differentiated cells are encoded within a placental-enriched polycomb group gene called *Sfmbt2*, suggesting an important role for these miRNAs in extraembryonic development. Inhibition of one of the miRNAs that was expressed at higher levels in undifferentiated cells, *miR-92a*, was shown to have an effect on ES cell numbers.

3.4.1 The majority of miRNAs are not differentially represented during trophectoderm differentiation

In order to explore the roles of miRNAs during early developmental transitions, an ES cell based model of trophectoderm differentiation was used. ZHBTc4 ES cells can be induced to differentiate down the trophectoderm lineage by conditional repression of the ES cell regulatory gene, *Oct4* (Niwa *et al.*, 2000). Characterisation of this model system confirmed previous observations that expression of the TE marker genes *Cdx2* and *Hand1* increased and expression of pluripotency marker genes *Oct4* and *Sox2* decreased upon downregulation of *Oct4* expression

(Niwa *et al.*, 2000, Hay *et al.*, 2004). Gene expression analysis was extended to further show that expression of *Nanog* and *Esrrß* was downregulated and expression of *Eomes*, *Mash2* and *Id2* was upregulated following *Oct4* downregulation. Similar changes in gene expression were observed following knockdown of *Oct4* expression by siRNA, which suggests that changes in expression of these genes is due to changes in expression level of *Oct4*, rather than a result of doxycycline addition to ES cells or an anomaly of the ZHBTc4 ES cell system.

miRNA profiling by cloning and sequencing of small RNAs during ZHBTc4 ES cell differentiation identified miRNAs that were present in undifferentiated mouse ES cells and in cells that had been induced to differentiate into trophectoderm. 461 individual sequences were identified; 160 sequences had been previously identified as miRNAs and 31 were identified as novel miRNAs. A total of 270 sequences were not previously identified as miRNAs or characterised as novel miRNAs in this analysis, which corresponded to just fewer than 60% of the total sequences. The origin of these short RNA sequences is not known. By way of comparison, in a deep sequencing analysis of short RNA in three human cell lines, the non-miRNA population made up between 23% and 82% of the total sequences (Vaz *et al.*, 2010), which is consistent with the sequencing data presented in this chapter.

The majority (90%) of the miRNAs did not change in representation during differentiation. This finding would be expected if those miRNAs were involved in general housekeeping of a cell or in regulating processes unique to the early embryo in its entirety. There is much evidence for roles of miRNAs in early developmental processes (Suh & Blelloch, 2011), but little evidence to suggest that miRNAs regulate genes involved in cell homeostasis. In fact, it has been reported that genes involved in basic cellular processes avoid targeting by miRNAs due to short 3' UTRs that are specifically depleted for miRNA binding sites (Stark *et al.*, 2005). Therefore, based on current evidence, it is more likely that the large proportion of miRNAs that were not differentially represented in this analysis are involved in regulating processes specific to the early embryo rather than regulating processes related to cell homeostasis.

A single genomic cluster encoding the miRNAs, *miR-294* and *miR-295* accounted for over 40% of the total number of RNAs sequenced in this analysis. Although previously described as ES cell-specific and shown to be down regulated during embryoid body and retinoic acid induced differentiation (Houbaviy *et al.*,

2003), the representation levels of *miR-294* and *miR-295* were not significantly affected by differentiation into trophoblast. This finding agrees with more recent reports that these miRNAs are also expressed in trophoblast stem (TS) cells (Houbaviy *et al.*, 2005). Rather than facilitating the first perceivable developmental transition in the mammalian embryo, our data points to a more general role for *miR-294* and *miR-295* in early development. Supporting this is the embryonic phenotype of mice deficient for the *miR-290-295* miRNA cluster, which encodes *miR-294* and *miR-295*. Loss of the *miR-290-295* cluster in mice results in partially penetrant embryonic lethality with 50-60% of embryos exhibiting one of two abnormal phenotypes: either partial or complete localisation of mutant embryos outside the yolk sac or general developmental delays (Medeiros *et al.*, 2011). However, *miR-290-295*-null blastocysts appear morphologically normal, which suggests that normal ICM and TE compartments have been formed, and supports the proposal that these miRNAs are not required to facilitate the first perceivable developmental transition in the mammalian embryo.

3.4.2 Some miRNAs are differentially represented during trophectoderm differentiation

In this analysis, 19 miRNAs were characterised as being differentially represented during the differentiation of ES cells down the trophectoderm lineage; the levels of 12 miRNAs increased and the levels of 7 miRNAs decreased. The differential expression of a selection of these miRNAs was confirmed by northern analysis, which validated the sequencing data. Interestingly, several differentially represented miRNAs contained binding sites for key ES cell transcription factors within their promoter regions, according to data retrieved from Marson *et al.* (Marson *et al.*, 2008), which suggests that expression of these miRNAs may be directly controlled by these factors.

As well as identifying differentially represented miRNAs, this analysis also identified protein-coding genes that were differentially expressed during trophectoderm differentiation. It was found that, in most cases when a differentially represented miRNA was encoded within a gene, the host gene and the encoded miRNA had the same expression pattern. Baskerville and Bartel (2005) made the same observation in a paper that investigated the expression patterns of 175 human miRNAs.

However, they also concluded that proximal pairs of miRNAs are generally coexpressed as long as the sequences are separated by less than 50kb, a finding that disagrees with our sequencing data. It is interesting to note that when miRNAs are encoded as a cluster, the individual members of the cluster are not always represented at similar levels. For example, *miR-106a* was detected in undifferentiated and differentiated cell cultures, and, according to the sequencing data, was represented at twice the level in undifferentiated cultures compared to differentiated cultures. But another miRNA that is encoded within the same miRNA cluster, *miR-363*, was not detected in the profiling analysis. Differences in representation levels between different miRNAs encoded within the same cluster could be due to a variety of reasons including post-transcriptional regulation of miRNAs during miRNA biogenesis or miRNA turnover (Newman & Hammond, 2010). Supporting this finding, a recent paper has shown that DROSHA cleaves hundred of different pri-miRNA substrates with different efficiencies, including those present within the same miRNA cluster (Feng *et al.*, 2011).

Since commencing this study, several other groups have investigated expression of miRNAs during trophectoderm differentiation, both *in vitro* and *in vivo* (Viswanathan *et al.*, 2009, Spruce *et al.*, 2010, Ohnishi *et al.*, 2010). Comparing the profiling data presented here with that of other groups may provide insight into general miRNAs required for trophectoderm differentiation, and those that are specific to the *in vitro* system that was used in this analysis. Viswanathan *et al.* (2009) investigated expression of miRNAs in ES cells that were induced to differentiate into trophectoderm by ectopic expression of HRas/Q61L and during pre-implantation mouse development. As in our analysis, Viswanathan *et al.* found that during TE differentiation from ES cells, expression of *Sfmbt2* miRNAs increased and expression of the *miR-290* cluster was largely unchanged. The authors then collated their *in vitro* ES cell differentiation data with data generated from miRNA profiling of various stages of pre-implantation embryogenesis to identify 'candidate miRNAs involved in trophectoderm specification'. Of the 8 candidate miRNAs identified as being involved in trophectoderm specification in their study, only *miR-297* was also found in the analysis presented here. Additionally, the single miRNA that Viswanathan *et al.* identified as being pluripotency specific (*miR-367a*) was not differentially represented in our analysis. However, the *in vivo* data from the Viswanathan study was comparing different stages of embryo development, rather than separate ES and TE compartments. It could

be argued that this approach would not identify miRNAs involved in TE differentiation *per se*, but rather miRNAs involved in all aspects of transition through embryonic development.

Ohnishi *et al.* (2010) examined miRNA expression in ICM and whole blastocysts and, from this analysis, estimated miRNAs that may be expressed asymmetrically between ICM and TE. They examined the expression of two of these miRNAs, *miR-99b* and *miR-210*, by QPCR of material from isolated ICM and TE lineages, and reported that both of these miRNAs had higher expression in the TE than the ICM, which is consistent with the finding reported in this chapter that these miRNAs increased in expression during TE differentiation *in vitro*. However, other miRNAs that Ohnishi *et al.* estimated to be differentially expressed showed differences as well as similarities with the results presented here. Notably, Ohnishi *et al.* defined *miR-467a** as ICM associated and *miR-92a* as TE associated, which was directly opposite to the results presented here, as well as those reported in the study by Viswanathan *et al.*

Spruce *et al.* (2010) profiled miRNA expression by microarray in ES cells and TS cells. The authors reported that expression of the *miR-290* cluster of miRNAs is not different between ES cells and TS cells, which supports our finding that these miRNAs do not change during TE differentiation. They also reported that the majority of *Sfmbt2* cluster miRNAs were associated with TS cells rather than ES cells, which also supports our data. However, Spruce *et al.* also reported findings that disagreed with the results presented here. For example, the *Sfmbt2* cluster miRNA *miR-669*a, which was upregulated on TE differentiation in our analysis, was reported to be associated with both ES cells and TS cells in the Spruce *et al.* analysis.

Collectively, comparison of our data with the data from other similar analyses show the greatest similarity with other *in vitro* studies, although this could be explained by the fact that the majority of the *in vivo* data compared ICM and blastocyst miRNA expression profiles, rather than comparing expression profiles from isolated ICM and TE. Overall, the results presented here and the findings of others suggest that expression of the *Sfmbt2* miRNA cluster is TE associated, the *miR-17-92a* cluster appears to be ES cell associated and the *miR-290* cluster is not differentially expressed between the two cell lineages.

3.4.3 A differentially represented miRNA cluster is encoded within a trophectoderm-associated gene

Of the 12 miRNAs that increased in representation during trophectoderm differentiation, 8 were located within a single multi-miRNA cluster located in intron 10 of the Polycomb group gene, *Sfmbt2* (Scm-like with four mbt domains 2). Expression analysis of *Sfmbt2* confirmed that it increased in expression during *in vitro* trophectoderm differentiation and was expressed more highly in the mouse placenta than in the mouse embryo. This latter finding was supported by a previous analysis, which reported that, in E5.5-E8.0 embryos, *Sfmbt2* was expressed highly and continuously in the extraembryonic ectoderm, which is essential for formation of the placenta and maintenance of trophoblast stem cells (Frankenberg *et al.*, 2007). In an analysis of miRNA expression, Viswanathan *et al.* found that expression of *Sfmbt2* miRNAs increased during trophectoderm differentiation (Viswanathan *et al.*, 2009). Interestingly, Viswanathan *et al.* extended this analysis to investigate expression of miRNAs in TS cells that were induced to differentiate by removal of FGF4. Here they found that expression of *Sfmbt2* miRNAs significantly decreased. This may suggest that, during development, the *Sfmbt2* cluster of miRNAs are associated with trophoblast stem cells specifically, rather than being associated with terminally differentiated trophectoderm (Viswanathan *et al.*, 2009). In the adult, *Sfmbt2* is expressed in many tissues including testis, brain, lung, spleen and thymus. Interestingly, while it is expressed from both alleles in somatic tissues after E7.5, *Sfmbt2* is only expressed from the paternal allele in the early embryo and extraembryonic tissues (Kuzmin *et al.*, 2008). This makes *Sfmbt2* an imprinted gene in the early embryo, increasing the likelihood of an important role for *Sfmbt2*, and its encoded miRNA cluster, during early development. Intriguingly, it has been proposed that imprinting of the *Sfmbt2* gene is a direct result of the insertion of the large cluster of miRNAs in intron 10 of this gene (Wang *et al.*, 2011). This has been proposed because while mouse and rat *Sfmbt2* genes contain the miRNA cluster and are imprinted, *Sfmbt2* in other mammals does not encode a miRNA cluster, and expression of *Sfmbt2* in these mammals is biallelic (Wang *et al.*, 2011).

The function of *Sfmbt2* and its encoded miRNA cluster is not yet clear. However, one study suggests a possible function of this cluster in the downregulation of pluripotency genes during differentiation. One of the differentially expressed

miRNAs within the *Sfmbt2* cluster, *miR-669b*, has been shown to exhibit target enrichment for ESC-specific pluripotency genes (Tang *et al.*, 2010) suggesting a potential function of this cluster may be to suppress pluripotency during the transition from ES cell to trophoblast stem cell. A further report suggests that another *Sfmbt2* cluster miRNA, *miR-467a**, functionally overlaps with the *miR-290-295* miRNAs to promote growth and survival of mouse ES cells (Zheng *et al.*, 2011). The authors speculate that the *Sfmbt2* miRNA cluster may function to promote growth and survival of the early TE lineage, while the *miR-290-295* cluster fulfils a similar function in the ICM. Interestingly in the context of differential regulation of *Sfmbt2* during early development, there is a putative binding site for the key ES cell transcription factor *Oct4* in the promoter of *Sfmbt2* (Loh *et al.*, 2006), which is consistent with the finding that *Sfmbt2* expression increases upon suppression of *Oct4* expression, and suggests that there may be a direct interaction between *Oct4* and the *Sfmbt2* promoter. However, the fact that Marson *et al.* did not report *Oct4* binding at the *Sfmbt2* promoter in a ChIP-seq analysis of promoter binding by key ES cell transcription factors, argues against this possibility (Marson *et al.*, 2008). Together, these findings show that expression of *Sfmbt2* and the miRNAs encoded within *Sfmbt2* increase upon suppression of *Oct4* expression. It is speculated that *Sfmbt2* expression is enriched in the trophoblast stem cells of the early embryo, and possible functions include the regulation of pluripotency factors associated with the ICM or promotion of cell proliferation and survival in early embryonic lineages.

3.4.4 *miR-92a* and related miRNAs do not regulate expression of ES cell and trophectoderm associated genes

The results presented here show that expression of *miR-92a* is suppressed as a result of trophoblast differentiation. However, conflicting reports on the expression of *miR-92a* have suggested that levels of this miRNA both increase (Houbaviy *et al.*, 2003) and decrease (Viswanathan *et al*, 2009) as a result of ES cell differentiation. The analysis that suggested that the levels of *miR-92a* increased during differentiation (Houbaviy *et al.*, 2003) was an early miRNA profiling study, with only a few sequencing reads per miRNA. Additionally, the authors did not induce the ES cells to differentiate down the trophectoderm lineage specifically. These details may explain why Houbaviy *et al.* reached a different conclusion about the

expression of *miR-92a* during differentiation than our findings indicate. *miR-92a* is a member of two paralogous miRNA clusters: the *miR-17-92a* miRNA cluster and the *miR-106a-363* cluster and is closely related in sequence to *miR-25* and *miR-363*. Additionally, another paralogous miRNA cluster, *miR-106b-25*, also exists. These three miRNA clusters are highly conserved across vertebrates, and are believed to have arisen through a series of duplication and deletion events during early vertebrate evolution (Tanzer & Stadler, 2004). As well as *miR-92a*, the sequence related *miR-25* was also differentially regulated during trophectoderm differentiation, as shown by miRNA northern analysis. However, *miR-363* was not detected in the profiling data and expression was barely detectable by miRNA northern analysis. This finding is consistent with a previous analysis that failed to detect expression of *miR-363* in ES cells by RNase protection assay (Ventura *et al.*, 2008). Low expression of *miR-363* in ES cells may suggest that the entire *miR-106a-363* cluster is expressed at low levels in ES cells, which would also suggest that expression of *miR-92a* originates predominantly from the *miR-17-92a* cluster in ES cells. However, data from the miRNA profiling analysis suggested robust expression of another miRNA from the *miR-106a-363* cluster, *miR-106a*, in ES cells. For this reason it is not known if the *miR-92a* expression observed in ES cells originates from a single genomic region or from both miRNA clusters that encode this miRNA.

Despite the sequence similarity between the *miR-106a-363*, *miR-106b-25* and the *miR-17-92a* clusters, only the *miR-17-92a* cluster is essential for normal murine development. Mice deficient for *miR-17-92a* are smaller than their wild-type counterparts and die soon after birth (Ventura *et al.*, 2008), whereas loss of *miR-106a-363* or *miR-106b-25* has no obvious phenotypic effect. However, mice deficient for both the *miR-17-92a* and *miR-106b-25* miRNA clusters die before E15.0 with a much more severe phenotype than the *miR-17-92a* knockout alone suggesting some functional cooperation between the two clusters during embryonic development. miRNA northern analysis of *miR-92a* and *miR-25* expression in adult mouse tissues showed expression of both miRNAs in numerous tissues including lung, heart and bone marrow. The expression of *miR-92a* in lung and heart tissue is consistent with the phenotype of *miR-17-92a*-null mice, which exhibit lung hypoplasia and ventricular septal defects (Ventura *et al.*, 2008). Additionally, if *miR-17-92a* null bone marrow is used to reconstitute the bone marrow of irradiated mice, a significant and specific reduction in circulating B-cells is

observed (Ventura *et al.*, 2008). The fact that *miR-92a* is highly expressed in the bone marrow, and loss of the *miR-17-92a* cluster leads to a reduction in B-cells that normally mature in the bone marrow, suggests that *miR-92a* may play a role in the bone marrow to promote B-cell development.

miR-92 has been shown to have roles in endoderm formation and regulation of left-right asymmetry in zebrafish embryos (Li *et al.*, 2011), and in regulation of angiogenesis in mice (Bonauer *et al.*, 2009). Because of these functional associations with tissue growth and differentiation, and because *miR-92a* was differentially expressed during trophectoderm differentiation of ES cells, it was hypothesised that this miRNA, and sequence-related miRNAs, may be functionally involved in regulating genes involved in the differentiation process. However, depletion of *miR-92a*, *miR-25* and *miR-363* in ES cells had no effect on the expression of key ES cell genes or markers of trophectoderm differentiation. This is in accordance with the finding that mice deficient for the *miR-17-92a* cluster do not exhibit placental defects and embryos deficient for both the *miR-17-92a* and *miR-106b-25* clusters survive past the blastocyct stage (Ventura *et al.*, 2008). However, it cannot be ruled out that *miR-92a*, *miR-25* and *miR-363* are involved in fine-tuning the expression of other genes involved in trophectoderm differentiation or maintenance of the ES cell state that were not investigated in this study. Additionally, the possibility that these miRNAs might be required for the regulation of cellular events coincident with differentiation, such as modulating the cell cycle, was considered.

3.4.5 *miR-92a* regulates ES cell proliferation

Investigation of the role of *miR-92a*, and sequence-related miRNAs, in cell proliferation showed that knockdown of *miR-92a*, but not *miR-25* or *miR-363*, caused a significant reduction in ES cell numbers assayed both by cell counts and by MTT assay. Because the miRNA 'seed' region is thought to be the most important region for targeting, it is assumed that miRNAs with similar seed sequences will target similar transcripts (Bartel, 2009). The results presented here do not fully support this conclusion, because *miR-92a*, *miR-25* and *miR-363* have identical seed sequences, but only knockdown of *miR-92a* has an affect on cell number. These data suggest that regions outside the seed sequence may be important for miRNA function. This finding is supported by the observation that, rather than

just across the seed region, miRNA sequences show conservation across their entire length (Brenneck et al., 2005). Additionally, two broad categories of miRNA target sites are thought to exist: one class that requires perfect seed complementarity, and another class that has imperfect seed binding, and depends on strong 3' compensatory binding (Brennecke *et al.*, 2005). It is possible that the function of *miR-92a* in cell cycle regulation is mediated through regulation of targets that are not reliant solely on perfect seed complementarity, but require binding in other regions. This may explain the reason why inhibition of *miR-25* and *miR-363* has no effect on ES cell numbers.

Knockdown of *miR-92a* in ES cells led to a decrease in cell numbers. This suggests that *miR-92a* may function to promote the proliferation of ES cells, a conclusion supported by the demonstration that expression of the *miR-17-92a* cluster is regulated by the cell proliferation-related transcription factor, Myc, which is highly expressed in ES cells (Aguda *et al.*, 2008). Indeed, Cloonan *et al.* reported that the *miR-17-92a* cluster is differentially expression during different phases of the HeLa cell cycle suggesting that this locus is in itself cell cycle regulated, and increasing the likelihood that the effect of *miR-92a* inhibition on cell number is mediated through regulation of the cell cycle (Cloonan *et al.*, 2008). Additionally, when introduced into miRNA-deficient DGCR8$^{-/-}$ ES cells, members of the *miR-17-92a* and *miR-106a-363* clusters (as well as members of the related cluster *miR-106b-25*) have been shown to rescue the proliferation defect observed in these cells (Wang *et al.*, 2008). However, it should be mentioned that the introduction of *miR-92a* into DGCR8$^{-/-}$ ES cells was not reported to rescue the proliferation defects of these cells. DGCR8$^{-/-}$ ES cells are believed to show proliferation defects because of an accumulation of cells in the G1 phase of the cell cycle, and re-introduction of certain miRNAs rescues this defect. The finding that introduction of *miR-92a* into DGCR8$^{-/-}$ ES cells does not rescue the proliferation defect is surprising because the related miRNA, *hsa-miR-92b* is thought to control the G1/S checkpoint in human ES cells through direct regulation of p57, which inhibits G1/S phase progression (Sengupta *et al.*, 2009).

A recent analysis comparing gene expression between the inner cell mass and embryonic stem cells showed that *miR-92a* was expressed at a significantly higher level in embryonic stem cells relative to the inner cell mass (Tang *et al.*, 2010). Differential expression of *miR-92a* between the ICM and ES cells suggests that this miRNA may contribute to the ES cell state through maintenance of pluripo-

tency or the ability to self-renew indefinitely, two characteristics that are not part of the normal developmental program. High levels of *miR-92a* may be maintained in ES cells through BMP-mediated regulation. The date presented here show that BMP inhibition lead to a reduction in expression of *miR-92a* in ES cells. The finding that *pre-miR-17-92a* expression was reduced in the absence of BMP was previously demonstrated by Wang *et al.*, who showed a reduction in expression of *pre-miR-17-92a* in BMP knockout embryos relative to wild-type embryos (Wang *et al.*, 2011). They further showed that BMP directly regulated expression of this cluster through binding of SMAD1/5 proteins to a region within *miR-17-92a*. Our finding that *miR-92a* expression decreases upon BMP inhibition shows that BMP regulation of *miR-17-92a* expression also occurs in mouse ES cells. However, it must be noted that while dorsomorphin is an inhibitor of BMP signalling, it has also been shown to have off-target effects, for example inhibition of AMP kinase (Zhou *et al.*, 2001). Therefore while the results seen here are likely to be a result of BMP inhibition, the possibility exists that changes in *miR-92a* expression are due to modulation of factors other than BMP signalling. BMP already has a known role in ES cells – it promotes self-renewal by inhibiting neural differentiation through induction of Id proteins (Ying *et al.*, 2003). The data presented here suggest that BMP may have a secondary role to promote ES cell proliferation or survival by regulating expression of *miR-92a*.

In keeping with a role in proliferation of certain cell types, both the *miR-17-92a* and the *miR-106a-363* clusters have been associated with oncogenesis (Petrocca *et al.*, 2008, Landais *et al.*, 2007). Both ES cells and cancer cells proliferate rapidly, and may share some common features that are thought to be predominantly mediated through high expression of myc in both cell types (Kim *et al.*, 2010a). The *miR-17-92a* cluster is regulated by myc, which suggests a potential overlapping function for *miR-92a* in promoting the rapid proliferation of both cancer cells and ES cells. Of note is the finding that inhibition of *miR-92a* expression did not affect the cell numbers of the mouse fibroblast cell line, 3T3. This suggests that the effect of *miR-92a* on cell numbers may be specific to features of the ES cell cycle, which lacks the G1 phase and is not dependent on Cdk4/6-CyclinD-mediated phosphorylation of Rb protein (Savatier *et al.* 1994). Speculatively, perhaps similarities between the cell cycle of ES cells and cancer cells result in *miR-92a*-mediated regulation of ES and cancer cell proliferation, but not regulation of cell proliferation in other cell types. During ES cell differ-

entiation, the cell cycle regains the G1 phase and regulation by Cdk4/6-CyclinD, and cell proliferation slows (Savatier *et al.*, 2002). The reduction in ES cell numbers observed upon *miR-92a* inhibition is similar to that seen when mouse ES cells are induced to differentiate into trophectoderm. Perhaps these data suggest that *miR-92a* functions as a regulator of the ES cell cycle, but not the cell cycle of differentiated cell types. It is possible that *miR-92a* is downregulated during TE differentiation from ES cells in order to facilitate the changes in cell cycle regulation that occur with ES cell differentiation.

While it is believed, for reasons described above, that the decrease in cell number observed upon knockdown of *miR-92a* is a result of a reduction in cell proliferation, it cannot be ruled out that *miR-92a* may have a cell survival function and therefore its reduction may result in an increase in apoptosis and a subsequent decrease in cell numbers. Indeed, enforced expression of the *miR-17-92a* cluster in a mouse model of B-lymphoma has been reported to lead to a reduction in apoptosis in c-myc induced lymphomas, suggesting that expression of this miRNA cluster may be anti-apoptotic (He *et al.*, 2005), and it has recently been shown that *miR-92a* directly targets the pro-apoptotic protein, Bim (Tsuchida *et al.*, 2011). Although an obvious increase in cell death was not observed following inhibition of *miR-92a* in the experiments presented in this chapter, further analysis would be required to rule out the possibility that increased cell death is the cause of a reduction in cell numbers following *miR-92a* inhibition in ES cells.

3.4.6 Summary and Future Work

miRNAs have been shown to be essential for correct embryonic development, and it was hypothesised that they are involved in regulating individual developmental transitions. The data presented here show that miRNAs are differentially represented during a cell-based model of a developmental transition, and that differentially regulated miRNAs can be functionally associated with the differentiation process. This analysis identified miRNAs that are differentially represented during ES cell differentiation into trophoblast, being either induced during differentiation, or principally associated with the pluripotent state. Several of the miRNAs that were represented at higher levels in differentiated cells are encoded within a gene called *Sfmbt2* that has been shown in this analysis, and in previous analyses, to be associated with the trophectoderm lineage. Intriguingly,

the large miRNA cluster that is encoded within *Sfmbt2* is thought to be specific to only some rodent species. Further examination of *Sfmbt2* and its encoded miRNA cluster may provide insight into possible functional reasons behind rapid evolution of this cluster in rodents, but not other mammals.

This analysis also identified miRNAs that were represented more highly in undifferentiated cells than differentiated cells, and therefore may be associated with maintenance of the ES cell state. ES cell-associated miRNAs may be downregulated during the transition from ES cell to trophoblast in order to facilitate the many cellular changes required for differentiation to occur. For example, downregulation of *miR-92a* may facilitate the reduction in proliferation from the ES cell state to that of a differentiated cell. Further analysis in this area would determine the exact nature of the effect of *miR-92a* inhibition on ES cell number. Additionally, expression analysis during embryogenesis may indicate if *miR-92a* regulation of proliferation or survival is likely to occur in pluripotent cells of the early mouse embryo.

Chapter 4

Methods to identify miRNA target gene transcripts in ES cells

4.1 Introduction

4.1.1 Sequence-based methods for identifying miRNA targets

miRNAs function through post-transcriptional regulation of target genes, and therefore identification of miRNA targets is a key goal in any miRNA-focused work. Because of the imperfect binding between a miRNA and its target mRNA molecule, the identification of miRNA targets has proved challenging. Nucleotides 2-8 from the 5' end of the miRNA, which make up the seed region, generally pair perfectly with the target mRNA, but a 7 nucleotide sequence is too short to allow specific identification of targets in the *mus musculus* genome, and the involvement of the 3' end of the miRNA in targeting is still under debate. In addition, secondary and tertiary RNA structures must be taken into consideration, and obviously the miRNA and mRNA must be expressed at the same time in the same tissue for an interaction to be biologically significant. In this part of the project, various means of identifying mRNA targets of miRNAs were assessed.

Some key requirements for miRNA function have been identified, and these have been utilised in computational methods to predict miRNA targets. The most general feature of miRNA regulation is the perfect complementarity between the

seed region of the miRNA and regions within the mRNA target. Together with other criteria, such as conservation of the target site and site accessibility, this makes up the basis of the majority of target prediction programs. Examples of commonly used algorithms are miRanda (John *et al.*, 2004), TargetScan (Lewis *et al.*, 2003) and PicTar (Krek *et al.*, 2005).

A major limitation of computation approaches for miRNA target prediction is that they cannot highlight novel aspects of mRNA target recognition because the programs predict miRNA targets based on only a few established rules. The weakness of target prediction software is highlighted by the interaction between let-7 and the oncogene *Ras*: *Ras* has been validated as a target of *let-7* in several systems (Johnson *et al.*, 2005, Kumar *et al.*, 2007), but at least three major prediction programs (TargetScan, miRanda and Diana-microT v3.0) fail to identify this interaction because the interaction includes GU pairings (Johnson *et al.*, 2005). Another limitation of target prediction algorithms is the number of false-positives. Many predicted targets do not show miRNA regulation in validation experiments. For example, Jiang *et al.* (2009) used luciferase reporter assays to test miRNAs predicted to target *CyclinD1*. Of the 45 miRNAs tested, only 7 recapitulated regulation in the validation experiments. Another analysis showed a similar result for miRNAs predicted to target p21 (Wu *et al.*, 2010), with only 28 out of 266 miRNA predictions standing up to experimental validation. The list of 266 miRNAs was compiled from four of the most widely used target prediction algorithms (TargetScan, miRanda, PicTar and RNA22). The authors observed that overall the programs produced very different sets of miRNA predictions but those that considered conservation of target sites performed the best (although clearly these programs would not be effective for species-specific miRNAs). Although the large numbers of false positive results generated by target prediction programs could in theory be determined by experimental analysis, what is difficult to estimate is the number of real miRNA targets that these programs miss. The targeting of *Ras* by *let-7*, mentioned above, is an example of a miRNA target that was not predicted by target prediction software.

Combining data from experimental approaches for identifying miRNA targets may help to overcome the limitations of computer based target prediction. Numerous attempts have been made to try and identify miRNA targets experimentally both directly and indirectly. Direct methods would include biochemical approaches involving pulldown of RISC complex proteins, while indirect meth-

ods include transcriptome-based methods (such as microarrays) and proteome analysis methods (including SILAC).

4.1.2 Direct experimental methods to identify miRNA targets

Biochemical approaches involve pull-down of endogenous or tagged AGO proteins (Beitzinger *et al.*, 2007, Easow *et al.*, 2007) or tagged miRNA molecules (Orom & Lund., 2007) followed by analysis of associated mRNAs. A study by Easow *et al.* (2007) in Drosophila identified 89 mRNAs that were reproducibly enriched in an HA tagged AGO1 immunoprecipitated sample compared with control samples. Although they showed that the mRNAs were enriched for miRNA seed matches, the authors did not confirm if they were direct miRNA targets. Beitzinger *et al.* (2007) performed direct immunoprecipitation of AGO1 and AGO2 using monoclonal antibodies in HEK293 human cells. They identified 95 RNAs that co-purified with AGO1 and 49 RNAs that co-purified with AGO2, and validated that 5 out of 6 of the AGO1-bound miRNAs tested were direct miRNA targets. It is notable that only 60% of AGO1-bound mRNAs and 50% of AGO2-bound mRNAs were predicted as miRNA targets by computation prediction programs. While it is possible that a substantial proportion of the AGO-associated mRNAs identified in this study are not direct miRNA targets suggesting a high level of background noise in the system, it seems more likely that the prediction programs are lacking. The immunoprecipitation methods mentioned above yield lower numbers of potential targets than trancriptome-based methods and have the benefit of identifying transcripts more likely to be directly involved in miRNA-mediated regulation because the transcripts identified have a direct association with RNA silencing machinery.

Crosslinking of RNA-protein complexes has recently been shown to increase the number of potential target transcripts being identified in immunoprecipitation-based methods (Chi *et al.*, 2009). Hafner *et al.* (2010a&b) developed the PAR-CLIP method, which involves UV crosslinking of photoactivatable nucleoside analogs (following incorporation into cellular RNAs) to their associated proteins followed by immunoprecipitation of the desired protein and recovery of associated RNAs. Using this method the group identified approximately 4000 clusters of sequence reads that overlapped between AGO proteins 1-4 in HEK293 cells. Because the different AGO proteins bound similar sets of transcripts, the authors

amalgamated the data sets yielding 17,319 clusters corresponding to 4647 different transcripts of which 557 were miRNAs. Enrichment for 7-mer sequences was investigated in the combined data set of sequences relative to random sequences. The authors found that the most significantly enriched 7-mers corresponded to the reverse complement of the seed region of the most abundant HEK293 miRNAs. These data show that the PAR-CLIP method yields large numbers of AGO-associated mRNAs, which are likely to represent miRNA targets. For these reasons the PAR-CLIP protocol was investigated as a possible method for identifying direct miRNA targets in ES cells.

4.1.3 Indirect experimental methods to identify miRNA targets

Alternative to the immunoprecipitation-based direct methods of identifying miRNAs targets, are indirect methods of target identification. These methods yield data on changes in expression of mRNA or proteins. The relationship between these changes and changes in miRNA expression are then correlated and used to identify potential miRNA targets. These methods identify both direct and indirect miRNA targets and so may provide insight into the broader effects on the transcriptome and/or proteome resulting from changes in miRNA expression.

When it was realised that as well as translational repression, animal miRNAs may also down-regulate the level of their target transcripts (Bagga *et al.*, 2005, Lim *et al.*, 2005), transcriptome-based methods began to be used to identify miRNA targets. As a result several studies have attempted to identify miRNA targets through overexpression or inhibition of a miRNA, followed by transcriptome analysis by microarray (Lim *et al.*, 2005). More recently, studies have combined the transcriptome analyses from both overexpression and inhibition of a miRNA in order to gain a more accurate indication of direct targets (Nicolas *et al.*, 2008). Transcriptome-based methods are successful at producing putative lists of direct and indirect miRNA targets, but are limited to identifying those mRNAs that are degraded by detectable levels by their targeting miRNA.

Proteome analysis has also been used to identify miRNA targets through stable isotope labelling by amino acids in cell culture (SILAC) (Vinther *et al.*, 2006, Yang *et al.*, 2010). The SILAC method involves analysing the protein content

of two differently labelled cell culture extracts using mass-spectrometry. One advantage of this method lies in the fact that the two samples are combined prior to cell lysis and sample preparation. This is a complex, multi-step procedure and combining samples prior to protein extraction avoids the introduction of artificial differences frequently observed with this type of proteomic analysis. Samples can be combined in this way because the proteins in one sample have incorporated heavy isotopes of particular amino acids, allowing them to be distinguished from proteins from the other sample during mass-spectrometry analysis.

Proteome-based methods have an advantage over transcriptome-based methods in the context of identifying miRNA targets. This is because the SILAC method will detect changes in putative target levels as a result of both methods of miRNA regulation: mRNA destabilisation and translational inhibition. The SILAC technique was investigated as a method to identify miRNA targets in ES cells by correlating differential expression of proteins during trophoblast differentiation with differential expression of miRNAs identified in Chapter 3 of this thesis.

4.2 Aims of Chapter 4

The aim of this chapter was to compare different methods of miRNA target identification. Current predictive software approaches and direct and indirect biochemical approaches were evaluated.

4.3 Results

4.3.1 Investigation of sequence-based methods of identifying miRNA target transcripts

4.3.1.1 Sequence-based prediction of *miR-92a* targets

There are numerous programs currently available for computational prediction of miRNA targets but these methods have varying degrees of reliability (Baek *et al.*, 2008, Selbach *et al.*,2008) as a result of using different parameters for target

prediction. In order to increase the power of computer-based target predictions there are now online utilities that amalgamate the target prediction results from multiple prediction programs. One such resource is miRecords (Xiao *et al.*, 2009, http://mirecords.biolead.org/). miRecords shows prediction miRNA targets from up to 11 different target predictions programs. A summary of the target prediction programs represented, and their criteria for target identification, is shown in *Table 4.1.*

Assuming that the criteria adopted by different prediction algorithms have merit, it was reasoned that miRNAs targets predicted by a number of programs would be more likely to represent genuine targets than those predicted by a single algorithm. For this reason miRecords was used for all target prediction in this chapter. Firstly, miRecords was used to predict targets for *miR-92a.* This analysis generated 6331 predicted interactions for at least one of the 8 target prediction programs that yielded predictions for *miR-92a. Table 4.2* shows the 69 targets for *miR-92a* that were predicted by at least 5 of the 8 target prediction programs.

According to the database of 'Validated Targets' available through miRecords, only one gene listed in *Table 4.2* has been validated as a direct *miR-92a* target. This target is *Mylip* (Myosin regulatory light chain interacting protein), which is also the only target to be predicted by 6 of the 8 algorithms. *Mylip* interacts with myosin regulatory light chain B and affects cytoskeleton interaction regulating cell motility, such as neurite growth (Olsson *et al.*, 1999). The only other validated target listed for *miR-92a* is LOC100048439, which is predicted as an *miR-92a* target by 3 of the 8 target prediction programs in miRecords. Both of these targets were validated in a paper by Landais *et al.* (2007) who showed that the protein levels of MYLIP and LOC100048439 (similar to HIPK3) decreased upon *miR-92a* expression. In addition, the 3' UTRs of *Mylip* and LOC100048439 were cloned downstream of firefly luciferase and luciferase reporter activity was shown to decrease as a result of *miR-92a* expression.

In order to determine if *Mylip* may be a functional target of *miR-92a* during trophectoderm differentiation of ZHBTc4 ES cells, expression of *Mylip* was investigated during ZHBTc4 ES cell differentiation and in ES cells following knockdown of *miR-92a* expression. A QPCR assay was designed and optimised for *Mylip* (*Figure 4.1a*). Primer optimisation was carried out on pooled cDNA derived from undifferentiated and differentiated ES cell cultures. Optimisation QPCR

Program	Prediction Criteria	Reference
Diana-MicroT	7, 8 or 9 nt seed match accepted Allows G:U wobble Sites ranked according to site type and conservation Adds weight for multiple sites in the 3'UTR Compares interactions to mock miRNAs - signal-to-noise ratio	M. Maragkakis et al (2009) **DIANA-microT web server: elucidating microRNA functions through target prediction.** *Nucleic Acids Research*
MicroInspector	analyse a user defined mRNA sequence for miRNA binding allows: variation of temperature setting of energy values selection of different miRNA databases visual inspection of secondary structures	Ventsislav Rusinov et al (2005) **MicroInspector: a web tool for detection of miRNA binding sites in an RNA sequence.** *Nucleic Acids Research*
miRanda	sequence complementarity - position-weighted local alignment takes into account G-U wobble pairs allows moderate insertions and deletions weighting scheme rewards complementarity at the miRNA 5' end free energies of RNA-RNA duplexes conservation of target sites in related genomes	Enright AJ et al (2003) **MicroRNA targets in Drosophila.** *Genome Biology*
MirTarget2	Training-based algorithm. Training based on following criteria: Seed conservation Seed type Base composition RNA Secondary structure Location in 3' UTR	Xiaowei Wang and Issam M. El Naqa (2008) **Prediction of both conserved and nonconserved microRNA targets in animals.** *Bioinformatics*
miTarget	Uses a support vector machine classifier for miRNA target prediction Takes into account: structural features - seed binding and binding of 3' of miRNA Thermodynamic features – free energy of total alignment, seed part and 3' part Position-based features – specific nucleotide pairs at specific positions Does not take into account sequence conservation	Kim et al (2006) **miTarget: microRNA target gene prediction using a support vector machine.** *BMC Bioinformatics*
NBmiRTar	machine learning by a naïve Bayes classifier uses miRanda output applies Naïve Bayes classifier to miRanda output Uses 'seed' and 'out-seed' regions of complementarity to refine target list Does not rely on conservation	Yousef et al (2007) **Naïve Bayes for microRNA target predictions- machine learning for microRNA targets**. *Bioinformatics*
PicTar	Calculates 7nt perfect matches (starting at position 1 or 2 of miRNA) Filters results based on free energy Considers conservation of target sites Accounts for synergistic effects of multiple binding sites	Krek et al (2005) **Combinatorial microRNA target predictions.** *Nature Genetics*
PITA	Considers seeds of 6-8 nt beginning at position 2 of the miRNA No mismatches or loops allowed G:U wobble allowed in 7 & 8-mers Calculates energy based score factoring in secondary structure Does not rely on conservation of target sites	Kertesz et al (2007) **The role of site accessibility in microRNA target recognition.** *Nature Genetics*
RNA22	Looks for intra- and inter-species patterns of conserved sequence features Generates reverse complement and identifies putative miRNA binding sites Determine identity of miRNA Predict secondary structure and free energy of interaction Filter results based on user-defined parameters	Miranda et al (2006) **A Pattern-Based Method for the Identification of MicroRNA Binding Sites and Their Corresponding Heteroduplexes.** *Cell*
RNAhybrid	miRNA hybridised to target in an energetically optimal way Forbids intramolecular base pairs and branching structures Restricts bulge loops and internal loops to maximum length Additional binding sites are found by masking previously reported sites Uninterrupted helical binding is required in the seed region	Rehmsmeier et al (2004) **Fast and effective prediction of microRNA/target duplexes.** *RNA*
TargetScan	Searches for seed binding (bases 2-8 of miRNA) Extends seed with additional base pairs (including G:U) as far as possible Optimises base pairing of the region around the seed Assigns a folding free-energy to each miRNA:target site interaction Considers number of target sits in 3'UTR Considers species conservation	Lewis et al (2003) **Prediction of Mammalian MicroRNA Targets.** *Cell*

Table 4.1 – *Target prediction programs.* Table showing prediction criteria for the miRNA target prediction programs used by miRecords.

Refseq	Symbol	total
NM_153789	Mylip	6
NM_173426	1700012H17Rik	5
NM_028720	3930401K13Rik	5
NM_146217	Aars	5
NM_011780	Adam23	5
NM_009627	Adm	5
NM_001003909	Ankib1	5
NM_001001182	Baz2b	5
NM_199149	BC024868	5
NM_133889	Bsdc1	5
NM_007570	Btg2	5
NM_009806	Cask	5
NM_001033261	Ccdc131	5
NM_016746	Ccnc	5
NM_146040	Cdca7l	5
NM_001039154	Cdh8	5
NM_009876	Cdkn1c	5
NM_027882	Cic	5
NM_198300	Cpeb3	5
NM_172513	D1Ertd53e	5
NM_026821	D4Bwg0951e	5
NM_013505	Dsc2	5
NM_138677	Edem1	5

Refseq	Symbol	total
NM_172591	Fcho2	5
NM_173182	Fndc3b	5
NM_008037	Fosl2	5
NM_175284	Fzd10	5
NM_008083	Gap43	5
NM_021434	Gpr180	5
NM_026240	Gramd3	5
NM_145890	Grhl1	5
NM_016886	Gria3	5
NM_020259	Hhip	5
NM_010577	Itga5	5
NM_010637	Klf4	5
NM_015771	Lats2	5
NM_008549	Man2a1	5
NM_009157	Map2k4	5
NM_010863	Myo1b	5
NM_010864	Myo5a	5
NM_026267	Necap1	5
NM_001039386	Nelf	5
NM_001039387	Nelf	5
NM_020276	Nelf	5
NM_008699	Nkx2-3	5
NM_015760	Nox4	5

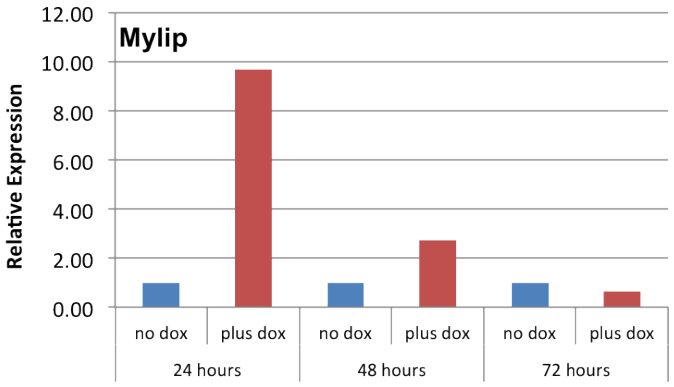
Refseq	Symbol	total
NM_013625	Pafah1b1	5
NM_011041	Pax9	5
NM_183028	Pcmtd1	5
NM_011866	Pde10a	5
NM_198600	Pols	5
NM_001038621	Rabgap1l	5
NM_011264	Rev3l	5
NM_027897	Rhpn2	5
NM_001038993	Rnf38	5
NM_175201	Rnf38	5
NM_011278	Rnf4	5
NM_028385	Setd5	5
NM_178392	Snapc1	5
NM_026130	Srpr	5
NM_021314	Tacc2	5
NM_153484	Tef	5
NM_009372	Tgif1	5
NM_009427	Tob1	5
NM_011632	Traf3	5
NM_031998	Tsga14	5
NM_175482	Usp28	5
NM_030215	Wrnip1	5
NM_026417	Yipf4	5

Table 4.2 – *Predicted targets of miR-92a.* Table showing Refseq identifiers and gene symbols for *mus musculus* targets of *miR-92a* as predicted by at least 5 different target prediction programs. 'Total' column indicates the number of programs that predict each target.

(a)

Gene	Single Product?	Effective amplification range and Ct spread	RSq	Amplification efficiency (%)
Mylip	Yes	1,000 fold (Ct 26-35.5)	0.996	105.6
Trim33	Yes	10,000 fold (Ct 23.5-36.5)	0.990	103.1
Dnmt3a	Yes	100 fold (Ct 23-30)	0.989	94.6
DGCR8	Yes	10,000 fold (25-38)	0.978	104.8
ActB	Yes	100,000 fold (Ct 16-33)	0.999	101.1

(b)



(c)

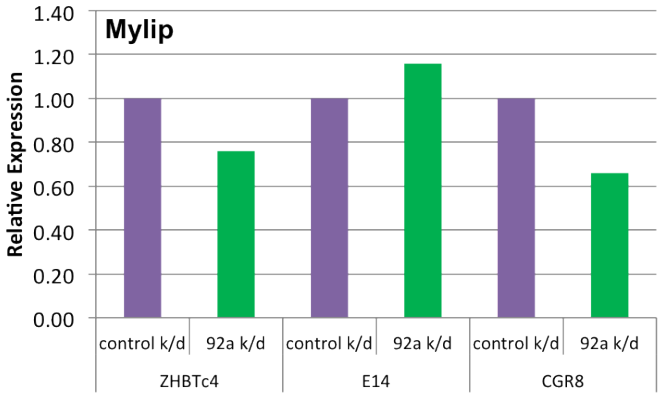


Figure 4.1 – *Expression analysis of Mylip following Oct4 knockdown and miR-92a knockdown.* (a) Optimisation details for QPCR primers used in Chapter 4. (b) ZHBTc4 ES cells were grown without doxycycline or in the presence of 1µg/ml doxycycline for 24, 48 or 72 hours. RNA was extracted from untreated and doxycycline-treated cell cultures, DNAse treated and used to generate cDNA for QPCR analysis of *Mylip* expression. QPCR Ct values were normalised to *β-Actin* expression. Untreated cultures are represented by a value of 1.0 and results show gene expression in doxycycline treated cultures relative to non-treated cultures at each timepoint. N=1. (c) ZHBTc4, E14 and CGR8 ES cells were transfected with a combination of LNAs targeting *miR-92a*, *miR-25* and *miR-363* or an LNA targeting DM1 (*gga-miR-2954* - control transfection). RNA was extracted 96 hours after transfection. RNA was DNAse treated, converted to cDNA and used for QPCR analysis of expression of *Mylip*. QPCR Ct values were normalised to *β-Actin* and expression is shown for the *miR-92a*, *miR-25* and *miR-363* knockdown relative to expression in the control transfection. N=1.

was performed on a series of cDNA dilutions over a range of $1x10^6$ fold. Primers were considered suitable for use if they amplified a single product over >100-fold range with 95-105% efficiency and a 'goodness of fit' of >0.95. Primer sequences can be found in Materials and Methods, *Figure 4.1a* shows the amplification efficiencies, 'goodness of fit' and amplification range for primer pairs used in Chapter 4 of this thesis. ZHBTc4 ES cells were grown without doxycycline, or in the presence of 1μg/ml doxycycline, and RNA was harvested after 24, 48 and 72 hours of growth and used to generate cDNA for QPCR analysis of *Mylip* expression (*Figure 4.1b*). After 24 and 48 hours of growth an increase in *Mylip* expression was observed in doxycycline-treated cultures relative to untreated cultures. This expression pattern is consistent with a decrease in *miR-92a* expression at these timepoints. However, no difference in *Mylip* expression was observed 72 hours after addition of doxycycline. In order to assess if *miR-92a* and related miRNAs had any direct effect on *Mylip* expression in ES cells, QPCR analysis was carried out on ES cells after combined inhibition of *miR-92a*, *miR-25* and *miR-363* using 100nM of LNA oligonucleotides complementary to the three miRNAs. Control cultures were transfected with an LNA designed to inhibit a novel chicken miRNA (DM3) that is not found in mice. *Figure 4.1c* shows that there is no difference in the expression of *Mylip* following miRNA knockdown in ZHBTc4 ES cells. These data suggest that *Mylip* is not a functionally significant target of *miR-92a* in ES cells. However, it must be noted that this analysis investigates mRNA expression and significant differences in *Mylip* expression at the protein level as a result of ZHBTc4 differentiation or miRNA knockdown cannot be ruled out. Additionally, because this data is from a single experiment, conclusions about *Mylip* expression are not definitive

Of the 69 potential *miR-92a* targets predicted by 5 prediction programs, none have been validated as direct targets. However, one of the predicted targets, *Cdkn1c*, has been validated as a target of the closely related miRNA *hsa-miR-92b*, (Sengupta *et al.*, 2009) a miRNA with a single nucleotide difference to *miR-92a*. *Cdkn1c*, also known as *P57*, would be an interesting potential target of *miR-92a*, as *Cdkn1c* is a tumour suppressor, and has a role in cell proliferation in mouse embryos (Zhang *et al.*, 1997). Reduction in the expression of *Cdkn1c* is associated with Beckwith-Wiedemann syndrome in humans. This syndrome results in overgrowth, a predisposition to tumours and congenital malformations (Higashiomoto *et al.*, 2006). The function of *miR-92a* in promoting an increase

in cell numbers could partly be explained by its suppression of the inhibitor of cell proliferation, *Cdkn1c*.

4.3.1.2 Gene ontology analysis of *miR-92a* predicted targets

In order to assess whether the predicted targets of *miR-92a* show enrichment for functional associations, the list of targets predicted by at least 3 target prediction programs was subjected to analysis for enrichment of GO terms (*Figure 4.2a*) and KEGG pathways (*Figure 4.2b*) using DAVID Bioinformatics Resources 6.7 (http://david.abcc.ncifcrf.gov/home.jsp). Initially, the list of targets predicted by 5 prediction programs (shown in *Table 4.2*) was used for analysis but this did not yield significant enrichment for any GO terms or KEGG pathways, presumably because the number of targets was relatively low. 1238 targets are predicted for *miR-92a* by at least 3 target prediction programs, and it was this list of targets that was used for GO term and KEGG pathway analysis.

Enrichment values were calculated as –log2(P-Value) and a cut-off P-Value of <0.00003 (corresponding to an enrichment score of >15) was assigned. The list of predicted targets showed enrichment in multiple GO terms (*Figure 4.2a*) and KEGG pathways (*Figure 4.2b*). The enriched GO terms contain multiple terms involving positive regulation of metabolic processes, biosynthetic processes and gene expression. The enriched KEGG pathways contain a large proportion of cancer-associated pathways. Both of these findings are of interest in the context of *miR-92a* functioning as a regulator of cell proliferation. However, a similar analysis of targets of the unrelated *miR-155* also showed enrichment of targets in GO terms associated with gene expression and regulation of metabolic processes, but not biosynthetic processes (data not shown).

4.3.1.3 Sequence-based prediction of Sfmbt2 miRNA cluster targets

The *Sfmbt2* miRNA cluster was identified as being upregulated during trophectoderm differentiation of ES cells and contains multiple highly related miRNAs within intron 10 of *Sfmbt2* (see *Figure 3.7*). Target predictions for the *Sfmbt2* miRNA cluster were investigated to see if they would give insight into the function of this miRNA cluster. miRecords was used to compile a list of targets predicted

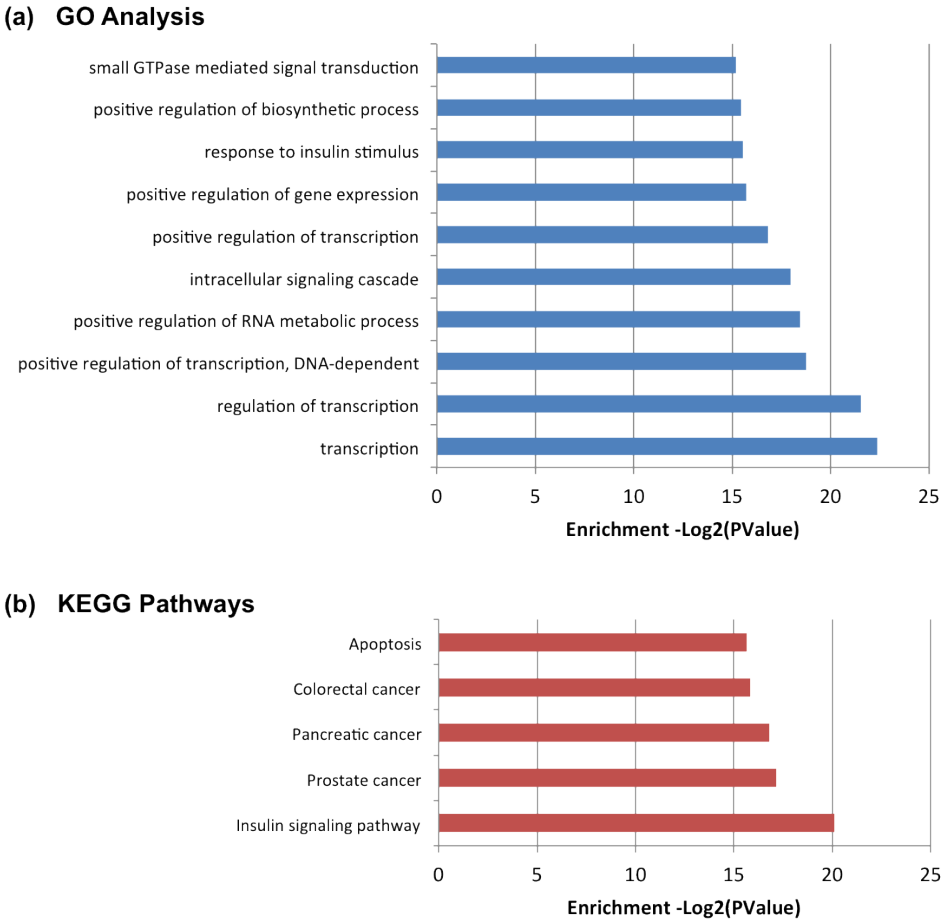


Figure 4.2 – *Analysis of functional associations of predicted miR-92a targets with GO Terms and KEGG pathways.* Enrichment of GO Terms (a) and KEGG pathways (b) in list of targets for *miR-92a* predicted by at least 3 different target prediction programs. Names of GO Terms and KEGG pathways are shown that had an enrichment score of >15, which corresponds to a P-Value of <0.00003. Enrichment for GO Terms and KEGG pathways was assessed using DAVID Bioinformatics Resources 6.7 (http://david.abcc.ncifcrf.gov/home.jsp).

by at least 3 target prediction programs for each of the 7 differentially represented *Sfmbt2*-cluster miRNAs identified in the profiling analysis. Targets predicted for miR-669d however are from a maximum of 2 target prediction programs due to predictions being unavailable from the other prediction programs. Sequences of the 7 *Sfmbt2* miRNAs used in this analysis are shown in *Figure 4.3a*.

The 7 lists of predicted targets were then compiled using the free-ware BioLayout Express 3D (Theocharidis *et al.*, 2009) program for network visualisation and analysis. This analysis resulted in a list of potential targets of all 7 of the differentially represented miRNAs from the *Sfmbt2* cluster. The list was then filtered to show targets shared by multiple different miRNAs. *Figure 4.3b* shows a list of targets predicted for at least 5 of the 7 differentially represented *Sfmbt2* cluster miRNAs. According to the database of 'Validated Targets' available through miRecords, none of these predicted targets have yet been validated as real targets of any of the miRNAs in this analysis. Interestingly, there were 6 genes that were predicted as targets of all 7 of the *Sfmbt2*-cluster miRNAs in this analysis (*Figure 4.3c*). A level of similarity between target lists is unsurprising because some of the 7 *Sfmbt2* cluster miRNAs belong to the same family and have very similar or identical seed sequences (*Figure 4.3a*). These miRNAs would therefore be expected to be predicted to target the same transcripts by the programmes. However, the seed sequences of other miRNAs within the cluster are different. The dissimilar miRNAs cannot be targeting the same sites within their shared targets, meaning there must be several predicted target sites for *Sfmbt2* miRNAs within these shared targets. This increases the likelihood of the 6 genes being real targets of the *Sfmbt2*-cluster miRNAs and suggests that their repression is important during trophoblast differentiation.

The 6 targets predicted for all 7 of the differentially represented *Sfmbt2*-cluster miRNAs are *Rabgap1l*, *Ogfod1*, *Trim33*, *Btbd14a* (also known as NACC2), *Bcl11b* and *Dnmt3a*. Little is known about the function of NACC2 (NACC family member 2, BEN and BTB (POZ) domain containing) although interestingly the related NACC1 interacts with NANOG in mouse ES cells to regulate proliferation via the Eras/PI3K/Akt pathway (Ma *et al.*, 2009). *Bcl11b* (B-cell CLL/lymphoma 11B (zinc finger protein)) is a transcriptional repressor that is essential for developmental processes of the immune and central nervous systems as well as skin and tooth development (Golonzhka *et al.*, 2009). Germline deletion of *Bcl11b* leads to perinatal lethality (Wakabayashi *et al.*, 2003). *Ogfod1*

(2-oxoglutarate and iron-dependent oxygenase domain containing 1) is a protein found in stress granules that functions to link recovery from stress with translation regulation. Overexpression of *Ogfod1* leads to an increase in apoptosis in stressed cells and those recovering from stress. The pro-apoptotic roles of *Ogfod1* are mediated through its regulation of phosphorylation of eukaryotic translation initiation factor 2alpha (eIF2a) (Wehner *et al.*, 2010). The function of *Rabgap1l* (RAB GTPase activating protein 1-like) is not yet known although it is related to *Rabgap1*, which may have roles in mitosis by participating in a RAB6A-mediated pathway involved in the metaphase-anaphase transition (Miserey-Lenkei *et al.*, 2006). Numerous studies have been published regarding *Trim33* (tripartite motif containing 33) function. *Trim33* is thought act as a transcriptional co-repressor and is essential for embryonic development (Kim & Kaartinen, 2008). It also acts as a SMAD4 monoubiquitin ligase (Dupont *et al.*,2009) enabling cells to set their Nodal responsiveness. In this way *Trim33* is involved in negatively controlling Nodal activity to allow a balance between stem cell self-renewal and differentiation of trophoblast cells in the early embryo (Morsut *et al.*, 2010). In the presence of too much Nodal signalling trophoblast stem cells are not maintained. The *Sfmbt2* miRNA cluster may be required to regulate *Trim33* expression levels during trophoblast differentiation. This would ensure that *Trim33* ubiquitination of SMAD4 stayed within certain boundaries allowing a balance between trophoblast stem cell maintenance and terminal trophectoderm differentiation. *Dnmt3a* (DNA (cytosine-5-)-methytransferase 3 alpha) is a *de novo* DNA methyltransferase that is essential for mouse development (Okano *et al.*, 1999), and genomic imprinting (Hata *et al.*, 2002). *Dnmt3a* interacts with polycomb proteins, suggesting that these proteins may direct *de novo* DNA methylation (Rush *et al.*, 2009). These findings are interesting in the context of *Dnmt3a* regulation by *Sfmbt2*-cluster miRNAs because *Sfmbt2* is both a predicted polycomb-group gene and is imprinted in early embryos (Kuzmin *et al.*, 2008).

In order to determine if *Trim33* and *Dnmt3a* may be a functional targets of *Sfmbt2* miRNAs during trophectoderm differentiation of ZHBTc4 ES cells, expression of these genes was investigated during ZHBTc4 ES cell differentiation and in ES cells following knockdown of *miR-467a** and *miR-669b* expression. QPCR assays were designed and optimised for *Trim33* and *Dnmt3a* (*Figure 4.1a*). ZHBTc4 ES cells were grown without doxycycline, or in the presence of 1μg/ml doxycycline, and RNA was harvested after 24, 48 and 72 hours of growth and used

(a)

mmu-miR-669a	aguugugugugcauguucaugucu
mmu-miR-669b	aguuuugugugcaugugcaugu
mmu-miR-669d	acuugugugugcauguauaugu
mmu-miR-467d	uaagugcgcgcauguauaugcg
mmu-miR-467e	auaagugugagcauguauaugu
mmu-miR-297a	auguaugugugcaugugcaugu
*mmu-miR-467a**	cauauacauacacacaccuaca

(b)

2510009E07Rik	Btbd14a	Epha4	Med14	Pofut1	Tcf4
2600010E01Rik	Cacna1d	Evi2b	Mef2a	Pten	Tcfap2b
2610207I05Rik	Cacnb4	Fbxl2	Mll1	Ptprd	Tet2
4933407H18Rik	Ceacam1	Frs2	Mllt3	Rabgap1	Tgfbr1
5031439G07Rik	Centg2	Gabpb2	Mobp	Rabgap1l	Tgfbrap1
A930041I02Rik	Clock	Gad2	Mxi1	Rp2h	Tiprl
Aak1	Cnr1	Glt8d3	Ncam1	Rreb1	Tmod2
Adam12	Col19a1	Gm467	Nfat5	Sbno1	Trim33
Adam22	Col4a4	Gnal	Nfic	Scn3b	Tsc22d2
Aebp2	Creb1	Gpc6	Nin	Scn8a	Ttbk2
Agl	Csta	Gpr116	Npr3	Sertad2	Vps37a
Amotl1	Cugbp2	Gria2	Nufip2	Sgk3	Wdr37
Ank2	Cxadr	Hist1h1d	Ogfod1	Slc10a5	Ypel2
Ankrd52	Cxxc4	Htr5a	Osbpl8	Slc24a2	Zfp68
Arhgef17	D730040F13Rik	Ids	Oxtr	Slc30a10	Zfp704
Armc2	Depdc6	Il28ra	Papola	Slc35b4	Zfx
Atg2b	Dicer1	Klhl14	Pcdh10	Slc6a6	Zfyve26
Atp5g1	Dnm1l	Klhl15	Pdik1l	Sorbs2	Zkscan1
B4galnt2	Dnmt3a	Lhx9	Pgm5	Sox6	Zxdb
BC026590	Dph3	Lin28b	Phactr1	St8sia4	Zyg11b
Bcl11b	EG545861	Lonrf2	Phf20l1	Stag2	Zzz3

(c)

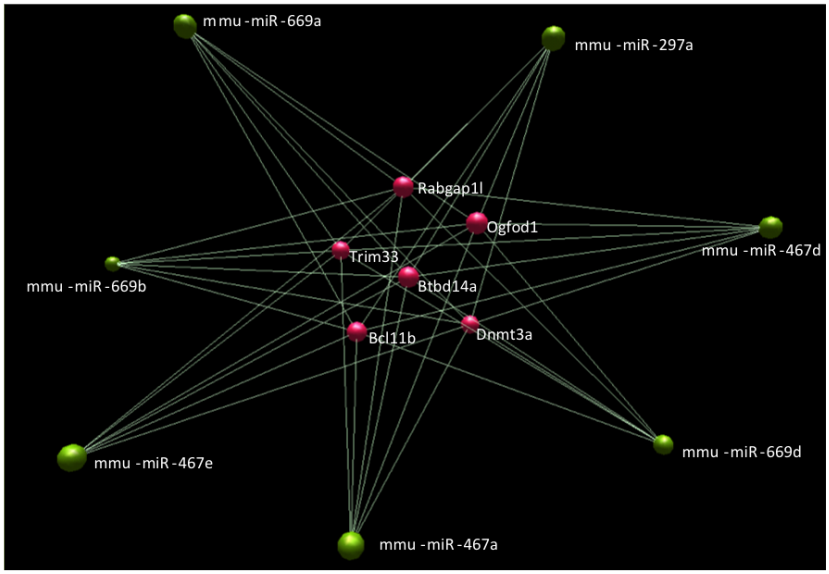


Figure 4.3 – *Predicted targets of the Sfmbt2 miRNA cluster.* (a) Sequences of 7 differentially represented miRNAs in *Sfmbt2* miRNA cluster. Seed sequences are highlighted in colour. (b) List of targets predicted by at least 3 target prediction programs using miRecords for at least 5 of the differentially represented *Sfmbt2* cluster miRNAs. (c) Biolayout Express 3D (http://www.biolayout.org/) schematic diagram showing targets that are predicted for all 7 of the differentially represented *Sfmbt2* cluster miRNAs.

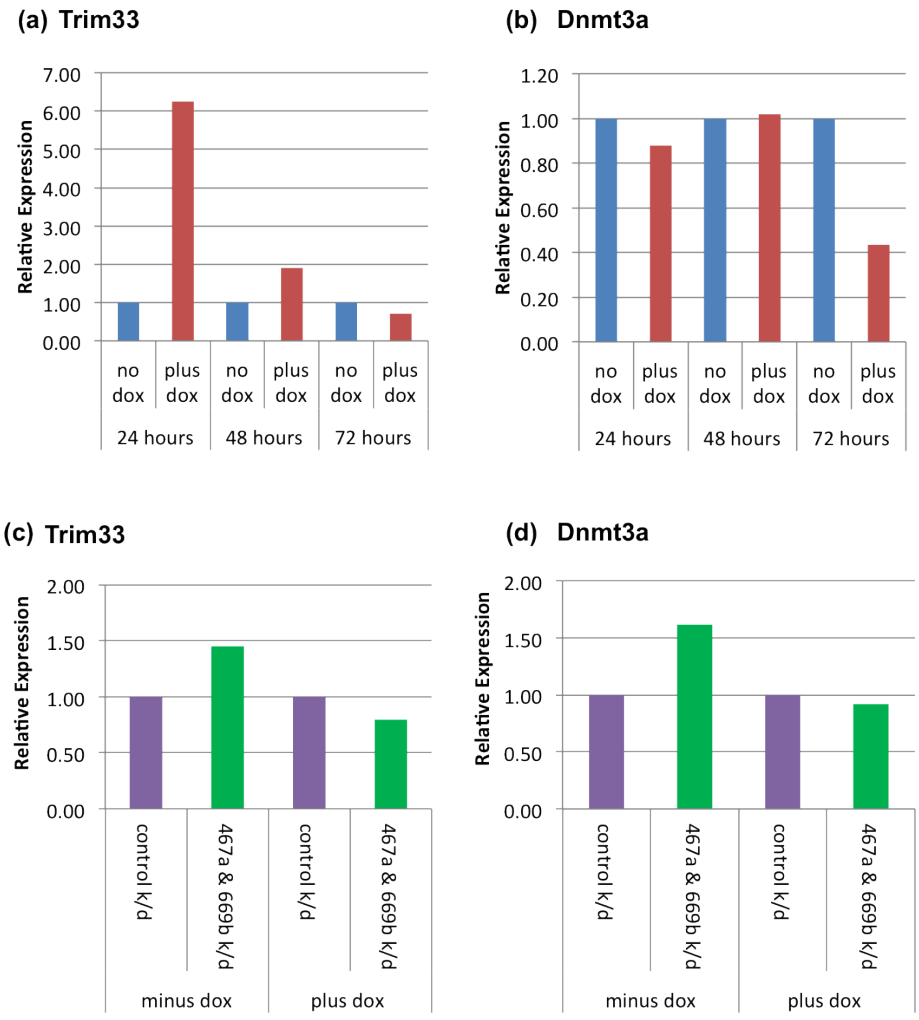


Figure 4.4 – *Expression analysis of Trim33 and Dnmt3a following Oct4 knockdown and miRNA knockdown.* (a & b) ZHBTc4 ES cells were grown without doxycycline or in the presence of 1μg/ml doxycyline for 24, 48 or 72 hours. RNA was extracted from untreated and doxycycline-treated cell cultures, DNAse treated and used to generate cDNA for QPCR analysis of *Trim33* (a) and *Dnmt3a* (b) expression. QPCR Ct values were normalised to *β-Actin* expression. Untreated cultures are represented by a value of 1.0 and results show gene expression in doxycycline treated cultures relative to non-treated cultures at each timepoint. N=1. (c & d) ZHBTc4 ES cells were grown without doxycycline or in the presence of 1μg/ml doxycycline for 24 hours then transfected with a combination of LNAs targeting *miR-467a* and *miR-669b* or an LNA targeting a chicken miRNA (control transfection). RNA was extracted 24 hours after transfection. RNA was DNAse treated, converted to cDNA and used for QPCR analysis of expression of *Trim33* (c) and *Dnmt3a* (b). QPCR Ct values were normalised to *β-Actin* and expression is shown for the *miR-467a*, *miR-669b* knockdown relative to expression in the control transfection. N=1.

to generate cDNA for QPCR analysis of gene expression. After 24 and 48 hours of growth an increase in *Trim33* expression was observed in doxycycline-treated cultures relative to untreated cultures (*Figure 4.4a*). This expression pattern is not consistent with regulation of *Trim33* by *Sfmbt2* miRNAs, which also increase in expression in doxycycline-treated cultures. Expression of *Dnmt3a* was similar in cultures grown without doxycycline and those grown in the presence of doxycycline for 24 and 48 hours (*Figure 4.4b*). However, at 72 hours expression of *Dnmt3a* was reduced in doxycycline-treated cultures relative to untreated cultures. This result is consistent with regulation of *Dnmt3a* by *Sfmbt2* miRNAs, which increase in expression in doxycycline-treated cultures.

In order to assess if *Sfmbt2* miRNAs had any direct effect on *Trim33* or *Dnmt3a* expression in undifferentiated and differentiated ZHBTc4 ES cells ES cells, QPCR analysis was carried out on cDNA from cells after combined inhibition of *miR-467a** and *miR-669b*. Control cultures were transfected with an LNA designed to inhibit a novel chicken miRNA. *Figures 4.4c* & *d* show that there is a slight increase in *Trim33* and *Dnmt3a* expression following *miR-467a** and *miR-669b* inhibition in undifferentiated ZHBTc4 ES cells relative to control knockdown, but no difference in differentiated cultures. The increase in expression of *Trim33* and *Dnmt3a* observed following knockdown of *miR-467a** and *miR-669b* in undifferentiated ES cells is consistent with the prediction that these mRNAs are targets of *Sfmbt2* miRNAs. However, the lack of an increase in expression following *miR-467a** and *miR-669b* knockdown in differentiated cells, where these miRNAs are expressed at higher levels and therefore may be exerting more repression on target transcripts, means that interactions between *Sfmbt2* miRNAs and *Trim33* or *Dnmt3a* cannot be confirmed in this system. It must be noted that this analysis investigates mRNA expression and therefore significant differences in *Trim33* and *Dnmt3a* expression at the protein level as a result of ZHBTc4 differentiation or miRNA knockdown cannot be ruled out. Additionally, these results arise from single experiments and replicate experiments would be required to make conclusions from expression data definitive.

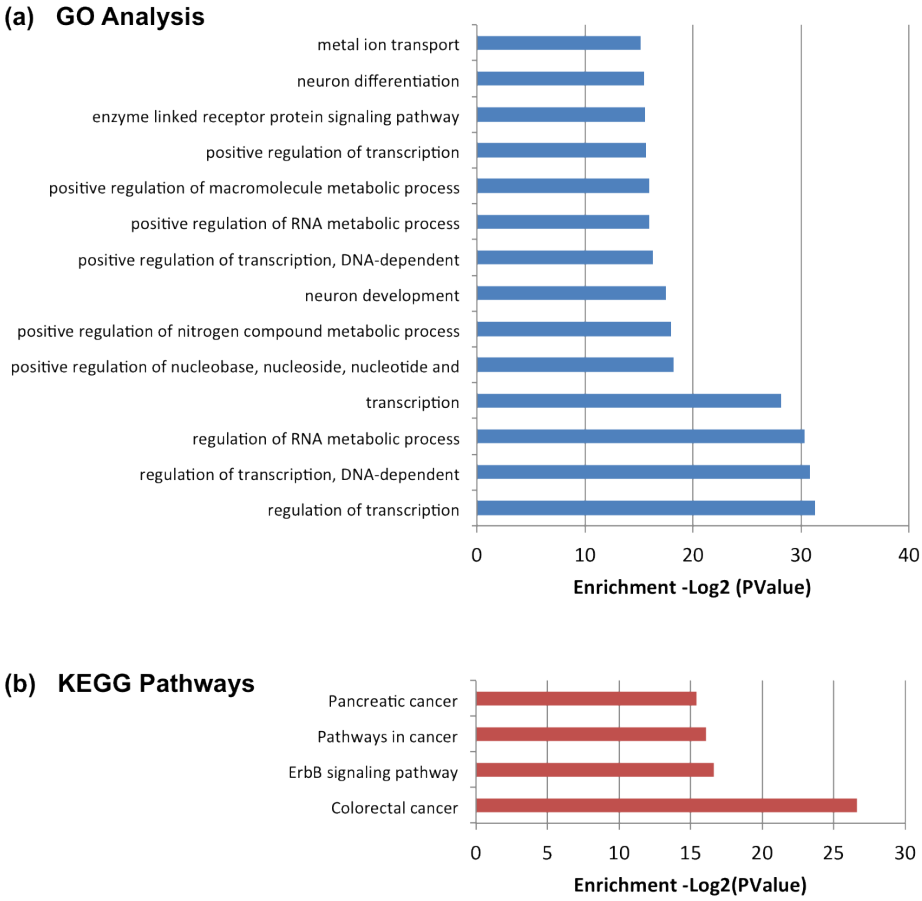


Figure 4.5 – *Analysis of functional associations of predicted Sfmbt2 cluster miRNA targets with GO Terms and KEGG pathways.* Enrichment of GO Terms (a) and KEGG pathways (b) in list of targets predicted for at least 3 different *Sfmbt2* cluster miRNAs by at least 3 target prediction programs (2 for miR-669d). Names of GO Terms and KEGG pathways are shown that had an enrichment score of >15, which corresponds to a P-Value of <0.00003. Enrichment for GO Terms and KEGG pathways was assessed using DAVID Bioinformatics Resources 6.7 (http://david.abcc.ncifcrf.gov/home.jsp).

4.3.1.4 Gene ontology analysis of predicted targets for Sfmbt2 cluster miRNAs

In order to assess whether the predicted targets of the *Sfmbt2* miRNA cluster show enrichment for functional associations, the list of targets predicted for at least 3 of the 7 miRNAs was subjected to analysis for enrichment of GO terms (*Figure 4.5a*) and KEGG pathways (*Figure 4.5b*) using DAVID Bioinformatics Resources 6.7 (Huang *et al.*, 2009, http://david.abcc.ncifcrf.gov/home.jsp). Enrichment values were calculated as –log2(P-Value) and cut-off P-Values were assigned as <0.00003 for GO term analysis and KEGG pathway analysis (corresponding to enrichment scores of >15, respectively).

Similar to results for *miR-92a* targets, GO terms showed enrichment for terms associated with positive regulation of metabolic processes and gene expression (*Figure 4.5a*). Similarly, the enriched KEGG pathways contained a large proportion of cancer-associated pathways (*Figure 4.5b*). *miR-92a* is a part of a known onco-miR cluster that has been associated with numerous cancers (Olive *et al.*, 2010), but no such association has yet been made for the *Sfmbt2* cluster. However, part of the function of this cluster may be to promote survival and proliferation of trophoblast cells in the placenta.

4.3.1.5 Investigating the limitations of sequence-based target prediction

As briefly discussed above, computer-based target prediction programs have multiple limitations. One major limitation is that each program will predict a largely different list of potential targets from the other programs. This is exemplified in *Figure 4.6*, which shows the number of targets predicted by miRecords by 1 or more target prediction programs. There are 6331 predictions made by 1 prediction program, but only 2049 potential targets shared by at least 2 target prediction programs. This means that a large proportion of targets within miRecords are predicted by a single algorithm.

Although the algorithms differ in the criteria they use for target analysis, the basic rules they employ are similar. Therefore, it is surprising that there is so little overlap between predicted lists of targets. Additionally, if we assume that

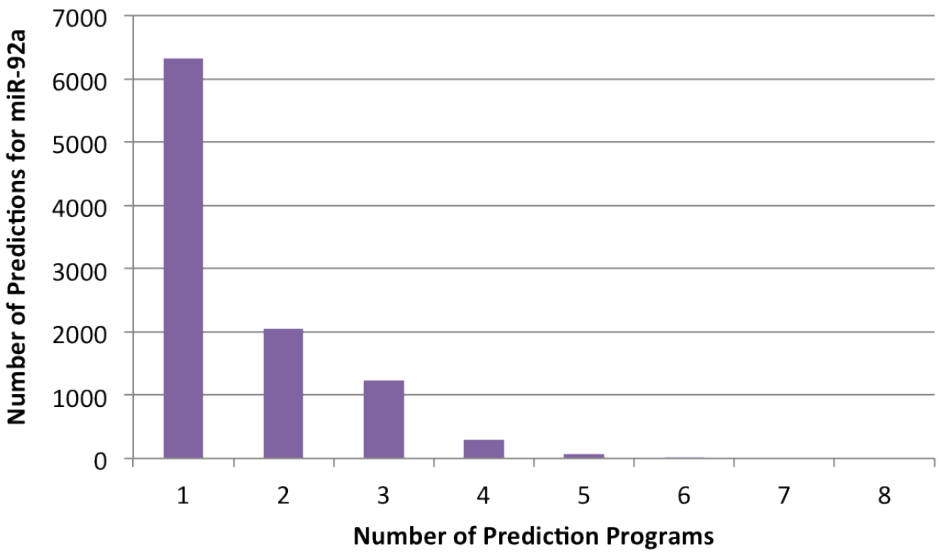


Figure 4.6 – *Comparison of the results from combinations of different target prediction programs.* Graph showing the number of predicted targets for *miR-92a* against the number of programs that predict them.

prediction of a miRNA target is more reliable if it occurs from more than one prediction program, then it follows that the majority of miRNA targets yielded by a single program are unreliable. The disagreement between target lists from different prediction programs makes it difficult for a user to determine which target prediction program to use or what proportion of the predictions are real. Because of the limitations of computational methods of target predictions, experimental methods were tested.

4.3.2 Investigation of PAR-CLIP as a method for identifying miRNA target transcripts

The PAR-CLIP (Photoactivatable-Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation) protocol (Hafner *et al.*, 2010) was used as part of a novel protocol to attempt to identify direct targets of a single miRNA in ES cells. The idea was to express a single miRNA in ES cells that were devoid of all other miRNAs. The PAR-CLIP protocol would then be used to isolate the miRNA targets from the RISC complex. In theory, any mRNAs identified should be targets of the

miRNA that was introduced to the cells. The full experimental protocol can be found in materials and methods (Section 1.3.2). Following optimisation of the PAR-CLIP method, PAR-CLIP analysis was performed on DGCR8$^{-/-}$ ES cells overexpressing *miR-467a**, as well as on DGCR8$^{-/-}$ ES cells (negative control) and on wt ES cells (positive control).

The resulting sequencing data was analysed for the presence of sequences originating from the PAR-CLIP method. 5 sequences >8nt were obtained from the DGCR8$^{-/-}$ untransfected library, 1 sequence from the DGCR8$^{-/-}$ miR-467a transfected library and 2 sequences from the E14 library. In contrast, the original PAR-CLIP paper showed thousands of different sequences from a single library, although this was as a result of deep sequencing of libraries rather than cloning and sequencing. Most sequences were from the untransfected DGCR8$^{-/-}$ library, which was the negative control in the experiment. This suggests the possibility of background noise in this system. However, a greater problem with this technique was the fact that only 2 sequences were recovered from the positive control libraries suggesting that aspects of the procedure were not successful. Further, only one sequence resulted in a 100% match to a mouse mRNA. PAR-CLIP is a complex, laborious, time-consuming set of procedures, mastery of which would obviously require a number of repetitions. While this analysis suggested that the majority of technical difficulties had been overcome, time limitation meant that final optimisation was not possible.

4.3.3 Investigating SILAC analysis as a method of identifying miRNA target transcripts

4.3.3.1 SILAC analysis of undifferentiated and differentiated ZHBTc4 cells

An indirect method for identifying potential miRNA targets was also investigated. The 'Stable isotope labelling of amino acids in cell culture' (SILAC) method (Ong *et al.*, 2002) was used to attempt to identify potential miRNA targets in ES cells. This method depends on metabolic incorporation of 'heavy' or 'light' isotopes of amino acids into newly synthesised proteins in cell culture. Two cell populations are grown in culture media that is identical except that one of them contains

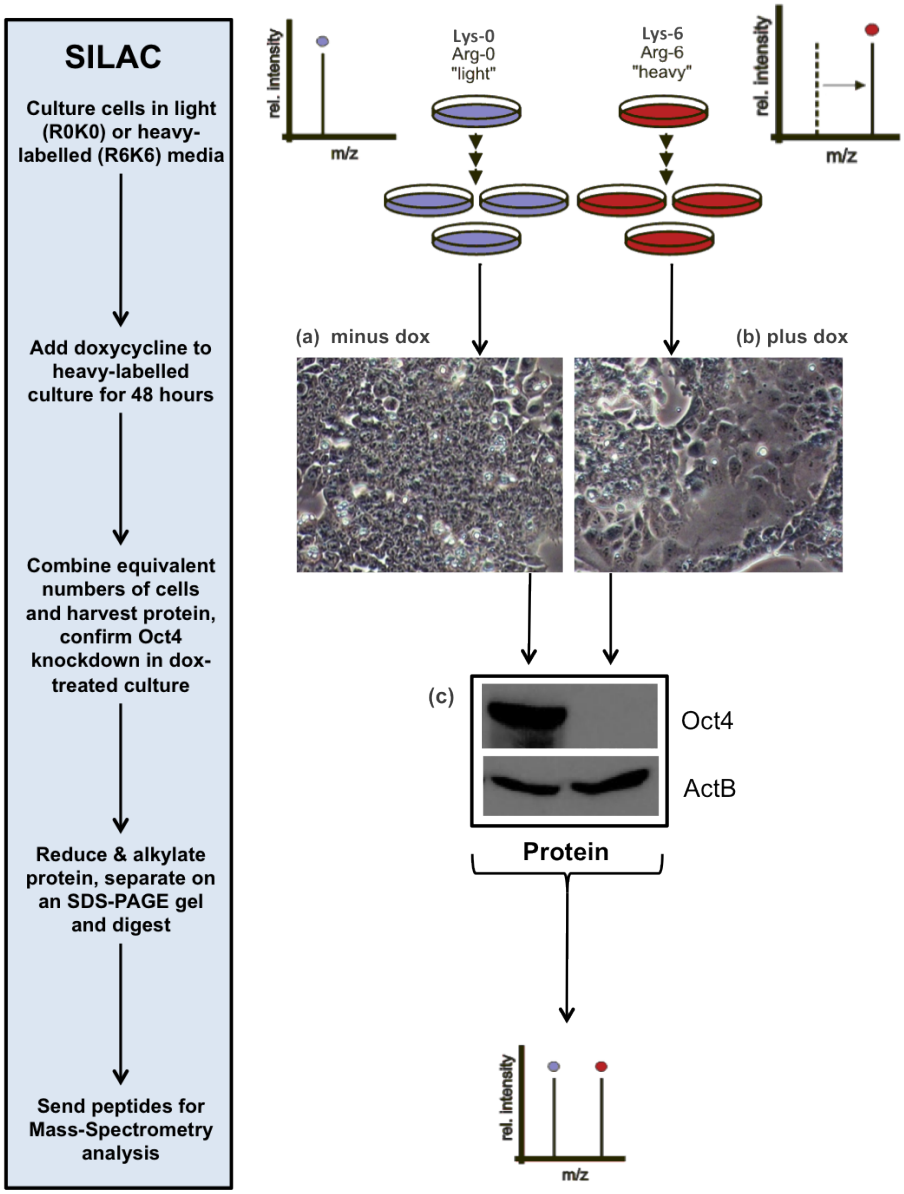


Figure 4.7 – *Overview of SILAC experiment.* Schematic diagram of the SILAC experimental procedure. ZHBTc4 ES cells were grown in the presence of 'light' media (R0K0, Lys-0, Arg-0) or 'heavy' media (R6K6, Lys-6, Arg-6) for 96 hours. The 'light' culture was subsequently grown without doxycycline and the 'heavy' culture was grown in the presence of 1μg/ml doxycycline for 48 hours. Bright field photographs were taken of undifferentiated (a) and differentiated (b) cell populations prior to SILAC analysis (x100 magnification). Protein was harvested from ZHBTc4 ES cells grown without doxycycline or in the presence of 1μg/ml doxycycline and used for western analysis of Oct4 expression (c). Western membrane was re-probed for β-actin expression to estimate protein-loading levels in individual lanes (loading control). Equal numbers of ZHBTc4 ES cells grown without doxycycline or in the presence of 1μg/ml doxycycline were combined and used for downstream SILAC analysis.

light forms of a particular amino acid and the other contains heavy forms. A key feature of the protocol is that the two cell populations are then mixed and used to prepare a single sample for mass spectrometry analysis. Because light and heavy labelled peptides have different m/z ratios, quantitative differences in protein expression between the two original cell populations can be determined. A visual overview of the SILAC experiment is shown in *Figure 4.13* and the full SILAC protocol can be found in Materials and Methods (Chapter 2).

Cells from a single ZHBTc4 ES cell culture were harvested and then cultured in either 'light' or 'heavy' medium for 96 hours in order to uniformly label all expressed proteins. The 'light' cell population was grown for a further 48 hours in 'light' media and remained morphologically undifferentiated (*Figure 4.13a*), whereas the 'heavy' cell population was induced to differentiate by the addition of 1μg/ml doxycycline and cultured for a further 48 hours in 'heavy' media. Doxycycline induces ZHBTc4 ES cells to switch off *Oct4* expression, which results in their rapid differentiation down the trophoblast lineage. Visual inspection of the cells showed that trophoblast differentiation was occurring in the doxycycline-treated population (*Figure 4.13b*). Western analysis of OCT4 expression confirmed that the 'heavy' cells had down-regulated the pluripotency marker OCT4 following doxycycline treatment (*Figure 4.13c*).

The cells from the undifferentiated and differentiated populations were counted and $5x10^6$ cells from each population were mixed and lysed by addition of T-Per buffer to the cell pellet. In addition, protein was extracted and quantitated from $1x10^6$ cells from the two uncombined populations to confirm that protein levels were approximately similar between the two populations (data not shown). Protein samples were taken from equal numbers of undifferentiated and differentiated ES cells grown in normal media and in 'light' or 'heavy' labelled media. These samples were separated on an SDS-PAGE gel, which was then silver-stained. Protein quantity appeared approximately equal between the 'light' and 'heavy' populations, confirming the protein quantitation data (data not shown).

4.3.3.2 Analysis of protein expression during the ES cell to trophoblast transition

Two of the ten gel pieces were selected for processing and analysis. The gel slices chosen were those that contained proteins of between 30kDa and 40kDa, and

between 40 kDa and 50kDa in size (estimated by position of gel slices relative to protein weight markers). 75% of peptides identified by mass spectrometry represented proteins of approximately the expected size (between 30kDa and 50kDa). Sequencing analysis was carried out by Dundee Cell Products and yielded almost 2000 unique peptide sequences corresponding to 444 different proteins. Raw data from SILAC analysis can be found in Appendix D. Proteins were classed as being differentially expressed if they had a heavy (R6K6)/light (R0K0) or a light (R0K0)/ heavy (R6K6) ratio of greater than or equal to 2. By this criterion, 12 proteins were predicted to be expressed at higher levels in the differentiated sample and 16 proteins were predicted to have higher expression in the undifferentiated sample (*Table 4.3*).

In order to validate the SILAC analysis, a protein identified as being up-regulated on differentiation, and a protein identified as being down-regulated on differentiation were chosen for further analysis. Rather than simply relying on predicted 'fold change', the differentially expressed proteins with the most significant PEP scores were selected. Accordingly, using protein samples previously prepared by T. Burdon, western analysis was used to establish the expression of Creatine Kinase B-type (CKB) and C-terminal binding protein 2 (CTBP2) in ZHBTc4 ES cells that had been cultured in the presence of doxycycline for various lengths of time compared to protein levels in ZHBTc4 ES cells that had been grown without doxycycline for the corresponding lengths of time (*Figure 4.14*). It is important to note that the SILAC analysis was carried out 48 hours after addition of doxycycline. At this timepoint the mass spectrometry analysis found that CKB was 3.46-fold higher in differentiated cells and CTBP2 was 2.97-fold higher in undifferentiated cells. Western blot analysis confirmed these results with both CKB and CTBP2 showing differential expression between undifferentiated and differentiated samples. Quantitation of western images showed expression differences of 3.5-fold for CKB and 13.3-fold for CTBP2. These values are preliminary however because data is from a single experimental replicate. CKB (Creatine kinase, brain) reversibly catalyses the transfer of phosphate between ATP and cellular phosphogens such as creatine phosphate (Morrison & O'Sullivan, 1965). The *Ctbp2* (C-terminal binding protein 2) gene produces two alternative transcripts, which produce distinct proteins. One protein is a transcriptional repressor (CTBP2) (Turner & Crossley, 1998), while the other is a component of specialised synapses called synaptic ribbons (RIBEYE) (Schmitz *et al.*, 2000). Interestingly,

(a)

Protein Name	Ratio R6K6/ R0K0	Peptides (seq)	Sequence Coverage [%]	Mol. Weight [kDa]	Sequence Length	PEP
Higher expression in differentiated cultures						
Creatine kinase B-type	3.46	13	55.1	42.713	381	4.90E-106
L-lactate dehydrogenase B chain	3.58	10	33.8	36.572	334	3.49E-83
Prostaglandin reductase 1	5.31	9	38.3	35.56	329	3.51E-56
Branched chain aminotransferase 1, cytosolic	2.38	9	22.4	57.306	519	6.93E-49
Serpin H1	4.91	5	18.9	46.589	417	1.50E-22
Quinone oxidoreductase	1.99	4	18.1	35.268	331	2.42E-20
Mitochondrial carnitine/acylcarnitine carrier protein	5.58	5	16.3	33.026	301	6.93E-15
Solute carrier family 2, facilitated glucose transporter member 3	4.87	2	6.1	53.478	493	1.22E-13
Heme oxygenase 1	2.21	2	10.7	32.928	289	3.20E-13
Four and a half LIM domains 1	2.41	2	7.7	36.261	323	3.69E-09
Annexin A7	2	6	12.8	52.251	485	8.88E-09
Glutathione synthetase	2.51	1	3.6	52.246	474	1.19E-01

(b)

Protein Name	Ratio R0K0/ R6K6	Peptides (seq)	Sequence Coverage [%]	Mol. Weight [kDa]	Sequence Length	PEP
Higher expression in undifferentiated cultures						
C-terminal-binding protein 2	2.97	10	33.3	48.956	445	3.07E-45
Serine hydroxymethyltransferase	2.33	8	18.7	55.76	504	4.43E-33
Purine nucleoside phosphorylase	2.23	8	34.9	32.277	289	2.28E-27
Aldose reductase	2.91	13	44	35.732	316	6.63E-24
Uridine phosphorylase 1	4.11	9	42.8	34.086	311	1.10E-21
L-threonine 3-dehydrogenase, mitochondrial	3.99	6	26.3	41.461	373	3.11E-20
Septin-1	3.28	2	11.5	42.019	366	1.93E-16
Pyrroline-5-carboxylate reductase 2	2.25	2	7.2	33.659	320	1.47E-11
Cysteine and histidine-rich domain-containing protein 1	1.98	3	12.4	37.35	331	2.64E-11
Lysophosphatidic acid phosphatase type 6	2.77	3	8.6	47.624	418	1.27E-09
Aldehyde dehydrogenase, mitochondrial	1.97	3	9.4	56.537	519	6.35E-09
Adult male testis cDNA, RIKEN full-length enriched library, clone:1700012H05	2.07	2	5.7	41.912	385	1.39E-08
Galectin-3	2.74	2	8.3	27.515	264	9.42E-08
Annexin A2	1.99	3	8.3	38.676	339	1.08E-04
Putative uncharacterized protein	3.10	2	5.9	42.866	371	1.24E-02
Uncharacterized protein FLJ11171 homolog	22.25	1	1.2	87.142	767	1.22E-01

Table 4.3 – *Results from SILAC analysis.* Table showing proteins with an R6K6/R0K0 ratio of >2.0 or <0.5. The R6K6/R0K0 ratios are the average ratio for all peptides sequenced for each protein. Also shown is the number of peptides sequenced, the percentage sequence coverage that these peptides represent, the molecular weight of the protein (kDa), the sequence length in amino acids and the PEP (posterior error probability) score, which is the probability of a false hit in protein identification.

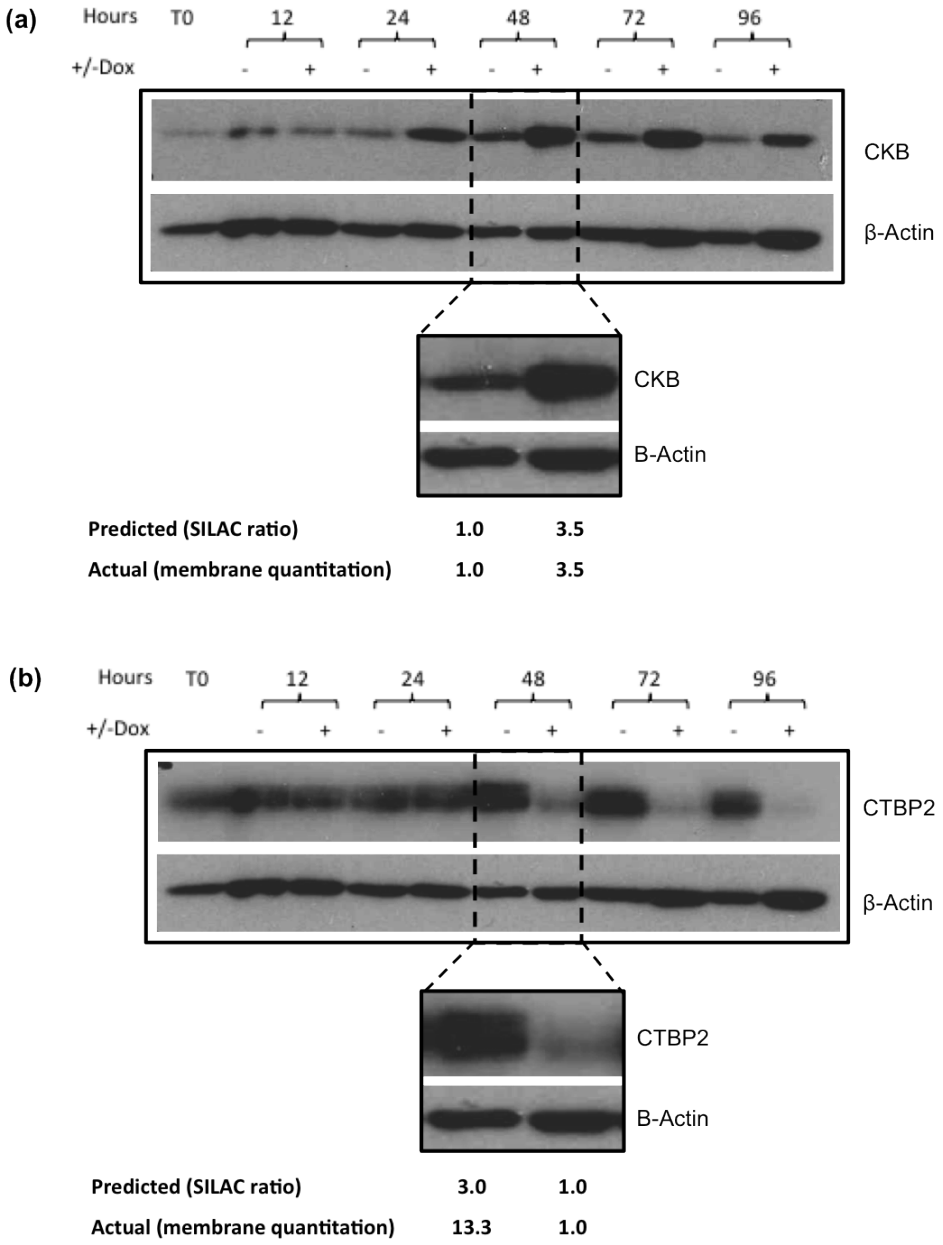


Figure 4.8 – *Western analysis of CKB and CTBP2 expression in undifferentiated and differentiated ZHBTc4 ES cells.* Western analysis of CKB (a) and CTBP2 (b) expression in ZHBTc4 ES cells at time of plating (Time 0) and ZHBTc4 ES cells grown without doxycycline (- dox) or in the presence of 1μg/ml doxycycline (+dox) for 12, 24, 48, 72 or 96 hours. Western membrane was re-probed for β-actin expression to estimate protein-loading levels in individual lanes (loading control). Lanes showing protein expression at 48 hours are expanded and the ratios of protein expression from SILAC analysis and quantitation of the autorad image are shown. N=1.

(a)

PROTEINS HIGHER IN DIFFERENTIATED SAMPLE		Ratio R6K6/R0K0	miR-672	miR-708	miR-30C	miR-124	miR-92a	miR-106a	miR-19a
CKB	Creatine kinase B-type	3.46	2	2	2	1	1	1	1
LDHB	L-lactate dehydrogenase B chain	3.58	2	3	3	1	2	1	4
PTGR1	Prostaglandin reductase 1	5.31	4	3	2	2	2	2	-
BCAT1	Branched chain aminotransferase 1, cytosolic	2.38	4	2	2	2	3	2	2
SERPH	Serpin H1	4.91	3	2	2	1	2	3	2
CRYZ	Quinone oxidoreductase	1.99	4	4	2	1	2	2	-
SLC25A20	Mitochondrial carnitine/acylcarnitine carrier protein	5.58	3	3	2	4	1	2	4
SLC2A3	Solute carrier family 2, facilitated glucose transporter member 3	4.87	2	3	3	2	2	4	2
HMOX1	Heme oxygenase 1	2.21	1	1	1	1	2	4	1
FHL1	Four and a half LIM domains 1	2.41	3	3	2	3	2	3	2
ANXA7	Annexin A7	2.00	3	2	3	3	3	3	5
GSS	Glutathione synthetase	2.51	2	2	1	3	1	1	1

(b)

PROTEINS HIGHER IN UNDIFFERENTIATED SAMPLE		Ratio R0K0/R6K6	miR-467a*	miR-23a	miR-467e	miR-669d	miR-191	miR-210	miR-99b	miR-297a	miR-669b	miR-669a
CTBP2	C-terminal-binding protein 2	2.97	-	2	1	3	2	4	1	2	3	2
SHMT1	Serine hydroxymethyltransferase	2.33	-	2	1	1	2	1	1	3	2	1
NP	Purine nucleoside phosphorylase	2.23	-	3	2	3	2	2	2	3	4	4
no data	Aldose reductase	2.91	no data									
UPP	Uridine phosphorylase 1	4.11	-	1	-	1	1	2	1	1	2	2
TDH	L-threonine 3-dehydrogenase, mitochondrial	3.99	-	1	-	2	2	3	1	1	3	3
*SEPT1	Septin-1	3.28	-	1	2	1	1	2	1	1	2	2
PYCR2	Pyrroline-5-carboxylate reductase 2	2.25	-	1	-	2	1	3	3	1	1	2
CHORDc1	Cysteine and histidine-rich domain-containing protein 1	1.98	-	3	2	1	3	1	2	1	2	2
ACP6	Lysophosphatidic acid phosphatase type 6	2.77	-	1	-	1	1	1	1	1	1	1
ALDH2	Aldehyde dehydrogenase, mitochondrial	1.97	-	1	1	1	1	3	1	2	1	3
no data	Adult male testis cDNA, RIKEN full-length enriched library, clone:1700012H05	2.07	no data									
LGALS3	Galectin-3	2.74	1	2	2	1	1	1	1	1	1	1
ANXA2	Annexin A2	1.99	-	3	2	4	2	2	2	3	4	4
no data	Putative uncharacterized protein	3.10	no data									
no data	Uncharacterized protein FLJ11171 homolog	22.25	no data									

Table 4.4 – *miRNA target predictions for differentially expressed proteins from SILAC analysis.* Table showing if proteins expressed at higher levels in differentiated cultures (a) or undifferentiated cultures (b) are predicted targets of miRNAs represented at higher levels in undifferentiated cultures or differentiated cultures, respectively. Values correspond to the number of prediction programs that predict an interaction between miRNAs and potential targets.

CTBP2 is mainly expressed during embryogenesis (Furusawa *et al.*, 1999) and CTBP2-null mice die during embryogenesis due to aberrant extraembryonic development (Hildebrand and Soriana, 2002). From these data the SILAC profiling appears reliable and has correctly identified proteins that are differentially expressed between ES cells and trophoblast-like cells.

4.3.3.3 Investigating interactions between proteins and miRNAs identified as being differentially expressed during ZHBTc4 ES cell differentiation

The purpose of the SILAC analysis was to identify potential miRNA targets during ES cell differentiation. As such the list of differentially expressed proteins was tested for sequence-based predictions of miRNA binding sites for the differentially represented miRNAs shown in *Table 3.3*. The online utility, miRecords (http://mirecords.biolead.org/), was used to determine the number of programs that predicted interactions between differentially expressed proteins and differentially represented miRNAs (*Table 4.4*). 19 out of the 28 differentially expressed proteins are predicted to be targets of at least one differentially represented miRNA by at least three different target prediction programs and 11 interactions are predicted by at least four target prediction programs.

4.4 Discussion

The primary aim of this chapter was to compare methods for the identification of miRNA targets. A sequence-based method was compared with two experimental methods: the direct experimental method, PAR-CLIP, and the indirect experimental method, SILAC. Lists of targets for *miR-92a* and *Sfmbt2* miRNAs were generated by sequence-based prediction, but the reliability of these target lists was questioned because none of the predicted targets could be experimentally validated. Direct miRNA targets were not successfully identified as a result of the PAR-CLIP method, probably due to lack of experience with the methodology. Use of the SILAC method identified proteins that were differentially expressed during the ES cell to trophoblast transition, some of which were predicted targets of miRNAs that were differentially represented during this transition.

4.4.1 Sequence-based methods predict targets of *miR-92a* and the Sfmbt2 miRNA cluster

The miRNA target database miRecords was used to predict targets of the *Sfmbt2* miRNA cluster and *miR-92a*. miRecords amalgamates target predictions from multiple prediction programs and results can be filtered to include only those targets predicted by a minimum number of programs (Xiao *et al.*, 2009). For example, there were 69 *miR-92a* targets predicted by five or more different target prediction programs. This list of predicted targets contained only one target that had been previously experimentally verified according to the miRecords validated targets database – *Mylip*. *Mylip* interacts with myosin regulatory light chain B, which in turn regulates the activity of the actomyosin complex important for many motile processes in eukaryotes (Olsson *et al.*, 1999). *Mylip* was identified as a direct target of *miR-92a* in 3T3 cells (Landais *et al.*, 2007). However, expression analysis presented here suggests that *Mylip* is not a target of *miR-92a* in ES cells. This disagreement could result from two possibilities: *Mylip* may be a target of *miR-92a* in 3T3 cells, but not in ES cells; or regulation by *miR-92a* may involve translational repression, rather than mRNA degradation. Indeed, Landais *et al.* reported that *miR-92a* affected *Mylip* expression at the protein level, but did not investigate *Mylip* mRNA expression.

The miRecords database was also used to predict targets of the seven differentially represented *Sfmbt2* cluster miRNAs, and results for all seven miRNAs were amalgamated using BioLayout Express. 126 targets were predicted for at least five of the *Sfmbt2* cluster miRNAs by at least three different target prediction programs. According to the database of 'Validated Targets' available through miRecords, none of these predicted targets have yet been validated as real targets of any of the miRNAs in this analysis. Significantly, six potential targets were predicted for all seven of the *Sfmbt2* cluster miRNAs: *Rabgap1l*, *Ogfod1*, *Trim33*, *Btbd14a*, *Bcl11b* and *Dnmt3a*. The fact that some of the seven miRNAs have different sequences, including different seed sequences, means that they must be predicted to target different sites within these genes. The presence of more than one predicted target site for *Sfmbt2* miRNAs increases the likelihood that these genes are targets of *Sfmbt2* miRNAs. Two of the potential targets, *Trim33* and *Dnmt3a*, were investigated further. Both *Trim33* and *Dnmt3a* increased in expression upon knockdown of *miR-467a** and *miR-669b* in undifferentiated ES

cells, but no increase in expression was observed upon miRNA knockdown in differentiated cells. *miR-467a** and *miR-669b* expression is higher in differentiated cells relative to undifferentiated cells, so it is surprising that an effect on gene expression was observed in the undifferentiated cultures, but not the differentiated cultures. While it is possible that *Trim33* and *Dnmt3a* are direct targets of *miR-467a** and/or *miR-669b* in ES cells but not in differentiated cells, this disagreement in results makes it difficult to infer a direct association between *Sfmbt2* miRNAs and *Trim33* or *Dnmt3a*. Analysis of expression of *Trim33* and *Dnmt3a* protein following miRNA suppression would confirm whether regulation by *Sfmbt2* miRNAs is evident at the protein level.

In order to assess if the predicted targets of the *Sfmbt2* cluster and *miR-92a* were associated with any particular biological functions, the enrichment of these targets in specific GO terms and KEGG pathways was investigated. Results showed an enrichment of *miR-92a* targets in GO terms associated with positive regulation of metabolic processes, biosynthetic processes and gene expression, and KEGG pathways associated with cancer. The targets of the *Sfmbt2* cluster miRNAs were also enriched in similar pathways. *miR-92a* and the *Sfmbt2* cluster miRNAs have different sequences including completely different seed sequences, and are shown to generate different target lists from the data presented here. The fact that *miR-92a* targets and targets of *Sfmbt2* miRNAs are associated with similar biological processes and pathways may be explained if these miRNAs have similar functions, perhaps in different systems. However, targets of the unrelated miRNA, *miR-155*, also show enrichment in GO terms and KEGG pathways that overlap with those of *miR-92a* targets and targets of the *Sfmbt2* cluster. While there is a possibility that *miR-155* has similar functional roles to *miR-92a* and *Sfmbt2* miRNAs, there is little experimental evidence to support this. The function of *miR-155* is normally associated with the immune system: loss of *miR-155* in mice results in immunodeficiency and *miR-155* is important for the function of B-cells, T-lymphocytes and dendritic cells (Rodriguez *et al.*, 2007). Therefore, it is believed more likely that these data reflect a limitation of the use of GO term and KEGG pathway enrichment for this type of analysis.

4.4.2 Estimating the reliability of target prediction programs

Investigation of a sequence-based method of target identification for *miR-92a* and *Sfmbt2* cluster miRNAs resulted in lists of predicted target genes. One out of the 69 predicted targets of *miR-92a* have been experimentally validated but none of the *Sfmbt2* cluster targets have any experimental verification. Because experimental verification of miRNA targets is a labour intensive process, very few predicted miRNA-mRNA interactions have been confirmed. Verification of targets should improve our understanding of the factors important for miRNA targeting and lead to improved sequence-based methods.

The lists of predicted targets of *miR-92a* and *Sfmbt2* miRNAs are likely to contain a significant proportion of false-positive predictions, which is a consequence of the challenges presented by predicting miRNA targets that show only partial complementarity with their regulatory miRNAs. The reason for a mechanism of regulation requiring partial, rather than full, complementarity between miRNA and mRNA target is not fully understood, but a number of possibilities have been proposed: that a central loop region improves miRNA-mediated regulation (Ye et al., 2008); or that extensive complementarity between miRNA and mRNA target site triggers miRNA destabilisation (Ameres et al., 2010). Large miRNA target lists are not unexpected however, because rather than existing as switches for a few targets, each miRNA is thought to regulate in the region of 200 target transcripts (Baek et al., 2008, Farh et al., 2005), and multiple miRNAs may act combinatorially to regulate expression of a single target (Stark *et al.*, 2005). Another limitation of several target prediction programs is that they are based on matching short complementary sequences without taking into consideration whether the target sequence is available in the cell. Many target prediction programs do not take into account the fact that a target site may be inaccessible due to RNA-folding or the presence of RNA-binding proteins.

Two proteome-based approaches of measuring the impact of miRNAs on protein levels examined the reliability of different target prediction programs based on their findings. Selbach *et al.* (2008) used a pulse-labelling SILAC approach to compare levels of newly synthesised proteins in cells with normal miRNA levels, and protein from cells where miRNA levels were perturbed, leading to a list of proteins that may be regulated by specific miRNAs. Comparison of this list of predicted miRNAs targets with targets predicted by sequence-based prediction

programs showed that all sequence-based prediction programs tested, apart from RNA22, achieved an accuracy that was greater than random chance. Prediction programs that took evolutionary conservation of seed sites into account achieved a further step up in accuracy (TargetScan, PicTar and Diana microT 3.0). Even the best prediction program however, only achieved an accuracy of approximately 66%. This means that even with the most reliable target prediction program, one third of all predicted targets are false positives. A similar analysis (Baek *et al.*, 2008) came to the same conclusion. They found that only taking into account the top 29 predictions improved the reliability of most of the programs by a small amount. However, while taking into account evolutionary conservation of seed sites clearly improves overall target prediction, it will bias against those miRNA targets that have arisen as a result of recent evolution.

Studies into the reliability of target prediction programs suggest that there is not one program that stands out as being the most reliable. Therefore it has been commonplace for researchers to consult several target prediction databases when searching for miRNA targets (Xiao *et al.*, 2009). It is presumed that by amalgamating results from multiple target prediction programs, the reliability of the predicted targets will be increased. This led to the development of a web-based utility called miRecords that combines the results from multiple target prediction programs (Xiao *et al.*, 2009). However, while it assumed, there is yet no experimental evidence that amalgamating target lists from different programs actually improves target prediction. An analysis of *miR-92a* targets, presented here, that were predicted by multiple target programs generated 6331 targets that were predicted by at least one target prediction program. This figure corresponds to almost one third of the entire transcriptome. Each miRNA is thought to regulate in the region of a few hundred target transcripts (Farh *et al.*, 2005), so it is highly unlikely that expression of the 6331 predicted targets is actually regulated by *miR-92a*. This means that the vast majority of predictions made by target prediction programs are not real targets, and it could therefore be argued that combining prediction from multiple packages may not actually improve prediction accuracy.

4.4.3 SILAC – a method to identify direct and indirect miRNA targets

The proteome-based method SILAC was investigated as a method of identifying miRNA targets. While the proteins identified in SILAC cannot be classed as direct miRNA targets without further experimental analysis, this method can identify proteins directly and indirectly regulated by miRNAs. A SILAC analysis, therefore, has greater power in determining the global proteome effects of perturbing miRNA expression than a direct method such as PAR-CLIP. The original purpose of the SILAC analysis presented in this thesis was to evaluate the method on the basic trophectoderm differentiation system used in Chapter 3 of this thesis. Following this, a further goal was to use the SILAC method to evaluate proteomic changes following missexpression of specific miRNAs. Due to time constraints the miRNA missexpression analysis was not carried out. Instead, proteins identified as being differentially expressed in the SILAC analysis of undifferentiated and differentiated cells were analysed as potential targets of miRNAs also identified as being differentially represented.

SILAC analysis was carried out on undifferentiated and differentiated populations of ZHBTc4 ES cells. From the SILAC analysis, mass spectrometry identified peptides corresponding to 444 different proteins. Considering that only one fifth of the total gel was processed for mass spectrometry analysis, it is possible that a full analysis could identify in the region of 1500 proteins. It is difficult to estimate the coverage of this analysis because the size of the mouse ES cell proteome is not known. For comparison, previous SILAC analyses in HeLa or mouse ES cells have resulted in quantitation of between 504 and 5111 individual proteins (Vinther *et al.*, 2006, Graumann *et al.*, 2008).

One potential limitation of the SILAC analysis is a lack of sensitivity, which would result in failure to identify proteins present in low cellular quantities. Generally, it is thought that proteins involved in cell signalling, including transcription factors, are present at lower quantities than housekeeping and structural proteins. The size range of proteins identified by this SILAC analysis should include the transcription factors OCT4 and CDX2, which are key for trophectoderm differentiation, and would be expected to show strong differential expression in the ZHBTc4 differentiation system. According to data presented in Chapter 3 of this thesis, OCT4 proteins levels should be drastically reduced upon trophectoderm

differentiation and CDX2 expression should increase, but these proteins were not identified in the SILAC analysis. This may suggest that SILAC analysis is limited because it is biased towards detecting changes in abundant cell proteins. However, the reason that differential expression of OCT4 and CDX2 may not have been reported may be if peptides from these proteins were not identified in one of the samples. Because automated SILAC analysis will only report the quantitative difference in protein expression, if a protein is expressed at low levels in one sample, and is therefore undetected by SILAC, then a ratio cannot be calculated and that protein will not be reported in the results. A previous SILAC analysis of mouse ES cells successfully identified the key ES cell transcription factors OCT4, SOX2 and NANOG suggesting that these proteins exist at a level detectable by SILAC analysis in ES cells (Graumann *et al.*, 2008). Further, Selbach *et al.* reported that, although mass spectrometry is biased to detect more highly expressed genes, this bias was mild in their SILAC analysis and did not affect the detection range (Selbach *et al.*, 2008).

4.4.4 SILAC analysis identified differentially expressed proteins

Proteins with differential expression levels between undifferentiated and differentiated ES cells were identified by SILAC analysis and the expression profiles of two of these proteins, CTBP2 and CKB, were confirmed by western analysis. This proved the reliability of the data and identified two proteins that are differentially expressed during trophectoderm differentiation. The list of differentially expressed proteins was then analysed for potential regulation by differentially represented miRNAs. The miRecords database was used to identify if differentially expressed proteins were predicted targets of differentially represented miRNAs, and how many target prediction programs predicted this interaction. For interactions predicted by at least 4 different target prediction programs: results showed that 8 out of 12 proteins that were expressed at higher levels in differentiated cells were predicted to be targets of miRNAs represented at higher levels in undifferentiated cells, and 3 out of 16 proteins that were expressed at higher levels in undifferentiated cells were predicted to be targets of miRNAs represented at higher levels in differentiated cells. This suggests the possibility that the expression patterns of these proteins could arise from direct regulation by differentially represented miRNAs. Experimental validation of predicted interactions between

differentially represented miRNAs and proteins would confirm whether the SILAC method can be used in combination with sequence-based prediction methods to identify direct miRNA targets.

Previous SILAC analyses have been used to identify potential targets of specific miRNAs in a variety of systems (Vinther *et al.*, 2006, Selbach *et al.*, 2008, Baek *et al.*, 2008, Yang *et al.*, 2010). In each of these analyses, miRNA expression was perturbed by miRNA expression or inhibition, and subsequent SILAC analysis was used to identify differentially expressed proteins. The SILAC analyses identified between 12 and 130 potential miRNA targets. Variation in the number of targets identified appeared to correlate with variation in the number of proteins identified from the analysis. Vinther *et al.* identified 12 proteins that were differentially regulated upon *miR-1* transfection in HeLa cells, and reported that 7 of these proteins were predicted as *miR-1* targets by several target prediction programs (Vinther *et al.*, 2006). This suggests two possibilities: that target prediction programs fail to identify a significant proportion of real miRNA targets, or that SILAC analysis identifies both direct and indirect miRNA targets. Both of these possibilities are likely to contribute to the discrepancy between predicted *miR-1* targets and targets identified by SILAC analysis. Further, sequence-based predictions identified 36 potential targets of *miR-1*, which were not found to be targets by SILAC analysis, suggesting that sequence-based programs predict a large number of false positives (Vinther *et al.*, 2006). Of the 12 proteins identified as potential *miR-1* targets, Vinther *et al.* validated 6 as direct *miR-1* targets by luciferase assay. In a similar analysis, Yang *et al.* identified 93 proteins as being downregulated by more than 2-fold upon expression of *miR-143*. Validation analysis of 34 of the potential targets by luciferase assay showed that 10 of them were likely to be direct targets of *miR-143* (Yang *et al.*, 2010). Together, these data suggest that SILAC analysis can be used to identify proteins that are directly or indirectly regulated by miRNAs.

4.4.5 Summary and Future Work

Three different methods for identifying miRNA targets were investigated: A sequence-based method led to predicted lists of targets for the *Sfmbt2* miRNA cluster and *miR-92a*; conclusions from the PAR-CLIP experimental method were limited by the technically challenging nature of this analysis; and the SILAC

method identified proteins that were differentially expressed during a developmental transition, many of which were predicted targets of miRNAs also differentially represented in the transition. Sequenced-based methods of target prediction are convenient, but normally generate very large lists of predicted targets. It is estimated that each miRNA probably targets several hundred different mRNAs (Farh *et al.*, 2005) so a large number of predicted targets are expected, but it is also known that a significant proportion of sequence-based predictions are false-positives (Selbach *et al.*, 2008, Baek *et al.*, 2008). Researchers that use sequence-based target predictions to select candidates for experimental investigation generally select targets that they believe will be functionally important in the system they are investigating, which introduces a degree of subjectivity into miRNA target analysis. This approach may miss miRNA targets that have not been previously associated with a function in the system, and/or lead to experimental investigation of targets that are not biologically relevant. Further investigation of the effect of missexpressing miRNAs on the protein levels of targets predicted by single or multiple prediction programs would provide insight into whether amalgamating prediction results actually improves the reliability of target prediction. Experimental methods can be used to narrow down the list of targets to those more likely to have biological relevance. The use of the PAR-CLIP method for the identification of direct miRNAs was not successful in this analysis, and the technically challenging nature of this protocol may render it unsuitable for routine use in the laboratory. Proteome-based methods such as SILAC are relatively straightforward, and have been previously used to identify direct and indirect miRNA targets. Further, SILAC analysis can be used to identify differential expression of proteins in cell-based models of developmental transitions. Lists of predicted miRNA targets can then be compared with analysis of differentially expressed proteins and allow focused investigation of potential miRNA targets that are likely to be biologically relevant. SILAC analysis could be further extended in this system by missexpression of differentially regulated miRNAs in undifferentiated or differentiated ZHBTc4 ES cells, and use of the SILAC method to identify biologically relevant targets of these miRNAs.

Chapter 5

The general role of miRNAs in ES cell differentiation

5.1 Introduction

The aim of this chapter was to investigate the overall role of the total miRNA population in developmental transitions. To study the role of miRNAs in general on cell differentiation, ES cells deficient in mature miRNAs were obtained and their differentiation potential was compared with the differentiation of wt ES cells. Differentiation is defined in The Chambers Dictionary as 'a change by which what was generalised or homogeneous became specialised or heterogeneous'. Although the signals and factors involved will differ depending on the differentiation method and lineage outcome, certain key aspects of ES cell differentiation are likely to remain the same, e.g. expression of pluripotency genes, including *Oct4*, *Sox2* and *Nanog*, will be downregulated. Concomitantly, up regulation of early markers of differentiation such as the epiblast marker *Fgf5* occurs. Cells will become increasingly specialised in terms of gene expression, growth requirements and functional capability.

In order to harness the potential of embryonic stem cells, differentiation protocols for the generation of specific cell types need to be developed. Many previous advances have resulted from *in vivo* and *in vitro* observation of the factors involved in differentiation. For this reason it is important to understand the factors that

drive differentiation into particular lineages, including the role of genetic regulators such as miRNAs.

5.1.1 miRNA deficient ES cells

miRNAs are essential for full mammalian development (Bernstein *et al.*, 2003, Wang *et al.*, 2007) and have been shown to be involved in a number of aspects of differentiation (reviewed in Ivey & Srivastava, 2010). In order to study the role of miRNAs in differentiation, a number of groups have generated miRNA-deficient ES cells. This was achieved by ablation of essential components of the miRNA biogenesis pathway such as Dicer (Murchison *et al.*, 2005, Kanellopoulou *et al.*, 2005) and DGCR8 (Wang *et al.*, 2007). Dicer functions during the biogenesis of miRNAs and siRNAs by cleaving short hairpin RNA molecules to double stranded RNA molecules in the cytoplasm. DGCR8 functions in the miRNA pathway with its partner Drosha, ensuring correct processing of pri-miRNAs to pre-miRNAs in the nucleus. Dicer is involved in processing of miRNAs and siRNAs while DGCR8 activity is thought to be restricted to miRNA processing. For this reason DGCR8$^{-/-}$ ES cells were used in this project to investigate the role of miRNAs in differentiation. However, it must be noted that examples of miRNAs that are not processed by the canonical pathway have now been documented (Babiarz *et al.*, 2008). These non-canonical miRNAs are rare but mean that the DGCR8$^{-/-}$ ES cells still contain a small number of miRNAs.

The DGCR8$^{-/-}$ ES cells were previously generated in the laboratory of Robert Blelloch at the University of California (Wang *et al.*, 2007). These cells were created by excision of exon 3 from DGCR8 by sequential targeting of mouse ES cells. Loss of exon 3 results in a frame shift in the open reading frame of the DGCR8 transcript and generates several stop codons in downstream exons. Wang *et al.* showed that the pre-miRNAs and the mature versions of selected miRNAs were not present in RNA extracted from DGCR8$^{-/-}$ ES cells. An accumulation of *pri-miR-293* was observed in the DGCR8$^{-/-}$ ES cells (not seen in wt ES cells or DGCR8+/- ES cells). This is in keeping with the reported function of DGCR8; processing of miRNAs from pri-miRNA to pre-miRNA.

Wang *et al.* reported that DGCR8$^{-/-}$ ES cells showed an extended population doubling time relative to wt ES cells although they reported that the cells appeared

morphologically normal and expressed ES cell marker genes. Additionally, they reported that the cells showed defects in differentiation following two non-directed differentiation protocols – EB differentiation and retinoic acid-induced differentiation. The differentiation defect was associated with a failure to fully downregulate pluripotency markers (Wang *et al.*, 2007). In these experiments, the cells were grown on a MEF feeder layer and removed from feeders one passage prior to experimental analysis. We asked the question of whether miRNA-deficient ES cells could be cultured long-term in the absence of feeder cells and maintain normal morphology and expression of ES cells markers. In addition, further investigation was carried out into the differentiation capabilities of miRNA-deficient ES cells, including their ability to differentiate when subjected to directed differentiation protocols compared to non-directed methods such as EB differentiation.

5.1.2 Systems for embryonic stem cell differentiation

In order to compare the differentiation capacity of DGCR8$^{-/-}$ ES cells and wt ES cells non-directed and directed differentiation methods were used. The differentiation methods mentioned below are some of the best-characterised protocols for ES cell differentiation. For this reason these protocols were used to investigate the differentiation capacity of miRNA-deficient ES cells.

Arguably the differentiation procedure that most closely replicates early embryonic development is 3D suspension culture of ES cells. Embryoid body (EB) differentiation was first documented by Doetschman *et al.* (1985), and involves the culture of ES cells in the absence of both feeders and Lif in suspension where cells preferentially adhere to one another rather than to the substrate. The ES cells aggregate to form spheres of cells, which may mimic aspects of early embryonic development including the formation of a visceral endoderm layer, which precedes formation of cavities called cysts - structures thought to be analogous to the visceral yolk sac (Doetschman *et al.*, 1985). EBs may also mimic elements of gastrulation because they show localised Wnt signalling, reminiscent of formation of a primitive streak-like structure (ten Berge *et al.*, 2008). After a fixed period of culture in suspension, EBs can be transferred to conditions allowing their adherence to a fixed surface. This allows observation of further differentiation events such as the differentiation of beating cardiomyocytes and neurons. During EB differentiation, cell types from all three germ layers are generated. As

a result this method is often used as an initial test to determine the differentiation potential of ES cells and iPS cells. Despite being non-directed, multiple studies of EB differentiation have concluded that it represents a relevant model of early embryonic development (Doetschman *et al.*, 1985, ten Berge *et al.*, 2008). As a result there are now numerous directed differentiation methods that utilise EB differentiation as part of the protocol (reviewed in Desbaillets *et al.*, 2000).

Over recent years protocols have been derived for the directed differentiation of ES cells into a broad range of cell types including cardiomyocytes, hematopoietic cells and dopaminergic neurons (reviewed in Murry and Keller, 2008). One of the best-characterised systems to date is that of neural differentiation. Bmp normally acts through induction of ID proteins to inhibit neural differentiation and neural differentiation can be induced in ES cells simply by removal of BMP signalling (Ying *et al.*, 2003). A standard method for neural differentiation is to transfer ES cells from growth in serum + Lif conditions to low-density growth in the serum free media, N2B27 (Lowell *et al.*, 2006). However, even in this well-characterised system, a significant proportion of ES cells still resist differentiation, which may be due to inherent heterogeneity of embryonic stem cell populations (Canham *et al.*, 2010). The question of why some cells resist commitment in stringent differentiation conditions needs to be answered for the generation of safe populations of cells for therapeutic use.

One differentiation system that does generate a relatively homogeneous population of differentiated cells is the ZHBTc4 ES cell model system. The ZHBTc4 ES cell system was the first system for the differentiation of trophectoderm from ES cell cultures and has been discussed in Chapter 3 (Niwa *et al.*, 2000). It relies on genetic modification of ES cells to allow rapid downregulation of *Oct4*, which results in upregulation of *Cdx2* and other trophectoderm marker genes. Downregulation of *Oct4* can also be achieved by transfection of *Oct4* specific short interfering RNA in ES cells (Hay *et al.*, 2004). This method has been shown to result in upregulation of the trophectoderm-associated gene *Cdx2* in mouse ES cells, providing a second means of inducing trophectoderm differentiation.

5.2 Aims of Chapter 5

In order to examine the role of the general miRNA population in ES cell differentiation the primary aim of this chapter was to compare the differentiation capabilities of DGCR8$^{-/-}$ ES cells with those of wt ES cells using non-directed and directed differentiation methods.

5.3 Results

5.3.1 Validation of the DGCR8$^{-/-}$ ES cell line

5.3.1.1 Transfer of DGCR8$^{-/-}$ ES cells to feeder-free culture conditions

DGCR8$^{-/-}$ ES cell were generated by Robert Blelloch's laboratory at the University of California and obtained from Novus Biologicals. Cells were grown on a mitotically inactivated layer of mouse embryonic fibroblasts, which were derived from E12.5 or E13.5 mouse embryos. Initially the cells grew very slowly and were passaged every 3 or 4 days. On the second passage a vial of cells was frozen in 10% DMSO, 50% FCS in ES cell media. Cells were further expanded for transfer to growth on gelatin and freezing of multiple stock vials of cells.

Initially, DGCR8$^{-/-}$ ES cells were grown on feeders, as suggested by the laboratory in which they originated (Wang *et al.*, 2007). However, growth of ES cells on feeder cells is undesirable for several reasons: potential contamination of experimental samples with fibroblast cells, undefined conditions and more labour intensive cell culture. For these reasons, DGCR8$^{-/-}$ ES cells were transferred to culture on gelatin-coated tissue culture plastic. A culture of DGCR8$^{-/-}$ ES cells growing on a feeder layer was dissociated and plated in a gelatin-coated tissue culture treated T25 flask. Initially, cells grew slowly and were passaged every 4 days. DGCR8$^{-/-}$ ES cells were expanded on gelatin for 3 passages before feeder-free stocks of cells were frozen for long-term storage. After removal from growth on a feeder layer, wt ES cells are frequently observed to go through a 'crisis' with increased cell death and high levels of differentiation. Although DGCR8$^{-/-}$ ES cells appeared to grow at a slightly slower rate immediately after transfer, cells

did not differentiate and no increase in cell death was observed suggesting that no 'crisis' occurred in these cells after transfer to growth on gelatin. The growth rate of DGCR8$^{-/-}$ appeared to be affected by the plating density. When plated at a similar density to wt ES cells the DGCR8$^{-/-}$ ES cells showed slowed growth compared to growth following plating at a higher density. Therefore, in general, DGCR8$^{-/-}$ cultures were passaged every 2 days at a 1:3 split ratio, in comparison to wt ES cells, which were passaged every 2 days at a 1:6 split ratio. DGCR8$^{-/-}$ ES cells could be continuously passaged in the absence of a feeder layer for at least 20 passages.

5.3.1.2 Generation of DGCR8$^{-/-}$ ES cell clones

It was necessary to subclone the DGCR8$^{-/-}$ ES cells because of recurring problems of wt contamination within the DGCR8$^{-/-}$ cultures. This problem occurred in two different batches of cells received from Novus Biologicals, but the source of the contamination is not known. Because wild type cells grow at a much-increased rate compared to DGCR8$^{-/-}$ ES cells (see *Figure 5.5*) they rapidly dominated the miRNA-deficient cultures. Early passage DGCR8$^{-/-}$ ES cells were plated at clonal density and left to grow. After >30 days 16 of 500 plated cells (3.2%) had grown into colonies of ES cells. This was a much lower proportion than the 116 clones that grew from 250 plated CGR8 ES cells (46.4%). Twelve DGCR8$^{-/-}$ ES cells clones were picked and expanded and when enough material was available, RNA was extracted from the cells and used to generate cDNA. QPCR analysis of synthesised cDNA from DGCR8$^{-/-}$ ES cell clones and wt ES cells showed that wt ES cells expressed DGCR8 but all 12 DGCR8$^{-/-}$ clones were deficient for DGCR8 expression (*Figure 5.1*). Further QPCR analysis for expression of ES cell marker genes showed that the clones had variable levels of expression of *Oct4* and *Nanog* relative to levels in wt ES cells (*Figure 5.1*). Four clones (DGCR8$^{-/-}$ clone 3, DGCR8$^{-/-}$ clone 4, DGCR8$^{-/-}$ clone 6 or DGCR8$^{-/-}$ clone 9) were selected for use in future experiments that represented the variation in expression of *Nanog* and *Oct4*. Further experimental repeats would be required to determine whether the variation in *Oct4* and *Nanog* expression between clones is stable or a result of slight differences in culture conditions of the different clones. To ensure that DGCR8$^{-/-}$ ES cell cultures were not contaminated with significant proportions of wt ES cells, all experiments within this chapter were carried out using either

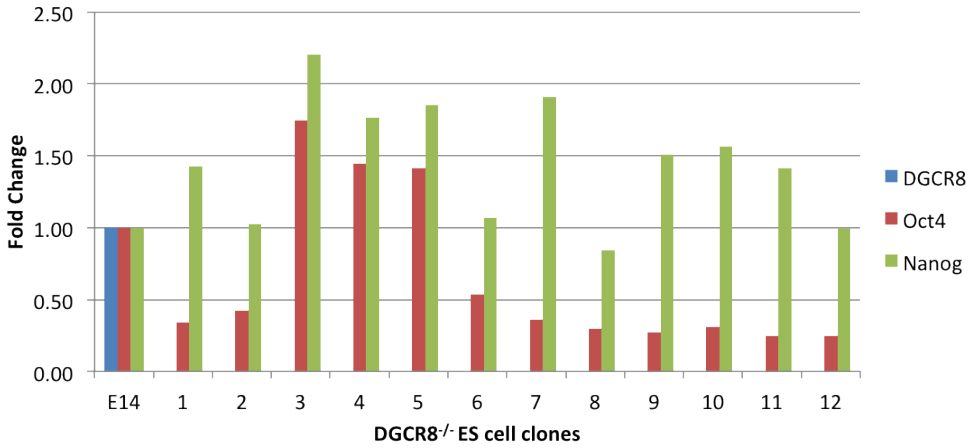


Figure 5.1 – *Gene expression in* DGCR8$^{-/-}$ *ES cell clones.* QPCR analysis of expression of *Dgcr8* (blue), *Oct4* (red) and *Nanog* (green) in E14 ES cells and 12 different DGCR8$^{-/-}$ ES cell clones (subclones). Gene expression was normalised to *β-Actin* expression and is shown relative to expression in E14 ES cells. N=1.

early passages of the original DGCR8$^{-/-}$ ES cell line (original stock) or the newly derived DGCR8$^{-/-}$ ES cells (subclones). For each experiment the figure legend indicates whether DGCR8$^{-/-}$ ES cells were from the original stock or newly derived subclones.

5.3.1.3 Analysis of DGCR8 and miRNAs expression in DGCR8$^{-/-}$ ES cells

Although the cell line was generated by disruption of exon 3 of DGCR8, Wang *et al.* (2007) did not confirm loss of DGCR8 protein directly, but indirectly through the loss of miRNAs in the DGCR8$^{-/-}$ cells. Therefore, the expression of DGCR8 protein was examined in wt and DGCR8$^{-/-}$ ES cell cultures. ES cells were cultured in conditions that would maintain the undifferentiated state and conditions that would induce differentiation (as embryoid bodies). Protein was extracted and used for Western analysis of DGCR8 expression. Western analysis confirmed that DGCR8 protein was present in undifferentiated wt ES cells and in wt ES cells that had been induced to differentiate. In contrast, no DGCR8 protein expression was detected in undifferentiated or differentiated DGCR8$^{-/-}$ ES cells although it must be noted that these data arise from a single experiment (*Figure 5.2a*). To ensure that DGCR8$^{-/-}$ cultures were free from contamination with wt

cells, DGCR8$^{-/-}$ cultures were routinely screened to ensure absence of DGCR8. In order to develop a simple and reliable method of screening cells for DGCR8 deficiency, a QPCR assay was developed with primers designed to amplify exon 3 of DGCR8. Primer optimisation was carried out on pooled cDNA derived from wt ES cell cultures. Optimisation QPCR was performed on a series of cDNA dilutions over a range of 1x10^6 fold. The primer pair was considered suitable for use because it amplified a single product over a 10,000-fold range with 104.8% efficiency and a 'goodness of fit' (RSq) of 0.978. Primer sequences can be found in Materials and Methods. *Table 5.1* contains a list of optimisation details for all primer pairs used in Chapter 5. Following primer optimisation, RNA was isolated from cultures of wt ES cells and DGCR8$^{-/-}$ ES cells and used to generate cDNA. QPCR amplification of cDNA from wt ES cells using the primers specific to exon 3 of DGCR8 showed product amplification (*Figure 5.2b*). In DGCR8$^{-/-}$ ES cells exon 3 of DGCR8 has been removed so no PCR amplification should occur. As expected, no product amplification was observed in DGCR8$^{-/-}$ ES cells (*Figure 5.2b*). In order to confirm that absence of amplification was not due to absence of cDNA within the DGCR8$^{-/-}$ ES cell sample, QPCR was also performed using primers specific to *β-Actin*. *β-Actin* was amplified from both wt and DGCR8$^{-/-}$ cDNA samples confirming that cDNA was present within the DGCR8$^{-/-}$ sample (data not shown).

DGCR8$^{-/-}$ ES cells should not express canonical miRNAs. In order to confirm loss of expression of miRNAs in DGCR8$^{-/-}$ ES cells, RNA was extracted from cultures of two wt ES cell lines (E14 and CGR8), DGCR8$^{-/-}$ ES cells, primary mouse embryonic fibroblasts (MEFs) and a fibroblast cell line called 3T3, and used for Northern analysis of miRNA expression. *MiR-294* is highly expressed in ES cells, but is not expressed in most differentiated cell types (Medeiros *et al.*, 2011). As expected, although *miR-294* was expressed in both wt ES cell lines, no miRNA expression was detected in DGCR8$^{-/-}$ ES cells, or in fibroblast cells (*Figure 5.2c*). It was possible that lack of expression of *miR-294* in DGCR8$^{-/-}$ ES cells was due to differences between wt ES cells and DGCR8$^{-/-}$ ES cells rather than a specific deficiency of miRNA processing in DGCR8$^{-/-}$ ES cells. Therefore, the expression of other miRNAs was also investigated. *Figure 5.2c* shows that *miR-199a* is expressed in fibroblast cells, but not in wt ES cells or in DGCR8$^{-/-}$ ES cells. This showed that DGCR8$^{-/-}$ ES cells do not express miRNAs associated with fibroblast cells and that there was no miRNA contamination from MEFs in the

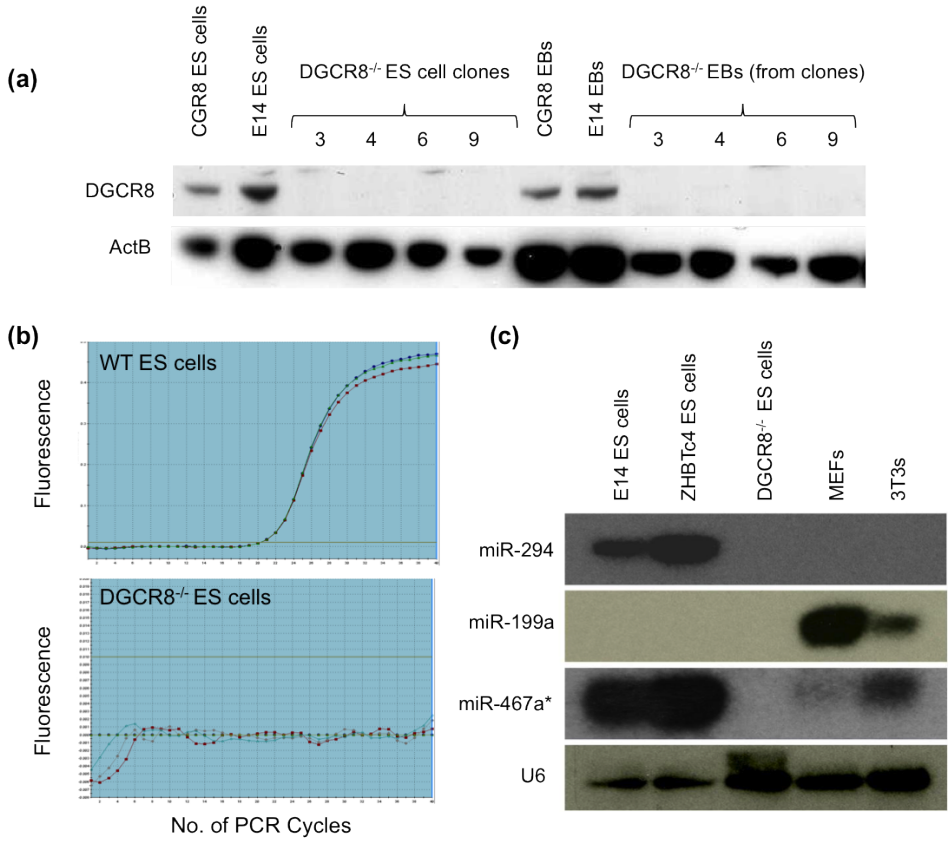


Figure 5.2 – *Expression of DGCR8 and miRNAs in* DGCR8$^{-/-}$ *ES cells.* (a) Western analysis for expression of DGCR8 and β-ACTIN in E14 and CGR8 ES cells and four DGCR8$^{-/-}$ ES cell clones (subclones), and in day 16 EBs generated from these cell lines. N=1. (b) QPCR amplification plots (MxPro) for expression of DGCR8 in E14 ES cells and DGCR8$^{-/-}$ ES cells (original stock). Triplicate reactions are shown. (c) miRNA northern analysis for expression of *miR-294*, *miR-199a*, *miR-467a** and *U6* using 5ug of total RNA from E14 ES cells, CGR8 ES cells, DGCR8$^{-/-}$ ES cells (original stock), primary mouse embryonic fibroblasts (MEFs) and 3T3 cells. N=1.

Gene	Single Product?	Effective amplification range and Ct spread	RSq	Amplification efficiency (%)
Oct4	Yes	10,000 fold (Ct 22-35.5)	0.994	99.9
Sox2	Yes	100,000 fold (Ct 16-32)	0.995	103.9
Nanog	Yes	100,000 fold (Ct 16-33)	0.995	96.5
Stella	Yes	1000 fold (Ct 22.5 – 32.5)	0.997	102.6
Rex1	Yes	10,000 fold (Ct 19 – 32)	0.997	103
Fgf5	Yes	1000 fold (Ct 26-36)	0.985	98.4
Cdx2	Yes	100,000 fold (Ct 22-38)	0.968	98.5
Nestin	Yes	1000 fold (Ct 25-34.5)	0.990	99.7
Gata4	Yes	1000 fold (Ct 26.5-36.5)	0.993	101.3
Gata6	Yes	1000 fold (Ct 26-36.5)	0.982	103.6
Brachyury	Yes	1000 fold (Ct 27.5-37)	0.970	101.6
Snail	Yes	1000 fold (Ct 26-36)	0.992	97.8
Goosecoid	Yes	100 fold (Ct 29-36)	0.971	95.4
DGCR8	Yes	10,000 fold (25-38)	0.978	104.8
ActB	Yes	100,000 fold (Ct 16-33)	0.999	101.1

Table 5.1 – *Optimisation details for QPCR primers used in Chapter 5*

feeder-free DGCR8$^{-/-}$ ES cell cultures. Additionally, *miR-467a** was expressed in both wt ES cells and fibroblasts, but no expression was detected in RNA from DGCR8$^{-/-}$ ES cells. While these data strongly suggest that DGCR8$^{-/-}$ ES cells do not express mature miRNAs it must be noted that these data are from a single experiment and conclusions are therefore preliminary. All together, these data suggest that DGCR8$^{-/-}$ ES cells do not express *Dgcr8* RNA, DGCR8 protein or mature miRNAs.

5.3.1.4 Can transient expression of DGCR8 rescue miRNA expression in DGCR8$^{-/-}$ ES cells?

The DGCR8$^{-/-}$ ES cells were obtained and cultured with the ultimate aim of investigating the role of the general miRNA population in ES cell differentiation. However, there is a possibility that differences may arise between different ES cell lines during their derivation from single cells (Wang *et al.*, 2007), which may introduce clonal variation. To evaluate whether differences between wt ES cells and DGCR8$^{-/-}$ ES cells were due to the absence of miRNAs alone, rescue of DGCR8 expression in DGCR8$^{-/-}$ ES cells was attempted.

Figure 5.3a shows a schematic map of a DGCR8 expression vector containing the cDNA for human DGCR8, which was obtained from Addgene (Addgene plasmid 10921, Landthaler *et al.*, 2004). Because this vector produced human DGCR8, it was necessary to first confirm that expression of human DGCR8 could rescue miRNA expression in DGCR8$^{-/-}$ mouse ES cells. The protein sequences of human and mouse DGCR8 were compared using ClustalW2 (Larkin *et al.*, 2007), which showed that the proteins were 95.6 % identical in amino acid sequence (Alignment is shown in Appendix E). In order to test whether this level of similarity was sufficient to allow human DGCR8 to compensate for a lack of mouse DGCR8, DGCR8$^{-/-}$ ES cell cultures were transiently transfected overnight with either the DGCR8 expression vector or an empty vector control. RNA was harvested from DGCR8$^{-/-}$ ES cells at the time of transfection (T0) and 24, 48, 72 and 96 hours after transfection as well as from wt ES cells (+ve), and was used for northern analysis of expression of *miR-294*, which is normally expressed in mouse ES cells (*Figure 5.3b*). Wt ES cells expressed *miR-294*, but DGCR8$^{-/-}$ ES cells at T0 showed no expression of *miR-294*. No expression of *miR-294* was observed at any timepoint following transfection of empty vector into DGCR8$^{-/-}$ ES cells. In

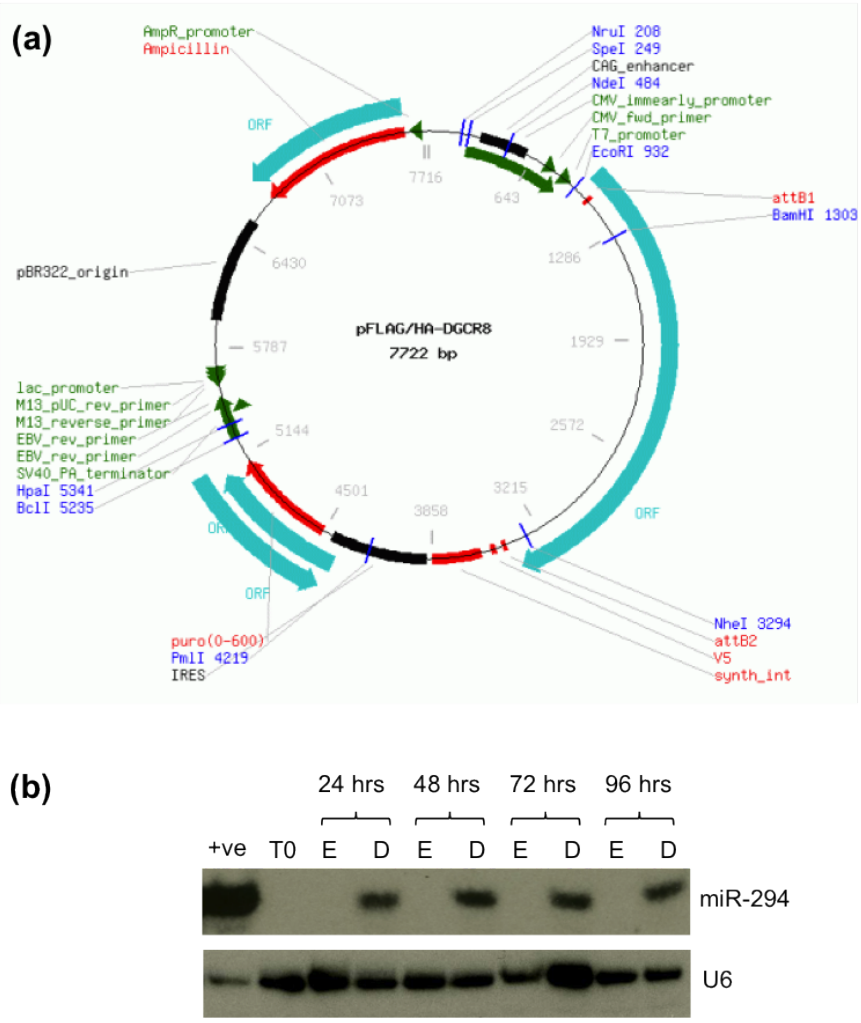


Figure 5.3 – *Expression analysis of miR-294 in* DGCR8$^{-/-}$ *ES cells transfected with a DGCR8 expression vector.* (a) Vector diagram of Addgene human DGCR8 expression vector 10921 (b) miRNA northern analysis for expression of *miR-294* and *U6* in 5ug of total RNA from E14 ES cells (+ve control) and DGCR8$^{-/-}$ ES cells (original stock) transfected with 2ug/ml empty vector (E) or DGCR8 expression vector (D) and harvested after increasing lengths of time. N=1.

contrast, transfection of the vector expressing human DGCR8 resulted in rescue of expression of *miR-294* in DGCR8$^{-/-}$ ES cells at all timepoints investigated (*Figure 5.3b*). This shows that transient expression of human DGCR8 is sufficient to rescue loss of miRNA expression in DGCR8$^{-/-}$ ES cells.

In order to attempt to generate a line of DGCR8$^{-/-}$ ES cells that stably expressed human DGCR8, DGCR8$^{-/-}$ ES cells were transfected with either linearised or circular DGCR8 expression vectors or with empty vector or a vector expressing EGFP. After 48 hours cells were replated at clonal density and 48 hours later puromycin was added to cultures in order to select for clones that had taken up the vector. After 15-17 days clones expressing the empty vector (>100 clones from 5x10^5 transfected cells = >0.02% efficiency) or the EGFP expression vector (>200 clones from 5x10^5 transfected cells = >0.04% efficiency) were large enough to transfer to 96-well plates. These clones were expanded and vials of cells were frozen for storage. No stable clones expressing the DGCR8 expression vector were recovered. The experiment was repeated but again no clones expressing the DGCR8 expression vector were recovered. The reason for the failure to generate a rescue clone is not clear; possible explanations include toxicity of high levels of expression of DGCR8 or a cloning efficiency that was too low to yield a clone from the number of cells initially transfected. Because a DGCR8$^{-/-}$ ES cell rescue clone was not generated, in this thesis chapter experimental results from DGCR8$^{-/-}$ ES cells were compared with results from two wt ES cell lines, with the caveat that some differences may be attributable to clonal variation of the DGCR8$^{-/-}$ ES cell line rather than absence of the general miRNA population.

5.3.2 Characterisation of DGCR8$^{-/-}$ ES cells

5.3.2.1 Morphological analysis of DGCR8$^{-/-}$ ES cells

It was noted that the morphology of the DGCR8$^{-/-}$ ES cells on gelatin was different to that of wt ES cells (compare *Figures 5.4a* & *Figure 5.4c* with *Figure 5.4d*) although when grown on a feeder layer DGCR8$^{-/-}$ ES cell cultures were morphologically similar to what is seen for wt ES cells (*Figure 5.4b*). DGCR8$^{-/-}$ ES cells had a flatter morphology and appeared to have a larger cytoplasmic to nuclear ratio than wt cells. Additionally, the DGCR8$^{-/-}$ cultures often had space between

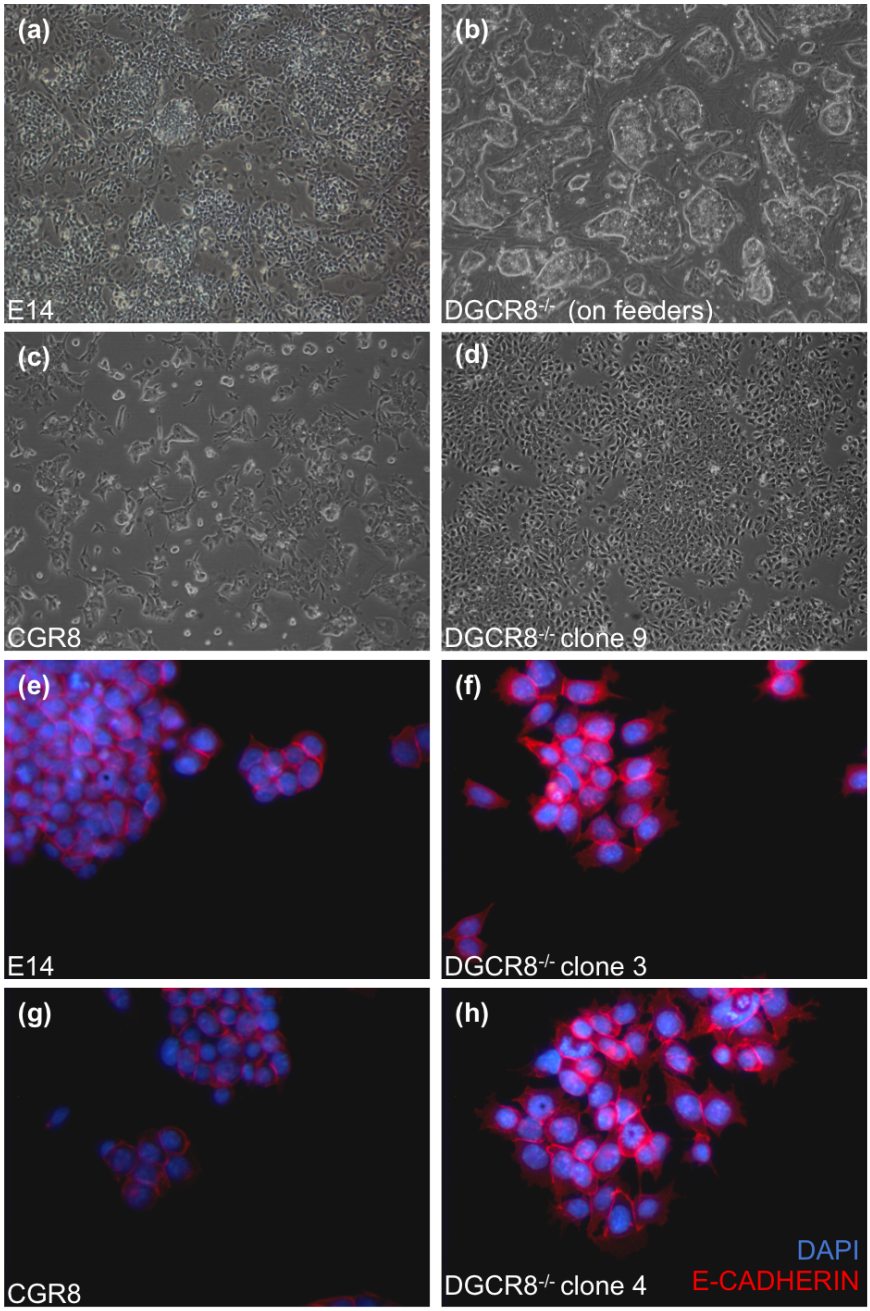


Figure 5.4 – *The morphology of DGCR8 k/o ES cells.* (a-d)Bright-field images of wt and DGCR8$^{-/-}$ ES cells (original stock) at x100 magnification. (a) E14 ES cells growing on gelatin-coated tissue culture treated plastic. (b) DGCR8$^{-/-}$ ES cells (original stock) growing on a mitotically inactivated layer of MEFs. (c) CGR8 ES cells growing on gelatin-coated tissue culture treated plastic. (d) DGCR8$^{-/-}$ ES cells (original stock) growing on gelatin-coated tissue culture treated plastic. (e-h) Images of IHC staining for E-CADHERIN (red) and DAPI (blue) on wt (e & g) and DGCR8$^{-/-}$ ES cells (subclones) (f & h) at x400 magnification. N=1.

cells, rather than existing as the tightly packed domed colonies characteristic of wt ES cells.

E-CADHERIN is expressed highly in undifferentiated mouse ES cells and is critical for their cell-cell adhesion (Larue *et al.*, 1996). E-CADHERIN null ES cells are defective in cell aggregation, reminiscent of the morphology of DGCR8$^{-/-}$ ES cells. To see if the morphological differences between wt and DGCR8$^{-/-}$ cells were reflected in changes in E-CADHERIN expression, DGCR8$^{-/-}$ ES cells and wt ES cells were grown in standard ES cell conditions and fixed in 4% PFA for immunohistochemical analysis of E-CADHERIN expression (*Figure 5.4e-h*). Although the distribution of E-CADHERIN seemed normal in DGCR8$^{-/-}$ ES cells, in that expression was cytoplasmic with enrichment at the boundaries between cells similar to what was seen in wt ES cells, the overall expression level appeared higher. While results were consistent across both wt cell lines and both DGCR8$^{-/-}$ ES cell lines it must be noted that conclusions from this data are preliminary because they arise from a single experiment.

5.3.2.2 Establishing the proliferation rate of DGCR8$^{-/-}$ ES cells

In addition to morphological differences, it has been reported that miRNA-deficient ES cells grow at a slower rate than wt cells (Wang *et al.*, 2007). To establish whether or not this difference was maintained under feeder-free conditions, cultures of wt ES cells and DGCR8$^{-/-}$ ES cells were grown in standard ES cell conditions in 96-well plates and were subjected to MTT analysis of cell proliferation 24, 48, 72 and 96 hours after plating. Cell numbers increased over time in wt and DGCR8$^{-/-}$ ES cell cultures, but the increase was much greater in the wt ES cell cultures than the DGCR8$^{-/-}$ ES cell cultures with wt ES cell doubling rates approximately twice that of DGCR8$^{-/-}$ ES cell doubling rates (*Figure 5.5*). Wild type doubling rates were consistent with that observed by Wang *et al.* (~16 hours) but Wang *et al.* reported doubling rates of DGCR8$^{-/-}$ ES cells to be ~21-23 hours whereas these data suggest a doubling time of >24 hours. The observed differences in doubling rates is difficult to rationalise because cell densities and sampling times were approximately similar between the two analyses. However, the cell proliferation analysis in the Wang *et al.* paper was carried out on cells that had been directly removed from a feeder layer whereas the cells in our analysis were grown permanently in the absence of feeder cells. Perhaps differences

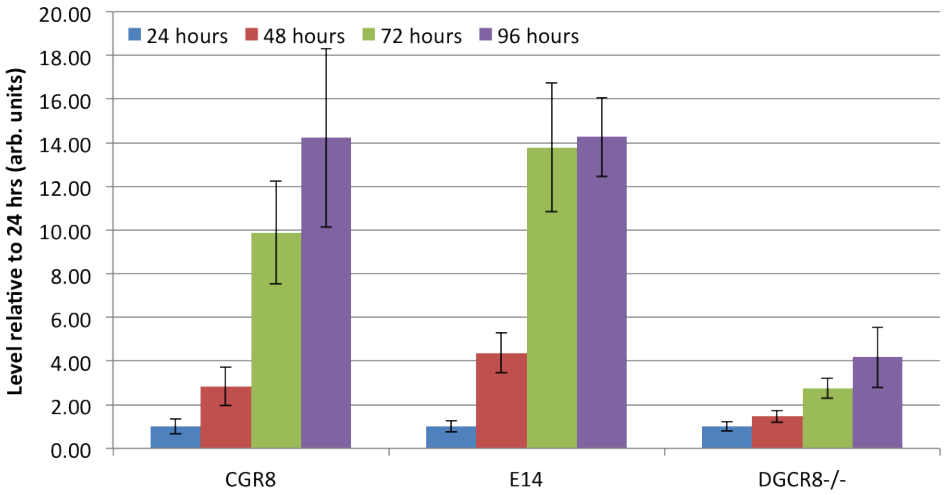


Figure 5.5 – *CYQUANT analysis of wt and* DGCR8$^{-/-}$ *ES cell proliferation.* Bar chart showing relative absorbance at 485nm of E14, CGR8 and DGCR8$^{-/-}$ ES cells (original stock) from CYQUANT analysis. Readings are an average of 12 replicates, were normalized to blank wells and are shown relative to absorbance at 24 hours. Error bars show standard deviation.

in proliferation rates are reflective of differences between cells that have recently been removed from a feeder layer, and those that have adapted to feeder-free growth.

5.3.2.3 Expression analysis of ES cell marker gene and signalling pathway components in DGCR8$^{-/-}$ ES cells

Wild type ES cells express characteristic markers of pluripotency. DGCR8$^{-/-}$ ES cells had previously been shown to express a limited set of ES cell markers. To determine how similar DGCR8$^{-/-}$ cells were to wt ES cells, this characterisation was extended at the RNA and protein levels. Cells were cultured in standard ES cell media and harvested for RNA or protein. RNA was used to generate cDNA for QPCR analysis of *Oct4*, *Sox2*, *Nanog*, *Stella* and *Rex1* expression. *Oct4*, *Sox2*, *Rex1* and *Stella* were expressed at similar levels in DGCR8$^{-/-}$ ES cells relative to levels in wt ES cells (*Figure 5.6a*). However, *Nanog* was expressed at almost twice the level in DGCR8$^{-/-}$ ES cells seen in wt ES cells. This finding was reproducible and statistically significant as assessed by Student T-Test ($p<0.01$).

Protein samples were also harvested from cell cultures and used for Western analysis of OCT4 expression and expression of key signalling pathway components. Western analysis confirmed that protein levels of the key ES cell pluripotency gene, OCT4, were equivalent between both wt ES cell lines and all 4 different DGCR8$^{-/-}$ ES cell clones (*Figure 5.6b*). The expression of the key signalling pathway components GSK3B, STAT3, AKT and ERK and their phosphorylated forms was compared by Western analysis of protein from wt and DGCR8$^{-/-}$ ES cells (*Figure 5.6b*). Western analysis was performed on protein from four different DGCR8$^{-/-}$ ES cell clones and from wt ES cell lines. The expression of phosphorylated forms of GSK3B, STAT3, AKT and ERK1/2 were compared with expression levels of the total protein. Analysis of ß-ACTIN expression was used to ensure that similar levels of total protein were being compared. Expression levels of phosphorylated forms of GSK3B, STAT3, AKT and ERK were similar in DGCR8$^{-/-}$ ES cells and wt ES cells. This was also the case for expression levels of the total protein. However, it must be noted that the antibody for phosphorylated STAT3 detected phosphorylation at serine residue 727 but phosphorylation of STAT3 normally associated with maintaining pluripotency is at tyrosine residue 705 (Niwa *et al.*, 1998). Therefore, differences in activation of the STAT3 pathway due to Lif signalling would not have been picked up by this analysis.

These results demonstrated that, in terms of marker gene expression and signalling pathway activation, DGCR8$^{-/-}$ ES cells are similar to wt mouse ES cells. This suggests that miRNAs are not required for the normal expression of key pluripotency markers or activation of signalling pathways associated with the pluripotent state. However, IHC analysis of E-CADHERIN expression suggested that expression of this protein may be higher in DGCR8$^{-/-}$ ES cells than wt ES cells, and *Nanog* RNA was expressed at higher levels in DGCR8$^{-/-}$ ES cells than in wt ES cells. Therefore, DGCR8$^{-/-}$ ES cells show differences in expression of some ES cell-related genes compared to wt ES cells.

Nanog was expressed at a consistently higher level in DGCR8$^{-/-}$ ES cells compared to wild type ES cells. This observation was explored further at the cellular level. While OCT4 is expressed uniformly in undifferentiated ES cells, NANOG expression is normally heterogeneous in ES cell cultures (Chambers *et al.*, 2003, Singh *et al.*, 2007). The possibility that the increased expression of *Nanog* mRNA in the DGCR8$^{-/-}$ ES cells may be due to a reduction in heterogeneity of expression

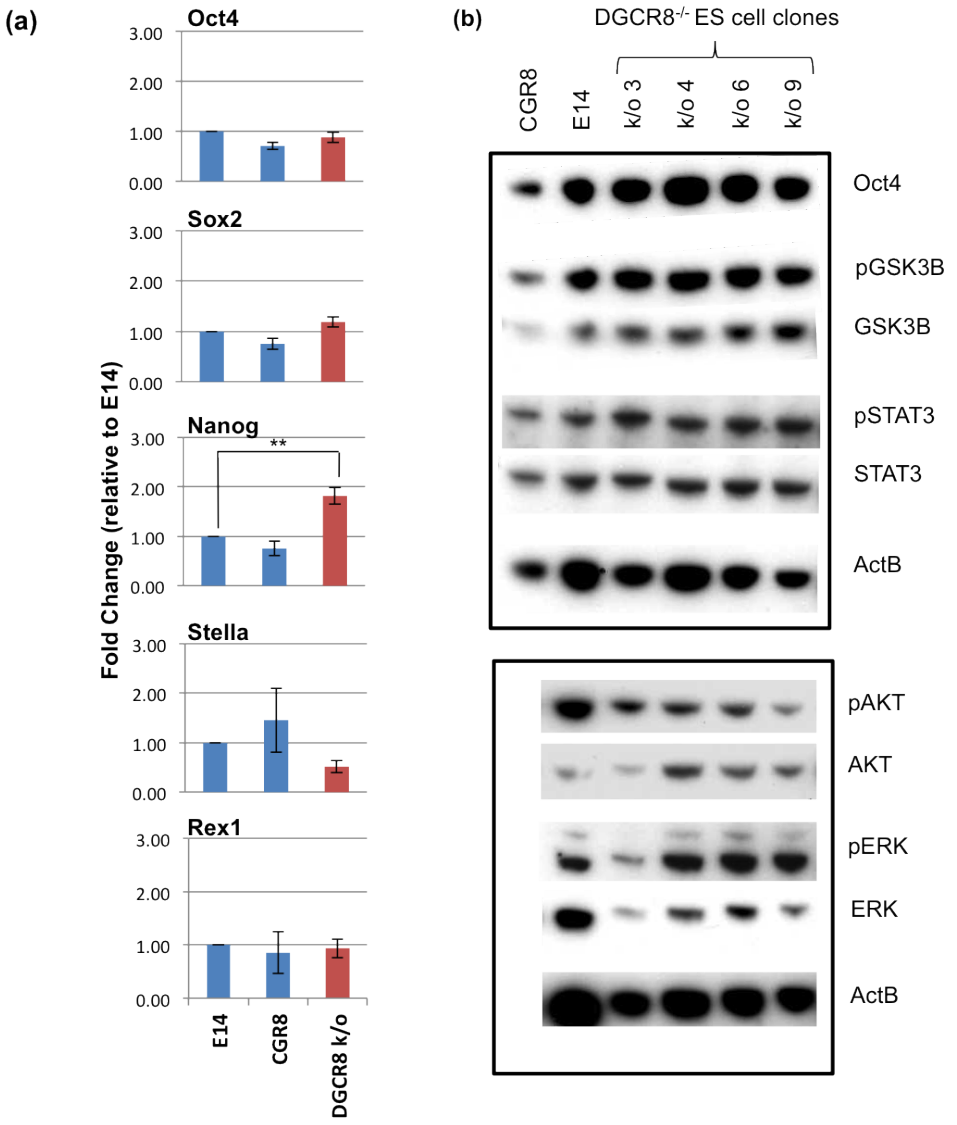


Figure 5.6 – *Expression of ES cell genes and signaling pathway components in* DGCR8$^{-/-}$ *ES cells and wt ES cells.* (a) QPCR analysis for ES cell genes in E14 ES cells, CGR8 ES cells and DGCR8$^{-/-}$ ES cells (subclones). QPCR readings were normalized to expression of the housekeeping gene, *β-Actin* and corrected values are shown relative to expression in E14 ES cells. Error bars are SEM. **=p<0.01. Data shows mean of at least 2 E14 and CGR8 replicates and at least 3 DGCR8$^{-/-}$ replicates. (b) Western analysis for expression of OCT4, signaling pathway components and β-ACTIN on protein from 1x10^6 CGR8 ES cells, E14 ES cells and 4 different DGCR8$^{-/-}$ clones (subclones).

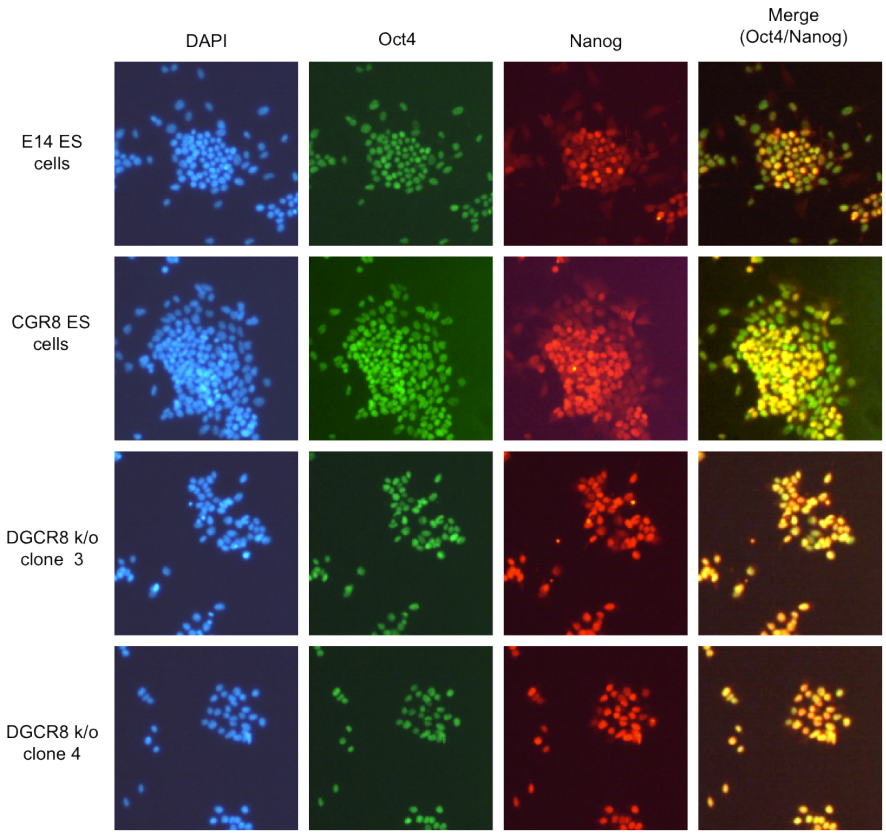


Figure 5.7 – *IHC analysis of OCT4 and NANOG expression in wt and* DGCR8$^{-/-}$ *ES cells.* Photographs of DAPI staining of nuclei (blue) and IHC for OCT4 (green) and NANOG (red) expression in E14 ES cells, CGR8 ES cells and 2 different clones of DGCR8$^{-/-}$ ES cells (subclones). Photographs are representative of two independent experiments. Magnification is x100.

in these cells was considered. In order to test this idea, immunohistochemistry to visualise NANOG expression was performed in DGCR8$^{-/-}$ ES cells and wt ES cells. Cultures were co-stained for OCT4 to identify undifferentiated cells. Heterogeneous NANOG expression could therefore be visualised by comparing the proportions of NANOG +ve/OCT4 +ve cells (yellow) with NANOG –ve/OCT4 +ve cells (green). *Figure 5.7* shows that there is greater heterogeneity in NANOG expression in wt cultures than in DGCR8$^{-/-}$ cultures indicated by a greater proportion of green OCT4 stained nuclei than yellow co-stained nuclei. *Figure 5.7* shows images representative of two experimental repeats of OCT4/NANOG immunohistochemistry staining of 2 wt ES cell lines and multiple DGCR8$^{-/-}$ ES cell clones. These data suggest that miRNAs may be required for maintaining the cyclic expression of NANOG that is normally seen in ES cell cultures.

5.3.3 Differentiation of DGCR8$^{-/-}$ ES cells

5.3.3.1 Analysis of gene expression following siRNA knockdown of *Oct4* expression

Chapter 3 of this thesis described the results of investigating the role of miRNAs during trophectoderm differentiation of ES cells. In the initial analysis, trophectoderm differentiation was induced in the ZHBTc4 ES cell line through rapid downregulation of *Oct4* (Niwa *et al.*, 2000) but can also be induced in wt ES cells by transfection of siRNAs targeting *Oct4* (Hay *et al.*, 2004). This method was used to test whether DGCR8$^{-/-}$ ES cells were capable of differentiating into trophectoderm, as assessed by upregulation of a panel of trophoblast marker genes.

siRNAs designed to inhibit expression of *Oct4* or EGFP (as control) were transfected into wt E14 ES cells, ZHBTc4 ES cells and DGCR8$^{-/-}$ ES cells. ZHBTc4 ES cells grown in the presence of doxycycline differentiate rapidly into trophectoderm and were used as positive controls. In ZHBTc4 cultures morphological changes were observed after as little as 24 hours of growth: areas of the cultures became flat and the boundaries between cells were less obvious. 72 hours after the initial transfection similar morphological changes were observed in the cultures transfected with *Oct4* siRNA but not in those transfected with a siRNA designed to inhibit EGFP expression (*Figure 5.8*).

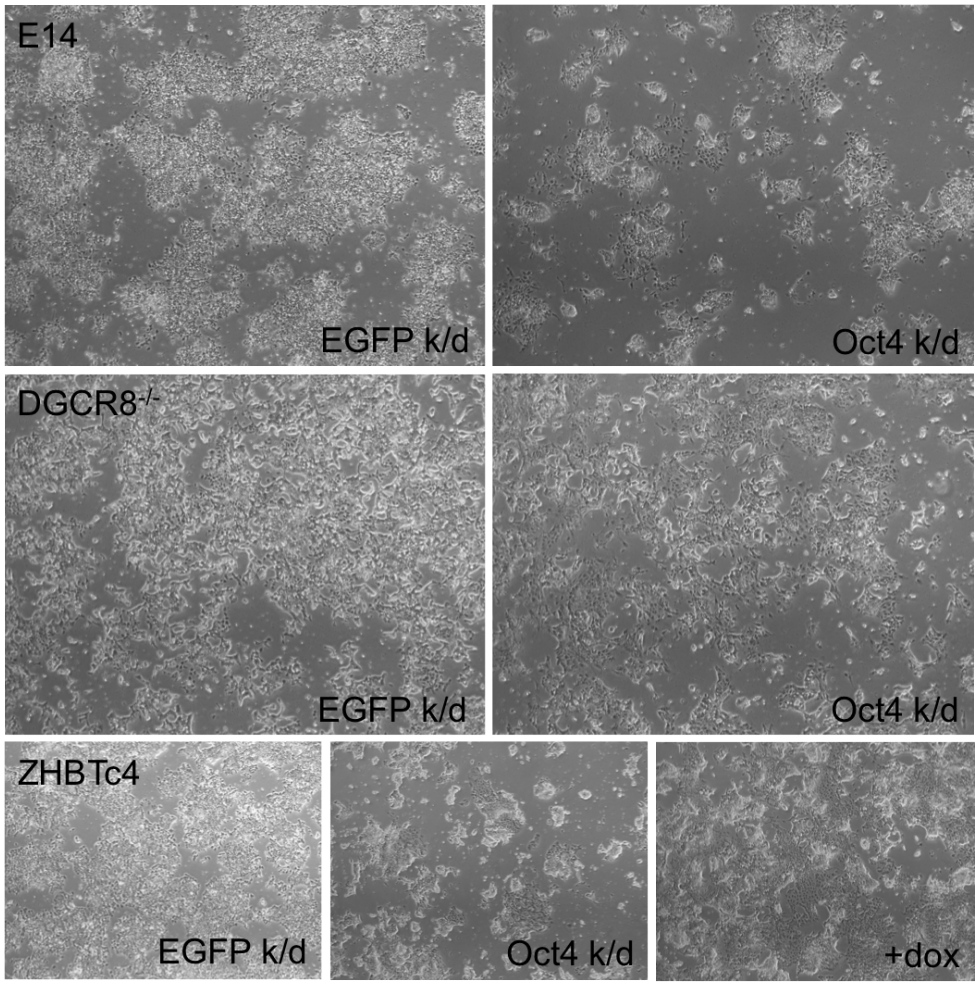


Figure 5.8 – *The morphology of wt and* DGCR8$^{-/-}$ *ES cells following siRNA knockdown of Oct4 expression.* Bright field photographs of E14 ES cells, ZHBTc4 ES cells and DGCR8$^{-/-}$ ES cells (original stock) 72 hours after transfection of 100ng/ml siRNA targeting *Oct4* or EGFP (control). Positive control for differentiation is ZHBTc4 ES cells grown in the presence of 1μM doxycycline. Photographs are representative of three replicates. Magnification is x100.

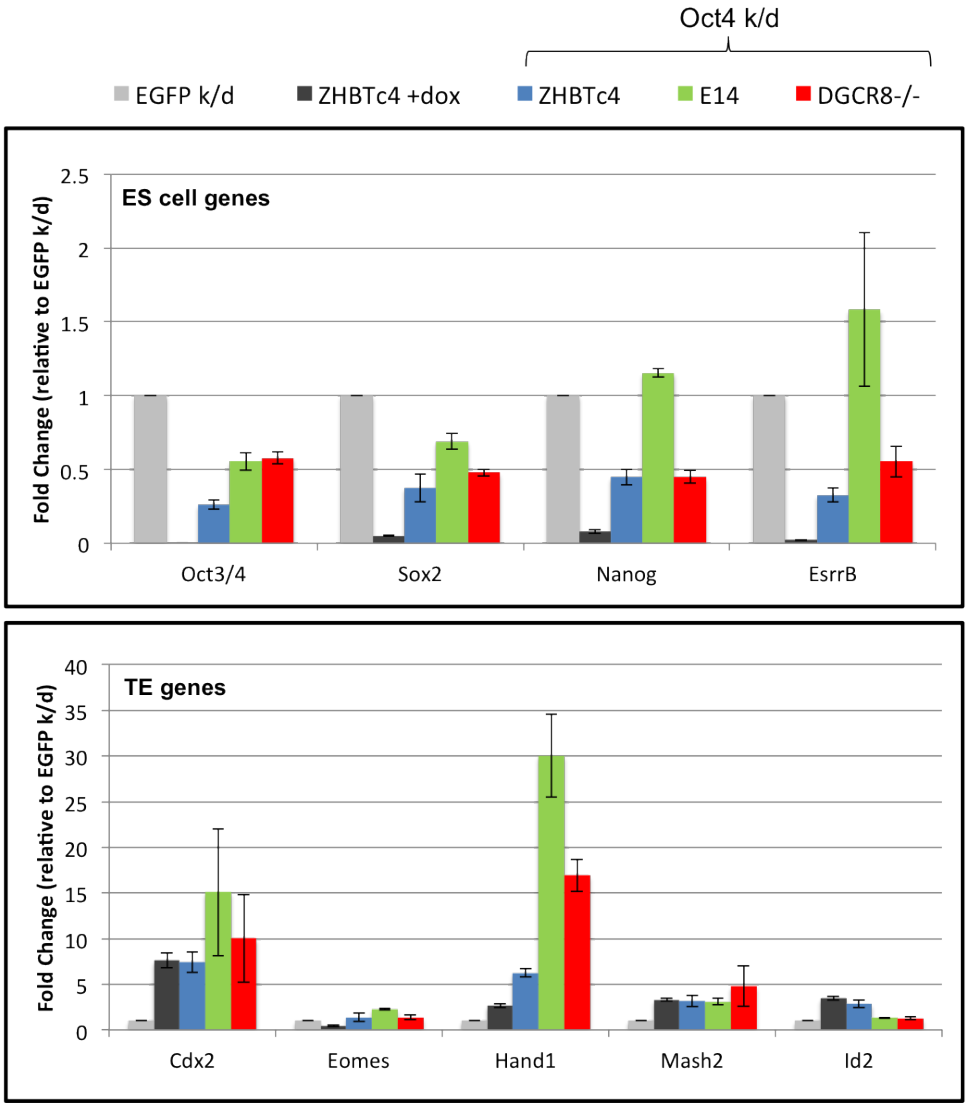


Figure 5.9 – *Expression of ES and TE genes following siRNA knockdown of Oct4 expression in ES cells.* QPCR analysis for markers of pluripotency (top) and trophectoderm differentiation (bottom) after EGFP knockdown (light grey), in ZHBTc4 ES cells grown in the presence of 1 μM doxycycline (dark grey), and in ZHBTc4 ES cells (blue), E14 ES cells (green) and DGCR8$^{-/-}$ ES cells (original stock - red) after transfection with 100ng/ml siRNA targeting *Oct4*. QPCR readings were normalized to expression of the housekeeping gene, *β-Actin* and corrected values are shown relative to expression after EGFP knockdown (light grey). Error bars are SEM, n=3.

RNA was extracted from the cultures 72 hours after transfection of siRNA oligonucleotides and used to make cDNA for QPCR analysis of expression of ES cell and TE marker genes (*Figure 5.9*). ZHBTc4 ES cells grown in the presence of doxycycline showed marked downregulation of ES cell genes compared to ZHBTc4 ES cell cultures transfected with siRNA targeting EGFP, which were grown in the absence of doxycycline. They also showed upregulation of the TE marker genes *Cdx2*, *Hand1*, *Mash2* and *Id2*. Following siRNA knockdown of *Oct4* expression, the ES cell markers *Oct4* and *Sox2* were downregulated in ZHBTc4, E14 and DGCR8$^{-/-}$ cultures, and *Nanog* and *Esrrβ* were downregulated in ZHBTc4 and DGCR8$^{-/-}$ cultures relative to cultures transfected with EGFP siRNA. Conversely, expression of multiple trophoblast markers (*Cdx2*, *Eomes*, *Hand1*, *Mash2* and *Id2*) was up regulated after transfection of *Oct4* siRNA relative to EGFP siRNA in wt and DGCR8$^{-/-}$ cultures. These data show that DGCR8$^{-/-}$ ES cells are capable of downregulating expression of ES cell marker genes and up regulating expression of trophoblast marker genes to the same extent as wt ES cells after *Oct4* knockdown and suggests that miRNAs are not required for initiation of TE differentiation. A recent paper from Robert Blelloch's laboratory, which showed that miRNAs are not required for pre-implantation blastocyst formation (and therefore initiation of TE formation), supported this finding (Suh *et al.*, 2010).

5.3.3.2 Examination of the morphology and gene expression in DGCR8$^{-/-}$ ES cells subjected to a neural differentiation protocol

The capacity for miRNA-deficient ES cells to undergo neural differentiation was investigated using a neural differentiation method that was developed in our laboratory. Two wt ES cell lines and 4 DGCR8$^{-/-}$ ES cell clones were grown on laminin in N2B27 media supplemented with 0.25uM PD0325901 and 1.5uM CHIR99021 for 7 days. After 7 days cultures were fixed in 4% PFA and expression of the early neural marker NESTIN, and the mature neural marker TUJ1 were visualised by immunohistochemistry (*Figure 5.10*). Tuj is a marker of mature neurons only when the cell expressing TUJ1 also displays a neural morphology, with long thin processes protruding from it. As can be seen from *Figure 5.10*, both wt and miRNA-deficient ES cells were capable of producing cells with a neural morphology that were positive for TUJ1. QPCR analysis of cDNA from

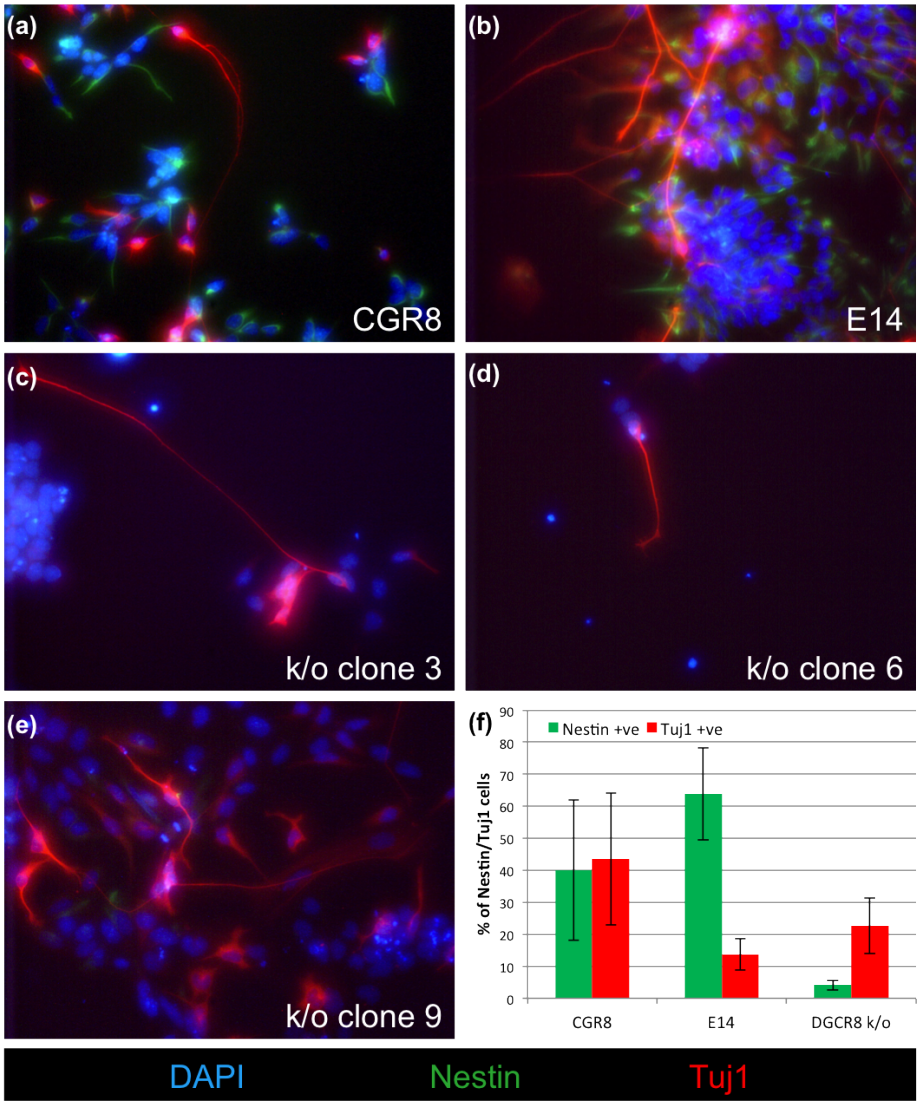


Figure 5.10 – *IHC analysis of NESTIN and* TUJ1 *expression in wt ES cells and* DGCR8$^{-/-}$ *ES cells subjected to a neural differentiation protocol.* (a-e) Photographs of IHC staining for DAPI (blue), NESTIN (green) and TUJ1 (red) expression in cultures grown in N2B27 with 0.25uM PD0325901 and 1.5uM CHIR99021 for 7 days. Photographs are from a single experiment and show cultures from 2 wild type ES cell lines - E14 (a) and CGR8 (b) and 3 different clones of DGCR8$^{-/-}$ ES cells (subclones) (c-e). Magnification is x400. (f) Quantitation of IHC photographs by cell counting. Bar graphs show percentage of NESTIN positive or TUJ1 positive cells relative to DAPI stained cells. 3 photographs were counted per well. DGCR8$^{-/-}$ result is average of clone 3, clone 6 and clone 9 counts. Error bars show standard deviation.

parallel cultures confirmed the DGCR8-deficient status of the DGCR8$^{-/-}$ cultures (data not shown).

It appeared that the DGCR8$^{-/-}$ cultures produced a smaller proportion of NESTIN-positive neural precursors than wt cultures. In order to quantitate this the number of Dapi stained cells, NESTIN positive cells and TUJ1 positive neural cells were counted. Three randomly selected areas were counted per well of differentiated cells. It should be noted that counts were only performed on areas of culture that contained TUJ1 positive neurons so will overestimate the proportion of NESTIN positive and TUJ1 positive cells in the total culture. *Figure 5.10f* shows the average percentages of cells that stained positive for either NESTIN or TUJ1. The two wild type cultures differed in their capacity to make mature neurons, E14 ES cells produced a greater proportion of NESTIN-positive neural precursors, but a lower proportion of TUJ1 positive neurons than CGR8 ES cells. The DGCR8$^{-/-}$ ES cell clones produced similar proportions of TUJ1 positive neurons but a much lower proportion of NESTIN-positive neural precursors than wild type cultures. These data suggest that some form of neurogenesis can occur in the absence of miRNAs, but that miRNAs may be involved in regulating aspects of normal neurogenesis including maintenance of neural progenitor populations. These conclusions are preliminary because data is from a single experimental analysis.

5.3.3.3 Comparison of the differentiation of embryoid bodies from wt and DGCR8$^{-/-}$ ES cells

DGCR8$^{-/-}$ ES cells appear to be able to undergo some level of differentiation down the trophectoderm and neuroectoderm lineages as a result of directed differentiation. Therefore the question of whether miRNAs were required for differentiation into other lineages using an undirected differentiation method was asked. Embryoid body (EB) differentiation is a general differentiation method, which is capable of generating cells belonging to the three germ lineages. EBs were made from wt ES cells and from miRNA deficient ES cells and grown in suspension for 8 days, followed by growth in adherent culture for a further 8 days. Cultures were harvested and RNA prepared from cells from the starting population of ES cells (T0) and from cells collected at days 4, 8, 12 and 16 of EB differentiation. Differences in morphology were observed between wild type and miRNA-deficient EBs

(*Figure 5.11*). DGCR8$^{-/-}$ EBs were generally smaller than wt EBs, and never formed cysts during suspension culture. It is thought that primitive endoderm forms on the surface of embryoid bodies and this allows an inner layer of columnar epithelium to form that surrounds a fluid filled cavity thought to be functionally analogous to the visceral yolk sac (Weitzer, 2006). The total absence of such 'cysts' in DGCR8$^{-/-}$ embryoid bodies suggests a problem with formation of this structure, which may be due to an inability to properly specify primitive endoderm. When transferred to adherent culture, cells from the EBs did not spread across the plate to the same extent as cells from wt EBs; the plated DGCR8$^{-/-}$ EBs remained small with a tight area of cell outgrowth containing cells that resembled those of an ES-like morphology or a much more flattened morphology akin to that of trophoblast-like cells. In contrast the wt EB outgrowths were very heterogeneous in appearance.

RNA was extracted from samples collected at various timepoints during EB differentiation. Following cDNA synthesis, QPCR was performed to determine levels of transcripts of markers of different lineages. In wt ES cells *Fgf5* levels peak in expression on day 4 and then decline throughout the remainder of the differentiation timecourse (*Figure 5.12*). This could be expected as *Fgf5* expression is seen to increase upon Lif withdrawal (Martinez-Ceballos *et al.*, 2005). In contrast, in DGCR8$^{-/-}$ ES cells the peak in *Fgf5* expression occurs at day 8 and levels remain high during differentiation. Maintenance of *Fgf5* expression may indicate that miRNA-deficient ES cells are initiating early differentiation but failing to progress to more advanced stages.

Previous experiments indicated that miRNA-deficient ES cells could initiate trophectoderm differentiation upon siRNA knockdown of *Oct4* expression displaying a 10-fold increase in the expression of the TE marker gene, *Cdx2*. During embryoid body differentiation of wt ES cells, *Cdx2* expression increased up to a peak at day 8 of differentiation and then declined through the remainder of the differentiation timecourse (*Figure 5.12*). In DGCR8$^{-/-}$ EBs *Cdx2* expression increased throughout differentiation, peaking at day 12 and remaining high at day 16. This suggests that DGCR8$^{-/-}$ EBs are capable of trophectoderm differentiation, but that differentiation may occur at a slower rate than in wt EBs.

Proper early embryonic development is characterised by the specification of the three germ lineages: ectoderm, endoderm and mesoderm. In order to assess the

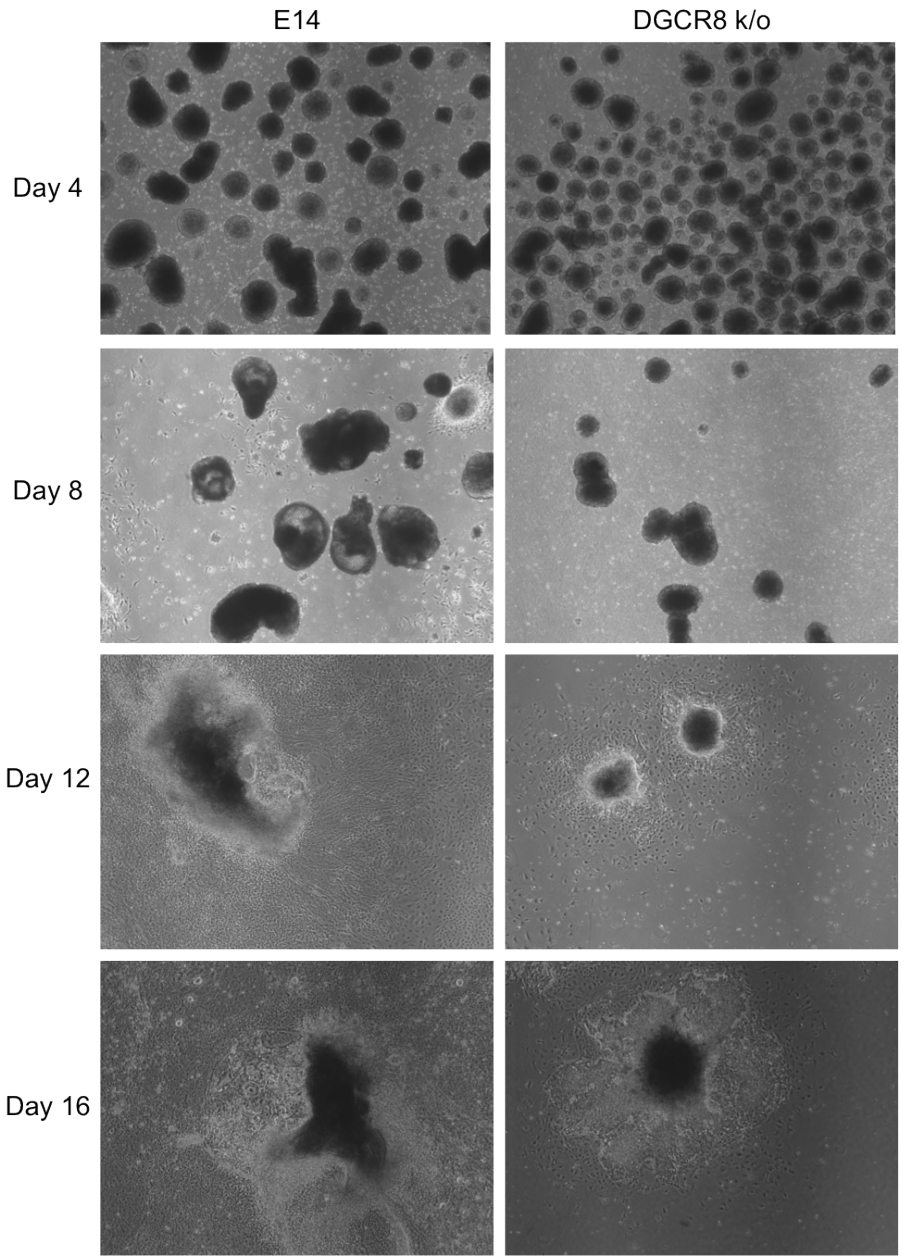


Figure 5.11 – *The morphology of wt and DGCR8 /- embryoid bodies.* Bright field photographs of embryoid bodies generated from E14 ES cells and DGCR8 /- ES cells (subclones). Photographs show embryoid bodies in suspension at day 4 and day 8 of differentiation, and in adherent culture at day 12 and day 16 of differentiation. Photographs are representative of 4 experimental replicates for E14 ES cells and 5 experimental replicates for DGCR8 /- ES cells. Magnification is x40.

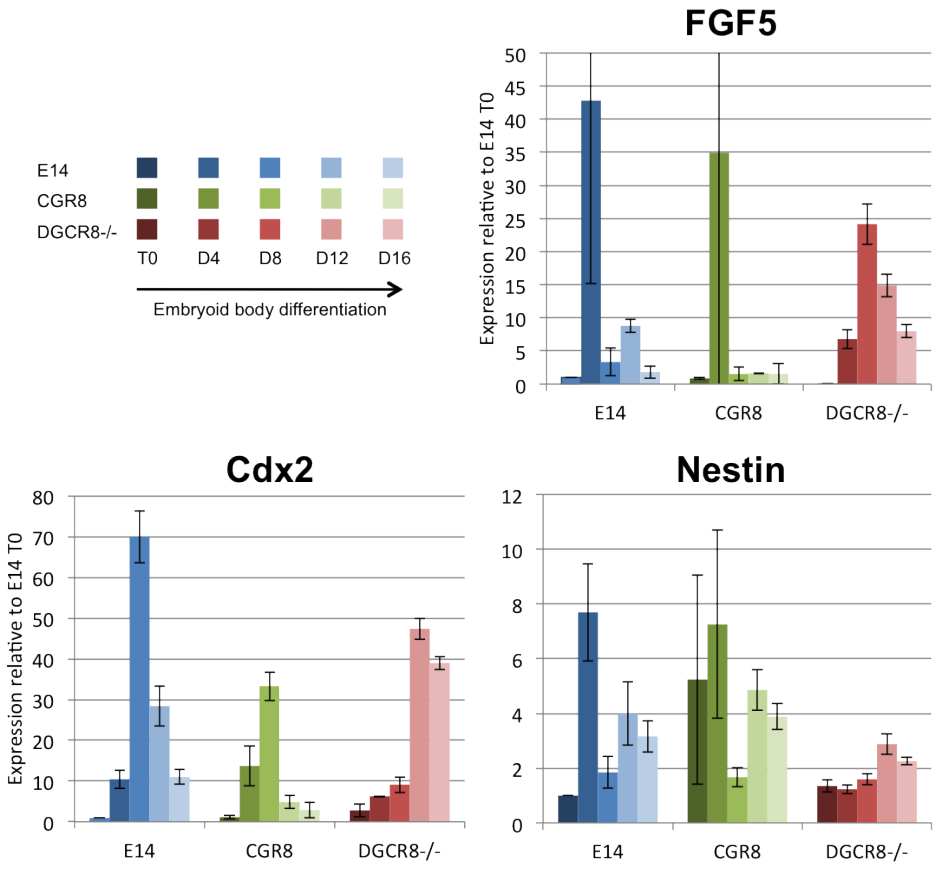


Figure 5.12 – *Expression of lineage marker genes during EB differentiation of wt ES cells and DGCR8 -/- ES cells.* QPCR analysis for expression of the epiblast marker *Fgf5*, the trophectoderm marker *Cdx2* and the neuroectoderm marker *Nestin* during embryoid body differentiation of E14 ES cells (blue) CGR8 ES cells (green) and DGCR8$^{-/-}$ ES cells (subclones - red). QPCR readings were normalized to expression of the housekeeping gene, *β-Actin* and corrected values are shown relative to expression in E14 ES cells at T0. Error bars are SEM, E14 n=4, CGR8 n=3, DGCR8$^{-/-}$ n=5.

overall differentiation capacity of DGCR8$^{-/-}$ ES cells, expression of markers of each of these lineages was assessed during EB differentiation, and compared to levels in wt ES cells. The expression of *Nestin*, a marker of neuroectodermal differentiation, was variable during differentiation of EBs from wt ES cells and induction of expression varied between approximately 2-fold and 8-fold. In contrast, *Nestin* expression remained relatively constant in EBs from DGCR8$^{-/-}$ ES cells and expression was not induced by greater than 3-fold during the differentiation timecourse (*Figure 5.12*). This was consistent with the finding in *Figure 5.10*, which showed that the proportion of NESTIN positive cells was lower in cultures of DGCR8$^{-/-}$ ES cells induced to differentiate down the neural lineage compared to the proportion of cells from wt cultures. However, cells with neural morphology were observed in later cultures of wt and DGCR8$^{-/-}$ EBs suggesting that some form of neural differentiation does occur in miRNA-deficient cultures following non-directed EB differentiation.

Gata4 and *Gata6* are genetic markers of the endoderm lineage. *Gata4* expression was rapidly induced and maintained at 50-80-fold above T0 levels in wt ES cells, but in DGCR8$^{-/-}$ ES cells *Gata4* levels were induced and maintained throughout the differentiation timecourse at only 20-fold above the level in E14 T0 cultures (*Figure 5.13*). More strikingly, while expression of *Gata6* was also rapidly induced in wt EBs and was maintained at levels 150-fold higher than in E14 T0 samples, expression in DGCR8$^{-/-}$ EBs did not reach E14 T0 levels until day 12 of differentiation, and then increased to an expression level that was only 30-fold greater than in the E14 T0 sample (*Figure 5.13*). These findings suggest a deficiency in proper endoderm specification in miRNA-deficient EBs. The observation that DGCR8$^{-/-}$ EBs never formed cysts supports this conclusion, because differentiation of primitive endoderm is thought to be essential for cyst formation.

The mesoderm marker *Brachyury* was dramatically induced early in wt EB differentiation, with a peak in expression at day 4 of differentiation. In contrast, *Brachyury* expression was never detected in DGCR8$^{-/-}$ ES cells and remained undetectable at day 4 of embryoid body differentiation (*Figure 5.14a*). Very low levels of *Brachyury* expression were present in 2 of the 5 DGCR8$^{-/-}$ EB differentiation replicate timecourses at day 8 of differentiation and in 4 of the 5 timecourses at days 12 and 16 of differentiation. The expression of Snail, a regulator of epithelial to mesenchymal transition, was induced and maintained in DGCR8$^{-/-}$ EBs

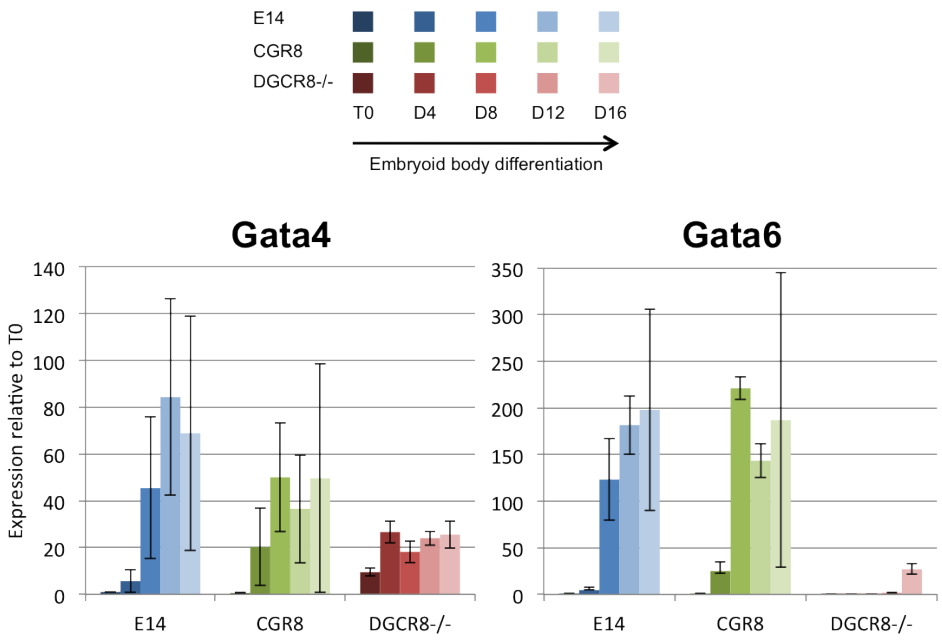


Figure 5.13 – *Expression of endoderm marker genes during EB differentiation of wt ES cells and DGCR8 /- ES cells.* QPCR analysis for expression of the endoderm markers *Gata4* and *Gata6* during embryoid body differentiation of E14 ES cells (blue) CGR8 ES cells (green) and DGCR8$^{-/-}$ ES cells (subclones - red). QPCR readings were normalized to expression of the housekeeping gene, *β-Actin* and corrected values are shown relative to expression in E14 ES cells at T0. Error bars are SEM, E14 n=4, CGR8 n=3, DGCR8$^{-/-}$ n=5.

similar to what was observed for wt EBs although the induction was of a smaller magnitude. In wt EBs, Snail expression peaked at 6-8-fold higher than the level in the E14 T0 sample, but in DGCR8$^{-/-}$ ES cells the induction peaked at only 3-fold (*Figure 5.14b*). Expression of the mesoderm marker, Goosecoid, was induced to a similar extent in EBs from E14 and DGCR8$^{-/-}$ ES cells although a greater induction in expression was observed in CGR8 EBs at day 4 of differentiation (*Figure 5.14b*). Overall, these data suggest that DGCR8$^{-/-}$ EBs may have a particular deficiency in aspects of mesoderm differentiation. This conclusion was supported by the observation that although between 40% and 50% of plated wt EBs contained beating cardiomyocytes, beating cardiomyocytes were never observed in DGCR8$^{-/-}$ EBs (*Figure 5.14b*).

In order to investigate the expression of signalling pathway components in wt and DGCR8$^{-/-}$ cultures induced to differentiate, EBs were made from wt ES cells and from DGCR8$^{-/-}$ ES cells and grown in suspension for 8 days, followed by adherent culture for a further 8 days. At day 16 of differentiation, protein was harvested from cultures and used for western analysis of expression of signalling pathway components (*Figure 5.15*). Expression of signalling molecules was similar between EBs from CGR8 ES cells and E14 ES cells. Additionally, the expression of pGSK3B and GSK3B was similar in EBs from wt ES cells and four clones of DGCR8$^{-/-}$ ES cells. In contrast, pSTAT3 (ser727) and STAT3 were expressed at lower levels in miRNA-deficient EBs than wt EBs. In ES cells, phosphorylated STAT3 is associated with the maintenance of pluripotency, but it is also associated with differentiation during embryonic development (Calo *et al.*, 2003). The fact there is less phosphorylated STAT3 in miRNA deficient EBs substantiates the finding that expression of differentiation markers is generally lower than in wt EBs. The reduced level of phosphorylated STAT3 is likely to be a result of reduced differentiation in DGCR8$^{-/-}$ EBs, but there is also evidence to suggest that it could represent the cause of the differentiation deficiency (Foshay *et al.*, 2005). The levels of AKT and ERK were similar between wt and DGCR8$^{-/-}$ EBs, but the active versions of these molecules, pAKT and pERK, were generally expressed at lower levels in DGCR8$^{-/-}$ EBs than in wt EBs. Signalling through the ERK pathway is known to have a pro-differentiative effect in ES cells (Burdon *et al.*, 1999), interesting in the context of reduced expression in DGCR8$^{-/-}$ EBs because of their reduced differentiation capacity. The function of AKT is normally associated with cell survival and proliferation (Brazil *et al.*, 2004). Perhaps reduced

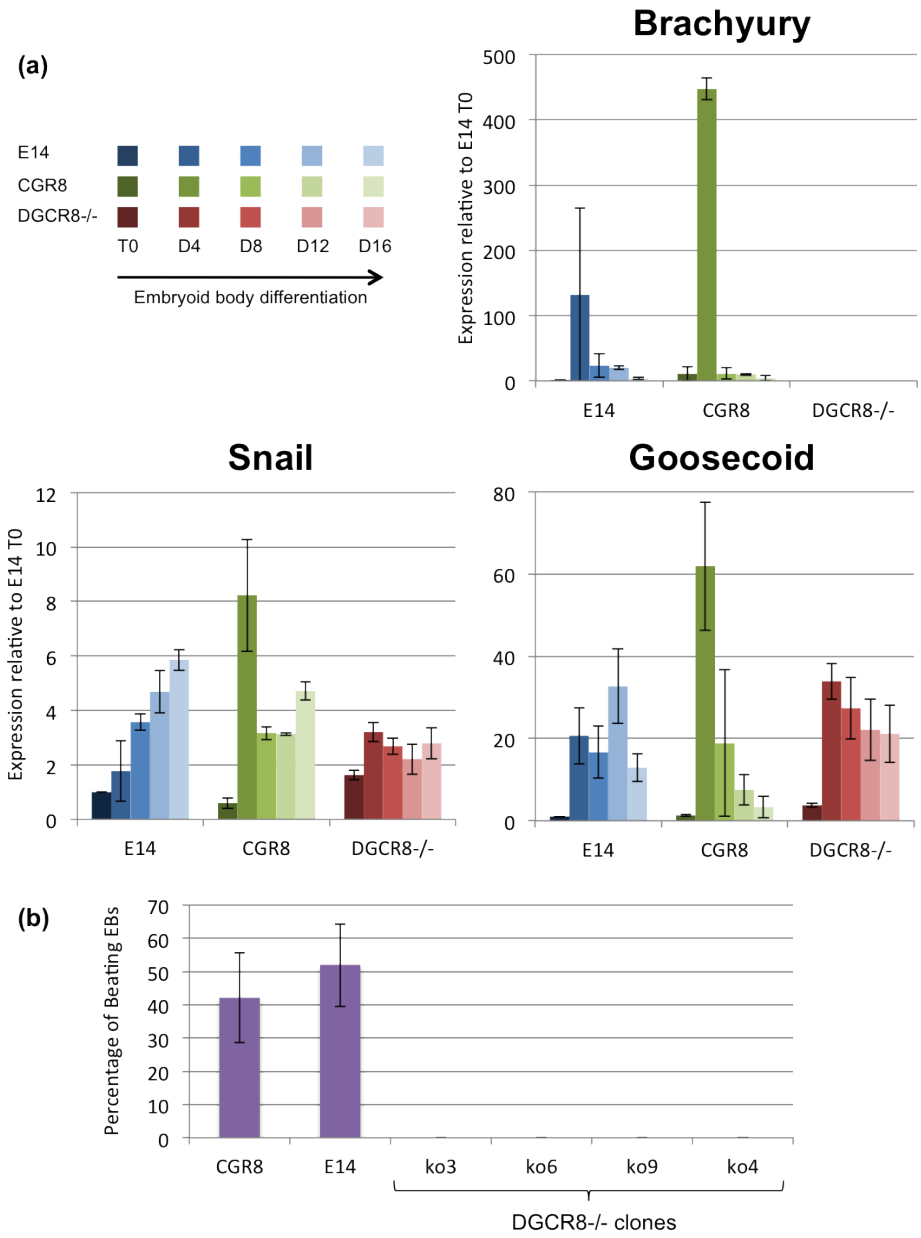


Figure 5.14 – *Expression of mesoderm marker genes during EB differentiation of wt ES cells and DGCR8 -/- ES cells.* (a) QPCR analysis for expression of the mesoderm markers *Brachyury*, Snail and Goosecoid during embryoid body differentiation of E14 ES cells (blue) CGR8 ES cells (green) and DGCR8$^{-/-}$ ES cells (subclones - red). QPCR readings were normalized to expression of the housekeeping gene, *β-Actin* and corrected values are shown relative to expression in E14 ES cells at T0. Error bars are SEM, E14 n=4, CGR8 n=3, DGCR8$^{-/-}$ n=5. (b) Bar graph showing percentage of embryoid bodies containing beating cardiomyocytes at day 12 of differentiation for E14 ES cells, CGR8 ES cells and 4 different clones of DGCR8$^{-/-}$ ES cells (subclones). Result is mean percentage from 6 different cultures containing >10 EBs per culture. Error bars show standard deviation.

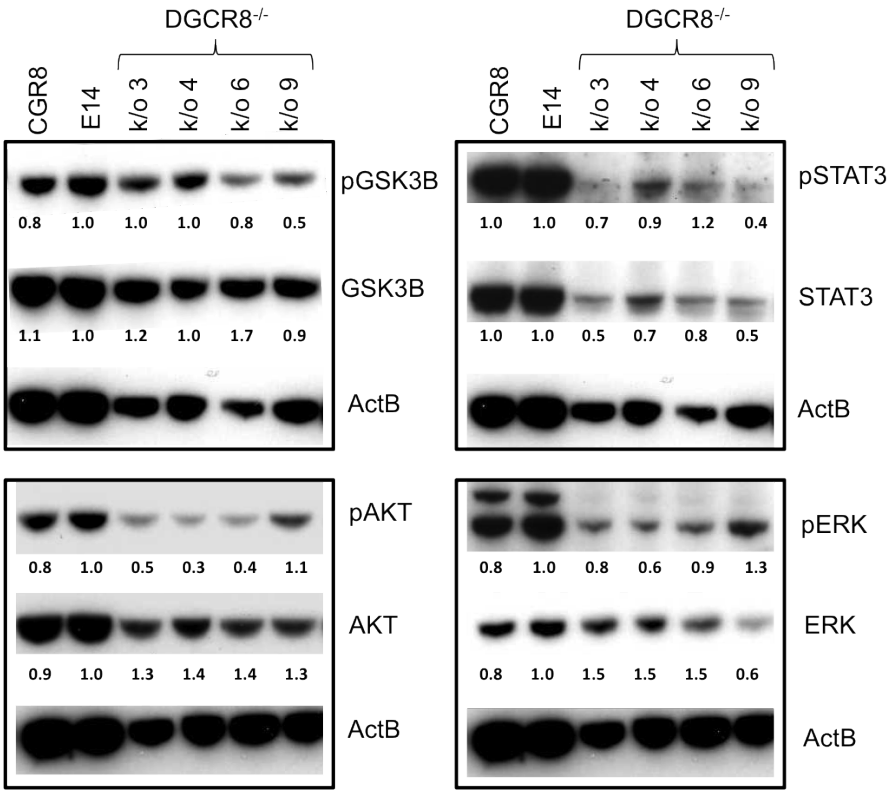


Figure 5.15 – *Expression of signaling pathway components in wt and* DGCR8$^{-/-}$ *EBs.* Western analysis for expression of signaling pathway components and β-ACTIN on protein from day 16 EBs generated from CGR8 ES cells, E14 ES cells and 4 different DGCR8$^{-/-}$ clones (subclones). Numbers show expression signal in each band relative to signal in E14 EBs after normalization to β-ACTIN expression.

activation of the AKT pathway partly explains the observation that DGCR8$^{-/-}$ EBs are smaller than wt EBs.

5.3.3.4 Analysis of expression of ES cell marker genes during embryoid body differentiation of wt and DGCR8$^{-/-}$ ES cells

In order to investigate the expression of pluripotency genes in differentiating wt and DGCR8$^{-/-}$ ES cells, EBs were made from wt ES cells and from DGCR8$^{-/-}$ ES cells and grown in suspension for 8 days, followed by adherent culture for a further 8 days. Cultures were harvested and RNA prepared from cells from the starting population of ES cells (T0) and from cells collected at days 4, 8, 12 and 16 of EB differentiation. RNA was used to generate cDNA for QPCR analysis of ES cell genes. *Figure 5.16a* shows that expression of *Oct4*, *Sox2* and *Nanog* is downregulated during EB differentiation of E14 ES and CGR8 ES cells. However, expression of *Oct4*, *Sox2* and *Nanog* is maintained during the differentiation timecourse in EBs from DGCR8$^{-/-}$ ES cells. Indeed, *Sox2* and *Nanog* expression increased relative to E14 T0 levels during EB differentiation of DGCR8$^{-/-}$ cultures. Protein was harvested from parallel cultures of day 16 EBs from CGR8, E14 and DGCR8$^{-/-}$ ES cells and used for western analysis of OCT4 expression. OCT4 protein was not detectable in day 16 EBs from either wt cell line but was readily detectable in day 16 EBs from 4 different DGCR8$^{-/-}$ ES cell clones (*Figure 5.16b*). This finding confirmed that OCT4 expression is not downregulated during DGCR8$^{-/-}$ EB differentiation.

This result suggested two possibilities: either differentiating DGCR8$^{-/-}$ ES cells co-express markers of differentiation and pluripotency, or the cells remain in an ES cell-like state and simply fail to differentiate. In order to distinguish between these possibilities IHC analysis was performed on cells from wt and DGCR8$^{-/-}$ EBs. EBs were formed from CGR8 ES cells, E14 ES cells and DGCR8$^{-/-}$ ES cells. At T0 and at day 2, 4, 6 and 8 of EB differentiation, cells were dissociated to a single-cell suspension, plated as a monolayer and fixed for IHC analysis of expression of the pluripotency marker, OCT4 and the differentiation marker, NESTIN. IHC analysis showed that a large proportion of DGCR8$^{-/-}$ cells were still expressing OCT4 protein after 8 days of EB differentiation, whereas most wt cells did not express OCT4 but did express the differentiation marker NESTIN (*Figure 5.17a*).

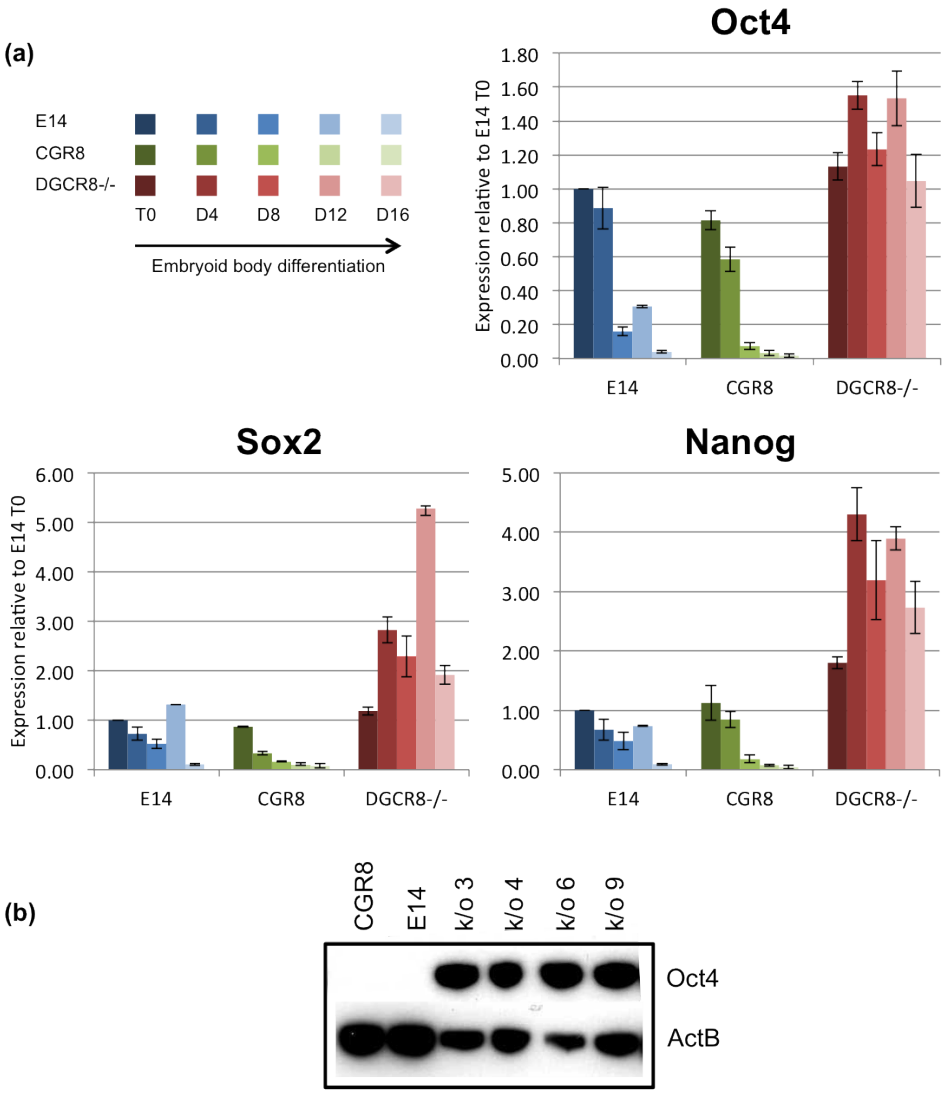


Figure 5.16 – *Expression of pluripotency marker genes during EB differentiation of wt and* DGCR8$^{-/-}$ *ES cells.* QPCR analysis for the pluripotency markers *Oct4*, *Sox2* and *Nanog* during embryoid body differentiation of E14 ES cells, CGR8 ES cells and DGCR8$^{-/-}$ ES cells (subclones). QPCR readings were normalized to expression of the housekeeping gene, *β-Actin* and corrected values are shown relative to expression in E14 ES cells at T0. Error bars are SEM, E14 n=4, CGR8 n=3, DGCR8$^{-/-}$n=5.

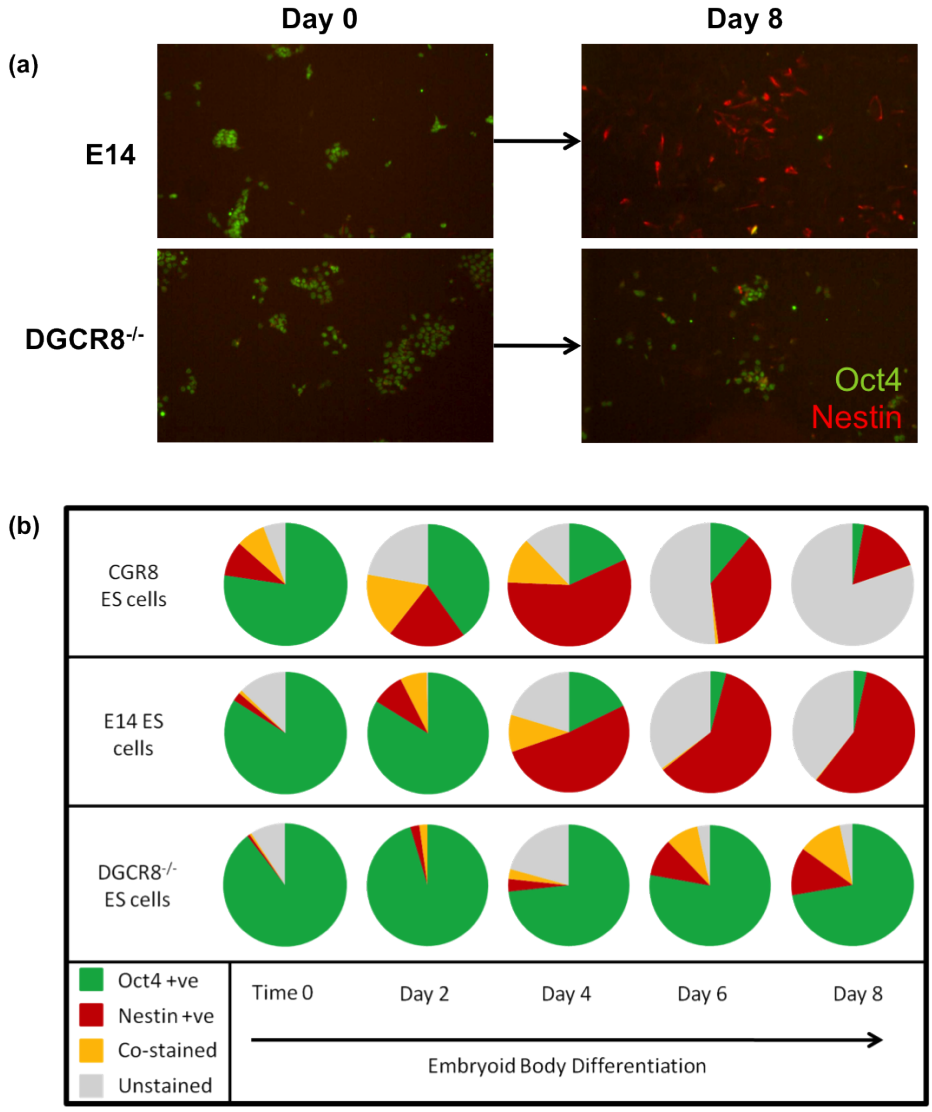


Figure 5.17 – *IHC analysis of OCT4 and NESTIN expression in differentiating EBs from wt and* DGCR8$^{-/-}$ *ES cells.* (a) Photographs of IHC for OCT4 (green) and NESTIN (red) in E14 ES cells and DGCR8$^{-/-}$ ES cells (subclones), and after 8 days of EB differentiation of each ES cell line. Magnification is x100. (b) Quantitiation of IHC analysis of OCT4 and NESTIN expression during embryoid body differentiation of E14 ES cells, CGR8 ES cells and DGCR8$^{-/-}$ ES cells (subclones). The number of OCT4 positive (green), NESTIN positive (red), co-stained (yellow) and unstained (grey) cells were counted at T0 and at 2, 4, 6 and 8 days of differentiation. Each pie chart segment represents the mean percentage of cells that stained for each marker. 3 photographs were counted per well of cells. N=2 experimental replicates.

IHC analysis of OCT4 and NESTIN expression was carried out on CGR8, E14 and DGCR8$^{-/-}$ cells at T0 and at day 2, 4, 6 and 8 of EB differentiation. Photographs of cells were taken and the numbers of DAPI positive, OCT4 positive, NESTIN positive, double positive and total cells were counted. *Figure 5.17b* shows pie charts of the proportion of the cell populations staining for different proteins through differentiation. In wt cultures, the proportion of OCT4 positive cells decreased during differentiation, and the proportion of NESTIN positive cells increased and then decreased (CGR8) or stabilised (E14). Additionally, CGR8 ES cells showed greater expression of NESTIN than E14 ES cells at earlier time points through differentiation. This may suggest that CGR8 ES cells cultures are pre-disposed to differentiate down the neural lineage because of a bias towards this lineage in the starting population of ES cells. In contrast to the findings from the wt EBs, a large proportion of cells from DGCR8$^{-/-}$ EBs remained OCT4 positive throughout the differentiation timecourse. Some DGCR8$^{-/-}$ cells became NESTIN positive showing that these cells are capable of differentiation although the proportion was lower than that of wt cultures. Cells co-stained for OCT4 and NESTIN expression were observed in wt and DGCR8$^{-/-}$ cultures. Importantly, the proportion of cells co-stained for OCT4 and NESTIN was not higher in DGCR8$^{-/-}$ cultures than wt cultures. In fact, the proportion of co-stained cells at day 6 and day 8 of DGCR8$^{-/-}$ differentiation was similar to that seen at day 2 and day 4 of wt differentiation. Together, these results show that maintenance of ES cell marker expression is not due to large proportions of DGCR8$^{-/-}$ ES cells co-expressing pluripotency and differentiation markers and that DGCR8$^{-/-}$ ES cells are capable of downregulating pluripotency markers and upregulating differentiation markers at the single cell level. However, even in differentiation inducing conditions a large proportion of DGCR8$^{-/-}$ ES cells remain undifferentiated, characterised by continued expression of pluripotency markers.

5.3.3.5 Expression of ES cell and differentiation marker genes in wt and DGCR8$^{-/-}$ ES cells cultured in different conditions

In order to investigate the reason behind impaired differentiation of the DGCR8$^{-/-}$ ES cells the ES cells themselves were assessed for expression of pluripotency and differentiation markers by QPCR after growth in different conditions for 24 hours. Following growth in standard ES cell media, wt and DGCR8$^{-/-}$ ES cell cultures

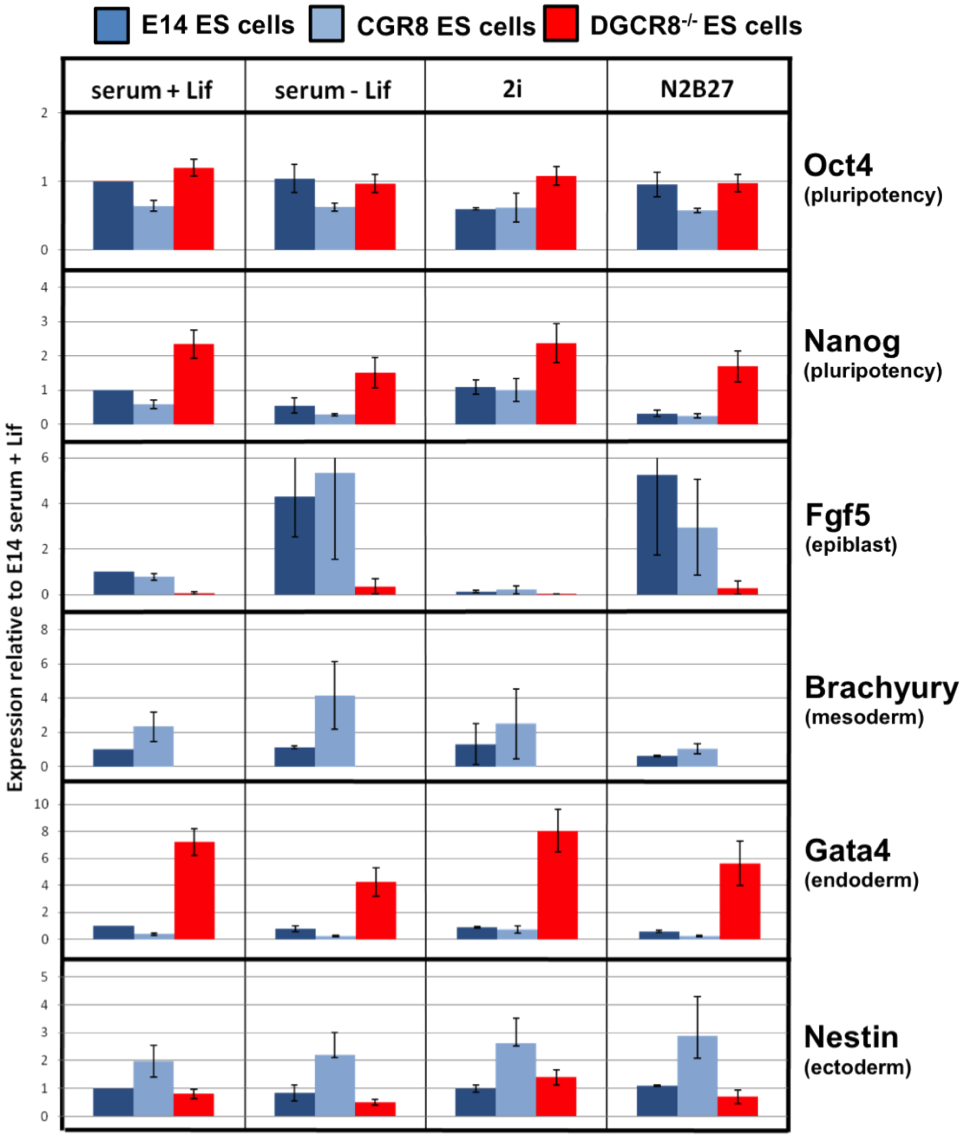


Figure 5.18 – *Expression of ES cell and lineage marker genes in wt and* DGCR8$^{-/-}$ *ES cells grown in different culture conditions.* QPCR analysis for the pluripotency markers *Oct4* and *Nanog*, the epiblast marker *Fgf5*, the endoderm marker *Gata4*, the mesoderm marker *Brachyury* (T) and the neuroectoderm marker *Nestin* in E14 ES cells (dark blue), CGR8 ES cells (light blue) and DGCR8$^{-/-}$ ES cells (subclones - red). Cells were cultured in normal ES cell media, washed twice with PBS and then grown for 24 hours in either normal ES cell media (serum + Lif), ES cell media without Lif (serum – Lif), N2B27 media supplemented with 1.0uM PD0325901 and 3.0uM CHIR99021 (2i) or N2B27 media alone (N2B27). QPCR readings were normalized to expression of the housekeeping gene, *β-Actin* and corrected values are shown relative to expression in E14 ES cells in serum + Lif. Error bars are SEM. In +lif: CGR8 n=3, E14 n=3, DGCR8$^{-/-}$ n=6, in 2i: CGR8 n=3, E14 n=3, DGCR8$^{-/-}$ n=5, in –lif and N2B27: CGR8 n=2, E14 n=2, DGCR8$^{-/-}$ n=3.

were washed twice in PBS and then one of four different medias was added to cultures. The cultures were either grown in standard ES cell media (serum + Lif) or in media with serum but without the Lif cytokine. Removal of Lif from the culture media normally induces differentiation of ES cells after several days. Cells were grown without Lif for 24 hours to see if early changes in marker gene expression could explain the defects in differentiation observed in DGCR8$^{-/-}$ ES cells. Cells were also grown in 2i media, which is supposed to maintain ES cells in a 'ground state', and in N2B27, which is the base media for 2i, and is reported to induce neural differentiation of ES cells. After 24 hours of growth in the different medias RNA was harvested from the cultures and used to generate cDNA for QPCR analysis of marker gene expression (*Figure 5.18*).

Oct4 expression was similar in CGR8 ES cells, E14 ES cells and DGCR8$^{-/-}$ ES cells in all four medias. This shows that even in conditions inducing differentiation (serum – Lif and N2B27) *Oct4* expression does not change after 24 hours. However, expression of *Nanog* decreased after 24 hours of growth in differentiation inducing conditions (serum – Lif and N2B27) compared to growth in conditions that maintained pluripotency (serum + Lif and 2i). The level of *Nanog* expression was consistently higher in DGCR8$^{-/-}$ ES cell cultures. Nevertheless, the decrease in *Nanog* expression upon transfer to differentiation inducing conditions was observed in CGR8, E14 and DGCR8$^{-/-}$ ES cell cultures. *Fgf5* is a marker of epiblast and expression increases on Lif withdrawal from ES cell cultures. As expected, *Fgf5* expression increased in wt ES cell cultures grown in serum – Lif and N2B27 media for 24 hours. Additionally, *Fgf5* expression decreased when cultures were transferred to growth in 2i media, consistent with the cells being in a 'ground state' and not expressing even early markers of differentiation. Interestingly, *Fgf5* expression was much lower in DGCR8 /- ES cells than in wt ES cells and expression did not increase significantly upon transfer to differentiation inducing conditions. In all conditions, the early mesoderm and epiblast marker *Brachyury* was undetectable in DGCR8$^{-/-}$ ES cells whereas it was expressed in wt ES cells. Expression of the endoderm marker *Gata4* remained constant in CGR8 and E14 ES cell cultures. Strikingly, expression of *Gata4* was significantly higher in DGCR8$^{-/-}$ ES cell cultures than both wt cultures in all conditions. This may suggest that *Gata4* expression is normally under miRNA control in ES cells or that DGCR8$^{-/-}$ ES cells contain some cells with an endodermal-like identity. Expression of the neural differentiation marker *Nestin* was similar between E14 ES

cells and DGCR8$^{-/-}$ ES cells in all four conditions. CGR8 ES cells showed consistently higher *Nestin* expression in all conditions, which may suggest that CGR8 ES cells cultures are biased towards differentiating down the neural lineage.

5.4 Discussion

The primary aim of this chapter was to examine the role of the general miRNA population in ES cell differentiation through comparison of DGCR8$^{-/-}$ ES cells and wt ES cells. When comparing DGCR8$^{-/-}$ ES cell cultures with wt ES cell cultures two differences were immediately apparent: a difference in morphology and a slower proliferation rate in the DGCR8$^{-/-}$ ES cells. In addition, it was shown that DGCR8$^{-/-}$ ES cell cultures express NANOG homogenously compared to heterogenous expression in wt ES cells, show deficiencies in silencing self-renewal when exposed to differentiation stimuli, and that markers of differentiation were expressed differently in DGCR8$^{-/-}$ ES cell and wt ES cells in standard culture conditions and during differentiation.

5.4.1 miRNA-deficient ES cells are morphologically different from wt ES cells

DGCR8$^{-/-}$ ES cells were transferred from growth on feeder cells (MEFs) to feeder-free conditions and it was observed that when grown in the absence of feeders DGCR8$^{-/-}$ ES colonies appeared much flatter than those of their wt counterparts and contained cells that were loosely packed together. It was surprising to find that both DGCR8$^{-/-}$ ES cells and Dicer$^{-/-}$ ES cells have been previously described as morphologically normal (Wang *et al.*, 2007, Kanellopoulou *et al.*, 2005). However, in these reports the cells were grown on a feeder layer rather than on gelatin-coated tissue culture plastic, which may account for the differences in morphology of the cells. Indeed, in this chapter it is observed that, when grown on a feeder layer, DGCR8$^{-/-}$ ES cells appeared similar in morphology to wt ES cells with tightly packed, domed colonies. It is probable that the physical constraints of feeder-associated culture, where ES cell colonies must grow between feeder cells, provides the simple explanation for this. In these conditions

DGCR8$^{-/-}$ ES cells may be forced to grow in close proximity to one another giving the appearance of tightly packed domed colonies. An alternative explanation is that the feeder layer is producing a factor required for proper cell-cell adhesion of ES cells in the absence of miRNAs.

It was hypothesised that the morphological difference observed between DGCR8$^{-/-}$ and wt ES cells, grown in the absence of feeders, may be due to aberrant expression of E-CADHERIN because E-CADHERIN null ES cells lose cell-cell contacts and grow more dispersed that wt ES cells, similar to the phenotype of DGCR8$^{-/-}$ ES cells (Larue *et al.*, 1996). Although E-CADHERIN had a normal expression pattern in the DGCR8$^{-/-}$ ES cells its expression level appeared higher in DGCR8$^{-/-}$ ES cell cultures than in wt ES cell cultures. This finding was difficult to rationalise in the context of a reduction of cell-cell contact in DGCR8$^{-/-}$ ES cells and could be explained by two possibilities: Differences in morphology between wt ES cells and DGCR8$^{-/-}$ ES cells may be explained by elevated expression of E-CADHERIN by an as yet unknown mechanism; or the morphological differences are due to defects in another, uncharacterised mechanism of cell adhesion in ES cells.

As well as differences in morphology, there was a clear difference in proliferation rate between DGCR8$^{-/-}$ ES cells and wt ES cells, with DGCR8$^{-/-}$ ES cells growing at a slower rate. This has been previously observed in both DGCR8$^{-/-}$ ES cells (Wang *et al.*, 2007) and Dicer$^{-/-}$ ES cells (Kanellopoulou *et al.*, 2007, Murchison *et al.*, 2005) showing a clear role for miRNAs in maintaining the rapid proliferation of ES cells. The reduction in proliferation rate in DGCR8$^{-/-}$ ES cells is associated with an accumulation of the cells in the G1 phase of the cell cycle (Wang *et al.*, 2007). Introduction of miRNAs belonging to the *miR-290* family was shown to rescue the slow proliferation rate of DGCR8$^{-/-}$ ES cells by suppressing key regulators of the G1/S transition (Wang *et al.*, 2008). In contrast to the finding that *miR-92a* is a regulator of ES cell proliferation (Chapter 3, this thesis), Wang *et al.* did not observe rescue of the proliferation defect upon introduction of this miRNA into DGCR8$^{-/-}$ ES. However, this experiment was carried out as part of a large scale screening strategy of 266 different miRNAs so perhaps subtle affects on cell proliferation of individual miRNAs may have been missed.

5.4.2 miRNAs are not required for maintenance of ES cell pluripotency

QPCR analysis of expression of the ES cell marker genes *Oct4*, *Sox2*, *Rex1* and *Stella* showed similar expression levels in DGCR8$^{-/-}$ ES cell and wt ES cell cultures suggesting that miRNAs are not required for maintenance of pluripotency gene expression. These data are supported by previous findings: *Oct4* has previously been reported to be expressed at similar levels in Dicer$^{-/-}$ ES cells compared to wt ES cells (Kanellopoulou *et al.*, 2005) and in DGCR8$^{-/-}$ ES cells compared to wt ES cells (Wang *et al.*, 2007). Additionally, the expression of *Sox2*, *Rex1* and *Nanog* has been demonstrated in DGCR8$^{-/-}$ ES cells although the analysis was non-quantitative (Wang *et al.*, 2007). The finding that the pluripotency gene *Oct4* shows normal expression in the absence of miRNAs also agrees with *in vivo* findings. Analysis of *Oct4* expression in DGCR8$^{-/-}$ blastocyst stage embryos showed normal expression in cells of the ICM (Suh *et al.*, 2010).

Interestingly, while the expression levels of the majority of key pluripotency genes were similar in DGCR8$^{-/-}$ ES cells and wt ES cells, *Nanog* was expressed at approximately twice the level in the DGCR8$^{-/-}$ ES cell cultures as in wt ES cell cultures. Immunohistochemistry for OCT4 and NANOG showed robust expression of both proteins in individual DGCR8$^{-/-}$ ES cells and wt ES cells. NANOG is normally expressed heterogeneously in ES cells but IHC analysis revealed that NANOG expression is homogeneous in DGCR8$^{-/-}$ ES cells than wt ES cells. These findings may be due to the absence of specific miRNAs, as several miRNAs are known to target *Nanog*: *miR-134* (Tay *et al.*, 2008b), *miR-296* and *miR-470* (Tay *et al.*, 2008a) in DGCR8$^{-/-}$ ES cells. Perhaps direct regulation by these miRNAs normally regulates the cyclical expression of NANOG that is observed in wt cells. On the other hand, homogeneous NANOG expression may be an indirect result of cellular change imposed by the loss of the miRNA population. This result has not been reported in any other analyses of miRNA-deficient ES cells although the presence of *Nanog* expression has been previously demonstrated in DGCR8$^{-/-}$ ES cells (Wang *et al.*, 2007).

In addition to expression of key ES cell pluripotency markers DGCR8$^{-/-}$ ES cells were found to activate similar signalling pathways as wt ES cells. These data support the proposition that DGCR8$^{-/-}$ ES cells retain key features of ES cells. Furthermore, DGCR8$^{-/-}$ ES cell colonies can be grown from a single cell (albeit at

much reduced frequency), which is a property of ES cells. These cells obviously provide a good model system for the analysis of the role of miRNAs in ES cell differentiation and developmental transitions.

These results suggest that, at least superficially, miRNAs are not required to maintain ES cell pluripotency per se. This conclusion comes from several observations: (i) the expression levels of key ES cell genes are similar between DGCR8$^{-/-}$ and wt ES cells; (ii) DGCR8$^{-/-}$ ES cells robustly express OCT4 and NANOG protein; (iii) DGCR8$^{-/-}$ ES cells activate similar signalling pathways to wt ES cells; and (iv) DGCR8$^{-/-}$ ES cells can be propagated clonally.

The likelihood that miRNAs are not required for the pluripotent state is interesting in the context of the 'ground state' hypothesis. Austin Smith's laboratory propose that ES cells exist in a ground state and can maintain a pluripotent state indefinitely if shielded from the appropriate pro-differentiative stimuli (Ying *et al.*, 2008). In the ground state hypothesis, ES cells are intrinsically self-maintaining. DGCR8$^{-/-}$ ES cells have lost expression of almost all mature miRNAs and yet they continue to proliferate and maintain expression of key pluripotency genes. The fact that ES cells devoid of miRNA regulation still behave largely like their wt counterparts suggests that complex genetic regulation is not required for the ES cell state and may support the hypothesis that ES cells represent an intrinsically self-maintaining state.

5.4.3 miRNAs are required for silencing self-renewal

The original description of DGCR8$^{-/-}$ ES cells pointed to an inability of these cells to downregulate pluripotency markers during EB differentiation. QPCR analysis of expression of pluripotency genes during embryoid body differentiation confirmed that, in miRNA-deficient ES cells, expression does not decrease during differentiation. This finding is supported by previous reports of an inability to downregulate pluripotency markers during EB differentiation of DGCR8$^{-/-}$ ES cells, DICER$^{-/-}$ ES cells and AGO2$^{-/-}$ ES cells (Wang *et al.*, 2007, Kanellopoulou *et al.*, 2005, Chandra Shekar *et al.*, 2011). A surprising finding was that the expression of the pluripotency markers *Sox2* and *Nanog* increased during EB differentiation of DGCR8$^{-/-}$ ES cells. If the process of exiting self-renewal is

blocked, as it appears to be in DGCR8$^{-/-}$ ES cells, perhaps the EB microenviroment is permissive or even preferable for the maintenance of pluripotency marker gene expression.

IHC analysis of individual cells from EBs showed that a failure to downregulate pluripotency gene expression during differentiation is not due to co-expression of differentiation and pluripotency markers in individual cells, but is due to failure of individual cells to escape pluripotency. This implies that the cells are not silencing self-renewal factors and remain in a pluripotent state for longer even in the absence of self-renewal stimuli. The conclusion that miRNAs are required for silencing self-renewal is supported by data from Wang *et al.*, which shows that, when plated at clonal density, DGCR8$^{-/-}$ ES cells form more alkaline positive colonies that wt ES cells both in the presence and absence of Lif (Wang *et al.*, 2007). Wang *et al.* reported that in the absence of Lif, almost 100% of DGCR8$^{-/-}$ colonies were alkaline positive compared to $<$20% of wt colonies suggesting, as for the data presented in this chapter, that individual DGCR8$^{-/-}$ cells were failing to silence self-renewal. The inability to silence self-renewal genes observed in DGCR8$^{-/-}$ ES cells also occurs in Dicer$^{-/-}$ ES cells (Kanellopolou *et al.*, 2005) and is thought to be at least partly due to an inability of the cells to methylate the *Oct4* promoter on differentiation due to upregulation of Rbl2, a repressor of de novo DNA methyltransferases (Dnmts) (Sinkkonen *et al.*, 2008).

However, other data presented in this chapter shows that it is functionally possible for DGCR8$^{-/-}$ ES cells to downregulate ES cell markers. Firstly, IHC analysis showed that not all miRNA-deficient ES cells were OCT4 positive after 8 days of embryoid body differentiation, suggesting that OCT4 expression had been silenced in these cells. Secondly, siRNA knockdown of *Oct4* in DGCR8$^{-/-}$ ES cells led to downregulation of not only *Oct4*, but also *Sox2* and *Nanog* by approximately 50%. In the event of forced downregulation of a key ES cell pluripotency gene it appears that expression of other key pluripotency genes can also be destabilised, even in the absence of miRNAs. This is perhaps unsurprising, given that the key ES cell genes are known to function as part of a regulatory network with feed-forward and autoregulatory loops (Loh *et al.*, 2006). However, it may also suggest that one function of miRNAs during differentiation is to destabilise the network of key ES cell regulatory genes, and so initiate their collective downregulation. While it has been previously published that miRNA-deficient ES cells fail to downregulate pluripotency markers during EB differentiation (Wang *et al.*,

2007, Kanellopoulou *et al.*, 2005, Chandra Shekar *et al.*, 2011), and retinoic acid induced differentiation (Wang *et al.*, 2007), the ability of miRNA-deficient ES cells to downregulate pluripotency genes upon knockdown of key pluripotency genes has not been reported. Because of this, and because of the report that miRNA-deficient ES cells cannot methylate the *Oct4* promoter (Sinkkonen *et al.*, 2008), it would be interesting to investigate whether the repression of pluripotency genes is stable following siRNA knockdown of *Oct4*.

Overall, it can be concluded that, compared to wt ES cells, pluripotency genes are not downregulated upon undirected differentiation of miRNA-deficient ES cells, and that this is due to an inability of individual cells to silence self-renewal. However, it is functionally possible for individual cells to downregulate expression of pluripotency markers and general downregulation of pluripotency marker gene expression occurs upon destabilisation of the pluripotency network. Although further work is needed to confirm that downregulation of pluripotency genes is stable in miRNA-deficient ES cells, these data suggest the conclusion that miRNA-deficient ES cells are incapable of silencing self-renewal is premature.

5.4.4 miRNAs are required for proper ES cell differentiation

The differentiation capacity of DGCR8$^{-/-}$ ES cell was examined in comparison to wt ES cells by three differentiation methods: trophectoderm differentiation induced by siRNA k/d of *Oct4*; neural differentiation induced by a serum-free culture method; and embryoid body differentiation. In both the trophectoderm and neural differentiation experiments, DGCR8$^{-/-}$ ES cells differentiated into the expected lineage as defined by analysis of morphology and marker gene expression. This may suggest that miRNAs are not required for the initiation of trophectoderm or neuroectoderm differentiation and/or that miRNA-deficient ES cells are capable of differentiating down particular lineages if directed differentiation methods are used. siRNA knockdown of *Oct4* and serum-free neural differentiation 'force' undifferentiated cells into a particular cell type so it is perhaps unsurprising that trophectoderm and neurons can be generated from miRNA-deficient ES cells. It has previously been shown that DGCR8$^{-/-}$ embryos can initiate trophectoderm differentiation *in vivo*, forming blastocysts with normal distributions of *Cdx2* positive TE and OCT4 positive ICM cells (Suh *et al.*, 2010). Later in development, however, Ago2-null mice have hypertrophic placentas with a marked

reduction in the thickness of the labyrinthine layer (Cheloufi *et al.*, 2010). Additionally, Dicer ablation in the mouse neocortex leads to defects in neural survival and differentiation (De Pietri Tonelli *et al.*, 2008). This shows that although it may be possible for some level of trophectoderm and neural differentiation to occur from DGCR8$^{-/-}$ ES cells, it is doubtful that these cells are functionally equivalent to cells generated from ES cells with miRNAs. Substantiating this, Spruce *et al.* found that Dicer null embryos could not maintain the TS cell compartment although differentiated trophoblast cell types were observed (Spruce *et al.*, 2010).

Embryoid body differentiation is largely undirected and should result in the differentiation of cells into the three germ lineages. It was found that DGCR8$^{-/-}$ ES cells expressed normal levels of the trophectoderm gene *Cdx2*, but induction of neuroectoderm, endoderm and mesoderm marker gene expression was either absent, delayed or markedly reduced in the DGCR8$^{-/-}$ embryoid bodies. Analysis of EB differentiation in miRNA-deficient ES cell has been reported previously and supports the data presented here: Dicer$^{-/-}$ EBs are reported to show no expression of the endoderm marker *Hnf4* or the mesodermal markers *Brachyury*, *Bmp4* and *Gata1* during differentiation (Kanellopoulou *et al.*, 2005); and Wang *et al.* reported that EBs from DGCR8$^{-/-}$ ES cells exhibited reduced or delayed expression of the endoderm marker *Hnf4*, the mesoderm marker *Brachyury* and the neuroectoderm marker *Sox1* (Wang *et al.*, 2007). Although some evidence of neural differentiation was observed in DGCR8$^{-/-}$ embryoid bodies, they did not form cysts, possibly suggesting a potential defect in primitive endoderm formation, and never formed beating bodies, illustrating an obvious defect in cardiomyocyte differentiation. These data showed that embryoid bodies do not undergo proper germ layer differentiation in the absence of miRNAs. This conclusion is supported by the findings of Spruce *et al.* who examined gene expression during early development of Dicer$^{-/-}$ embryos. They found that, in the absence of miRNAs, initiation of gastrulation and mesoderm formation did occur, but was delayed with respect to wt embryos and that there was a failure to elongate the primitive streak (Spruce *et al.*, 2010). Therefore, although some aspects of normal development are observed in Dicer$^{-/-}$ embryos, progression of patterning was abnormal.

Induction of expression of the early differentiation marker, *Fgf5*, was delayed in EBs from DGCR8$^{-/-}$ ES cells and expression was maintained throughout differen-

tiation, rather than being downregulated as it was in EBs from wt ES cells. This suggested that the observed defects in DGCR8-/- EB differentiation might be due to defects in early differentiation. In order to investigate the reason behind impaired differentiation of DGCR8$^{-/-}$ ES cells, the expression of pluripotency and differentiation markers was assessed in these cells after 24 hours of growth in conditions supporting pluripotency or inducing differentiation. *Fgf5* expression did not increase in DGCR8$^{-/-}$ ES cells immediately after Lif withdrawal as it did in wt ES cells. These data suggest that the differentiation deficiencies seen in DGCR8$^{-/-}$ ES cells may originate from a defective *Fgf5* response to the removal of self-renewal stimuli.

Interestingly, *Fgf5* expression was markedly lower in DGCR8$^{-/-}$ ES cells relative to wt ES cells. This supports the *Nanog* data in suggesting that miRNA-deficient ES cells may be in a more naïve state that wt ES cells. *Fgf5* is an early marker of primitive ectoderm differentiation, and its expression increases in wt ES cells as a result of Lif withdrawal (Martinez-Ceballos *et al.*, 2005). A low level of *Fgf5* expression may indicate that DGCR8$^{-/-}$ ES cells contain a smaller proportion of cells that are primed to differentiate. Furthermore, while *Fgf5* expression markedly decreased in wt ES cells upon transfer to 2i media, it did not decrease in DGCR8$^{-/-}$ ES cells. It is thought that ES cells grown in 2i media are in a more naïve state than those grown in serum + Lif conditions and the finding that wt expression of *Fgf5* decreases upon transfer from serum + Lif to 2i conditions supports this. The finding that *Fgf5* expression did not decrease upon transfer of DGCR8$^{-/-}$ ES cells from growth in serum + Lif to 2i conditions supports the proposition that these cells already exist in a more naïve state than wt ES cells.

Another gene expressed in the epiblast, called *Brachyury*, which also marks mesoderm, is easily detected in wt ES cells by QPCR, but is undetectable in DGCR8$^{-/-}$ ES cells. This may support the hypothesis that DGCR8$^{-/-}$ ES cells are in a more primitive state that wt ES cells, or it may point simply to a potential deficiency in mesoderm differentiation in these cells. However, while the neuroectoderm marker *Nestin* is expressed at similar levels in DGCR8$^{-/-}$ ES cells and wt ES cells, the endoderm marker *Gata4* is found at between 5 and 8 fold higher levels in DGCR8$^{-/-}$ ES cells. This may point to the DGCR8$^{-/-}$ ES cells being in a state that is predisposed to differentiate into endoderm, or it may suggest that normally *Gata4* is regulated by expression of a miRNA or miRNAs in ES cells. Our data point to the latter option because another marker of endoderm, *Gata6*, does

not show increased expression in DGCR8$^{-/-}$ ES cell relative to wt ES cells (from EB differentiation data). Additionally, the expression of *Gata4* in ES cells is not indicative of a state that is observed *in vivo*: GATA4 is not normally expressed in the embryo until the mid-blastocyst stage and then rarely co-stains with cells positive for the pluripotency marker, NANOG (Plusa *et al.*, 2008).

Together, these data suggest that the differentiation deficiencies observed in DGCR8$^{-/-}$ ES cells could be due to two possibilities: either the DGCR8$^{-/-}$ ES cell culture is in a naïve state that cannot respond appropriately to differentiation stimuli; or the culture exists in a state 'primed' for endoderm differentiation, which compromises its differentiation capability. For the following reasons it is believed that the former of these two options is more likely to be true: (i) expression of NANOG is elevated and homogeneous in DGCR8$^{-/-}$ cultures compared to wt cultures, reminiscent of 'naïve' wt ES cell cultures; (ii) expression of the primitive ectoderm marker *Fgf5* is markedly reduced in DGCR8$^{-/-}$ cultures compared to wt cultures and does not increase upon removal of Lif from the culture media; (iii) DGCR8$^{-/-}$ cultures show increased expression of only one differentiation marker gene, and another marker gene of the same lineage is not elevated in expression in DGCR8$^{-/-}$ cultures; and (iv) differentiation of DGCR8$^{-/-}$ cultures is not biased to one particular lineage.

5.4.5 Summary and Future Research

The data in this chapter suggest that miRNAs are not required for the maintenance of the ES cell state, but that they are required for efficient exit from self-renewal. This may be due to an inability of miRNA-deficient ES cell to destabilise the pluripotency network by silencing key pluripotency genes, particularly *Nanog*, during EB differentiation. Further investigation into the reason for homogeneous expression of NANOG in miRNA-deficient ES cells, and experiments forcing its downregulation prior to differentiation protocols, would help to determine whether elevated *Nanog* expression is the cause or an effect of the inability of miRNA-deficient ES cells to silence self-renewal.

These data also show that miRNAs are required for proper differentiation of lineages derived from all three germ layers during embryoid body differentiation, but that ES cells deficient for miRNAs can be induced to initiate trophectoderm

and neuroectoderm differentiation by stringent differentiation protocols. With the finding that differentiation into mesoderm and endoderm lineages was particularly compromised during embryoid body differentiation, it would be interesting to investigate if these lineages could be generated from miRNA-deficient ES cells by directed differentiation methods. The data presented here suggest that miRNA-deficient ES cells exist in a naïve state that is unable to respond to early differentiation stimuli. Further investigation into this possibility may provide insight into the differentiation defects observed in these cells.

Chapter 6

General Discussion

6.1 Introduction

The process of embryonic development, whereby an organism develops from a single cell to a complex form consisting of highly specialised structures, tissues and cell types, is truly remarkable and far from being understood. In order for correct development to proceed, tight temporal and spatial regulation of gene expression is essential. Investigation of the factors that control developmental transitions, including the regulation of gene expression, will not only provide insight into the processes of normal and abnormal embryonic developmental, but may also lead to the establishment of an unlimited source of specialised cell types and organs for therapeutic benefit.

Because the process of embryonic development is so complex, it is likely that gene expression is controlled at multiple levels to ensure that it remains tightly controlled throughout development, even if errors arise at one level of genetic regulation. Some means of regulating gene expression, such as control of transcription, translation and protein modification, have long been established. Recently, an entirely novel level of regulation has been described – post-transcriptional regulation of gene expression by microRNAs (miRNAs). miRNAs were initially discovered as regulators of development in nematode worms, and have since been shown to be essential for embryonic developmental of numerous organisms. Therefore, the primary aim of this thesis was to investigate the roles of miRNAs in regulating

mammalian developmental transitions. This final discussion will detail how the findings presented here, along with the findings of others, support a model where miRNAs primarily function to ensure canalisation during development.

6.2 Two possible roles for miRNAs in development

As post-transcriptional regulators of gene expression, miRNAs are likely to play roles in most, if not all, developmental transitions. Two of the possible ways in which miRNAs may function during development are as master regulators of developmental transitions, and to ensure canalisation during development. The term 'canalisation' was first coined by C. H. Waddington to describe a process "as to bring about one definite end-result regardless of minor variations in conditions during the course of the reaction" (Waddington, 1942). In other words, the process of canalisation ensures robust outcomes from developmental transitions in the face of perturbations arising from environmental changes and/or differences in gene expression.

While some miRNAs act as master regulators of developmental transitions, for example, *lin-4* regulation of progression from *C. elegans* larval stage 1 to larval stage 2 (Lee *et al.* 1993), it is believed that most miRNAs function during mammalian development to fine-tune developmental transitions. The general miRNA population is clearly essential for proper embryonic development in numerous species, including mice (Wienholds *et al.*, 2003, Reinhart *et al.*, 2000, Bernstein *et al.*, 2003, Brennecke *et al.*, 2003). However, mice deficient for single miRNAs or families of miRNAs often do not show early embryonic lethality, and several are viable into adulthood (Medeiros *et al.*, 2011, Ventura *et al.*, 2008, van Rooij *et al.*, 2007, Zhao *et al.*, 2007, Rodriguez *et al.*, 2007, Thai *et al.*, 2007, Boldin *et al.*, 2011). Although deletion of individual miRNAs or miRNA families in mice can lead to specific developmental abnormalities, the mice are often viable and show relatively normal phenotypes. This suggests that these individual miRNAs do not function as master regulators of development, but function to fine-tune and therefore ensure canalisation during developmental transitions.

6.3 miRNAs can be functionally associated with individual developmental transitions

Results from the third chapter of this thesis show that miRNAs are differentially expressed in a cell-based model of a developmental transition. Some of the differentially expressed miRNAs were expressed at higher levels in undifferentiated ES cells, and others were expressed at higher levels in cells that had been induced to differentiate down the trophectoderm lineage. It is possible that differentially expressed miRNAs could be functioning as master regulators of the ES cell: trophectoderm transition, or to ensure canalisation during this process. Analysis of predicted targets for *miR-92a* and *Sfmbt2* miRNAs did not show interactions between these differentially expressed miRNAs and key regulators of the ES cell: trophoblast transition such as *Oct4* and *Cdx2*. Based on target predictions, this suggests that these differentially expressed miRNAs may have roles in 'finessing' the ES: trophoblast transition, rather than directly regulating it. Furthermore, evidence from *in vivo* studies does not support miRNAs as master regulators initiating trophectoderm specification. It has been reported that mice deficient for the miRNA processing enzymes DICER or DGCR8 arrest around E7.5 (Bernstein *et al.*, 2003, Wang *et al.*, 2007), and although *Dicer*$^{-/-}$ embryos appear abnormal at this stage, embryonic and abembryonic regions are recognisable, suggesting that some level of trophectoderm differentiation has occurred (Bernstein *et al.*, 2003). Further, an analysis of *Dgcr8*$^{-/-}$ blastocyst stage embryos reported that they were morphologically normal, with similar numbers of cells and similar distributions of OCT4-positive ICM cells and CDX2-positive trophectoderm cells (Suh *et al.*, 2010). In contrast to some other analyses, this observation was made on embryos deficient for maternal and zygotic DGCR8, which means the embryos will exhibit a complete miRNA deficiency (Suh *et al.*, 2010). The Suh *et al.* study shows that loss of the general miRNA population does not result in loss of the trophectoderm lineage, which suggests that this transition is not dependent on the presence of miRNAs.

miR-92a showed higher expression in undifferentiated cells relative to differentiated cells, and was functionally associated with regulating the proliferation of ES cells. Inhibition of *miR-92a* expression resulted in a reduction in ES cell numbers and, along with data from other laboratories, suggested that *miR-92a* is involved in cell cycle regulation in ES cells. In contrast, suppression of *miR-*

92a expression had no effect on the expression of pluripotency marker genes or genes associated with trophectoderm differentiation showing that this miRNA is not critical for trophectoderm differentiation. It is believed that expression of *miR-92a* may be downregulated during ES cell differentiation to facilitate the transition between the different proliferation rates of ES cells and trophectoderm, a conclusion consistent with *miR-92a* functioning to ensure canalisation during this developmental transition. Suppression of *miR-92a* expression had no effect on 3T3 cell numbers, suggesting that it regulates the cell cycle of ES cells, but not all cell types. This may be explained by the fact that ES cells have an unusual cell cycle, which lacks the G1 phase (Savatier *et al.*, 1994), and illustrates that some miRNA functions are cell-type specific.

6.4 Intronic miRNAs may show cooperative or antagonistic functions to those of their host gene

Several of the miRNAs that were expressed at higher levels in differentiated cells were encoded in a miRNA cluster within intron 10 of the *Sfmbt2* gene. *Sfmbt2* is interesting in the context of early development because it shows continued expression in the extraembryonic ectoderm of E5.5-E8.0 mouse embryos, which is derived from the trophectoderm and contains the trophoblast stem cell population (Frankenberg *et al.*, 2008). Furthermore, *Sfmbt2* is expressed from only the paternal allele in extraembryonic lineages, but expression is biallelic in embryonic lineages, which means that *Sfmbt2* is an imprinted gene and increases the likelihood of roles for *Sfmbt2*, and its encoded miRNA cluster, in placental development (Kuzmin *et al.*, 2007). Several other differentially expressed miRNAs were also encoded within introns of protein-coding genes, and most were found to share similar expression profiles, a finding that is supported by previous reports (Baskerville and Bartel, 2005, Lutter *et al.*, 2010, Najafi-Shoushtari *et al.*, 2010). It is interesting to consider the possibility that intronic miRNAs may show some functional cooperativity with their host genes. Indeed, it has been reported that host genes show correlated expression patterns with encoded miRNA target genes, which have synergistic or antagonistic functions to that of the host gene

(Lutter *et al.*, 2010). For example, the sterol regulatory element–binding proteins (SREBPs) are key transcription regulators of genes involved in cholesterol biosynthesis and uptake, and the miRNA, *miR-33*, is encoded within an intron of SREBP. Investigation of the function of *miR-33* revealed that it directly targets adenosine triphosphate – binding cassette A1 (ABCA1), which is a key cholesterol transporter and controls intracellular cholesterol efflux. The authors propose that through targeting ABCA1, *miR-33* functionally cooperates with SREBP to control intracellular cholesterol levels (Najafi-Shoushtari *et al.*, 2010). In this sense, intronic miRNAs may be functioning to ensure robustness of events associated with host gene function, although functional cooperativity of host gene and encoded miRNA is not incompatible with the idea of miRNAs functioning as master regulators. As well as functional cooperativity, there are also examples of miRNAs that appear to exert antagonistic functions to that of their host gene. C-MYC is a transcription factor that regulates cell proliferation, growth and apoptosis, and directly regulates expression of the transcription factor E2F1, which promotes cell proliferation. It also directly regulates expression of the *miR-17-92* cluster of miRNAs, and it has been reported that two of the miRNA encoded within this cluster, *miR-17* and *miR-20a*, directly regulate expression of E2F1 (O'Donnell *et al.*, 2005). This reveals a mechanism whereby C-MYC directly activates expression of E2F1, and indirectly limits its translation, which is thought to allow tight control of this proliferative signal. This is an example of miRNAs functioning to maintain expression of key factors within particular limits, which supports a model of miRNAs functioning to ensure canalisation.

6.5 Related miRNAs may have different functions

Some miRNA belong to miRNA families that share the same seed sequence and, because the seed sequence is thought to be the most important region for target recognition, these miRNAs are often predicted to regulate the same targets (Lewis *et al.*, 2003). However, data presented here suggest that related miRNAs do not always share the same function. Inhibition of *miR-92a* expression, but not the expression of the related miRNAs, *miR-25* and *miR-363*, resulted in a decrease in ES cell numbers. These three miRNAs all have an identical seed sequence so these data suggest that sequences outside of the seed region must be important for miRNA function. Indeed, there is mounting evidence that 'non-seed' sites may

mediate miRNA-directed regulation. Computer predictions that do not rely on seed complementarity or 3'UTR bias have predicted functional binding sites for miRNAs in the coding regions of the pluripotency factors *Oct4*, *Sox2* and *Nanog*, most of which do not show perfect seed region binding (Miranda *et al.*, 2006, Tay *et al.*, 2008a). This illustrates that there is still much that we do not understand about the criteria for a functional miRNA target site, and suggests the possibility that miRNAs with identical seed regions may have different functions.

6.6 Do miRNAs function as master regulators during the switch from self-renewal to differentiation?

The findings that the expression of key ES cell genes are similar in wt and DGCR8$^{-/-}$ ES cells, but that ES cell marker genes are not downregulated during EB differentiation, and expression of lineage markers is absent, reduced or delayed during differentiation, suggests that while miRNAs may not be required for the maintenance of the ES cell state, they are required for both the efficient exit of ES cells from self-renewal, and for proper ES cell differentiation.

Maintenance of the ES cell state, as assessed by expression of key ES cell genes and activation of signalling pathways, appears to be largely independent of miRNA regulation. In fact, expression of *Nanog* was elevated in miRNA-deficient ES cells suggesting that these cells may be in a more naïve state that wt ES cells. However, the observation that the proliferation rate of DGCR8$^{-/-}$ cells was slower than that of wt ES cells may mean that, while the general miRNA population is not required for maintenance of the ES cell state *per se*, miRNAs are involved in fine-tuning of the ES cell state. Findings from other laboratories support these conclusions (Murchison *et al.*, 2005, Kanellopoulou *et al.*, 2005, Wang *et al.*, 2007, Shekar *et al.*, 2011).

The data presented here, together with reports from other laboratories, suggest that miRNAs are required for exit of ES cells from self-renewal (Kanellopoulou *et al.*, 2005, Wang *et al.*, 2007). This raises the possibility that miRNAs may function as master regulators to control the switch between self-renewal and differentiation of pluripotent cells in the early embryo. However, the gross phenotype of miRNA-deficient embryos does not support this. DGCR8$^{-/-}$ blastocysts

are reported to show OCT4 expression in the ICM by IHC analysis, but not in the trophectoderm, suggesting that pluripotency genes can be downregulated *in vivo* (although given the limits of the sensitivity of the IHC method, it cannot be concluded that OCT4 expression is completely absent in the trophectoderm of DGCR8$^{-/-}$ embryos) (Suh *et al.*, 2010). Similarly, at E6.5 Dicer$^{-/-}$ embryos do not show aberrant expression of *Oct4* compared to wt embryos, supporting the conclusion that loss of miRNAs does not prevent silencing of self-renewal *in vivo* (Spruce *et al.*, 2010). A caveat with the data from Spruce *et al.* is that processed miRNAs were found to be present in the Dicer$^{-/-}$ embryos, presumably due to the presence of maternal Dicer. Nonetheless, in combination with the report from Suh *et al.*, these data suggest that self-renewal can be silenced *in vivo* in the absence of miRNAs.

The discrepancy between *in vitro* and *in vivo* findings may arise from the strength of the signal driving differentiation in these systems. It would be expected that conditions for differentiation are optimal *in vivo*, and therefore miRNAs may function to fine-tune the switch between self-renewal and differentiation, rather than directly control it. *In vitro*, however, the differentiation-inducing conditions may not be strong enough to drive differentiation in the absence of additional 'support' from extra levels of genetic regulation such as miRNAs. Supporting this is the observation that DGCR8$^{-/-}$ ES cells are capable of downregulating expression of pluripotency marker genes when *Oct4* expression was suppressed using siRNAs. This shows that the DGCR8$^{-/-}$ ES cells are capable of silencing self-renewal, but that it happens at a reduced frequency during non-stringent differentiation in the absence of miRNAs. Rather than supporting the theory that miRNAs may function as master regulators during silencing of self-renewal, these findings suggest that miRNAs act to ensure canalisation during the switch between self-renewal and differentiation, and loss of this canalising activity is enough to disrupt the silencing of self-renewal in non-directed differentiation protocols.

6.7 miRNAs are required for robust developmental transitions

Interestingly, miRNA-deficient ES cells appeared to be able to initiate trophectoderm and neuroectoderm differentiation during stringent differentiation protocols,

but showed impaired differentiation capacity during embryoid body differentiation, with mesoderm and endoderm lineages appearing particularly compromised. The observation that miRNA-deficient ES cells appear to initiate trophectoderm and neuroectoderm differentiation may suggest that miRNAs are not required for these developmental transitions. However, mice deficient for *Ago2* have hypertrophic placentas with a marked reduction in the thickness of the labyrinthine layer, and *Dicer*$^{-/-}$ embryos cannot maintain the TS cell compartment (Cheloufi *et al.*, 2010, Spruce *et al.*, 2010). Additionally, *Dicer* ablation in the mouse neocortex leads to defects in neural survival and differentiation (De Pietri Tonelli *et al.*, 2008). This suggests that although miRNA-deficient ES cells may be capable of initiating trophectoderm and neuroectoderm differentiation in stringent differentiation conditions *in vitro*, it is doubtful that these cells are functionally equivalent to cells generated from ES cells with miRNAs. Furthermore, it is unlikely that miRNAs are functioning as master regulators during trophectoderm or neuroectoderm specification. In this context, miRNAs may be functioning to ensure canalisation during differentiation to ensure formation of fully functional differentiated lineages.

miRNA-deficient ES cells showed particular defects in mesoderm and endoderm formation during embryoid body differentiation, and never formed beating cardiomyocytes or cystic structures. Perhaps then miRNAs directly control mesoderm and endoderm formation? Indeed, in the absence of miRNAs, the key early marker of mesoderm formation, *brachyury*, is undetectable in ES cells and shows drastically impaired induction during embryoid body formation. Conflicting data exist around the expression of *brachyury* in early Dicer$^{-/-}$ embryos: At E7.5 Bernstein *et al.* reported that, while wt embryos showed robust *brachyury* expression denoting the primitive streak, Dicer$^{-/-}$ embryos did not express *brachyury*; In contrast, Spruce *et al.* reported expression of *brachyury* in the primitive streak region of Dicer$^{-/-}$ embryos as well as expression of *Eomes* and *Nodal* (Bernstein *et al.*, 2003, Spruce *et al.*, 2010). As previously mentioned, a caveat of the Spruce *et al.* data is the presence of mature miRNAs in early Dicer$^{-/-}$ embryos. Bernstein *et al.* demonstrated that the mutant DICER protein produced in their Dicer$^{-/-}$ embryos was incapable of processing siRNAs, but did not test expression of miRNAs in the embryos themselves.

The data presented in this thesis are consistent with the findings of Bernstein *et al.*, and suggest that miRNAs are required for proper expression of *brachyury*.

However, the conflicting report from Spruce *et al.* means that it cannot be concluded that miRNAs are essential for initiation of the mesoderm lineage. miRNAs are certainly required for aspects of proper mesoderm and endoderm formation however. In human ES cells, loss of *miR-302* function strongly inhibits mesodermal and endodermal lineages, which may be mediated through *miR-302* regulation of the *Nodal* inhibitor *Lefty*, and in *Xenopus* embryos *miR-427* is required for proper mesoderm and endoderm specification (Rosa *et al.*, 2009). While the Rosa *et al.* study shows that miRNAs are important for mesoderm and endoderm development, it does not show that miRNAs are required for mesoderm or endoderm specification *per se*. Taken together, it is still not clear whether the general miRNA population is essential for mesoderm and endoderm lineage specification, but it is thought that miRNAs are involved in proper mesoderm and endoderm development *in vivo* and *in vitro*.

6.8 Concluding remarks

The discovery, a decade ago, that miRNAs functioned as post-transcriptional regulators of gene expression in numerous species, and were not an idiosyncrasy of nematode development, led to their identification as critical regulators of proper mammalian development. The data presented in this thesis furthers our understanding of the roles of specific miRNAs, as well as the general miRNA population, in regulating developmental transitions. Rather than functioning as master regulators, these data support a model in which miRNAs primarily act to ensure canalisation during development.

Bibliography

[Aguda et al., 2008] Aguda, B. D., Kim, Y., Piper-Hunter, M. G., Friedman, A., and Marsh, C. B. (2008). MicroRNA regulation of a cancer network: consequences of the feedback loops involving miR-17-92, E2F, and myc. *Proceedings of the National Academy of Sciences of the United States of America*, 105(50):19678–19683. PMID: 19066217.

[Ameres et al., 2010] Ameres, S. L., Horwich, M. D., Hung, J., Xu, J., Ghildiyal, M., Weng, Z., and Zamore, P. D. (2010). Target RNA Directed trimming and tailing of small silencing RNAs. *Science*, 328(5985):1534–1539.

[Anokye-Danso et al., 2011] Anokye-Danso, F., Trivedi, C. M., Juhr, D., Gupta, M., Cui, Z., Tian, Y., Zhang, Y., Yang, W., Gruber, P. J., Epstein, J. A., and Morrisey, E. E. (2011). Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. *Cell Stem Cell*, 8(4):376–388. PMID: 21474102.

[Avilion et al., 2003] 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 & Development*, 17(1):126–140. PMID: 12514105.

[Babiarz et al., 2008] Babiarz, J. E., Ruby, J. G., Wang, Y., Bartel, D. P., and Blelloch, R. (2008). Mouse ES cells express endogenous shRNAs, siRNAs, and other microprocessor-independent, dicer-dependent small RNAs. *Genes & Development*, 22(20):2773–2785. PMID: 18923076.

[Baccarini et al., 2011] Baccarini, A., Chauhan, H., Gardner, T. J., Jayaprakash, A. D., Sachidanandam, R., and Brown, B. D. (2011). Kinetic analysis reveals the fate of a microRNA following target regulation in mammalian cells. *Current Biology: CB*, 21(5):369–376. PMID: 21353554.

[Baek et al., 2008] Baek, D., Villon, J., Shin, C., Camargo, F. D., Gygi, S. P., and Bartel, D. P. (2008). The impact of microRNAs on protein output. *Nature*, 455(7209):64–71. PMID: 18668037 PMCID: 2745094.

[Bagga et al., 2005] Bagga, S., Bracht, J., Hunter, S., Massirer, K., Holtz, J., Eachus, R., and Pasquinelli, A. E. (2005). Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. *Cell*, 122(4):553–563. PMID: 16122423.

[Bartel, 2009] Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell*, 136(2):215–233. PMID: 19167326.

[Baskerville and Bartel, 2005] Baskerville, S. and Bartel, D. P. (2005). Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA (New York, N.Y.)*, 11(3):241–247. PMID: 15701730.

[Beddington and Robertson, 1989] Beddington, R. S. and Robertson, E. J. (1989). An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. *Development (Cambridge, England)*, 105(4):733–737. PMID: 2598811.

[Behm-Ansmant et al., 2006] Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., and Izaurralde, E. (2006). mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes & Development*, 20(14):1885–1898. PMID: 16815998.

[Beitzinger et al., 2007] Beitzinger, M., Peters, L., Zhu, J. Y., Kremmer, E., and Meister, G. (2007). Identification of human microRNA targets from isolated argonaute protein complexes. *RNA Biology*, 4(2):76–84. PMID: 17637574.

[Bernstein et al., 2001] 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(6818):363–366. PMID: 11201747.

[Bernstein et al., 2003] 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. *Nature Genetics*, 35(3):215–217. PMID: 14528307.

[Blum et al., 1987] Blum, H., Beier, H., and Gross, H. J. (1987). Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *ELECTROPHORESIS*, 8(2):93–99.

[Boland et al., 2011] Boland, A., Huntzinger, E., Schmidt, S., Izaurralde, E., and Weichenrieder, O. (2011). Crystal structure of the MID-PIWI lobe of a eukaryotic argonaute protein. *Proceedings of the National Academy of Sciences of the United States of America*, 108(26):10466–10471. PMID: 21646546.

[Boldin et al., 2011] Boldin, M. P., Taganov, K. D., Rao, D. S., Yang, L., Zhao, J. L., Kalwani, M., Garcia-Flores, Y., Luong, M., Devrekanli, A., Xu, J., Sun, G., Tay, J., Linsley, P. S., and Baltimore, D. (2011). miR-146a is a significant brake on autoimmunity, myeloproliferation, and cancer in mice. *The Journal of Experimental Medicine*, 208(6):1189 –1201.

[Bonauer et al., 2009] Bonauer, A., Carmona, G., Iwasaki, M., Mione, M., Koyanagi, M., Fischer, A., Burchfield, J., Fox, H., Doebele, C., Ohtani, K., Chavakis, E., Potente, M., Tjwa, M., Urbich, C., Zeiher, A. M., and Dimmeler, S. (2009). MicroRNA-92a controls angiogenesis and functional recovery of ischemic tissues in mice. *Science (New York, N.Y.)*, 324(5935):1710–1713. PMID: 19460962.

[Borchert et al., 2006] Borchert, G. M., Lanier, W., and Davidson, B. L. (2006). RNA polymerase III transcribes human microRNAs. *Nat Struct Mol Biol*, 13(12):1097–1101.

[Boyerinas et al., 2010] Boyerinas, B., Park, S., Hau, A., Murmann, A. E., and Peter, M. E. (2010). The role of let-7 in cell differentiation and cancer. *Endocrine-Related Cancer*, 17(1):F19 –F36.

[Brazil et al., 2004] Brazil, D. P., Yang, Z., and Hemmings, B. A. (2004). Advances in protein kinase b signalling: AKTion on multiple fronts. *Trends in Biochemical Sciences*, 29(5):233–242. PMID: 15130559.

[Brennecke et al., 2003] Brennecke, J., Hipfner, D. R., Stark, A., Russell, R. B., and Cohen, S. M. (2003). bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in drosophila. *Cell*, 113(1):25–36. PMID: 12679032.

[Brennecke et al., 2005] Brennecke, J., Stark, A., Russell, R. B., and Cohen, S. M. (2005). Principles of MicroRNA Target recognition. *PLoS Biology*, 3(3). PMID: 15723116 PMCID: 1043860.

[Brodersen et al., 2008] Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y. Y., Sieburth, L., and Voinnet, O. (2008). Widespread translational inhibition by plant miRNAs and siRNAs. *Science (New York, N.Y.)*, 320(5880):1185–1190. PMID: 18483398.

[Brons et al., 2007] Brons, I. G. M., Smithers, L. E., Trotter, M. W. B., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S. M., Howlett, S. K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R. A., and Vallier, L. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature*, 448(7150):191–195. PMID: 17597762.

[Burdon et al., 2002] Burdon, T., Smith, A., and Savatier, P. (2002). Signalling, cell cycle and pluripotency in embryonic stem cells. *Trends in Cell Biology*, 12(9):432–438. PMID: 12220864.

[Burdon et al., 1999] Burdon, T., Stracey, C., Chambers, I., Nichols, J., and Smith, A. (1999). Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells. *Developmental Biology*, 210(1):30–43. PMID: 10364425.

[Cai et al., 2008] Cai, K. Q., Capo-Chichi, C. D., Rula, M. E., Yang, D., and Xu, X. (2008). Dynamic GATA6 expression in primitive endoderm formation and maturation in early mouse embryogenesis. *Developmental Dynamics: An Official Publication of the American Association of Anatomists*, 237(10):2820–2829. PMID: 18816845.

[Calabrese et al., 2007] 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. *Proceedings of the National Academy of Sciences of the United States of America*, 104(46):18097–18102. PMID: 17989215 PMCID: 2084302.

[Calin et al., 2002] Calin, G. A., Dumitru, C. D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., Aldler, H., Rattan, S., Keating, M., Rai, K., Rassenti, L., Kipps, T., Negrini, M., Bullrich, F., and Croce, C. M. (2002). Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proceedings of the National Academy of Sciences of the United States of America*, 99(24):15524–15529. PMID: 12434020.

[Calin et al., 2004] Calin, G. A., Sevignani, C., Dumitru, C. D., Hyslop, T., Noch, E., Yendamuri, S., Shimizu, M., Rattan, S., Bullrich, F., Negrini, M., and Croce, C. M. (2004). Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proceedings of the National Academy of Sciences of the United States of America*, 101(9):2999 –3004.

[Calo et al., 2003] Calo, V., Migliavacca, M., Bazan, V., Macaluso, M., Buscemi, M., Gebbia, N., and Russo, A. (2003). STAT proteins: from normal control of cellular events to tumorigenesis. *Journal of Cellular Physiology*, 197(2):157–168. PMID: 14502555.

[Canham et al., 2010] Canham, M. A., Sharov, A. A., Ko, M. S. H., and Brickman, J. M. (2010). Functional heterogeneity of embryonic stem cells revealed through translational amplification of an early endodermal transcript. *PLoS Biology*, 8(5):e1000379. PMID: 20520791.

[Capo-Chichi et al., 2005] Capo-Chichi, C. D., Rula, M. E., Smedberg, J. L., Vanderveer, L., Parmacek, M. S., Morrisey, E. E., Godwin, A. K., and Xu, X. (2005). Perception of differentiation cues by GATA factors in primitive endoderm lineage determination of mouse embryonic stem cells. *Developmental Biology*, 286(2):574–586. PMID: 16162334.

[Carmell et al., 2007] Carmell, M. A., Girard, A., van de Kant, H. J., Bourchis, D., Bestor, T. H., de Rooij, D. G., and Hannon, G. J. (2007). MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Developmental Cell*, 12(4):503–514.

[Cartwright et al., 2005] Cartwright, P., McLean, C., Sheppard, A., Rivett, D., Jones, K., and Dalton, S. (2005). LIF/STAT3 controls ES cell self-renewal and pluripotency by a myc-dependent mechanism. *Development*, 132(5):885 –896.

[Caudy et al., 2003] Caudy, A. A., Ketting, R. F., Hammond, S. M., Denli, A. M., Bathoorn, A. M. P., Tops, B. B. J., Silva, J. M., Myers, M. M., Hannon, G. J., and Plasterk, R. H. A. (2003). A micrococcal nuclease homologue in RNAi effector complexes. *Nature*, 425(6956):411–414. PMID: 14508492.

[Caudy et al., 2002] Caudy, A. A., Myers, M., Hannon, G. J., and Hammond, S. M. (2002). Fragile x-related protein and VIG associate with the RNA interference machinery. *Genes & Development*, 16(19):2491–2496. PMID: 12368260.

[Chambers et al., 2003] 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(5):643–655. PMID: 12787505.

[Chazaud et al., 2006] Chazaud, C., Yamanaka, Y., Pawson, T., and Rossant, J. (2006). Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. *Developmental Cell*, 10(5):615–624.

[Cheloufi et al., 2010] Cheloufi, S., Dos Santos, C. O., Chong, M. M. W., and Hannon, G. J. (2010). A dicer-independent miRNA biogenesis pathway that requires ago catalysis. *Nature*, 465(7298):584–589. PMID: 20424607 PMCID: 2995450.

[Chen et al., 2008] Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V. B., Wong, E., Orlov, Y. L., Zhang, W., Jiang, J., Loh, Y., Yeo, H. C., Yeo, Z. X., Narang, V., Govindarajan, K. R., Leong, B., Shahab, A., Ruan, Y., Bourque, G., Sung, W., Clarke, N. D., Wei, C., and Ng, H. (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell*, 133(6):1106–1117.

[Chendrimada et al., 2005] Chendrimada, T. P., Gregory, R. I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., and Shiekhattar, R. (2005). TRBP recruits the dicer complex to ago2 for microRNA processing and gene silencing. *Nature*, 436(7051):740–744.

[Chi et al., 2009] Chi, S. W., Zang, J. B., Mele, A., and Darnell, R. B. (2009). Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature*, 460(7254):479–486. PMID: 19536157.

[Choy et al., 2008] Choy, E. Y., Siu, K., Kok, K., Lung, R. W., Tsang, C. M., To, K., Kwong, D. L., Tsao, S. W., and Jin, D. (2008). An Epstein-Barr virus-encoded microRNA targets PUMA to promote host cell survival. *The Journal of Experimental Medicine*, 205(11):2551–2560. PMID: 18838543 PMCID: 2571930.

[Cifuentes et al., 2010] Cifuentes, D., Xue, H., Taylor, D. W., Patnode, H., Mishima, Y., Cheloufi, S., Ma, E., Mane, S., Hannon, G. J., Lawson, N. D., Wolfe, S. A., and Giraldez, A. J. (2010). A novel miRNA processing pathway independent of dicer requires argonaute2 catalytic activity. *Science (New York, N.Y.)*, 328(5986):1694–1698. PMID: 20448148.

[Cloonan et al., 2008] Cloonan, N., Brown, M. K., Steptoe, A. L., Wani, S., Chan, W. L., Forrest, A. R. R., Kolle, G., Gabrielli, B., and Grimmond, S. M. (2008). The miR-17-5p microRNA is a key regulator of the G1/S phase cell cycle transition. *Genome Biology*, 9(8):R127. PMID: 18700987.

[Cohen and Melton, 2011] Cohen, D. E. and Melton, D. (2011). Turning straw into gold: directing cell fate for regenerative medicine. *Nature Reviews. Genetics*, 12(4):243–252. PMID: 21386864.

[Cole et al., 2008] Cole, M. F., Johnstone, S. E., Newman, J. J., Kagey, M. H., and Young, R. A. (2008). Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells. *Genes & Development*, 22(6):746 –755.

[Conlon et al., 1994] Conlon, F. L., Lyons, K. M., Takaesu, N., Barth, K. S., Kispert, A., Herrmann, B., and Robertson, E. J. (1994). A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse. *Development (Cambridge, England)*, 120(7):1919–1928. PMID: 7924997.

[Cullen, 2009] Cullen, B. R. (2009). Viral and cellular messenger RNA targets of viral microRNAs. *Nature*, 457(7228):421–425. PMID: 19158788.

[Davis et al., 2005] Davis, E., Caiment, F., Tordoir, X., Cavaille, J., Ferguson-Smith, A., Cockett, N., Georges, M., and Charlier, C. (2005). RNAi-mediated allelic trans-interaction at the imprinted Rtl1/Peg11 locus. *Current Biology: CB*, 15(8):743–749. PMID: 15854907.

[De Pietri Tonelli et al., 2008] De Pietri Tonelli, D., Pulvers, J. N., Haffner, C., Murchison, E. P., Hannon, G. J., and Huttner, W. B. (2008). miRNAs are essential for survival and differentiation of newborn neurons but not for expansion of neural progenitors during early neurogenesis in the mouse embryonic neocortex. *Development*, 135(23):3911 –3921.

[Deng et al., 2007] Deng, H., Sun, Y., Zhang, Y., Luo, X., Hou, W., Yan, L., Chen, Y., Tian, E., Han, J., and Zhang, H. (2007). Transcription factor NFY globally represses the expression of the c. elegans hox gene Abdominal-B homolog egl-5. *Developmental Biology*, 308(2):583–592. PMID: 17574230.

[Denli et al., 2004] Denli, A. M., Tops, B. B. J., Plasterk, R. H. A., Ketting, R. F., and Hannon, G. J. (2004). Processing of primary microRNAs by the microprocessor complex. *Nature*, 432(7014):231–235.

[Desbaillets et al., 2000] Desbaillets, I., Ziegler, U., Groscurth, P., and Gassmann, M. (2000). Embryoid bodies: an in vitro model of mouse embryogenesis. *Experimental Physiology*, 85(6):645–651. PMID: 11187960.

[Diederichs and Haber, 2007] Diederichs, S. and Haber, D. A. (2007). Dual role for argonautes in microRNA processing and posttranscriptional regulation of microRNA expression. *Cell*, 131(6):1097–1108. PMID: 18083100.

[Doench and Sharp, 2004] Doench, J. G. and Sharp, P. A. (2004). Specificity of microRNA target selection in translational repression. *Genes & Development*, 18(5):504–511. PMID: 15014042.

[Doetschman et al., 1985] Doetschman, T. C., Eistetter, H., Katz, M., Schmidt, W., and Kemler, R. (1985). The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *Journal of Embryology and Experimental Morphology*, 87:27–45. PMID: 3897439.

[Dupont et al., 2009] Dupont, S., Mamidi, A., Cordenonsi, M., Montagner, M., Zacchigna, L., Adorno, M., Martello, G., Stinchfield, M. J., Soligo, S., Morsut, L., Inui, M., Moro, S., Modena, N., Argenton, F., Newfeld, S. J., and Piccolo, S. (2009). FAM/USP9x, a deubiquitinating enzyme essential for TGFbeta signaling, controls smad4 monoubiquitination. *Cell*, 136(1):123–135. PMID: 19135894.

[Easow et al., 2007] Easow, G., Teleman, A. A., and Cohen, S. M. (2007). Isolation of microRNA targets by miRNP immunopurification. *RNA (New York, N.Y.)*, 13(8):1198–1204. PMID: 17592038.

[Erlebacher et al., 2004] Erlebacher, A., Price, K. A., and Glimcher, L. H. (2004). Maintenance of mouse trophoblast stem cell proliferation by TGF-beta/activin. *Developmental Biology*, 275(1):158–169. PMID: 15464579.

[Eulalio et al., 2007] Eulalio, A., Behm-Ansmant, I., Schweizer, D., and Izaurralde, E. (2007). P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. *Molecular and Cellular Biology*, 27(11):3970–3981. PMID: 17403906.

[Evans, 1972] Evans, M. J. (1972). The isolation and properties of a clonal tissue culture strain of pluripotent mouse teratoma cells. *Journal of Embryology and Experimental Morphology*, 28(1):163–176. PMID: 4672577.

[Evans and Kaufman, 1981] Evans, M. J. and Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature*, 292(5819):154–156. PMID: 7242681.

[Fabian et al., 2009] Fabian, M. R., Mathonnet, G., Sundermeier, T., Mathys, H., Zipprich, J. T., Svitkin, Y. V., Rivas, F., Jinek, M., Wohlschlegel, J., Doudna, J. A., Chen, C. A., Shyu, A., Yates III, J. R., Hannon, G. J., Filipowicz, W., Duchaine, T. F., and Sonenberg, N. (2009). Mammalian miRNA RISC recruits CAF1 and PABP to affect PABP-Dependent deadenylation. *Molecular Cell*, 35(6):868–880.

[Farh et al., 2005] Farh, K. K., Grimson, A., Jan, C., Lewis, B. P., Johnston, W. K., Lim, L. P., Burge, C. B., and Bartel, D. P. (2005). The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. *Science*, 310(5755):1817 –1821.

[Favre et al., 1986] Favre, A., Moreno, G., Blondel, M. O., Kliber, J., Vinzens, F., and Salet, C. (1986). 4-Thiouridine photosensitized RNA-protein crosslinking in mammalian cells. *Biochemical and Biophysical Research Communications*, 141(2):847–854. PMID: 2432896.

[Fazi and Nervi, 2008] Fazi, F. and Nervi, C. (2008). MicroRNA: basic mechanisms and transcriptional regulatory networks for cell fate determination. *Cardiovascular Research*, 79(4):553 –561.

[Feng et al., 2002] Feng, K., Zhou, X., Oohashi, T., Morgelin, M., Lustig, A., Hirakawa, S., Ninomiya, Y., Engel, J., Rauch, U., and Fussler, R. (2002). All four members of the Ten-m/Odz family of transmembrane proteins form dimers. *Journal of Biological Chemistry*, 277(29):26128 –26135.

[Feng et al., 2011] Feng, Y., Zhang, X., Song, Q., Li, T., and Zeng, Y. (2011). Drosha processing controls the specificity and efficiency of global microRNA expression. *Biochimica Et Biophysica Acta*. PMID: 21683814.

[Fire et al., 1998] Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in caenorhabditis elegans. *Nature*, 391(6669):806–811. PMID: 9486653.

[Firulli et al., 1998] Firulli, A. B., McFadden, D. G., Lin, Q., Srivastava, D., and Olson, E. N. (1998). Heart and extra-embryonic mesodermal defects in mouse embryos lacking the bHLH transcription factor hand1. *Nature Genetics*, 18(3):266–270. PMID: 9500550.

[Fleming, 1987] Fleming, T. P. (1987). A quantitative analysis of cell allocation to trophectoderm and inner cell mass in the mouse blastocyst. *Developmental Biology*, 119(2):520–531. PMID: 3803716.

[Fleming et al., 2001] Fleming, T. P., Sheth, B., and Fesenko, I. (2001). Cell adhesion in the preimplantation mammalian embryo and its role in trophectoderm differentiation and blastocyst morphogenesis. *Frontiers in Bioscience: A Journal and Virtual Library*, 6:D1000–1007. PMID: 11487467.

[Foshay et al., 2005] Foshay, K., Rodriguez, G., Hoel, B., Narayan, J., and Gallicano, G. I. (2005). JAK2/STAT3 directs cardiomyogenesis within murine embryonic stem cells in vitro. *Stem Cells (Dayton, Ohio)*, 23(4):530–543. PMID: 15790774.

[Frankenberg et al., 2007] Frankenberg, S., Smith, L., Greenfield, A., and Zernicka-Goetz, M. (2007). Novel gene expression patterns along the proximo-distal axis of the mouse embryo before gastrulation. *BMC Developmental Biology*, 7:8. PMID: 17302988.

[Friedman et al., 2009] Friedman, R. C., Farh, K. K., Burge, C. B., and Bartel, D. P. (2009). Most mammalian mRNAs are conserved targets of microRNAs. *Genome Research*, 19(1):92–105. PMID: 18955434.

[Fujikura et al., 2002] Fujikura, J., Yamato, E., Yonemura, S., Hosoda, K., Masui, S., Nakao, K., Miyazaki Ji, J.-i., and Niwa, H. (2002). Differentiation of embryonic stem cells is induced by GATA factors. *Genes & Development*, 16(7):784–789. PMID: 11937486.

[Furusawa et al., 1999] Furusawa, T., Moribe, H., Kondoh, H., and Higashi, Y. (1999). Identification of CtBP1 and CtBP2 as corepressors of zinc finger-homeodomain factor deltaEF1. *Molecular and Cellular Biology*, 19(12):8581–8590. PMID: 10567582.

[Garzon et al., 2010] Garzon, R., Marcucci, G., and Croce, C. M. (2010). Targeting microRNAs in cancer: rationale, strategies and challenges. *Nature Reviews. Drug Discovery*, 9(10):775–789. PMID: 20885409.

[Giraldez et al., 2005] Giraldez, A. J., Cinalli, R. M., Glasner, M. E., Enright, A. J., Thomson, J. M., Baskerville, S., Hammond, S. M., Bartel, D. P., and Schier, A. F. (2005). MicroRNAs regulate brain morphogenesis in zebrafish. *Science (New York, N.Y.)*, 308(5723):833–838. PMID: 15774722.

[Giraldez et al., 2006] Giraldez, A. J., Mishima, Y., Rihel, J., Grocock, R. J., Van Dongen, S., Inoue, K., Enright, A. J., and Schier, A. F. (2006). Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science (New York, N.Y.)*, 312(5770):75–79. PMID: 16484454.

[Girard et al., 2006] Girard, A., Sachidanandam, R., Hannon, G. J., and Carmell, M. A. (2006). A germline-specific class of small RNAs binds mammalian piwi proteins. *Nature*, 442(7099):199–202. PMID: 16751776.

[Golonzhka et al., 2009] Golonzhka, O., Metzger, D., Bornert, J., Bay, B. K., Gross, M. K., Kioussi, C., and Leid, M. (2009). Ctip2/Bcl11b controls ameloblast formation during mammalian odontogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, 106(11):4278–4283. PMID: 19251658.

[Gottwein et al., 2007] Gottwein, E., Mukherjee, N., Sachse, C., Frenzel, C., Majoros, W. H., Chi, J. A., Braich, R., Manoharan, M., Soutschek, J., Ohler, U., and Cullen, B. R. (2007). A viral microRNA functions as an orthologue of cellular miR-155. *Nature*, 450(7172):1096–1099. PMID: 18075594.

[Graumann et al., 2008] Graumann, J., Hubner, N. C., Kim, J. B., Ko, K., Moser, M., Kumar, C., Cox, J., Scholer, H., and Mann, M. (2008). Stable isotope labeling by amino acids in cell culture (SILAC) and proteome quantitation of mouse embryonic stem cells to a depth of 5,111 proteins. *Molecular & Cellular Proteomics: MCP*, 7(4):672–683. PMID: 18045802.

[Griffiths-Jones, 2004] Griffiths-Jones, S. (2004). The microRNA registry. *Nucleic Acids Research*, 32(Database issue):D109–111. PMID: 14681370.

[Grimson et al., 2008] Grimson, A., Srivastava, M., Fahey, B., Woodcroft, B. J., Chiang, H. R., King, N., Degnan, B. M., Rokhsar, D. S., and Bartel, D. P. (2008). Early origins and evolution of microRNAs and piwi-interacting RNAs in animals. *Nature*, 455(7217):1193–1197.

[Guillemot et al., 1994] Guillemot, F., Nagy, A., Auerbach, A., Rossant, J., and Joyner, A. L. (1994). Essential role of mash-2 in extraembryonic development. *Nature*, 371(6495):333–336. PMID: 8090202.

[Gulyas, 1975] Gulyas, B. J. (1975). A reexamination of cleavage patterns in eutherian mammalian eggs: rotation of blastomere pairs during second cleavage in the rabbit. *The Journal of Experimental Zoology*, 193(2):235–248. PMID: 1176904.

[Guo et al., 2010] Guo, H., Ingolia, N. T., Weissman, J. S., and Bartel, D. P. (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature*, 466(7308):835–840. PMID: 20703300.

[Guo et al., 2008] Guo, Y., Mantel, C., Hromas, R. A., and Broxmeyer, H. E. (2008). Oct-4 is critical for survival/antiapoptosis of murine embryonic stem cells subjected to stress: effects associated with stat3/survivin. *Stem Cells (Dayton, Ohio)*, 26(1):30–34. PMID: 17932422.

[Hafner et al., 2010a] Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, Manuel, J., Jungkamp, A., Munschauer, M., Ulrich, A., Wardle, G. S., Dewell, S., Zavolan, M., and Tuschl, T. (2010a). Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell*, 141(1):129–141. PMID: 20371350.

[Hafner et al., 2010b] Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jungkamp, A., Munschauer, M., Ulrich, A., Wardle, G. S., Dewell, S., Zavolan, M., and Tuschl, T. (2010b). PAR-CliP–a method to identify transcriptome-wide the binding sites of RNA binding proteins. *Journal of Visualized Experiments: JoVE*, (41). PMID: 20644507.

[Han et al., 2006] Han, J., Lee, Y., Yeom, K., Nam, J., Heo, I., Rhee, J., Sohn, S. Y., Cho, Y., Zhang, B., and Kim, V. N. (2006). Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell*, 125(5):887–901. PMID: 16751099.

[Hata et al., 2002] Hata, K., Okano, M., Lei, H., and Li, E. (2002). Dnmt3L cooperates with the dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development (Cambridge, England)*, 129(8):1983–1993. PMID: 11934864.

[Haub and Goldfarb, 1991] Haub, O. and Goldfarb, M. (1991). Expression of the fibroblast growth factor-5 gene in the mouse embryo. *Development (Cambridge, England)*, 112(2):397–406. PMID: 1794310.

[Hay et al., 2004] Hay, D. C., Sutherland, L., Clark, J., and Burdon, T. (2004). Oct-4 knockdown induces similar patterns of endoderm and trophoblast differentiation markers in human and mouse embryonic stem cells. *Stem Cells (Dayton, Ohio)*, 22(2):225–235. PMID: 14990861.

[Hayashi et al., 2008a] Hayashi, K., Chuva de Sousa Lopes, S. M., Kaneda, M., Tang, F., Hajkova, P., Lao, K., O'Carroll, D., Das, P. P., Tarakhovsky, A., Miska, E. A., and Surani, M. A. (2008a). MicroRNA biogenesis is required for mouse primordial germ cell development and spermatogenesis. *PloS One*, 3(3):e1738. PMID: 18320056.

[Hayashi et al., 2008b] Hayashi, K., Lopes, S. M. C. d. S., Tang, F., and Surani, M. A. (2008b). Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. *Cell Stem Cell*, 3(4):391–401. PMID: 18940731.

[He et al., 2005] He, L., Thomson, J. M., Hemann, M. T., Hernando-Monge, E., Mu, D., Goodson, S., Powers, S., Cordon-Cardo, C., Lowe, S. W., Hannon, G. J., and Hammond, S. M. (2005). A microRNA polycistron as a potential human oncogene. *Nature*, 435(7043):828–833.

[Heo et al., 2008] Heo, I., Joo, C., Cho, J., Ha, M., Han, J., and Kim, V. N. (2008). Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. *Molecular Cell*, 32(2):276–284. PMID: 18951094.

[Higashimoto et al., 2006] Higashimoto, K., Soejima, H., Saito, T., Okumura, K., and Mukai, T. (2006). Imprinting disruption of the CDKN1C/KCNQ1OT1 domain: the molecular mechanisms causing Beckwith-Wiedemann syndrome and cancer. *Cytogenetic and Genome Research*, 113(1-4):306–312. PMID: 16575194.

[Hildebrand and Soriano, 2002] Hildebrand, J. D. and Soriano, P. (2002). Overlapping and unique roles for C-Terminal binding protein 1 (CtBP1) and CtBP2 during mouse development. *Molecular and Cellular Biology*, 22(15):5296–5307. PMID: 12101226 PMCID: 133942.

[Hobert et al., 1991] Hobert, J. M., Boyle, M., and Martin, G. R. (1991). mRNA localization studies suggest that murine FGF-5 plays a role in gastrulation. *Development (Cambridge, England)*, 112(2):407–415. PMID: 1794311.

[HOUBAVIY et al., 2005] HOUBAVIY, H. B., DENNIS, L., JAENISCH, R., and SHARP, P. A. (2005). Characterization of a highly variable eutherian microRNA gene. *RNA*, 11(8):1245–1257. PMID: 15987809 PMCID: 1370808.

[Houbaviy et al., 2003] Houbaviy, H. B., Murray, M. F., and Sharp, P. A. (2003). Embryonic stem cell-specific MicroRNAs. *Developmental Cell*, 5(2):351–358. PMID: 12919684.

[Huang et al., 2009] Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 4(1):44–57. PMID: 19131956.

[Humphreys et al., 2005] Humphreys, D. T., Westman, B. J., Martin, D. I. K., and Preiss, T. (2005). MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. *Proceedings of the National Academy of Sciences of the United States of America*, 102(47):16961–16966. PMID: 16287976.

[Huntzinger and Izaurralde, 2011] Huntzinger, E. and Izaurralde, E. (2011). Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nature Reviews. Genetics*, 12(2):99–110. PMID: 21245828.

[Hutvagner et al., 2001] Hutvagner, G., McLachlan, J., Pasquinelli, A. E., Blint, 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(5531):834 –838.

[Ivanova et al., 2006] Ivanova, N., Dobrin, R., Lu, R., Kotenko, I., Levorse, J., DeCoste, C., Schafer, X., Lun, Y., and Lemischka, I. R. (2006). Dissecting self-renewal in stem cells with RNA interference. *Nature*, 442(7102):533–538.

[Ivey et al., 2008] Ivey, K. N., Muth, A., Arnold, J., King, F. W., Yeh, R., Fish, J. E., Hsiao, E. C., Schwartz, R. J., Conklin, B. R., Bernstein, H. S., and Srivastava, D. (2008). MicroRNA regulation of cell lineages in mouse and human embryonic stem cells. *Cell Stem Cell*, 2(3):219–229. PMID: 18371447.

[Ivey and Srivastava, 2010] Ivey, K. N. and Srivastava, D. (2010). MicroRNAs as regulators of differentiation and cell fate decisions. *Cell Stem Cell*, 7(1):36–41. PMID: 20621048.

[Janowski et al., 2007] Janowski, B. A., Younger, S. T., Hardy, D. B., Ram, R., Huffman, K. E., and Corey, D. R. (2007). Activating gene expression in mammalian cells with promoter-targeted duplex RNAs. *Nat Chem Biol*, 3(3):166–173.

[Jiang et al., 2009] Jiang, Q., Feng, M., and Mo, Y. (2009). Systematic validation of predicted microRNAs for cyclin d1. *BMC Cancer*, 9:194. PMID: 19538740.

[Jin et al., 2004] Jin, P., Zarnescu, D. C., Ceman, S., Nakamoto, M., Mowrey, J., Jongens, T. A., Nelson, D. L., Moses, K., and Warren, S. T. (2004). Biochemical and genetic interaction between the fragile x mental retardation protein and the microRNA pathway. *Nat Neurosci*, 7(2):113–117.

[John et al., 2004] John, B., Enright, A. J., Aravin, A., Tuschl, T., Sander, C., and Marks, D. S. (2004). Human MicroRNA targets. *PLoS Biology*, 2(11):e363. PMID: 15502875.

[Johnson et al., 2005] Johnson, S. M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K. L., Brown, D., and Slack, F. J. (2005). RAS is regulated by the let-7 microRNA family. *Cell*, 120(5):635–647. PMID: 15766527.

[Judson et al., 2009] Judson, R. L., Babiarz, J. E., Venere, M., and Blelloch, R. (2009). Embryonic stem cell-specific microRNAs promote induced pluripotency. *Nature Biotechnology*, 27(5):459–461. PMID: 19363475.

[Kamatani and Carson, 1981] Kamatani, N. and Carson, D. A. (1981). Dependence of adenine production upon polyamine synthesis in cultured human lymphoblasts. *Biochimica Et Biophysica Acta*, 675(3-4):344–350. PMID: 6791702.

[Kanellopoulou et al., 2005] 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 & Development*, 19(4):489–501. PMID: 15713842.

[Kawahara et al., 2011] Kawahara, H., Okada, Y., Imai, T., Iwanami, A., Mischel, P. S., and Okano, H. (2011). Musashi1 cooperates in abnormal cell lineage protein 28 (Lin28)-mediated let-7 family microRNA biogenesis in early neural differentiation. *The Journal of Biological Chemistry*, 286(18):16121–16130. PMID: 21378162.

[Kim and Kaartinen, 2008] Kim, J. and Kaartinen, V. (2008). Generation of mice with a conditional allele for trim33. *Genesis (New York, N.Y.: 2000)*, 46(6):329–333. PMID: 18543301.

[Kim et al., 2010a] Kim, J., Woo, A. J., Chu, J., Snow, J. W., Fujiwara, Y., Kim, C. G., Cantor, A. B., and Orkin, S. H. (2010a). A myc network accounts for similarities between embryonic stem and cancer cell transcription programs. *Cell*, 143(2):313–324. PMID: 20946988.

[Kim et al., 2010b] Kim, J. W., Mori, S., and Nevins, J. R. (2010b). Myc-Induced MicroRNAs integrate Myc-Mediated cell proliferation and cell fate. *Cancer Research*, 70(12):4820 –4828.

[Kim et al., 2010c] Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., Kim, J., Aryee, M. J., Ji, H., Ehrlich, L. I. R., Yabuuchi, A., Takeuchi, A., Cunniff, K. C., Hongguang, H., McKinney-Freeman, S., Naveiras, O., Yoon, T. J., Irizarry, R. A., Jung, N., Seita, J., Hanna, J., Murakami, P., Jaenisch, R., Weissleder, R., Orkin, S. H., Weissman, I. L., Feinberg, A. P., and Daley, G. Q. (2010c). Epigenetic memory in induced pluripotent stem cells. *Nature*, 467(7313):285–290. PMID: 20644535.

[Kim et al., 2010d] Kim, S., Lim, B., and Kim, J. (2010d). EWS-Oct-4B, an alternative EWS-Oct-4 fusion gene, is a potent oncogene linked to human epithelial tumours. *British Journal of Cancer*, 102(2):436–446. PMID: 20051954.

[Klattenhoff and Theurkauf, 2008] Klattenhoff, C. and Theurkauf, W. (2008). Biogenesis and germline functions of piRNAs. *Development*, 135(1):3 –9.

[Krek et al., 2005] Krek, A., Grun, D., Poy, M. N., Wolf, R., Rosenberg, L., Epstein, E. J., MacMenamin, P., da Piedade, I., Gunsalus, K. C., Stoffel, M., and Rajewsky, N. (2005). Combinatorial microRNA target predictions. *Nature Genetics*, 37(5):495–500. PMID: 15806104.

[Krichevsky et al., 2006] Krichevsky, A. M., Sonntag, K., Isacson, O., and Kosik, K. S. (2006). Specific microRNAs modulate embryonic stem cell-derived neurogenesis. *Stem Cells (Dayton, Ohio)*, 24(4):857–864. PMID: 16357340.

[Krol et al., 2010a] Krol, J., Busskamp, V., Markiewicz, I., Stadler, M. B., Ribi, S., Richter, J., Duebel, J., Bicker, S., Fehling, H. J., Schobeler, D., Oertner, T. G., Schratt, G., Bibel, M., Roska, B., and Filipowicz, W. (2010a). Characterizing light-regulated retinal microRNAs reveals rapid turnover as a common property of neuronal microRNAs. *Cell*, 141(4):618–631. PMID: 20478254.

[Krol et al., 2010b] Krol, J., Loedige, I., and Filipowicz, W. (2010b). The widespread regulation of microRNA biogenesis, function and decay. *Nature Reviews. Genetics*, 11(9):597–610. PMID: 20661255.

[Kuhn et al., 2008] Kuhn, D. E., Martin, M. M., Feldman, D. S., Terry, Alvin V, J., Nuovo, G. J., and Elton, T. S. (2008). Experimental validation of miRNA targets. *Methods (San Diego, Calif.)*, 44(1):47–54. PMID: 18158132.

[Kumar et al., 2007] Kumar, M. S., Lu, J., Mercer, K. L., Golub, T. R., and Jacks, T. (2007). Impaired microRNA processing enhances cellular transformation and tumorigenesis. *Nat Genet*, 39(5):673–677.

[Kunath et al., 2005] Kunath, T., Arnaud, D., Uy, G. D., Okamoto, I., Chureau, C., Yamanaka, Y., Heard, E., Gardner, R. L., Avner, P., and Rossant, J. (2005). Imprinted x-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts. *Development (Cambridge, England)*, 132(7):1649–1661. PMID: 15753215.

[Kunath et al., 2007] Kunath, T., Saba-El-Leil, M. K., Almousailleakh, M., Wray, J., Meloche, S., and Smith, A. (2007). FGF stimulation of the erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development (Cambridge, England)*, 134(16):2895–2902. PMID: 17660198.

[Kuramochi-Miyagawa et al., 2004] Kuramochi-Miyagawa, S., Kimura, T., Ijiri, T. W., Isobe, T., Asada, N., Fujita, Y., Ikawa, M., Iwai, N., Okabe, M., Deng, W., Lin, H., Matsuda, Y., and Nakano, T. (2004). Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. *Development*, 131(4):839 –849.

[Kuzmin et al., 2008] Kuzmin, A., Han, Z., Golding, M. C., Mann, M. R. W., Latham, K. E., and Varmuza, S. (2008). The PcG gene sfmbt2 is paternally expressed in extraembryonic tissues. *Gene Expression Patterns: GEP*, 8(2):107–116. PMID: 18024232.

[Lagos-Quintana et al., 2001] Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. *Science (New York, N.Y.)*, 294(5543):853–858. PMID: 11679670.

[Landais et al., 2007] Landais, S., Landry, S., Legault, P., and Rassart, E. (2007). Oncogenic potential of the miR-106-363 cluster and its implication in human t-cell leukemia. *Cancer Research*, 67(12):5699–5707. PMID: 17575136.

[Landthaler et al., 2004] 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. *Current Biology: CB*, 14(23):2162–2167. PMID: 15589161.

[Larkin et al., 2007] Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007). Clustal w and clustal x version 2.0. *Bioinformatics (Oxford, England)*, 23(21):2947–2948. PMID: 17846036.

[Larue et al., 1996] Larue, L., Antos, C., Butz, S., Huber, O., Delmas, V., Dominis, M., and Kemler, R. (1996). A role for cadherins in tissue formation. *Development (Cambridge, England)*, 122(10):3185–3194. PMID: 8898231.

[Lau et al., 2001] Lau, N. C., Lim, L. P., Weinstein, E. G., and Bartel, D. P. (2001). An abundant class of tiny RNAs with probable regulatory roles in caenorhabditis elegans. *Science (New York, N.Y.)*, 294(5543):858–862. PMID: 11679671.

[Lawson et al., 1991] Lawson, K. A., Meneses, J. J., and Pedersen, R. A. (1991). Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. *Development (Cambridge, England)*, 113(3):891–911. PMID: 1821858.

[Lee et al., 2010] Lee, J., Go, Y., Kang, I., Han, Y., and Kim, J. (2010). Oct-4 controls cell-cycle progression of embryonic stem cells. *The Biochemical Journal*, 426(2):171–181. PMID: 19968627.

[Lee et al., 2006] Lee, J., Kim, H. K., Rho, J., Han, Y., and Kim, J. (2006). The human OCT-4 isoforms differ in their ability to confer self-renewal. *The Journal of Biological Chemistry*, 281(44):33554–33565. PMID: 16951404.

[Lee et al., 2007] Lee, J., Kim, J. Y., Kang, I. Y., Kim, H. K., Han, Y., and Kim, J. (2007). The EWS-Oct-4 fusion gene encodes a transforming gene. *The Biochemical Journal*, 406(3):519–526. PMID: 17564582.

[Lee and Ambros, 2001] Lee, R. C. and Ambros, V. (2001). An extensive class of small RNAs in caenorhabditis elegans. *Science (New York, N.Y.)*, 294(5543):862–864. PMID: 11679672.

[Lee et al., 1993] 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(5):843–854. PMID: 8252621.

[Lee et al., 2003] Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S., and Kim, V. N. (2003). The nuclear RNase III drosha initiates microRNA processing. *Nature*, 425(6956):415–419. PMID: 14508493.

[Lee et al., 2002] Lee, Y., Jeon, K., Lee, J., Kim, S., and Kim, V. N. (2002). MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J*, 21(17):4663–4670.

[Lee et al., 2004] Lee, Y., Kim, M., Han, J., Yeom, K., Lee, S., Baek, S. H., and Kim, V. N. (2004). MicroRNA genes are transcribed by RNA polymerase II. *The EMBO Journal*, 23(20):4051–4060. PMID: 15372072.

[Leung et al., 2011] Leung, A. K. L., Young, A. G., Bhutkar, A., Zheng, G. X., Bosson, A. D., Nielsen, C. B., and Sharp, P. A. (2011). Genome-wide identification of ago2 binding sites from mouse embryonic stem cells with and without mature microRNAs. *Nature Structural & Molecular Biology*, 18(2):237–244. PMID: 21258322.

[Lewis et al., 2005] 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(1):15–20. PMID: 15652477.

[Lewis et al., 2003] Lewis, B. P., Shih, I.-h., Jones-Rhoades, M. W., Bartel, D. P., and Burge, C. B. (2003). Prediction of mammalian MicroRNA targets. *Cell*, 115(7):787–798.

[Lewis and Steel, 2010] Lewis, M. A. and Steel, K. P. (2010). MicroRNAs in mouse development and disease. *Seminars in Cell & Developmental Biology*, 21(7):774–780.

[Li et al., 2006] Li, L., Okino, S. T., Zhao, H., Pookot, D., Place, R. F., Urakami, S., Enokida, H., and Dahiya, R. (2006). Small dsRNAs induce transcriptional activation in human cells. *Proceedings of the National Academy of Sciences of the United States of America*, 103(46):17337–17342. PMID: 17085592.

[Li et al., 2011] Li, N., Wei, C., Olena, A. F., and Patton, J. G. (2011). Regulation of endoderm formation and left-right asymmetry by miR-92 during early zebrafish development. *Development (Cambridge, England)*, 138(9):1817–1826. PMID: 21447552.

[Lim et al., 2005] Lim, L. P., Lau, N. C., Garrett-Engele, P., Grimson, A., Schelter, J. M., Castle, J., Bartel, D. P., Linsley, P. S., and Johnson, J. M. (2005). Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature*, 433(7027):769–773. PMID: 15685193.

[Lin et al., 2008] Lin, S., Chang, D. C., Chang-Lin, S., Lin, C., Wu, D. T., Chen, D. T., and Ying, S. (2008). Mir-302 reprograms human skin cancer cells into a pluripotent ES-cell-like state. *RNA*, 14(10):2115 –2124.

[Liu et al., 2004] Liu, J., Carmell, M. A., Rivas, F. V., Marsden, C. G., Thomson, J. M., Song, J., Hammond, S. M., Joshua-Tor, L., and Hannon, G. J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. *Science (New York, N.Y.)*, 305(5689):1437–1441. PMID: 15284456.

[Liu et al., 2005] Liu, J., Valencia-Sanchez, M. A., Hannon, G. J., and Parker, R. (2005). MicroRNA-dependent localization of targeted mRNAs to mammalian p-bodies. *Nature cell biology*, 7(7):719–723. PMID: 15937477 PMCID: 1855297.

[Liu et al., 1999] Liu, P., Wakamiya, M., Shea, M. J., Albrecht, U., Behringer, R. R., and Bradley, A. (1999). Requirement for wnt3 in vertebrate axis formation. *Nature Genetics*, 22(4):361–365. PMID: 10431240.

[Loh and Lim, 2011] Loh, K. M. and Lim, B. (2011). A precarious balance: pluripotency factors as lineage specifiers. *Cell Stem Cell*, 8(4):363–369. PMID: 21474100.

[Loh et al., 2006] Loh, Y., Wu, Q., Chew, J., Vega, V. B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., Wong, K., Sung, K. W., Lee, C. W. H., Zhao, X., Chiu, K., Lipovich, L., Kuznetsov, V. A., Robson, P., Stanton, L. W., Wei, C., Ruan, Y., Lim, B., and Ng, H. (2006). The oct4 and nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nature Genetics*, 38(4):431–440. PMID: 16518401.

[Loh et al., 2007] Loh, Y., Zhang, W., Chen, X., George, J., and Ng, H. (2007). Jmjd1a and jmjd2c histone h3 lys 9 demethylases regulate self-renewal in embryonic stem cells. *Genes & Development*, 21(20):2545–2557. PMID: 17938240.

[Lossie et al., 2005] Lossie, A. C., Nakamura, H., Thomas, S. E., and Justice, M. J. (2005). Mutation of l7Rn3 shows that odz4 is required for mouse gastrulation. *Genetics*, 169(1):285 –299.

[Lowell et al., 2006] Lowell, S., Benchoua, A., Heavey, B., and Smith, A. G. (2006). Notch promotes neural lineage entry by pluripotent embryonic stem cells. *PLoS Biology*, 4(5):e121. PMID: 16594731.

[Lu et al., 2005] Lu, J., Getz, G., Miska, E. A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B. L., Mak, R. H., Ferrando, A. A., Downing, J. R., Jacks, T., Horvitz, H. R., and Golub, T. R. (2005). MicroRNA expression profiles classify human cancers. *Nature*, 435(7043):834–838. PMID: 15944708.

[Luo et al., 1997] Luo, J., Sladek, R., Bader, J. A., Matthyssen, A., Rossant, J., and Gigure, V. (1997). Placental abnormalities in mouse embryos lacking the orphan nuclear receptor ERR-beta. *Nature*, 388(6644):778–782. PMID: 9285590.

[Lutter et al., 2010] Lutter, D., Marr, C., Krumsiek, J., Lang, E. W., and Theis, F. J. (2010). Intronic microRNAs support their host genes by mediating synergistic and antagonistic regulatory effects. *BMC Genomics*, 11:224. PMID: 20370903.

[Lutticken et al., 1994] Lutticken, C., Wegenka, U. M., Yuan, J., Buschmann, J., Schindler, C., Ziemiecki, A., Harpur, A. G., Wilks, A. F., Yasukawa, K., and Taga, T. (1994). Association of transcription factor APRF and protein kinase jak1 with the interleukin-6 signal transducer gp130. *Science (New York, N.Y.)*, 263(5143):89–92. PMID: 8272872.

[Ma et al., 2009] Ma, T., Wang, Z., Guo, Y., and Pei, D. (2009). The c-terminal pentapeptide of nanog tryptophan repeat domain interacts with nac1 and regulates stem cell proliferation but not pluripotency. *Journal of Biological Chemistry*, 284(24):16071 –16081.

[Marson et al., 2008] 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., Calabrese, J. M., Dennis, L. M., Volkert, T. L., Gupta, S., Love, J., Hannett, N., Sharp, P. A., Bartel, D. P., Jaenisch, R., and Young, R. A. (2008). Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell*, 134(3):521–533. PMID: 18692474.

[Martin, 1980] Martin, G. R. (1980). Teratocarcinomas and mammalian embryogenesis. *Science (New York, N.Y.)*, 209(4458):768–776. PMID: 6250214.

[Martin, 1981] Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, 78(12):7634–7638. PMID: 6950406.

[Martin and Evans, 1975] Martin, G. R. and Evans, M. J. (1975). Differentiation of clonal lines of teratocarcinoma cells: formation of embryoid bodies in vitro. *Proceedings of the National Academy of Sciences of the United States of America*, 72(4):1441–1445. PMID: 1055416.

[Martinez-Ceballos et al., 2005] Martinez-Ceballos, E., Chambon, P., and Gudas, L. J. (2005). Differences in gene expression between wild type and hoxa1 knockout embryonic stem cells after retinoic acid treatment or leukemia inhibitory factor (LIF) removal. *The Journal of Biological Chemistry*, 280(16):16484–16498. PMID: 15722554.

[Martins-Taylor and Xu, 2010] Martins-Taylor, K. and Xu, R. (2010). Determinants of pluripotency: from avian, rodents, to primates. *Journal of Cellular Biochemistry*, 109(1):16–25. PMID: 19937733.

[Masui et al., 2007] Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A. A., Ko, M. S. H., and Niwa, H. (2007). Pluripotency governed by sox2 via regulation of oct3/4 expression in mouse embryonic stem cells. *Nat Cell Biol*, 9(6):625–635.

[Medeiros et al., 2011] 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., and Jaenisch, R. (2011). Mir-290-295 deficiency in mice results in partially penetrant embryonic lethality and germ cell defects. *Proceedings of the National Academy of Sciences of the United States of America*, 108(34):14163–14168. PMID: 21844366.

[Melton et al., 2010] Melton, C., Judson, R. L., and Blelloch, R. (2010). Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. *Nature*, 463(7281):621–626. PMID: 20054295.

[Mendell, 2008] Mendell, J. T. (2008). miRiad roles for the miR-17-92 cluster in development and disease. *Cell*, 133(2):217–222. PMID: 18423194.

[Meshorer and Misteli, 2006] Meshorer, E. and Misteli, T. (2006). Chromatin in pluripotent embryonic stem cells and differentiation. *Nature Reviews. Molecular Cell Biology*, 7(7):540–546. PMID: 16723974.

[Mineno et al.,] Mineno, J., Okamoto, S., Ando, T., Sato, M., Chono, H., Izu, H., Takayama, M., Asada, K., Mirochnitchenko, O., Inouye, M., and Kato, I. The expression profile of microRNAs in mouse embryos. *Nucleic Acids Research*, 34(6):1765 –1771.

[Miranda et al., 2006] Miranda, K. C., Huynh, T., Tay, Y., Ang, Y., Tam, W., Thomson, A. M., Lim, B., and Rigoutsos, I. (2006). A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell*, 126(6):1203–1217. PMID: 16990141.

[Miserey-Lenkei et al., 2006] Miserey-Lenkei, S., Coudel-Courteille, A., Del Nery, E., Bardin, S., Piel, M., Racine, V., Sibarita, J., Perez, F., Bornens, M., and Goud, B. (2006). A role for the Rab6A' GTPase in the inactivation of the mad2-spindle checkpoint. *The EMBO Journal*, 25(2):278–289. PMID: 16395330.

[Mitsui et al., 2003] 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(5):631–642. PMID: 12787504.

[Monteys et al., 2010] Monteys, A. M., Spengler, R. M., Wan, J., Tecedor, L., Lennox, K. A., Xing, Y., and Davidson, B. L. (2010). Structure and activity of putative intronic miRNA promoters. *RNA (New York, N.Y.)*, 16(3):495–505. PMID: 20075166.

[MORRISON and O'SULLIVAN, 1965] MORRISON, J. F. and O'SULLIVAN, W. J. (1965). KINETIC STUDIES OF THE REVERSE REACTION CATALYSED BY ADENOSINE TRIPHOSPHATE-CREATINE PHOSPHOTRANSFERASE. THE INHIBITION BY MAGNESIUM IONS AND ADENOSINE DIPHOSPHATE. *The Biochemical Journal*, 94:221–235. PMID: 14342234.

[Morsut et al., 2010] Morsut, L., Yan, K., Enzo, E., Aragona, M., Soligo, S. M., Wendling, O., Mark, M., Khetchoumian, K., Bressan, G., Chambon, P., Dupont, S., Losson, R., and Piccolo, S. (2010). Negative control of smad activity by ectodermin/Tif1gamma patterns the mammalian embryo. *Development (Cambridge, England)*, 137(15):2571–2578. PMID: 20573697.

[Mu et al., 2009] Mu, P., Han, Y., Betel, D., Yao, E., Squatrito, M., Ogrodowski, P., de Stanchina, E., D'Andrea, A., Sander, C., and Ventura, A. (2009). Genetic dissection of the miR-17-92 cluster of microRNAs in myc-induced b-cell lymphomas. 23(24):2806–2811. PMID: 20008931 PMCID: 2800095.

[Murchison et al., 2005] Murchison, E. P., Partridge, J. F., Tam, O. H., Cheloufi, S., and Hannon, G. J. (2005). Characterization of dicer-deficient murine embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, 102(34):12135–12140. PMID: 16099834.

[Murry and Keller, 2008] Murry, C. E. and Keller, G. (2008). Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. *Cell*, 132(4):661–680. PMID: 18295582.

[Nahid et al., 2009] Nahid, M. A., Pauley, K. M., Satoh, M., and Chan, E. K. L. (2009). miR-146a is critical for endotoxin-induced tolerance: IMPLICATION IN INNATE IMMUNITY. *The Journal of Biological Chemistry*, 284(50):34590–34599. PMID: 19840932.

[Najafi-Shoushtari et al., 2010] Najafi-Shoushtari, S. H., Kristo, F., Li, Y., Shioda, T., Cohen, D. E., Gerszten, R. E., and Naar, A. M. (2010). MicroRNA-33 and the SREBP host genes cooperate to control cholesterol homeostasis. *Science*, 328(5985):1566 –1569.

[Newman and Hammond, 2010] Newman, M. A. and Hammond, S. M. (2010). Emerging paradigms of regulated microRNA processing. *Genes & Development*, 24(11):1086–1092. PMID: 20516194.

[Newman et al., 2008] Newman, M. A., Thomson, J. M., and Hammond, S. M. (2008). Lin-28 interaction with the let-7 precursor loop mediates regulated microRNA processing. *RNA (New York, N.Y.)*, 14(8):1539–1549. PMID: 18566191.

[Nichols et al., 2001] Nichols, J., Chambers, I., Taga, T., and Smith, A. (2001). Physiological rationale for responsiveness of mouse embryonic stem cells to gp130 cytokines. *Development (Cambridge, England)*, 128(12):2333–2339. PMID: 11493552.

[Nichols et al., 2009] Nichols, J., Silva, J., Roode, M., and Smith, A. (2009). Suppression of erk signalling promotes ground state pluripotency in the mouse embryo. *Development (Cambridge, England)*, 136(19):3215–3222. PMID: 19710168.

[Nichols et al., 1998] 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(3):379–391. PMID: 9814708.

[Nicolas et al., 2008] Nicolas, F. E., Pais, H., Schwach, F., Lindow, M., Kauppinen, S., Moulton, V., and Dalmay, T. (2008). Experimental identification of microRNA-140 targets by silencing and overexpressing miR-140. *RNA (New York, N.Y.)*, 14(12):2513–2520. PMID: 18945805.

[Nishioka et al., 2009] Nishioka, N., Inoue, K.-i., Adachi, K., Kiyonari, H., Ota, M., Ralston, A., Yabuta, N., Hirahara, S., Stephenson, R. O., Ogonuki, N., Makita, R., Kurihara, H., Morin-Kensicki, E. M., Nojima, H., Rossant, J., Nakao, K., Niwa, H., and Sasaki, H. (2009). The hippo signaling pathway components lats and yap pattern tead4 activity to distinguish mouse trophectoderm from inner cell mass. *Developmental Cell*, 16(3):398–410. PMID: 19289085.

[Nishioka et al., 2008] Nishioka, N., Yamamoto, S., Kiyonari, H., Sato, H., Sawada, A., Ota, M., Nakao, K., and Sasaki, H. (2008). Tead4 is required for specification of trophectoderm in pre-implantation mouse embryos. *Mechanisms of Development*, 125(3-4):270–283. PMID: 18083014.

[Niwa, 2010] Niwa, H. (2010). Mouse ES cell culture system as a model of development. *Development, Growth & Differentiation*, 52(3):275–283. PMID: 20148924.

[Niwa et al., 1998] Niwa, H., Burdon, T., Chambers, I., and Smith, A. (1998). Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes & Development*, 12(13):2048 –2060.

[Niwa et al., 2000] Niwa, H., Miyazaki, J., and Smith, A. G. (2000). Quantitative expression of oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nature Genetics*, 24(4):372–376. PMID: 10742100.

[Niwa et al., 2005] Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., and Rossant, J. (2005). Interaction between oct3/4 and cdx2 determines trophectoderm differentiation. *Cell*, 123(5):917–929.

[O'Donnell et al., 2005] O'Donnell, K. A., Wentzel, E. A., Zeller, K. I., Dang, C. V., and Mendell, J. T. (2005). c-Myc-regulated microRNAs modulate E2F1 expression. *Nature*, 435(7043):839–843. PMID: 15944709.

[Ohnishi et al., 2010] Ohnishi, Y., Totoki, Y., Toyoda, A., Watanabe, T., Yamamoto, Y., Tokunaga, K., Sakaki, Y., Sasaki, H., and Hohjoh, H. (2010). Small RNA class transition from siRNA/piRNA to miRNA during pre-implantation mouse development. *Nucleic Acids Research*, 38(15):5141–5151. PMID: 20385573.

[Okamoto et al., 1990] Okamoto, K., Okazawa, H., Okuda, A., Sakai, M., Muramatsu, M., and Hamada, H. (1990). A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells. *Cell*, 60(3):461–472. PMID: 1967980.

[Okamura et al., 2008] Okamura, D., Tokitake, Y., Niwa, H., and Matsui, Y. (2008). Requirement of oct3/4 function for germ cell specification. *Developmental Biology*, 317(2):576–584. PMID: 18395706.

[Okano et al., 1999] Okano, M., Bell, D. W., Haber, D. A., and Li, E. (1999). DNA methyltransferases dnmt3a and dnmt3b are essential for de novo methylation and mammalian development. *Cell*, 99(3):247–257. PMID: 10555141.

[Olive et al., 2010] Olive, V., Jiang, I., and He, L. (2010). mir-17-92, a cluster of miRNAs in the midst of the cancer network. *The International Journal of Biochemistry & Cell Biology*, 42(8):1348–1354. PMID: 20227518.

[Olsen and Ambros, 1999] Olsen, P. H. and Ambros, V. (1999). The lin-4 regulatory RNA controls developmental timing in caenorhabditis elegans by blocking LIN-14 protein synthesis after the initiation of translation. *Developmental Biology*, 216(2):671–680. PMID: 10642801.

[Olsson et al., 1999] Olsson, P. A., Korhonen, L., Mercer, E. A., and Lindholm, D. (1999). MIR is a novel ERM-like protein that interacts with myosin regulatory light chain and inhibits neurite outgrowth. *The Journal of Biological Chemistry*, 274(51):36288–36292. PMID: 10593918.

[Ong et al., 2002] Ong, S., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular & Cellular Proteomics*, 1(5):376 –386.

[Orom and Lund, 2007] Orom, U. A. and Lund, A. H. (2007). Isolation of microRNA targets using biotinylated synthetic microRNAs. *Methods (San Diego, Calif.)*, 43(2):162–165. PMID: 17889804.

[Ozsolak et al., 2008] Ozsolak, F., Poling, L. L., Wang, Z., Liu, H., Liu, X. S., Roeder, R. G., Zhang, X., Song, J. S., and Fisher, D. E. (2008). Chromatin structure analyses identify miRNA promoters. *Genes & Development*, 22(22):3172 –3183.

[Pedersen et al., 2007] Pedersen, I. M., Cheng, G., Wieland, S., Volinia, S., Croce, C. M., Chisari, F. V., and David, M. (2007). Interferon modulation of cellular microRNAs as an antiviral mechanism. *Nature*, 449(7164):919–922. PMID: 17943132 PMCID: 2748825.

[Pegg and Williams-Ashman, 1969] Pegg, A. E. and Williams-Ashman, H. G. (1969). On the role of S-Adenosyl-l-methionine in the biosynthesis of spermidine by rat prostate. *Journal of Biological Chemistry*, 244(3):682 –693.

[Petersen et al., 2006] Petersen, C. P., Bordeleau, M., Pelletier, J., and Sharp, P. A. (2006). Short RNAs repress translation after initiation in mammalian cells. *Molecular Cell*, 21(4):533–542. PMID: 16483934.

[Petrocca et al., 2008] Petrocca, F., Vecchione, A., and Croce, C. M. (2008). Emerging role of miR-106b-25/miR-17-92 clusters in the control of transforming growth factor beta signaling. *Cancer Research*, 68(20):8191–8194. PMID: 18922889.

[Pfeffer et al., 2004] Pfeffer, S., Zavolan, M., Grasser, F. A., Chien, M., Russo, J. J., Ju, J., John, B., Enright, A. J., Marks, D., Sander, C., and Tuschl, T. (2004). Identification of virus-encoded microRNAs. *Science (New York, N.Y.)*, 304(5671):734–736. PMID: 15118162.

[Piko and Clegg, 1982] Piko, L. and Clegg, K. B. (1982). Quantitative changes in total RNA, total poly(A), and ribosomes in early mouse embryos. *Developmental Biology*, 89(2):362–378. PMID: 6173273.

[Pillai et al., 2005] Pillai, R. S., Bhattacharyya, S. N., Artus, C. G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E., and Filipowicz, W. (2005). Inhibition of translational initiation by let-7 MicroRNA in human cells. *Science (New York, N.Y.)*, 309(5740):1573–1576. PMID: 16081698.

[Place et al., 2008] Place, R. F., Li, L., Pookot, D., Noonan, E. J., and Dahiya, R. (2008). MicroRNA-373 induces expression of genes with complementary promoter sequences. *Proceedings of the National Academy of Sciences*, 105(5):1608 –1613.

[Plath and Lowry, 2011] Plath, K. and Lowry, W. E. (2011). Progress in understanding reprogramming to the induced pluripotent state. *Nature Reviews. Genetics*, 12(4):253–265. PMID: 21415849.

[Plusa et al., 2008] Plusa, B., Piliszek, A., Frankenberg, S., Artus, J., and Hadjantonakis, A. (2008). Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. *Development (Cambridge, England)*, 135(18):3081–3091. PMID: 18725515.

[Polo et al., 2010] Polo, J. M., Liu, S., Figueroa, M. E., Kulalert, W., Eminli, S., Tan, K. Y., Apostolou, E., Stadtfeld, M., Li, Y., Shioda, T., Natesan, S., Wagers, A. J., Melnick, A., Evans, T., and Hochedlinger, K. (2010). Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nature Biotechnology*, 28(8):848–855. PMID: 20644536.

[Prosser et al., 2011] Prosser, H. M., Koike-Yusa, H., Cooper, J. D., Law, F. C., and Bradley, A. (2011). A resource of vectors and ES cells for targeted deletion of microRNAs in mice. *Nat Biotech*, advance online publication.

[Rehwinkel et al., 2005] Rehwinkel, J., Behm-Ansmant, I., Gatfield, D., and Izaurralde, E. (2005). A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing. *RNA (New York, N.Y.)*, 11(11):1640–1647. PMID: 16177138.

[Reinhart et al., 2000] Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., Horvitz, H. R., and Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in caenorhabditis elegans. *Nature*, 403(6772):901–906. PMID: 10706289.

[Riley et al., 1998] Riley, P., Anson-Cartwright, L., and Cross, J. C. (1998). The hand1 bHLH transcription factor is essential for placentation and cardiac morphogenesis. *Nature Genetics*, 18(3):271–275. PMID: 9500551.

[Rinkenberger et al., 1997] Rinkenberger, J. L., Cross, J. C., and Werb, Z. (1997). Molecular genetics of implantation in the mouse. *Developmental Genetics*, 21(1):6–20. PMID: 9291576.

[Robertson, 1987] Robertson, E. (1987). *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*. Oxford University Press.

[Rodriguez et al., 2007] Rodriguez, A., Vigorito, E., Clare, S., Warren, M. V., Couttet, P., Soond, D. R., van Dongen, S., Grocock, R. J., Das, P. P., Miska, E. A., Vetrie, D., Okkenhaug, K., Enright, A. J., Dougan, G., Turner, M., and Bradley, A. (2007). Requirement of bic/microRNA-155 for normal immune function. *Science*, 316(5824):608 –611.

[Rosa et al., 2009] 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. *Developmental Cell*, 16(4):517–527. PMID: 19386261.

[Rosner et al., 1990] Rosner, M. H., Vigano, M. A., Ozato, K., Timmons, P. M., Poirier, F., Rigby, P. W., and Staudt, L. M. (1990). A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. *Nature*, 345(6277):686–692. PMID: 1972777.

[Rossant and Cross, 2001] Rossant, J. and Cross, J. C. (2001). Placental development: lessons from mouse mutants. *Nature Reviews. Genetics*, 2(7):538–548. PMID: 11433360.

[Rossant and Tam, 2009] Rossant, J. and Tam, P. P. L. (2009). Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. *Development (Cambridge, England)*, 136(5):701–713. PMID: 19201946.

[Ruby et al., 2007] Ruby, J. G., Jan, C. H., and Bartel, D. P. (2007). Intronic microRNA precursors that bypass drosha processing. *Nature*, 448(7149):83–86. PMID: 17589500.

[Rush et al., 2009] Rush, M., Appanah, R., Lee, S., Lam, L. L., Goyal, P., and Lorincz, M. C. (2009). Targeting of EZH2 to a defined genomic site is sufficient for recruitment of dnmt3a but not de novo DNA methylation. *Epigenetics: Official Journal of the DNA Methylation Society*, 4(6):404–414. PMID: 19717977.

[Rybak et al., 2008] Rybak, A., Fuchs, H., Smirnova, L., Brandt, C., Pohl, E. E., Nitsch, R., and Wulczyn, F. G. (2008). A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. *Nat Cell Biol*, 10(8):987–993.

[Saunders et al., 2002] Saunders, C. M., Larman, M. G., Parrington, J., Cox, L. J., Royse, J., Blayney, L. M., Swann, K., and Lai, F. A. (2002). PLCÎ¶: a sperm-specific trigger of ca2+ oscillations in eggs and embryo development. *Development*, 129(15):3533 –3544.

[Savatier et al., 1994] Savatier, P., Huang, S., Szekely, L., Wiman, K. G., and Samarut, J. (1994). Contrasting patterns of retinoblastoma protein expression in mouse embryonic stem cells and embryonic fibroblasts. *Oncogene*, 9(3):809–818. PMID: 8108123.

[Savatier et al., 2002] Savatier, P., Lapillonne, H., Jirmanova, L., Vitelli, L., and Samarut, J. (2002). Analysis of the cell cycle in mouse embryonic stem cells. *Methods in Molecular Biology (Clifton, N.J.)*, 185:27–33. PMID: 11768996.

[Schmitz et al., 2000] Schmitz, F., Konigstorfer, A., and Sodhof, T. C. (2000). RIBEYE, a component of synaptic ribbons: a protein's journey through evolution provides insight into synaptic ribbon function. *Neuron*, 28(3):857–872. PMID: 11163272.

[Scholer et al., 1990] Scholer, H. R., Ruppert, S., Suzuki, N., Chowdhury, K., and Gruss, P. (1990). New type of POU domain in germ line-specific protein oct-4. *Nature*, 344(6265):435–439. PMID: 1690859.

[Schwarz et al., 2003] 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(2):199–208.

[Seitz and Zamore, 2006] Seitz, H. and Zamore, P. D. (2006). Rethinking the microprocessor. *Cell*, 125(5):827–829.

[Selbach et al., 2008] Selbach, M., Schwanhusser, B., Thierfelder, N., Fang, Z., Khanin, R., and Rajewsky, N. (2008). Widespread changes in protein synthesis induced by microRNAs. *Nature*, 455(7209):58–63. PMID: 18668040.

[Sengupta et al., 2009] Sengupta, S., Nie, J., Wagner, R. J., Yang, C., Stewart, R., and Thomson, J. A. (2009). MicroRNA 92b controls the G1/S checkpoint gene p57 in human embryonic stem cells. *Stem Cells (Dayton, Ohio)*, 27(7):1524–1528. PMID: 19544458.

[Shekar et al., 2011] Shekar, P. C., Naim, A., Sarathi, D. P., and Kumar, S. (2011). Argonaute-2-null embryonic stem cells are retarded in self-renewal and differentiation. *Journal of Biosciences*, 36(4):649–657. PMID: 21857111.

[Shimosato et al., 2007] Shimosato, D., Shiki, M., and Niwa, H. (2007). Extra-embryonic endoderm cells derived from ES cells induced by GATA factors acquire the character of XEN cells. *BMC Developmental Biology*, 7:80. PMID: 17605826.

[Shimozaki et al., 2003] Shimozaki, K., Nakashima, K., Niwa, H., and Taga, T. (2003). Involvement of oct3/4 in the enhancement of neuronal differentiation of ES cells in neurogenesis-inducing cultures. *Development (Cambridge, England)*, 130(11):2505–2512. PMID: 12702663.

[Sikand et al., 2009] Sikand, K., Slane, S. D., and Shukla, G. C. (2009). Intrinsic expression of host genes and intronic miRNAs in prostate carcinoma cells. *Cancer Cell International*, 9:21. PMID: 19674469.

[Singh et al., 2007] Singh, A. M., Hamazaki, T., Hankowski, K. E., and Terada, N. (2007). A heterogeneous expression pattern for nanog in embryonic stem cells. *Stem Cells (Dayton, Ohio)*, 25(10):2534–2542. PMID: 17615266.

[Sinha et al., 1996] Sinha, S., Maity, S. N., Seldin, M. F., and de Crombrugghe, B. (1996). Chromosomal assignment and tissue expression of CBF-C/NFY-C, the third subunit of the mammalian CCAAT-binding factor. *Genomics*, 37(2):260–263. PMID: 8921405.

[Sinkkonen et al., 2008] 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. *Nature Structural & Molecular Biology*, 15(3):259–267. PMID: 18311153.

[Skalsky et al., 2007] Skalsky, R. L., Samols, M. A., Plaisance, K. B., Boss, I. W., Riva, A., Lopez, M. C., Baker, H. V., and Renne, R. (2007). Kaposi's sarcoma-associated herpesvirus encodes an ortholog of miR-155. *Journal of Virology*, 81(23):12836–12845. PMID: 17881434.

[Smith et al., 1988] Smith, A. G., Heath, J. K., Donaldson, D. D., Wong, G. G., Moreau, J., Stahl, M., and Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature*, 336(6200):688–690. PMID: 3143917.

[Smith and Hooper, 1987] Smith, A. G. and Hooper, M. L. (1987). Buffalo rat liver cells produce a diffusible activity which inhibits the differentiation of murine embryonal carcinoma and embryonic stem cells. *Developmental Biology*, 121(1):1–9. PMID: 3569655.

[Smith and Hooper, 1983] Smith, T. A. and Hooper, M. L. (1983). Medium conditioned by feeder cells inhibits the differentiation of embryonal carcinoma cultures. *Experimental Cell Research*, 145(2):458–462. PMID: 6861901.

[Solter et al., 1970] Solter, D., Skreb, N., and Damjanov, I. (1970). Extrauterine growth of mouse egg-cylinders results in malignant teratoma. *Nature*, 227(5257):503–504. PMID: 5428470.

[Song et al., 2004] Song, J., Smith, S. K., Hannon, G. J., and Joshua-Tor, L. (2004). Crystal structure of argonaute and its implications for RISC slicer activity. *Science (New York, N.Y.)*, 305(5689):1434–1437. PMID: 15284453.

[Spruce et al., 2010] 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. *Developmental Cell*, 19(2):207–219. PMID: 20708584.

[Stark et al., 2005] Stark, A., Brennecke, J., Bushati, N., Russell, R. B., and Cohen, S. M. (2005). Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution. *Cell*, 123(6):1133–1146. PMID: 16337999.

[Stark et al., 2003] Stark, A., Brennecke, J., Russell, R. B., and Cohen, S. M. (2003). Identification of drosophila MicroRNA targets. *PLoS Biology*, 1(3):E60. PMID: 14691535.

[Stavridis et al., 2007] Stavridis, M. P., Lunn, J. S., Collins, B. J., and Storey, K. G. (2007). A discrete period of FGF-induced erk1/2 signalling is required for vertebrate neural specification. *Development (Cambridge, England)*, 134(16):2889–2894. PMID: 17660197.

[Stewart et al., 1992] Stewart, C. L., Kaspar, P., Brunet, L. J., Bhatt, H., Gadi, I., KÃ¶ntgen, F., and Abbondanzo, S. J. (1992). Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor. *Nature*, 359(6390):76–79. PMID: 1522892.

[Strumpf et al., 2005] Strumpf, D., Mao, C., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F., and Rossant, J. (2005). Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. *Development (Cambridge, England)*, 132(9):2093–2102. PMID: 15788452.

[Suh et al., 2004] Suh, M., Lee, Y., Kim, J. Y., Kim, S., Moon, S., Lee, J. Y., Cha, K., Chung, H. M., Yoon, H. S., Moon, S. Y., Kim, V., and Kim, K. (2004). Human embryonic stem cells express a unique set of microRNAs. *Developmental Biology*, 270(2):488–498.

[Suh et al., 2010] Suh, N., Baehner, L., Moltzahn, F., Melton, C., Shenoy, A., Chen, J., and Blelloch, R. (2010). MicroRNA function is globally suppressed in mouse oocytes and early embryos. *Current Biology: CB*, 20(3):271–277. PMID: 20116247.

[Suh and Blelloch, 2011] Suh, N. and Blelloch, R. (2011). Small RNAs in early mammalian development: from gametes to gastrulation. *Development (Cambridge, England)*, 138(9):1653–1661. PMID: 21486922.

[Swann, 1990] Swann, K. (1990). A cytosolic sperm factor stimulates repetitive calcium increases and mimics fertilization in hamster eggs. *Development (Cambridge, England)*, 110(4):1295–1302. PMID: 2100264.

[Tada et al., 1997] Tada, M., Tada, T., Lefebvre, L., Barton, S. C., and Surani, M. A. (1997). Embryonic germ cells induce epigenetic reprogramming of somatic nucleus in hybrid cells. *The EMBO Journal*, 16(21):6510–6520. PMID: 9351832.

[Taganov et al., 2006] Taganov, K. D., Boldin, M. P., Chang, K., and Baltimore, D. (2006). NF-kappaB-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses. *Proceedings of the National Academy of Sciences of the United States of America*, 103(33):12481–12486. PMID: 16885212.

[Takada et al., 2006] Takada, S., Berezikov, E., Yamashita, Y., Lagos-Quintana, M., Kloosterman, W. P., Enomoto, M., Hatanaka, H., Fujiwara, S.-i., Watanabe, H., Soda, M., Choi, Y. L., Plasterk, R. H. A., Cuppen, E., and Mano, H. (2006). Mouse microRNA profiles determined with a new and sensitive cloning method. *Nucleic Acids Research*, 34(17):e115–e115. PMID: 16973894 PMCID: 1635289.

[Takahashi and Yamanaka, 2006] Takahashi, K. and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, 126(4):663–676. PMID: 16904174.

[Takahashi-Tezuka et al., 1998] Takahashi-Tezuka, M., Yoshida, Y., Fukada, T., Ohtani, T., Yamanaka, Y., Nishida, K., Nakajima, K., Hibi, M., and Hirano, T. (1998). Gab1 acts as an adapter molecule linking the cytokine receptor gp130 to ERK mitogen-activated protein kinase. *Molecular and Cellular Biology*, 18(7):4109–4117. PMID: 9632795.

[Takeda et al., 1997] Takeda, K., Noguchi, K., Shi, W., Tanaka, T., Matsumoto, M., Yoshida, N., Kishimoto, T., and Akira, S. (1997). Targeted disruption of the mouse stat3 gene leads to early embryonic lethality. *Proceedings of the National Academy of Sciences of the United States of America*, 94(8):3801–3804. PMID: 9108058.

[Tam et al., 2008] Tam, O. H., Aravin, A. A., Stein, P., Girard, A., Murchison, E. P., Cheloufi, S., Hodges, E., Anger, M., Sachidanandam, R., Schultz, R. M., and Hannon, G. J. (2008). Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature*, 453(7194):534–538. PMID: 18404147 PMCID: 2981145.

[Tanaka et al., 1998] Tanaka, S., Kunath, T., Hadjantonakis, A. K., Nagy, A., and Rossant, J. (1998). Promotion of trophoblast stem cell proliferation by FGF4. *Science (New York, N.Y.)*, 282(5396):2072–2075. PMID: 9851926.

[Tang et al., 2010] Tang, F., Barbacioru, C., Bao, S., Lee, C., Nordman, E., Wang, X., Lao, K., and Surani, M. A. (2010). Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell RNA-Seq analysis. *Cell Stem Cell*, 6(5):468–478. PMID: 20452321.

[Tanzer and Stadler, 2004] Tanzer, A. and Stadler, P. F. (2004). Molecular evolution of a microRNA cluster. *Journal of Molecular Biology*, 339(2):327–335. PMID: 15136036.

[Tay et al., 2008a] Tay, Y., Zhang, J., Thomson, A. M., Lim, B., and Rigoutsos, I. (2008a). MicroRNAs to nanog, oct4 and sox2 coding regions modulate embryonic stem cell differentiation. *Nature*, 455(7216):1124–1128.

[Tay et al., 2008b] Tay, Y. M., Tam, W., Ang, Y., Gaughwin, P. M., Yang, H., Wang, W., Liu, R., George, J., Ng, H., Perera, R. J., Lufkin, T., Rigoutsos, I., Thomson, A. M., and Lim, B. (2008b). MicroRNA-134 modulates the differentiation of mouse embryonic stem cells, where it causes post-transcriptional attenuation of nanog and LRH1. *Stem Cells (Dayton, Ohio)*, 26(1):17–29. PMID: 17916804.

[ten Berge et al., 2008] ten Berge, D., Koole, W., Fuerer, C., Fish, M., Eroglu, E., and Nusse, R. (2008). Wnt signaling mediates self-organization and axis formation in embryoid bodies. *Cell Stem Cell*, 3(5):508–518. PMID: 18983966.

[Tesar et al., 2007] 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. G. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature*, 448(7150):196–199. PMID: 17597760.

[Thai et al., 2007] Thai, T., Calado, D. P., Casola, S., Ansel, K. M., Xiao, C., Xue, Y., Murphy, A., Frendewey, D., Valenzuela, D., Kutok, J. L., Schmidt-Supprian, M., Rajewsky, N., Yancopoulos, G., Rao, A., and Rajewsky, K. (2007). Regulation of the germinal center response by MicroRNA-155. *Science*, 316(5824):604 –608.

[Theocharidis et al., 2009] Theocharidis, A., van Dongen, S., Enright, A. J., and Freeman, T. C. (2009). Network visualization and analysis of gene expression data using BioLayout Express(3D). *Nature Protocols*, 4(10):1535–1550. PMID: 19798086.

[Thomson et al., 2004] Thomson, J. M., Parker, J., Perou, C. M., and Hammond, S. M. (2004). A custom microarray platform for analysis of microRNA gene expression. *Nat Meth*, 1(1):47–53.

[Thomson et al., 2011] Thomson, M., Liu, S. J., Zou, L., Smith, Z., Meissner, A., and Ramanathan, S. (2011). Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. *Cell*, 145(6):875–889. PMID: 21663792.

[Tili et al., 2007] Tili, E., Michaille, J., Cimino, A., Costinean, S., Dumitru, C. D., Adair, B., Fabbri, M., Alder, H., Liu, C. G., Calin, G. A., and Croce, C. M. (2007). Modulation of miR-155 and miR-125b levels following lipopolysaccharide/TNF-alpha stimulation and their possible roles in regulating the response to endotoxin shock. *Journal of Immunology (Baltimore, Md.: 1950)*, 179(8):5082–5089. PMID: 17911593.

[Toyooka et al., 2008] Toyooka, Y., Shimosato, D., Murakami, K., Takahashi, K., and Niwa, H. (2008). Identification and characterization of subpopulations in undifferentiated ES cell culture. *Development (Cambridge, England)*, 135(5):909–918. PMID: 18263842.

[Trang et al., 2008] Trang, P., Weidhaas, J. B., and Slack, F. J. (2008). MicroRNAs as potential cancer therapeutics. *Oncogene*, 27 Suppl 2:S52–57. PMID: 19956180.

[Tropepe et al., 2001] Tropepe, V., Hitoshi, S., Sirard, C., Mak, T. W., Rossant, J., and van der Kooy, D. (2001). Direct neural fate specification from embryonic stem cells: a primitive mammalian neural stem cell stage acquired through a default mechanism. *Neuron*, 30(1):65–78. PMID: 11343645.

[Tsuchida et al., 2011] Tsuchida, A., Ohno, S., Wu, W., Borjigin, N., Fujita, K., Aoki, T., Ueda, S., Takanashi, M., and Kuroda, M. (2011). miR-92 is a key oncogenic component of the miR-17-92 cluster in colon cancer. *Cancer Science*. PMID: 21883694.

[Tucker and Chiquet-Ehrismann, 2006] Tucker, R. P. and Chiquet-Ehrismann, R. (2006). Teneurins: a conserved family of transmembrane proteins involved in intercellular signaling during development. *Developmental Biology*, 290(2):237–245. PMID: 16406038.

[Turner and Crossley, 1998] Turner, J. and Crossley, M. (1998). Cloning and characterization of mCtBP2, a co-repressor that associates with basic krÃŒppel-like factor and other mammalian transcriptional regulators. *The EMBO Journal*, 17(17):5129–5140. PMID: 9724649.

[Umbach et al., 2008] Umbach, J. L., Kramer, M. F., Jurak, I., Karnowski, H. W., Coen, D. M., and Cullen, B. R. (2008). MicroRNAs expressed by herpes simplex virus 1 during latent infection regulate viral mRNAs. *Nature*, 454(7205):780–783. PMID: 18596690 PMCID: 2666538.

[van Rooij et al., 2007] van Rooij, E., Sutherland, L. B., Qi, X., Richardson, J. A., Hill, J., and Olson, E. N. (2007). Control of Stress-Dependent cardiac growth and gene expression by a MicroRNA. *Science*, 316(5824):575 –579.

[Vaz et al., 2010] Vaz, C., Ahmad, H. M., Sharma, P., Gupta, R., Kumar, L., Kulshreshtha, R., and Bhattacharya, A. (2010). Analysis of microRNA transcriptome by deep sequencing of small RNA libraries of peripheral blood. *BMC Genomics*, 11:288. PMID: 20459673.

[Ventura et al., 2008] 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., Jaenisch, R., Sharp, P. A., and Jacks, T. (2008). Targeted deletion reveals essential and overlapping functions of the miR-17 through 92 family of miRNA clusters. *Cell*, 132(5):875–886. PMID: 18329372.

[Vinther et al., 2006] Vinther, J., Hedegaard, M. M., Gardner, P. P., Andersen, J. S., and Arctander, P. (2006). Identification of miRNA targets with stable isotope labeling by amino acids in cell culture. *Nucleic Acids Research*, 34(16):e107. PMID: 16945957.

[Viswanathan et al., 2008] Viswanathan, S. R., Daley, G. Q., and Gregory, R. I. (2008). Selective blockade of microRNA processing by lin28. *Science (New York, N.Y.)*, 320(5872):97–100. PMID: 18292307.

[Viswanathan et al., 2009] Viswanathan, S. R., Mermel, C. H., Lu, J., Lu, C., Golub, T. R., and Daley, G. Q. (2009). microRNA expression during trophectoderm specification. *PloS One*, 4(7):e6143. PMID: 19582159.

[Volinia et al., 2006] Volinia, S., Calin, G. A., Liu, C., Ambs, S., Cimmino, A., Petrocca, F., Visone, R., Iorio, M., Roldo, C., Ferracin, M., Prueitt, R. L., Yanaihara, N., Lanza, G., Scarpa, A., Vecchione, A., Negrini, M., Harris, C. C., and Croce, C. M. (2006). A microRNA expression signature of human solid tumors defines cancer gene targets. *Proceedings of the National Academy of Sciences of the United States of America*, 103(7):2257–2261. PMID: 16461460.

[Waddington, 1942] Waddington, C. H. (1942). Canalization of development and the inheritance of acquired characters. *Nature*, 150:563–565.

[Wakabayashi et al., 2003] Wakabayashi, Y., Watanabe, H., Inoue, J., Takeda, N., Sakata, J., Mishima, Y., Hitomi, J., Yamamoto, T., Utsuyama, M., Niwa, O., Aizawa, S., and Kominami, R. (2003). Bcl11b is required for differentiation and survival of alphabeta t lymphocytes. *Nature Immunology*, 4(6):533–539. PMID: 12717433.

[Wang et al., 2010] Wang, J., Greene, S. B., Bonilla-Claudio, M., Tao, Y., Zhang, J., Bai, Y., Huang, Z., Black, B. L., Wang, F., and Martin, J. F. (2010). Bmp signaling regulates myocardial differentiation from cardiac progenitors through a MicroRNA-mediated mechanism. *Developmental Cell*, 19(6):903–912. PMID: 21145505.

[Wang et al., 2011] Wang, Q., Chow, J., Hong, J., Smith, A. F., Moreno, C., Seaby, P., Vrana, P., Miri, K., Tak, J., Chung, E. D., Mastromonaco, G., Caniggia, I., and Varmuza, S. (2011). Recent acquisition of imprinting at the rodent sfmbt2 locus correlates with insertion of a large block of miRNAs. *BMC Genomics*, 12:204. PMID: 21510876.

[Wang et al., 2008] 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. *Nature Genetics*, 40(12):1478–1483. PMID: 18978791.

[Wang et al., 2007] 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. *Nature Genetics*, 39(3):380–385. PMID: 17259983.

[Watanabe et al., 2006] Watanabe, T., Takeda, A., Tsukiyama, T., Mise, K., Okuno, T., Sasaki, H., Minami, N., and Imai, H. (2006). Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. *Genes & Development*, 20(13):1732 –1743.

[Watanabe et al., 2008] Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T., Surani, M. A., Sakaki, Y., and Sasaki, H. (2008). Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature*, 453(7194):539–543.

[Wehner et al., 2010] Wehner, K. A., Schetz, S., and Sarnow, P. (2010). OGFOD1, a novel modulator of eukaryotic translation initiation factor 2a phosphorylation and the cellular response to stress. 30(8):2006–2016. PMID: 20154146 PMCID: 2849474.

[Weitzer, 2006] Weitzer, G. (2006). Embryonic stem cell-derived embryoid bodies: an in vitro model of eutherian pregastrulation development and early gastrulation. *Handbook of Experimental Pharmacology*, (174):21–51. PMID: 16370323.

[Wellner et al., 2009] Wellner, U., Schubert, J., Burk, U. C., Schmalhofer, O., Zhu, F., Sonntag, A., Waldvogel, B., Vannier, C., Darling, D., zur Hausen, A., Brunton, V. G., Morton, J., Sansom, O., Scholer, J., Stemmler, M. P., Herzberger, C., Hopt, U., Keck, T., Brabletz, S., and Brabletz, T. (2009). The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs. *Nature Cell Biology*, 11(12):1487–1495. PMID: 19935649.

[Wheeler et al., 2006] Wheeler, G., Ntounia-Fousara, S., Granda, B., Rathjen, T., and Dalmay, T. (2006). Identification of new central nervous system specific mouse microRNAs. *FEBS Letters*, 580(9):2195–2200. PMID: 16566924.

[Wiebe et al., 2000] Wiebe, M. S., Wilder, P. J., Kelly, D., and Rizzino, A. (2000). Isolation, characterization, and differential expression of the murine sox-2 promoter. *Gene*, 246(1-2):383–393. PMID: 10767561.

[Wienholds et al., 2003] Wienholds, E., Koudijs, M. J., van Eeden, F. J. M., Cuppen, E., and Plasterk, R. H. A. (2003). The microRNA-producing enzyme dicer1 is essential for zebrafish development. *Nature Genetics*, 35(3):217–218. PMID: 14528306.

[Williams et al., 1988] Williams, R. L., Hilton, D. J., Pease, S., Willson, T. A., Stewart, C. L., Gearing, D. P., Wagner, E. F., Metcalf, D., Nicola, N. A., and Gough, N. M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature*, 336(6200):684–687. PMID: 3143916.

[Wilmut et al., 1997] Wilmut, I., Schnieke, A. E., McWhir, J., Kind, A. J., and Campbell, K. H. (1997). Viable offspring derived from fetal and adult mammalian cells. *Nature*, 385(6619):810–813. PMID: 9039911.

[Wilson et al., 1995] Wilson, V., Manson, L., Skarnes, W. C., and Beddington, R. S. (1995). The t gene is necessary for normal mesodermal morphogenetic cell movements during gastrulation. *Development (Cambridge, England)*, 121(3):877–886. PMID: 7720590.

[Wu et al., 2010] Wu, S., Huang, S., Ding, J., Zhao, Y., Liang, L., Liu, T., Zhan, R., and He, X. (2010). Multiple microRNAs modulate p21Cip1/Waf1 expression by directly targeting its 3' untranslated region. *Oncogene*, 29(15):2302–2308. PMID: 20190813.

[Xiao et al., 2009] Xiao, F., Zuo, Z., Cai, G., Kang, S., Gao, X., and Li, T. (2009). miRecords: an integrated resource for microRNA-target interactions. *Nucleic Acids Research*, 37(Database issue):D105–110. PMID: 18996891.

[Xu et al., 2009] Xu, N., Papagiannakopoulos, T., Pan, G., Thomson, J. A., and Kosik, K. S. (2009). MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. *Cell*, 137(4):647–658. PMID: 19409607.

[Yan et al., 2004] Yan, K., Dollo, P., Mark, M., Lerouge, T., Wendling, O., Chambon, P., and Losson, R. (2004). Molecular cloning, genomic structure, and expression analysis of the mouse transcriptional intermediary factor 1 gamma gene. *Gene*, 334:3–13. PMID: 15256250.

[Yang et al., 2010] Yang, Y., Chaerkady, R., Kandasamy, K., Huang, T., Selvan, L. D. N., Dwivedi, S. B., Kent, O. A., Mendell, J. T., and Pandey, A. (2010). Identifying targets of miR-143 using a SILAC-based proteomic approach. *Molecular bioSystems*, 6(10):1873–1882. PMID: 20544124.

[Ye et al., 2008] Ye, W., Lv, Q., Wong, C. A., Hu, S., Fu, C., Hua, Z., Cai, G., Li, G., Yang, B. B., and Zhang, Y. (2008). The effect of central loops in miRNA:MRE duplexes on the efficiency of miRNA-mediated gene regulation. *PloS One*, 3(3):e1719. PMID: 18320040.

[Yekta et al., 2004] Yekta, S., Shih, I., and Bartel, D. P. (2004). MicroRNA-directed cleavage of HOXB8 mRNA. *Science (New York, N.Y.)*, 304(5670):594–596. PMID: 15105502.

[Yi et al., 2008] Yi, R., Poy, M. N., Stoffel, M., and Fuchs, E. (2008). A skin microRNA promotes differentiation by repressing 'stemness'. *Nature*, 452(7184):225–229. PMID: 18311128.

[Ying and Smith, 2003] Ying, Q. and Smith, A. G. (2003). Defined conditions for neural commitment and differentiation. *Methods in Enzymology*, 365:327–341. PMID: 14696356.

[Ying et al., 2008] Ying, Q., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. *Nature*, 453(7194):519–523. PMID: 18497825.

[Ying et al., 2003] Ying, Q. L., Nichols, J., Chambers, I., and Smith, A. (2003). BMP induction of id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell*, 115(3):281–292. PMID: 14636556.

[Ying and Lin, 2005] Ying, S. and Lin, S. (2005). Intronic microRNAs. *Biochemical and Biophysical Research Communications*, 326(3):515–520. PMID: 15596130.

[Yoshida et al., 1994] Yoshida, K., Chambers, I., Nichols, J., Smith, A., Saito, M., Yasukawa, K., Shoyab, M., Taga, T., and Kishimoto, T. (1994). Maintenance of the pluripotential phenotype of embryonic stem cells through direct activation of gp130 signalling pathways. *Mechanisms of Development*, 45(2):163–171. PMID: 8199053.

[Yu et al., 2008] Yu, P. B., Hong, C. C., Sachidanandan, C., Babitt, J. L., Deng, D. Y., Hoyng, S. A., Lin, H. Y., Bloch, K. D., and Peterson, R. T. (2008). Dorsomorphin inhibits BMP signals required for embryogenesis and iron metabolism. *Nature Chemical Biology*, 4(1):33–41. PMID: 18026094.

[Zamore et al., 2000] 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(1):25–33. PMID: 10778853.

[Zhang et al., 1997] Zhang, P., Ligeois, N. J., Wong, C., Finegold, M., Hou, H., Thompson, J. C., Silverman, A., Harper, J. W., DePinho, R. A., and Elledge, S. J. (1997). Altered cell differentiation and proliferation in mice lacking p57KIP2 indicates a role in Beckwith-Wiedemann syndrome. *Nature*, 387(6629):151–158. PMID: 9144284.

[Zhao et al., 2007] 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(2):303–317.

[Zhao et al., 2011] Zhao, Y., Xu, H., Yao, Y., Smith, L. P., Kgosana, L., Green, J., Petherbridge, L., Baigent, S. J., and Nair, V. (2011). Critical role of the Virus-Encoded MicroRNA-155 ortholog in the induction of marek's disease lymphomas. 7(2). PMID: 21383974 PMCID: 3044692.

[Zhao et al., 2009] Zhao, Y., Yao, Y., Xu, H., Lambeth, L., Smith, L. P., Kgosana, L., Wang, X., and Nair, V. (2009). A functional MicroRNA-155 ortholog encoded by the oncogenic marek's disease virus. *Journal of Virology*, 83(1):489–492. PMID: 18945769 PMCID: 2612317.

[Zheng et al., 2011] Zheng, G. X. Y., Ravi, A., Gould, G. M., Burge, C. B., and Sharp, P. A. (2011). Genome-wide impact of a recently expanded microRNA cluster in mouse. *Proceedings of the National Academy of Sciences of the United States of America*, 108(38):15804–15809. PMID: 21911408.

[Zheng et al., 2010] Zheng, K., Li, H., Zhu, Y., Zhu, Q., and Qiu, M. (2010). MicroRNAs are essential for the developmental switch from neurogenesis to gliogenesis in the developing spinal cord. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 30(24):8245–8250. PMID: 20554876.

[Zhou et al., 2001] Zhou, G., Myers, R., Li, Y., Chen, Y., Shen, X., Fenyk-Melody, J., Wu, M., Ventre, J., Doebber, T., Fujii, N., Musi, N., Hirshman, M. F., Goodyear, L. J., and Moller, D. E. (2001). Role of AMP-activated protein kinase in mechanism of metformin action. *The Journal of Clinical Investigation*, 108(8):1167–1174. PMID: 11602624.

Appendix A - QPCR primer optimisation

Gene	Single Product?	Effective Amplification Range and Ct Spread	RSq	Amplification Efficiency (%)
Oct4	Yes	10,000 fold (Ct 22-35.5)	0.994	99.9
Sox2	Yes	100,000 fold (Ct 16-32)	0.995	103.9
Nanog	Yes	100, 000 fold (Ct 16-33)	0.995	96.5
Esrrß	Yes	10,000 fold (Ct 22-36)	0.992	100.1
Cdx2	Yes	100,000 fold (Ct 22-38)	0.968	98.5
Eomes	Yes	10,000 fold (Ct 23.5-36)	0.928	106.8
Hand1	Yes	10,000 fold (Ct 22-35.5)	0.988	103.2
Mash2	Yes	1000 fold (Ct 24-34.5)	0.996	98.8
Id2	Yes	1000 fold (Ct 23-33.5)	0.999	94.7
Sfmbt2	Yes	100,000 fold (Ct 20-36)	0.998	101.1
Gata4	Yes	1000 fold (Ct 26.5 - 36.5)	0.993	101.3
Gata6	Yes	1000 fold (Ct 26-35.5)	0.982	103.6
Fgf5	Yes	1000 fold (Ct 26-36)	0.985	98.4
Brachyury	Yes	1000 fold (Ct 27.5-37)	0.970	101.6
Snail	Yes	1000 fold (Ct 26-36)	0.992	97.8
Goosecoid	Yes	100 fold (Ct 29-36)	0.971	95.4
Nestin	Yes	1000 fold (Ct 25-34.5)	0.990	99.7
Odz4	Yes	1000 fold (Ct 27-37)	0.991	96.5
Nfyc	Yes	100,000 (Ct 21-38)	0.997	96.8
C77370	Yes	1000 fold (Ct 27-37)	0.995	102.3
Trim33	Yes	10,000 fold (Ct 23.5-36.5)	0.990	103.1
Dnmt3a	Yes	100 fold (Ct 23-30)	0.989	94.6
Mylip	Yes	1,000 fold (Ct 26-35.5)	0.996	105.6
Actß	Yes	100,000 fold (Ct 16-33)	0.999	101.1
Dgcr8	Yes	10,000 fold (Ct 25-38)	0.978	104.8

Appendix B - miRNA profiling data

miRNAname	Total	Total C	Total T	C0	C24	C48	T24	T48	T72	miRNA_seq	Blast geno	mm RNAfold 70%
mmu-let-7c	2	1	1	0	0	1	1	0	0	UGAGGUAGUAGGUUGUAUGGUU	mmu-let-7c,1,a|mm|chr16:775	mmu-let-7c.1.a|m
mmu-let-7d	1	1	0	0	1	0	0	0	0	AGAGGUAGUAGGUUGCAUAGUU	mmu-let-7d,1,a|mm|chr13:48(	mmu-let-7d.1.a|m
mmu-let-7e	1	1	0	0	0	1	0	0	0	UGAGGUAGGAGGUUGUAUAGUU	mmu-let-7e,1,a|mm|chr17:17(	mmu-let-7e.1.a|m
mmu-let-7f	1	0	1	0	0	0	1	0	0	UGAGGUAGUAGAUUGUAUAGUU	mmu-let-7f,1,a|mm|chrX:1483	mmu-let-7f.1.a|mr
mmu-let-7g	1	0	1	0	0	0	1	0	0	UGAGGUAGUAGUUUGUACAGUU	mmu-let-7g,1,a|mm|chr9:106(	mmu-let-7g.1.a|m
mmu-miR-103	6	2	4	0	1	1	0	3	1	AGCAGCATTGTACAGGGCTAT	mmu-miR-103,1,a|mm|chr2:1	mmu-miR-103.1.b
mmu-miR-106a	15	10	5	3	4	3	4	1	0	CAAAGTGCTAACAGTGCAGGTA	mmu-miR-106a,1,a|mm|chrX:	No Folding
mmu-miR-106b	41	18	23	3	9	6	5	10	8	TAAAGTGCTGACAGTGCAGATA	mmu-miR-106b,1,a|mm|chr5:	No Folding
mmu-miR-124	23	17	6	5	8	4	5	0	1	TAAGGCACGCGGTGAATGCCAA	mmu-miR-124,1,a|mm|chr3:1	mmu-miR-124.1.b
mmu-miR-125a-5p	4	3	1	0	1	2	0	1	0	TCCCTGAGACCCTTTAACCTGT	mmu-miR-125a-5p,1,a|mm|ch	mmu-miR-125a-5
mmu-miR-125b-5p	2	2	0	0	2	0	0	0	0	UCCCUGAGACCCUAACUUGUGA	mmu-miR-125b-5p,1,a|mm|ch	mmu-miR-125b-5
mmu-miR-128	1	0	1	0	0	0	0	0	1	UCACAGUGAACCGGUCUCUUU	mmu-miR-128,1,a|mm|chr9:1	mmu-miR-128.1.b
mmu-miR-130a	60	29	31	9	10	10	14	8	9	CAGUGCAAUGUUAAAAGGGCAU	mmu-miR-130a,1,a|mm|chr2:	mmu-miR-130a.1
mmu-miR-130b	6	3	3	0	0	3	1	0	2	CAGUGCAAUGAUGAAAGGGCAU	mmu-miR-130b,1,a|mm|chr16	mmu-miR-130b.1
mmu-miR-135b	4	3	1	1	0	2	1	0	0	UAUGGCUUUUCAUUCCUAUGUGA	mmu-miR-135b,1,a|mm|chr1:	mmu-miR-135b.1
mmu-miR-138	6	2	4	0	0	2	1	1	2	AGCUGGUGUUGUGAAUCAGGCCG	mmu-miR-138,1,a|mm|chrX_r	mmu-miR-138.1.a
mmu-miR-140*	25	12	13	0	4	8	3	6	4	ACCACAGGGTAGAACCACGGAC	mmu-miR-140*,1,a|mm|chr8:1	mmu-miR-140*.1.
mmu-miR-142-3p	4	3	1	2	1	0	1	0	0	TGTAGTGTTTCCTACTTTATGG	mmu-miR-142-3p,1,a|mm|chr	mmu-miR-142-3p
mmu-miR-143	1	1	0	1	0	0	0	0	0	UGAGAUGAAGCACUGUAGCUC	mmu-miR-143,1,a|mm|chr18:	mmu-miR-143.1.b
mmu-miR-146b	1	1	0	0	0	1	0	0	0	UGAGAACUGAAUUCCAUAGGCU	mmu-miR-146b,1,a|mm|chr19	mmu-miR-146b.3
mmu-miR-148b	1	0	1	0	0	0	1	0	0	UCAGUGCAUCACAGAACUUUGU	mmu-miR-148b,1,a|mm|chr15	No Folding
mmu-miR-149	1	1	0	0	1	0	0	0	0	TCTGGCTCCGTGTCTTCACT	mmu-miR-149,1,a|mm|chr1:9	mmu-miR-149.1.a
mmu-miR-154	1	1	0	0	1	0	0	0	0	UAGGUUAUCCGUGUUGCCUUCG	mmu-miR-154,1,a|mm|chr12:	mmu-miR-154.1.a
mmu-miR-15a	3	2	1	0	0	2	0	0	1	UAGCAGCACAUAAUGGUUUGUG	mmu-miR-15a,1,a|mm|chr14:	mmu-miR-15a.1.a
mmu-miR-15b	54	27	27	5	11	11	14	5	8	UAGCAGCACAUCAUGGUUUACA	mmu-miR-15b,1,a|mm|chr3:6	mmu-miR-15b.1.a
mmu-miR-16	87	42	45	14	11	17	11	14	20	UAGCAGCACGUAAAUAUUGGCG	mmu-miR-16,1,a|mm|chr3:68	mmu-miR-16.1.a|
mmu-miR-17	61	28	33	9	11	8	10	12	11	CAAAGTGCTTACAGTGCAGGTA	mmu-miR-17,1,a|mm|chr14:1	No Folding
mmu-miR-17*	1	0	1	0	0	0	1	0	0	ACUGCAGUGAGGGCACUUGUAG	mmu-miR-17*,1,a|mm|chr14:1	No Folding
mmu-miR-181c	1	0	1	0	0	0	1	0	0	AACAUUCAACCUGUCGGUGAGU	mmu-miR-181c,1,a|mm|chr8:2	mmu-miR-181c.1
mmu-miR-182	2	2	0	0	1	1	0	0	0	TTTGGCAATGGTAGAACTCAC	mmu-miR-182,1,a|mm|chr6:3	mmu-miR-182.1.a
mmu-miR-183	4	1	3	0	1	0	2	0	1	UAUGGCACUGGUAGAAUUCACU	mmu-miR-183,1,a|mm|chr6:3	mmu-miR-183.1.a
mmu-miR-187	1	1	0	0	0	1	0	0	0	UCGUGUCUUGUGUUGCAGCCGG	mmu-miR-187,1,a|mm|chr18:	mmu-miR-187.1.b
mmu-miR-188-5p	1	1	0	0	0	1	0	0	0	CAUCCCUUGCAUGGUGGAGGG	mmu-miR-188-5p,1,a|mm|chr	mmu-miR-188-5p
mmu-miR-18a	33	14	19	8	5	1	9	5	5	TAAGGTGCATCTAGTGCAGATA	mmu-miR-18a,1,a|mm|chr14:	mmu-miR-18a.1.a
mmu-miR-191	15	4	11	2	2	0	5	3	3	CAACGGAATCCCAAAAGCAGCT	mmu-miR-191,1,a|mm|chr9:1	mmu-miR-191.1.a
mmu-miR-192	1	1	0	1	0	0	0	0	0	CUGACCUAUGAAUUGACAGCC	mmu-miR-192,1,a|mm|chr19:	mmu-miR-192.1.a
mmu-miR-193	2	0	2	0	0	0	0	2	0	AACUGGCCUACAAAGUCCCAGU	mmu-miR-193,1,a|mm|chr11:	mmu-miR-193.1.b
mmu-miR-196a	2	1	1	0	0	1	1	0	0	UAGGUAGUUUCAUGUUGUUGGG	mmu-miR-196a,1,a|mm|chr15	mmu-miR-196a.1
mmu-miR-199a-5p	1	1	0	0	0	0	1	0	0	CCCAGUGUUCAGACUACCUGUUC	mmu-miR-199a-5p,1,a|mm|ch	mmu-miR-199a-5
mmu-miR-19a	18	12	6	5	4	3	2	2	2	TGTGCAAATCTATGCAAAACTG	mmu-miR-19a,1,a|mm|chr14:	mmu-miR-19a.1.b
mmu-miR-19b	61	34	27	12	8	14	11	5	11	TGTGCAAATCCATGCAAAACTG	mmu-miR-19b,1,a|mm|chrX:5	mmu-miR-19b.1.b
mmu-miR-200a	6	1	5	0	0	1	3	0	2	UAACACUGUCUGGUAACGAUGU	mmu-miR-200a,1,a|mm|chr4:	mmu-miR-200a.1
mmu-miR-200b	4	1	3	0	1	0	0	1	2	UAAUACUGCCUGGUAAUGAUGA	mmu-miR-200b,1,a|mm|chr4:	mmu-miR-200b.1

mmu-miR-200c	2	0	2	0	0	0	0	0	2	TAATACTGCCGGGTAATGATGG	mmu-miR-200c,1,a|mm|chr6:	mmu-miR-200c.1
mmu-miR-203	4	0	4	0	0	0	0	2	2	GUGAAAUGUUUAGGACCACUAG	mmu-miR-203,1,a|mm|chr12:	mmu-miR-203.1.b
mmu-miR-20a	268	133	135	39	51	43	35	40	60	TAAAGTGCTTATAGTGCAGGTA	mmu-miR-20a,1,a|mm|chr14:	No Folding
mmu-miR-20b	8	6	2	2	1	3	1	1	0	CAAAGTGCTCATAGTGCAGGTA	mmu-miR-20b,1,a|mm|chrX:5	mmu-miR-20b.1.a
mmu-miR-21	22	11	11	3	5	3	5	4	2	UAGCUUAUCAGACUGAUGUUGA	mmu-miR-21,1,a|mm|chr11:8	mmu-miR-21.1.a|
mmu-miR-210	13	4	9	1	3	0	1	2	6	CUGUGCGUGUGACAGCGGCUGA	mmu-miR-210,1,a|mm|chr7:1	mmu-miR-210.1.b
mmu-miR-221	2	2	0	1	0	1	0	0	0	AGCTACATTGTCTGCTGGGTTT	mmu-miR-221,1,a|mm|chrX:1	mmu-miR-221.1.b
mmu-miR-222	2	2	0	0	2	0	0	0	0	AGCUACAUCUGGCUACUGGGU	mmu-miR-222,1,a|mm|chrX:1	mmu-miR-222.1.b
mmu-miR-23a	15	2	13	0	2	0	2	3	8	ATCACATTGCCAGGGATTTCCA	mmu-miR-23a,1,a|mm|chr8:8	mmu-miR-23a.1.b
mmu-miR-24	8	3	5	1	1	1	0	2	3	UGGCUCAGUUCAGCAGGAACAG	mmu-miR-24,1,a|mm|chr8:86	mmu-miR-24.1.b|
mmu-miR-25	18	9	9	7	2	0	6	1	2	CAUUGCACUUGUCUCGGUCUGA	mmu-miR-25,1,a|mm|chr5:13	No Folding
mmu-miR-26a	4	2	2	0	2	0	1	0	1	UUCAAGUAAUCCAGGAUAGGCU	mmu-miR-26a,1,a|mm|chr9:1	mmu-miR-26a.1.a
mmu-miR-26b	1	0	1	0	0	0	0	1	0	TTCAAGTAATTCAGGATAGGTT	mmu-miR-26b,1,a|mm|chr1:7	mmu-miR-26b.1.a
mmu-miR-27a	2	2	0	1	0	1	0	0	0	UUCACAGUGGCUAAGUUCCGC	mmu-miR-27a,1,a|mm|chr8:8	mmu-miR-27a.1.b
mmu-miR-27b	2	2	0	0	0	2	0	0	0	UUCACAGUGGCUAAGUUCUGC	mmu-miR-27b,1,a|mm|chr13:	mmu-miR-27b.1.b
mmu-miR-290-5p	78	50	28	18	23	9	14	9	4	ACUCAAACUAUGGGGGCACUUU	mmu-miR-290-5p,1,a|mm|chr	mmu-miR-290-5p
mmu-miR-291a-3p	308	162	146	39	50	73	57	49	40	AAAGUGCUUCCACUUUGUGUGC	mmu-miR-291a-3p,1,a|mm|ch	mmu-miR-291a-3
mmu-miR-291a-5p	47	25	22	6	10	9	7	9	6	CAUCAAAGUGGAGGCCCUCUCU	mmu-miR-291a-5p,1,a|mm|ch	mmu-miR-291a-5
mmu-miR-291b-3p	2	1	1	0	1	0	1	0	0	AAAGUGCAUCCAUUUUGUUUGU	mmu-miR-291b-3p,1,a|mm|ch	mmu-miR-291b-3
mmu-miR-291b-5p	1	1	0	0	1	0	0	0	0	GAUCAAAGUGGAGGCCCUCUCC	mmu-miR-291b-5p,1,a|mm|ch	mmu-miR-291b-5
mmu-miR-292-3p	348	178	170	43	68	67	69	64	37	AAAGTGCCGCCAGGTTTTGAGT	mmu-miR-292-3p,1,a|mm|chr	mmu-miR-292-3p
mmu-miR-292-5p	90	59	31	17	31	11	19	7	5	ACUCAAACUGGGGGCUCUUUUG	mmu-miR-292-5p,1,a|mm|chr	mmu-miR-292-5p
mmu-miR-293	41	19	22	8	6	5	12	7	3	AGUGCCGCAGAGUUUGUAGUGU	mmu-miR-293,1,a|mm|chr7:3	mmu-miR-293.1.b
mmu-miR-294	1469	826	643	262	270	294	236	202	205	AAAGUGCUUCCCUUUUGUGUGU	mmu-miR-294,1,a|mm|chr7:3	mmu-miR-294.1.b
mmu-miR-295	882	446	436	145	136	165	159	136	141	AAAGTGCTACTACTTTTGAGTC	mmu-miR-295,1,a|mm|chr7:3	mmu-miR-295.1.b
mmu-miR-296-5p	1	0	1	0	0	0	0	1	0	AGGGCCCCCCCUCAAUCCUGU	No Blast hit	No Blast hit
mmu-miR-297a	12	3	9	1	2	0	3	0	6	AUGUAUGUGUGCAUGUGCAUGU	mmu-miR-297a,1,a|mm|chr7:	mmu-miR-297a.1
mmu-miR-297b-5p	1	0	1	0	0	0	0	0	1	AUGUAUGUGUGCAUGAACAUGU	mmu-miR-297b-5p,1,a|mm|ch	mmu-miR-297b-5
mmu-miR-29a	1	0	1	0	0	0	0	1	0	UAGCACCAUCUGAAAUCGGUUA	mmu-miR-29a,1,a|mm|chr6:3	mmu-miR-29a.1.b
mmu-miR-29b	2	0	2	0	0	0	0	1	1	UAGCACCAUUUGAAAUCAGUGUU	mmu-miR-29b,1,a|mm|chr6:3	mmu-miR-29b.1.b
mmu-miR-301a	3	1	2	0	0	1	0	1	1	CAGUGCAAUAGUAUUGUCAAAGC	mmu-miR-301a,1,a|mm|chr11	mmu-miR-301a.1
mmu-miR-302c	2	1	1	0	1	0	0	1	0	AAGUGCUUCCAUGUUUCAGUGG	mmu-miR-302c,1,a|mm|chr3:	mmu-miR-302c.1
mmu-miR-30a	2	1	1	1	0	0	1	0	0	UGUAAACAUCCUCGACUGGAAG	mmu-miR-30a,1,a|mm|chr1:2	mmu-miR-30a.1.a
mmu-miR-30b	6	3	3	1	2	0	0	1	2	UGUAAACAUCCUACACUCAGCU	mmu-miR-30b,1,a|mm|chr15:	mmu-miR-30b.1.a
mmu-miR-30c	14	11	3	4	3	4	2	1	0	TGTAAACATCCTACACTCTCAG	mmu-miR-30c,1,a|mm|chr4:1	mmu-miR-30c.2.a
mmu-miR-30d	1	0	1	0	0	0	0	0	1	UGUAAACAUCCCCGACUGGAAG	mmu-miR-30d,1,a|mm|chr15:	mmu-miR-30d.1.a
mmu-miR-30e	1	0	1	0	0	0	0	0	1	UGUAAACAUCCUUGACUGGAAG	mmu-miR-30e,1,a|mm|chr4:1	mmu-miR-30e.1.a
mmu-miR-320	2	1	1	0	1	0	0	1	0	AAAAGCUGGGUUGAGAGGGCGA	mmu-miR-320,1,a|mm|chr14:	mmu-miR-320.1.b
mmu-miR-324-3p	1	0	1	0	0	0	0	1	0	CCACTGCCCCAGGTGCTGCTGG	mmu-miR-324-3p,1,a|mm|chr	mmu-miR-324-3p
mmu-miR-324-5p	1	1	0	0	1	0	0	0	0	CGCAUCCCCUAGGGCAUUGGUGU	mmu-miR-324-5p,1,a|mm|chr	mmu-miR-324-5p
mmu-miR-331-3p	1	1	0	0	0	1	0	0	0	GCCCCUGGGCCUAUCCUAGAA	mmu-miR-331-3p,1,a|mm|chr	mmu-miR-331-3p
mmu-miR-342-3p	1	1	0	0	1	0	0	0	0	UCUCACACAGAAAUCGCACCCGU	mmu-miR-342-3p,1,a|mm|chr	mmu-miR-342-3p
mmu-miR-345-5p	2	1	1	1	0	0	0	0	1	GCUGACCCCUAGUCCAGUGCUU	mmu-miR-345-5p,1,a|mm|chr	No Folding
mmu-miR-34a	121	70	51	31	28	11	17	19	15	UGGCAGUGUCUUAGCUGGUUGU	mmu-miR-34a,1,a|mm|chr4:1	mmu-miR-34a.1.a
mmu-miR-350	2	2	0	0	1	1	0	0	0	UUCACAAAGCCCAUACACUUUC	mmu-miR-350,1,a|mm|chr1:1	mmu-miR-350.1.b

mmu-miR-362-5p	3	0	3	0	0	0	0	0	3	AATCCTTGGAACCTAGGTGTGA	mmu-miR-362-5p,1,a|mm|chr	mmu-miR-362-5p
mmu-miR-409-3p	1	1	0	0	0	1	0	0	0	GAAUGUUGCUCGGUGAACCCCU	mmu-miR-409-3p,1,a|mm|chr	mmu-miR-409-3p
mmu-miR-429	1	1	0	0	1	0	0	0	0	UAAUACUGUCUGGUAAUGCCGU	mmu-miR-429,1,a|mm|chr4:1	mmu-miR-429.1.b
mmu-miR-450a-5p	1	1	0	0	1	0	0	0	0	UUUUGCGAUGUGUUCCUAAUAU	mmu-miR-450a-5p,1,a|mm|ch	mmu-miR-450a-5
mmu-miR-466a-3p	7	4	3	2	2	0	0	1	2	TATACATACACGCACACATAAG	mmu-miR-466a-3p,1,a|mm|ch	mmu-miR-466a-3
mmu-miR-467a	9	2	7	1	1	0	1	2	4	UAAGUGCCUGCAUGUAUAUGCG	mmu-miR-467a,1,a|mm|chr2:	mmu-miR-467a.1
mmu-miR-467a-1*	22	2	20	1	1	0	3	4	13	AUAUACAUACACACACCUACAC	mmu-miR-467a-1*,1,a|mm|ch	mmu-miR-467a-1
mmu-miR-484	1	1	0	0	0	1	0	0	0	UCAGGCUCAGUCCCCUCCCGAU	mmu-miR-484,1,a|mm|chr16:	No Folding
mmu-miR-497	1	1	0	1	0	0	0	0	0	CAGCAGCACACUGUGGUUUGUA	mmu-miR-497,1,a|mm|chr11:7	mmu-miR-497.1.a
mmu-miR-500	1	0	1	0	0	0	0	1	0	AAUGCACCUGGGCAAGGGUUCA	mmu-miR-500,1,a|mm|chrX:6	mmu-miR-500.1.b
mmu-miR-501-3p	1	0	1	0	0	0	1	0	0	AAUGCACCCGGGCAAGGAUUUG	mmu-miR-501-3p,1,a|mm|chr	mmu-miR-501-3p
mmu-miR-503	6	0	6	0	0	0	0	2	4	TAGCAGCGGGAACAGTACTGCA	mmu-miR-503,1,a|mm|chrX:5	mmu-miR-503.1.a
mmu-miR-532-5p	1	0	1	0	0	0	1	0	0	CAUGCCUUGAGUGUAGGACCGU	mmu-miR-532-5p,1,a|mm|chr	mmu-miR-532-5p
mmu-miR-669a	13	4	9	0	3	1	3	2	4	AGUUGUGUGUGCAUGUUCAUGU	mmu-miR-669a,1,a|mm|chr2:	mmu-miR-669a.1
mmu-miR-669b	12	4	8	1	2	1	1	4	3	AGUUUUGUGUGCAUGUGCAUGU	mmu-miR-669b,1,a|mm|chr2:	mmu-miR-669b.1
mmu-miR-669c	8	1	7	1	0	0	2	2	3	AUAGUUGUGUGUGGAUGUGUGU	mmu-miR-669c,1,a|mm|chr2:1	mmu-miR-669c.1
mmu-miR-672	10	9	1	1	4	4	0	1	0	TGAGGTTGGTGTACTGTGTGT	mmu-miR-672,1,a|mm|chrX:1	mmu-miR-672.1.a
mmu-miR-674	5	4	1	1	1	2	1	0	0	GCACUGAGAUGGGAGUGGUGUA	mmu-miR-674,1,a|mm|chr2:1	mmu-miR-674.1.a
mmu-miR-676	1	1	0	0	0	1	0	0	0	CCGUCCUGAGGUUGUUGAGCU	mmu-miR-676,1,a|mm|chrX:9	mmu-miR-676.1.b
mmu-miR-690	1	0	1	0	0	0	0	0	1	AAAGGCUAGGCUCACAACCAAA	mmu-miR-690,1,a|mm|chrY:5	mmu-miR-690.27
mmu-miR-7a	55	36	19	15	15	6	9	4	6	TGGAAGACTAGTGATTTTGTTG	mmu-miR-7a,1,a|mm|chr7:86	mmu-miR-7a.1.a|
mmu-miR-92a	23	16	7	9	4	3	4	1	2	TATTGCACTTGTCCCGGCCTGA	mmu-miR-92a,1,a|mm|chrX:5	mmu-miR-92a.1.b
mmu-miR-93	104	63	41	24	19	20	10	21	10	CAAAGTGCTGTTCGTGCAGGTA	mmu-miR-93,1,a|mm|chr5:13	mmu-miR-93.1.a|
mmu-miR-96	3	1	2	1	0	0	0	1	1	UUUGGCACUAGCACAUUUUUGCU	mmu-miR-96,1,a|mm|chr6:30	mmu-miR-96.1.a|
mmu-miR-99b	20	6	14	4	0	2	2	4	8	CACCCGUAGAACCGACCUUGCG	mmu-miR-99b,1,a|mm|chr17:	mmu-miR-99b.1.a
mmu-miR-106b*	10	5	5	2	0	3	2	2	1	CCGCACUGUGGGUACUUGCUGC	mmu-miR-106b*,1,a|mm|chr5	mmu-miR-106b*.1
mmu-miR-124*	1	1	0	0	1	0	0	0	0	CGUGUUCACAGCGGACCUUGAU	mmu-miR-124*,1,a|mm|chr3:1	mmu-miR-124*.1.
mmu-miR-125a-3p	1	0	1	0	0	0	1	0	0	ACAGGUGAGGUUCUUGGGAGCC	mmu-miR-125a-3p,1,a|mm|ch	mmu-miR-125a-3
mmu-miR-130b*	1	1	0	0	0	1	0	0	0	ACUCUUUCCCUGUUGCACUACU	mmu-miR-130b*,1,a|mm|chr1	mmu-miR-130b*.1
mmu-miR-15a*	1	0	1	0	0	0	0	1	0	CAGGCCAUACUGUGCUGCCUCA	mmu-miR-15a*,1,a|mm|chr14	mmu-miR-15a*.1.
mmu-miR-15b*	1	1	0	0	0	1	0	0	0	CGAAUCAUUAUUUGCUGCUCUA	mmu-miR-15b*,1,a|mm|chr3:6	mmu-miR-15b*.1.
mmu-miR-181d	1	1	0	1	0	0	0	0	0	AACAUUCAUUGUUGUCGGUGGGU	mmu-miR-181d,1,a|mm|chr8:8	mmu-miR-181d.1
mmu-miR-183*	1	1	0	1	0	0	0	0	0	GUGAAUUACCGAAGGGCCAUAA	mmu-miR-183*,1,a|mm|chr6:3	mmu-miR-183*.1.
mmu-miR-1839-3p	1	1	0	0	1	0	0	0	0	AGACCTACTTATCTACCAACA	mmu-miR-1839-3p,1,a|mm|ch	mmu-miR-1839-3
mmu-miR-18a*	2	1	1	1	0	0	0	0	1	ACUGCCCUAAGUGCUCCUUCUG	mmu-miR-18a*,1,a|mm|chr14	mmu-miR-18a*.1.
mmu-miR-1965	2	2	0	1	1	0	0	0	0	AAGCCGGGCCGTAGTGGCGCAT	No Blast hit	No Blast hit
mmu-miR-1983	5	4	1	1	1	2	0	0	1	CUCACCUGGAGCAUGUUUUCU	mmu-miR-1983,1,a|mm|chr13	mmu-miR-1983.1
mmu-miR-200a*	1	0	1	0	0	0	1	0	0	CAUCUUACCGGACAGUGCUGGA	mmu-miR-200a*,1,a|mm|chr4	mmu-miR-200a*.1
mmu-miR-20a*	1	0	1	0	0	0	0	1	0	ACUGCAUUACGAGCACUUAAAG	mmu-miR-20a*,1,a|mm|chr14	mmu-miR-20a*.1.
mmu-miR-27b*	2	1	1	1	0	0	1	0	0	AGAGCUUAGCUGAUUGGUGAAC	mmu-miR-27b*,1,a|mm|chr13	mmu-miR-27b*.1.
mmu-miR-290-3p	1	0	1	0	0	0	1	0	0	AAAGUGCCGCCUAGUUUUAAGCCC	mmu-miR-290-3p,1,a|mm|chr	mmu-miR-290-3p
mmu-miR-293*	19	10	9	1	7	2	2	6	1	ACUCAAACUGUGUGACAUUUUG	mmu-miR-293*,1,a|mm|chr7:3	No Folding
mmu-miR-294*	9	5	4	0	2	3	2	2	0	ACUCAAAAUGGAGGCCCUAUCU	mmu-miR-294*,1,a|mm|chr7:3	mmu-miR-294*.1.
mmu-miR-297b-3p	2	0	2	0	0	0	0	1	1	UAUACAUACACACAUACCCAUA	mmu-miR-297b-3p,1,a|mm|ch	mmu-miR-297b-3
mmu-miR-29a*	1	0	1	0	0	0	0	0	1	ACUGAUUUCUUUUGGUGUUCAG	mmu-miR-29a*,1,a|mm|chr6:3	mmu-miR-29a*.1.

mmu-miR-301b	1	1	0	0	1	0	0	0	0	CAGTGCAATGGTATTGTCAAAG	mmu-miR-301b,1,a|mm|chr1	No Folding
mmu-miR-362-3p	1	1	0	0	1	0	0	0	0	AACACACCUGUUCAAGGAUUCA	mmu-miR-362-3p,1,a|mm|chr	mmu-miR-362-3p
mmu-miR-425	4	4	0	2	1	1	0	0	0	AATGACACGATCACTCCCGTTG	mmu-miR-425,1,a|mm|chr9:1	mmu-miR-425.1.a
mmu-miR-466a-5p	3	0	3	0	0	0	1	1	1	UAUGUGUGUGUACAUGUACAUA	mmu-miR-466a-5p,1,a|mm|ch	mmu-miR-466a-5
mmu-miR-466d-3p	1	0	1	0	0	0	0	1	0	UAUACAUACACGCACACAUAG	mmu-miR-466d-3p,1,a|mm|ch	mmu-miR-466d-3
mmu-miR-466d-5p	2	0	2	0	0	0	0	2	0	UGUGUGUGCGUACAUGUACAUG	mmu-miR-466d-5p,1,a|mm|ch	mmu-miR-466d-5
mmu-miR-466e-5p	1	0	1	0	0	0	0	1	0	GAUGUGUGUGUACAUGUACAUA	mmu-miR-466e-5p,1,a|mm|ch	mmu-miR-466e-5
mmu-miR-466f	8	5	3	2	2	1	0	1	2	ACGUGUGUGUGCAUGUGCAUGU	mmu-miR-466f,1,a|mm|chr9:4	mmu-miR-466f.2.
mmu-miR-466f-3p	1	0	1	0	0	0	1	0	0	CAUACACACACACAUACACAC	No Blast hit	No Blast hit
mmu-miR-466h	9	3	6	1	0	2	0	6	0	UGUGUGCAUGUGCUUGUGUGUA	mmu-miR-466h,1,a|mm|chr8:5	mmu-miR-466h.6
mmu-miR-466j	2	1	1	0	1	0	0	1	0	UGUGUGCAUGUGCAUGUGUGUAA	mmu-miR-466j,1,a|mm|chr7:1	mmu-miR-466j.1.l
mmu-miR-467c	13	2	11	1	1	0	3	4	4	UAAGUGCGUGCAUGUAUAUGUG	mmu-miR-467c,1,a|mm|chr2:1	No Folding
mmu-miR-467d	5	1	4	0	0	1	0	0	4	UAAGUGCGCGCAUGUAUAUGCG	mmu-miR-467d,1,a|mm|chr2:	No Folding
mmu-miR-467e	16	3	13	2	1	0	4	2	7	AUAAGUGUGAGCAUGUAUAUGU	mmu-miR-467e,1,a|mm|chr2:	No Folding
mmu-miR-669d	10	2	8	1	1	0	0	5	3	ACUUGUGUGUGCAUGUAUAUGU	mmu-miR-669d,1,a|mm|chr2:	mmu-miR-669d.1
mmu-miR-669l	3	0	3	0	0	0	1	1	1	AGUUGUGUGUGCAUGUAUAUGU	mmu-miR-669l,1,a|mm|chr2:1	mmu-miR-669l.1.a
mmu-miR-669o	3	1	2	1	0	0	0	2	0	UAGUUGUGUGUGCAUGUUUAUGU	mmu-miR-669o,1,a|mm|chr2:	mmu-miR-669o.1
mmu-miR-708	25	20	5	9	4	7	2	3	0	AAGGAGCTTACAATCTAGCTGG	mmu-miR-708,1,a|mm|chr7:1	mmu-miR-708.1.a
mmu-miR-744	3	1	2	0	0	1	2	0	0	UGCGGGGCUAGGGCUAACAGCA	mmu-miR-744,1,a|mm|chr11:	mmu-miR-744.1.a
mmu-miR-7a*	1	1	0	0	1	0	0	0	0	CAACAAAUCACAGUCUGCCAUA	mmu-miR-7a*,1,a|mm|chr13:5	mmu-miR-7a*.1.b
mmu-miR-872	1	0	1	0	0	0	0	1	0	AAGGUUACUUGUUAGUUCAGG	mmu-miR-872,1,a|mm|chr4:9	mmu-miR-872.1.a
mmu-miR-466b-5p	1	1	0	0	1	0	0	0	0	GAUGUGUGUGUACAUGUACAUG	mmu-miR-466b-5p,1,a|mm|ch	mmu-miR-466b-5
mmu-miR-466f-5p	1	0	1	0	0	0	0	1	0	UACGUGUGUGUGCAUGUGCAUG	mmu-miR-466f-5p,1,a|mm|ch	mmu-miR-466f-5p
mmu-miR-466k	6	4	2	3	0	1	0	1	1	TGTGTGTGTACATGTACATGTG	mmu-miR-466k,1,a|mm|chrX:	mmu-miR-466k.1
miR-181c*	2	1	1	0	0	1	1	0	0	AACCAUCGACCGUUGAGUGGAC	miR-181c*,1,a|mm|chr8:8670	miR-181c*.1.b|mr
miR-19b-1*	1	1	0	0	1	0	0	0	0	AGUUUUGCAGGUUUGCAUCCAGC	miR-19b-1*,1,a|mm|chr14:11	miR-19b-1*.1.a|m
miR-2889	1	1	0	1	0	0	0	0	0	GUCCCCGGGGCUCCCGCCGGC	No Blast hit	No Blast hit
miRNA_plate1_ES-C0_46-1	1	1	0	1	0	0	0	0	0	AAATGTGTAGCCCAGGCTGG	miRNA_plate1_ES-C0_46-1,1	miRNA_plate1_ES
miRNA_plate10_ES-T48_57-5	1	0	1	0	0	0	0	1	0	CGGGGTCCGGTGCGGAGAGCCG	miRNA_plate10_ES-T48_57-	miRNA_plate10_E
miRNA_plate11_ES-T72_47-7	1	0	1	0	0	0	0	0	1	ACTGACTGCTCTTCTGAAGGTCT	miRNA_plate11_ES-T72_47-	miRNA_plate11_E
miRNA_plate12_ES-T72_40-2	5	3	2	1	0	2	0	0	2	AGGCCAGCCTGGGCTACACATTTA	miRNA_plate12_ES-T72_40-	miRNA_plate12_E
miRNA_plate12_ES-T72_66-4	1	0	1	0	0	0	0	0	1	ATATACATACACACACCTATAC	miRNA_plate12_ES-T72_66-	miRNA_plate12_E
miRNA_plate13_ES-T24_14-1	1	0	1	0	0	0	1	0	0	CCGAGCCATCTCTTGTCCCTCAGA	miRNA_plate13_ES-T24_14-	miRNA_plate13_E
miRNA_plate4_ES-C24_43-12	1	1	0	0	1	0	0	0	0	TGTGTGGAAGGCAGGGGTGCGG	miRNA_plate4_ES-C24_43-1	miRNA_plate4_ES
miRNA_plate5_ES-C48_21-7	1	1	0	0	0	1	0	0	0	AGCCACTGCCCACCGCACACTG	miRNA_plate5_ES-C48_21-7	miRNA_plate5_ES
miRNA_plate6_ES-C48_14-5	1	1	0	0	0	1	0	0	0	TCGGGGTTTCGTACGTAGCA	miRNA_plate6_ES-C48_14-5	miRNA_plate6_ES
miRNA_plate6_ES-C48_70-3	1	1	0	0	0	1	0	0	0	AGCCCTGAAAATGGATGGCGCTG	miRNA_plate6_ES-C48_70-3	miRNA_plate6_ES
miRNA_plate7_ES-T24_70-3	1	0	1	0	0	0	1	0	0	TCGAGCCCCACGTTGGGCGCCA	miRNA_plate7_ES-T24_70-3	miRNA_plate7_ES
miRNA_plate7_ES-T24_71-1	10	2	8	0	1	1	1	5	2	ACATAACATACACACACACGTA	miRNA_plate7_ES-T24_71-1	miRNA_plate7_ES
miRNA_plate8_ES-T24_72-2	1	0	1	0	0	0	1	0	0	AGGATATGATGTGAGGCCTGATC	miRNA_plate8_ES-T24_72-2	miRNA_plate8_ES
miRNA_plate9_ES-T48_40-4	5	3	2	2	1	0	0	1	1	AGTTGTGTGTGCATGTGCATGT	miRNA_plate9_ES-T48_40-4	miRNA_plate9_ES
miRNA_plate9_ES-T48_57-4	1	0	1	0	0	0	0	1	0	TAAAATGTGCAAGTTCCCAAGTTC	miRNA_plate9_ES-T48_57-4	miRNA_plate9_ES
miRNA_plate10_ES-T48_23-2	1	0	1	0	0	0	0	1	0	GCGGCGTCCGGCAGTGTC	miRNA_plate10_ES-T48_23-	miRNA_plate10_E
miRNA_plate10_ES-T48_64-5	3	0	3	0	0	0	0	2	1	TATGCCTGGGCAGGGCGAA	miRNA_plate10_ES-T48_64-	miRNA_plate10_E
miRNA_plate11_ES-T72_06-1	1	0	1	0	0	0	0	0	1	ATCCCGGACGAGCCCCCA	miRNA_plate11_ES-T72_06-	miRNA_plate11_E

miRNA_plate11_ES-T72_28-3	2	0	2	0	0	0	0	0	2	CTGGCGCTAAACCATTCG	miRNA_plate11_ES-T72_28-	miRNA_plate11_I
miRNA_plate11_ES-T72_95-2	1	0	1	0	0	0	0	0	1	CCCGGGTGGAGCCGCCGCA	No Blast hit	No Blast hit
miRNA_plate12_ES-T72_69-1	6	3	3	1	0	2	1	1	1	TGGTGGGTGCATATGTTTT	miRNA_plate12_ES-T72_69-	miRNA_plate12_I
miRNA_plate2_ES-C0_06-1	1	1	0	1	0	0	0	0	0	AGGCAACTCAGAGGTCCG	miRNA_plate2_ES-C0_06-1,'	miRNA_plate2_E
miRNA_plate3_ES-C24_05-4	1	1	0	0	1	0	0	0	0	TCCCTGTCCTCCAGGAGCTCACAA	No Blast hit	No Blast hit
miRNA_plate3_ES-C24_64-2	2	2	0	0	1	1	0	0	0	TTGAAAGTCAGCCCTCGACACAA	miRNA_plate3_ES-C24_64-2	No Folding
miRNA_plate5_ES-C48_94-4	7	5	2	1	1	3	0	1	1	GTCCCTGTTCGGGCGCCA	miRNA_plate5_ES-C48_94-4	miRNA_plate5_E
miRNA_plate6_ES-C48_01-4	1	1	0	0	0	1	0	0	0	ACCGTGTTCGCAGTGCCCG	miRNA_plate6_ES-C48_01-4	miRNA_plate6_E
miRNA_plate6_ES-C48_52-1	1	1	0	0	0	1	0	0	0	TTGACCTCCAGCCTGTGAT	miRNA_plate6_ES-C48_52-1	miRNA_plate6_E
miRNA_plate7_ES-T24_32-6	2	0	2	0	0	0	1	0	1	CTCCCCGCGGGGCCTCGCTCCA	No Blast hit	No Blast hit
miRNA_plate8_ES-T24_26-1	1	0	1	0	0	0	1	0	0	TCTGATCGAGGCCCAGCC	miRNA_plate8_ES-T24_26-1	miRNA_plate8_E
miRNA_plate8_ES-T24_32-5	1	0	1	0	0	0	1	0	0	CGGCCCCGCCATGCCCGACC	miRNA_plate8_ES-T24_32-5	No Folding
miRNA_plate1_ES-C0_55-2	134	73	61	26	22	25	15	21	25	GAAGCGGGTGCTTCACTTTTTT	miRNA_plate1_ES-C0_55-2,'	No Folding
miRNA_plate1_ES-C0_82-5	2	2	0	2	0	0	0	0	0	AGTGCCGCAAAGTTTGTAGTGT	No Blast hit	No Blast hit
miRNA_plate10_ES-T48_15-1	6	4	2	4	0	0	0	2	0	ACTCAAACTGGGGGCTCTTTTT	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_03-9	3	0	3	0	0	0	0	1	2	TCGCCTCTCCCGAGGTGCG	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_19-4	74	35	39	14	7	14	5	20	14	TACCTGGTTGATCCTGCCAGTA	miRNA_plate11_ES-T72_19-	No Folding
miRNA_plate11_ES-T72_27-2	4	1	3	0	0	1	0	0	3	CACCCCCCTCCGTCGCCTCTC	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_46-6	5	1	4	0	1	0	1	0	3	CAAAGTGCTGTTCGTGCAGGTAA	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_49-1	3	2	1	2	0	0	0	0	1	ACCGCGTGCATTCCTTTT	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_79-2	2	0	2	0	0	0	0	0	2	AGCGCCGAAGGAGCCTCG	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_80-1	2	0	2	0	0	0	0	0	2	CGGTGGCCGGACCATTGC	No Blast hit	No Blast hit
miRNA_plate12_ES-T72_04-3	2	1	1	0	0	1	0	0	1	CCCGGGGAGCCCGGCGGG	No Blast hit	No Blast hit
miRNA_plate12_ES-T72_19-1	5	1	4	0	0	1	1	1	2	AAGGAAGGCAGCAGGCGCGC	miRNA_plate12_ES-T72_19-	No Folding
miRNA_plate12_ES-T72_21-3	2	0	2	0	0	0	1	0	1	AAAGGGCTTCCCTTTGGG	No Blast hit	No Blast hit
miRNA_plate12_ES-T72_32-7	5	1	4	0	1	0	0	0	4	ATACAGACACATGCACACACTA	No Blast hit	No Blast hit
miRNA_plate12_ES-T72_91-3	3	1	2	0	0	1	1	0	1	ACTCAAACTATGGGGGCACTTA	No Blast hit	No Blast hit
miRNA_plate13_ES-T24_20-3	2	0	2	0	0	0	2	0	0	AAAGGGCTTCCCTTTTGGGTGTCG	No Blast hit	No Blast hit
miRNA_plate13_ES-T24_20-4	2	0	2	0	0	0	2	0	0	AAAGGGCTTCCCTTTTGGGGG	No Blast hit	No Blast hit
miRNA_plate13_ES-T72_44-3	3	1	2	0	0	1	1	0	1	CATCAAAGTGGAGGCCCTCTCA	No Blast hit	No Blast hit
miRNA_plate15_ES_C48_19-8	6	4	2	1	2	1	0	2	0	AAGTGCCGCCAGGTTTTGAGTGA	No Blast hit	No Blast hit
miRNA_plate15_ES_C48_27-5	2	2	0	0	0	2	0	0	0	ACACTCAAAAATAATAACTCTTT	No Blast hit	No Blast hit
miRNA_plate2_ES-C0_11-5	4	4	0	1	2	1	0	0	0	CAGTGCAATGTTAAAAGGGCAC	No Blast hit	No Blast hit
miRNA_plate2_ES-C0_16-9	4	3	1	1	1	1	1	0	0	CGCCACTACTGCCCGGCTCCTT	miRNA_plate2_ES-C0_16-9,'	No Folding
miRNA_plate2_ES-C0_73-6	4	2	2	1	0	1	1	1	0	GTGGTGTGCTAGTTAATTTTT	miRNA_plate2_ES-C0_73-6,'	No Folding
miRNA_plate3_ES-C24_24-3	5	2	3	0	2	0	3	0	0	GCCCACCCAGGGACGCCA	No Blast hit	No Blast hit
miRNA_plate3_ES-C24_48-12	2	2	0	0	1	1	0	0	0	ACTAATCCTAGCCCTAGCCCTACA	miRNA_plate3_ES-C24_48-1	No Folding
miRNA_plate4_ES-C24_18-8	2	1	1	0	1	0	0	1	0	TTGGGACTGGTACTCCTTTATTTT	No Blast hit	No Blast hit
miRNA_plate4_ES-C24_65-1	2	2	0	0	2	0	0	0	0	GAAGCGTTTACTTTGAAA	No Blast hit	No Blast hit
miRNA_plate4_ES-C24_75-2	3	3	0	1	2	0	0	0	0	CAAAGTGCTTACAGTGCAGGTAT	No Blast hit	No Blast hit
miRNA_plate4_ES-C24_88-4	3	3	0	1	1	1	0	0	0	CGGTGGATCACTCGGCTCGTG	No Blast hit	No Blast hit
miRNA_plate5_ES-C48_50-3	2	2	0	0	0	2	0	0	0	ACATAGCATTGCGATGGTCAG	No Blast hit	No Blast hit
miRNA_plate5_ES-C48_71-3	54	31	23	6	7	18	8	7	8	CGCGACCTCAGATCAGACGTGG	No Blast hit	No Blast hit
miRNA_plate5_ES-C48_94-12	3	2	1	0	0	2	0	0	1	GTACGCGTTAGGACCCGA	No Blast hit	No Blast hit
miRNA_plate6_ES-C48_22-8	2	1	1	0	0	1	1	0	0	AAAGTGCTTCCCTTTGTGTGT	No Blast hit	No Blast hit

miRNA_plate6_ES-C48_50-1	5	3	2	1	0	2	1	0	1	TATCGTCCTTCTGCCCAGCGAT	miRNA_plate6_ES-C48_50-1	No Folding
miRNA_plate6_ES-C48_69-1	2	2	0	0	0	2	0	0	0	GGCCTTTGGTTTCAGTACTTGGG	No Blast hit	No Blast hit
miRNA_plate7_ES-T24_02-1	24	9	15	1	3	5	5	4	6	GACTCTTAGCGGTGGATCACTC	No Blast hit	No Blast hit
miRNA_plate7_ES-T24_09-4	2	1	1	0	1	0	1	0	0	AAACTGCTTCCCTTTTGTGTGT	No Blast hit	No Blast hit
miRNA_plate7_ES-T24_34-1	2	0	2	0	0	0	1	1	0	TGGCAGTGTCTTAACTGGTTGT	No Blast hit	No Blast hit
miRNA_plate7_ES-T24_56-6	2	1	1	0	1	0	1	0	0	AAAGTGCTTCCCTTTTGTGTGA	No Blast hit	No Blast hit
miRNA_plate7_ES-T24_63-4	10	7	3	1	4	2	2	0	1	TAAAGTGCTTATAGTGCAGGTAT	No Blast hit	No Blast hit
miRNA_plate7_ES-T24_85-1	18	15	3	5	6	4	2	0	1	AGCCTTTAATTTCAGTACTTGG	miRNA_plate7_ES-T24_85-1,	No Folding
miRNA_plate8_ES-T24_03-2	2	1	1	1	0	0	1	0	0	TCCCGTGGATCGCCTCAGC	No Blast hit	No Blast hit
miRNA_plate9_ES-T48_26-5	2	0	2	0	0	0	0	2	0	AAAGTGGTTCCTTTTTGTGTGTCG	No Blast hit	No Blast hit
miRNA_plate9_ES-T48_26-6	2	0	2	0	0	0	0	2	0	AAAGTGGTACTACTTTTGAGTC	No Blast hit	No Blast hit
miRNA_plate9_ES-T48_57-8	2	0	2	0	0	0	1	1	0	ATCCTGCCGACTACGCCA	No Blast hit	No Blast hit
miRNA_plate9_ES-T48_66-6	6	0	6	0	0	0	0	1	5	CTTCGATGTCGGCTCTTCCTA	miRNA_plate9_ES-T48_66-6,	No Folding
miRNA_plate9_ES-T48_71-3	3	2	1	1	0	1	0	1	0	AACGGCTACCGCCTACCTTT	miRNA_plate9_ES-T48_71-3,	No Folding
miRNA_plate9_ES-T48_73-1	2	0	2	0	0	0	0	1	1	CCGCCGGTGAAATACCACTAC	miRNA_plate9_ES-T48_73-1,	No Folding
miRNA_plate9_ES-T48_80-4	2	1	1	0	1	0	0	1	0	ACCCTATGATGATGACTGAGGGTC	No Blast hit	No Blast hit
miRNA_plate9_ES-T48_88-3	21	13	8	2	5	6	2	3	3	AAGTGCTACTACTTTTGAGTCC	No Blast hit	No Blast hit
miRNA_plate1_ES-C0_12-2	1	1	0	1	0	0	0	0	0	TCCCAGAACAGGGTCTCCTTG	miRNA_plate1_ES-C0_12-2,1	No Folding
miRNA_plate1_ES-C0_17-4	1	1	0	1	0	0	0	0	0	AAAGTGCTTCCCTTTTGAGTCT	No Blast hit	No Blast hit
miRNA_plate1_ES-C0_24-3	1	1	0	1	0	0	0	0	0	CGAAGGAGCCTCGGTTGGCC	miRNA_plate1_ES-C0_24-3,1	No Folding
miRNA_plate1_ES-C0_52-2	1	1	0	1	0	0	0	0	0	CGAGCCGGGCCCTTCCCGTGGATT	No Blast hit	No Blast hit
miRNA_plate1_ES-C0_54-4	1	1	0	1	0	0	0	0	0	TTAGTGACGCGCATGAATGGA	miRNA_plate1_ES-C0_54-4,1	No Folding
miRNA_plate1_ES-C0_55-3	1	1	0	1	0	0	0	0	0	CCCACTCCCTCCCCTCCGTGCAG	miRNA_plate1_ES-C0_55-3,1	No Folding
miRNA_plate1_ES-C0_59-1	1	1	0	1	0	0	0	0	0	AAAGTGCATCCATTTTGTTTAA	No Blast hit	No Blast hit
miRNA_plate1_ES-C0_75-3	1	1	0	1	0	0	0	0	0	ACTCAAACTGGGGGCTCTTTAAA	No Blast hit	No Blast hit
miRNA_plate1_ES-C0_78-3	1	1	0	1	0	0	0	0	0	ATCACCTGGCATGTGCCTCTTTA	No Blast hit	No Blast hit
miRNA_plate1_ES-C0_94-4	1	1	0	1	0	0	0	0	0	GACGATGATGATGACGATGAGTCC	miRNA_plate1_ES-C0_94-4,1	No Folding
miRNA_plate10_ES-T48_03-3	1	0	1	0	0	0	0	1	0	TATCAGCACATCATGGTTTAC	No Blast hit	No Blast hit
miRNA_plate10_ES-T48_10-2	1	0	1	0	0	0	0	1	0	CATTCGTAGACGACCTGC	No Blast hit	No Blast hit
miRNA_plate10_ES-T48_11-1	1	0	1	0	0	0	0	1	0	GGCAGCAGGCGCGCAAATT	No Blast hit	No Blast hit
miRNA_plate10_ES-T48_26-4	1	0	1	0	0	0	0	1	0	CTGGGTGTAGGCCAGAACCA	No Blast hit	No Blast hit
miRNA_plate10_ES-T48_27-2	1	0	1	0	0	0	0	1	0	TGGCAGTGTCTTAGCTGGTTAA	No Blast hit	No Blast hit
miRNA_plate10_ES-T48_39-4	1	0	1	0	0	0	0	1	0	TAGCTTATCAGACTGGTGTTGA	No Blast hit	No Blast hit
miRNA_plate10_ES-T48_42-3	1	0	1	0	0	0	0	1	0	GCGCCACTGGTGTACTGGAA	No Blast hit	No Blast hit
miRNA_plate10_ES-T48_42-4	1	0	1	0	0	0	0	1	0	TAACAGCACGTAAATATTGGCG	No Blast hit	No Blast hit
miRNA_plate10_ES-T48_50-7	1	0	1	0	0	0	0	1	0	ACCATGGTGACCACGGGTGACG	miRNA_plate10_ES-T48_50-7	No Folding
miRNA_plate10_ES-T48_61-8	1	0	1	0	0	0	0	1	0	CATCCCACTCAGTCGTGCCC	miRNA_plate10_ES-T48_61-8	No Folding
miRNA_plate10_ES-T48_75-1	1	0	1	0	0	0	0	1	0	CCAGAAAAGGTGTTGGTTGATA	miRNA_plate10_ES-T48_75-	No Folding
miRNA_plate10_ES-T48_76-8	1	0	1	0	0	0	0	1	0	CTCGCCTCGGCCGGCGCC	No Blast hit	No Blast hit
miRNA_plate10_ES-T48_78-6	1	0	1	0	0	0	0	1	0	TCCCCAGCATCTCCACCA	No Blast hit	No Blast hit
miRNA_plate10_ES-T48_81-1	1	0	1	0	0	0	0	1	0	CGCGCATGAATGGATGAACGAGAT	miRNA_plate10_ES-T48_81-	No Folding
miRNA_plate10_ES-T48_84-3	1	0	1	0	0	0	0	1	0	AAAGTGCTTCCCCTTTGTGTGT	No Blast hit	No Blast hit
miRNA_plate10_ES-T48_84-5	1	0	1	0	0	0	0	1	0	CTCCAAGGTGAACAGCCTCTGGCA	miRNA_plate10_ES-T48_84-5	No Folding
miRNA_plate11_ES-T72_01-6	1	0	1	0	0	0	0	0	1	CCCGACCCCTCCACCCGCGC	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_03-1	1	0	1	0	0	0	0	0	1	AAGGGAGTCGGGTTCAGA	No Blast hit	No Blast hit

miRNA_plate11_ES-T72_11-6	1	0	1	0	0	0	0	0	1	TGGAAGCTGCAAAACCAACCA	miRNA_plate11_ES-T72_11-	No Folding
miRNA_plate11_ES-T72_15-5	1	0	1	0	0	0	0	0	1	GGATATGATGCCTTCTTCTGATC	miRNA_plate11_ES-T72_15-	No Folding
miRNA_plate11_ES-T72_20-1	1	0	1	0	0	0	0	0	1	TCGAATCCTGCTCACAGCGCCA	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_24-2	1	0	1	0	0	0	0	0	1	GTAAACTGCTTAGATTTT	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_26-2	1	0	1	0	0	0	0	0	1	ACCCTAAAGAGATGTTTG	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_30-4	1	0	1	0	0	0	0	0	1	GTGAAATGTTTAGGACCACTAA	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_41-1	1	0	1	0	0	0	0	0	1	CACTGCAATGATGAAAGGCCAT	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_41-2	1	0	1	0	0	0	0	0	1	AAAGTGCTACGTACTTTTGAGTC	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_46-4	1	0	1	0	0	0	0	0	1	TCGACTCCCGGTATGGGAACCA	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_56-5	1	0	1	0	0	0	0	0	1	AGGATGCGGAATTGATTA	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_60-3	1	0	1	0	0	0	0	0	1	CCAGGGTGGCGGTGGTGGCA	miRNA_plate11_ES-T72_60-	No Folding
miRNA_plate11_ES-T72_69-4	1	0	1	0	0	0	0	0	1	CCGCCGGGGTCGGCCCACG	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_70-6	1	0	1	0	0	0	0	0	1	CGGCGGCGGCGACTCTGGACGCGA	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_73-1	1	0	1	0	0	0	0	0	1	CTGTGCCCGTTTTATTTT	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_78-9	1	0	1	0	0	0	0	0	1	GAAGCGGGTGCTTCACCTT	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_88-2	1	0	1	0	0	0	0	0	1	AAAGTACTACTACTTTTGAGTCT	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_94-2	1	0	1	0	0	0	0	0	1	CAAAGTGCTGTTCGTGCATGA	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_94-3	1	0	1	0	0	0	0	0	1	ACTCGAAATGGACGCCCTATC	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_94-4	1	0	1	0	0	0	0	0	1	TAAAGTGCTTATACTGCAGGTAG	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_94-5	1	0	1	0	0	0	0	0	1	TAATACTGTCTGGTAATGACGA	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_94-6	1	0	1	0	0	0	0	0	1	CAGTGCAATGTTAAAAAGTGAT	No Blast hit	No Blast hit
miRNA_plate12_ES-T72_18-3	1	0	1	0	0	0	0	0	1	AGTCTATGAGAGTTACGA	No Blast hit	No Blast hit
miRNA_plate12_ES-T72_21-2	1	0	1	0	0	0	0	0	1	AAAGGGCTTCCTCTTTGGGGGCCG	No Blast hit	No Blast hit
miRNA_plate12_ES-T72_23-1	1	0	1	0	0	0	0	0	1	CCGAGTGTTACAGCCCTC	No Blast hit	No Blast hit
miRNA_plate12_ES-T72_32-1	1	0	1	0	0	0	0	0	1	TTCTCACTACTGCACTTGACTA	miRNA_plate12_ES-T72_32-	No Folding
miRNA_plate12_ES-T72_33-8	1	0	1	0	0	0	0	0	1	CTCTTATCGGTGGATCAC	No Blast hit	No Blast hit
miRNA_plate12_ES-T72_35-4	1	0	1	0	0	0	0	0	1	ACTTGTGTGTGAATGTATATGT	No Blast hit	No Blast hit
miRNA_plate12_ES-T72_35-5	1	0	1	0	0	0	0	0	1	GAAGCGTTTGCTTCACTTT	No Blast hit	No Blast hit
miRNA_plate12_ES-T72_35-6	1	0	1	0	0	0	0	0	1	AAAGTGCTACTACTTTTGATTCT	No Blast hit	No Blast hit
miRNA_plate12_ES-T72_40-6	1	0	1	0	0	0	0	0	1	TCCCCGAACGGGCCACCA	No Blast hit	No Blast hit
miRNA_plate12_ES-T72_42-2	1	0	1	0	0	0	0	0	1	CGCGGTGAACCCCGGCGAC	No Blast hit	No Blast hit
miRNA_plate12_ES-T72_49-2	1	0	1	0	0	0	0	0	1	CCCGACGAAGCCGAGCGC	No Blast hit	No Blast hit
miRNA_plate12_ES-T72_61-1	1	0	1	0	0	0	0	0	1	TTGGTGTCAGGCTAGTTTT	No Blast hit	No Blast hit
miRNA_plate12_ES-T72_79-4	1	0	1	0	0	0	0	0	1	CGACCATAAACGATGCCGAC	miRNA_plate12_ES-T72_79-	No Folding
miRNA_plate12_ES-T72_80-1	1	0	1	0	0	0	0	0	1	TGGCAGTGTCTTACCTGGTTGT	No Blast hit	No Blast hit
miRNA_plate12_ES-T72_83-7	1	0	1	0	0	0	0	0	1	AACCTCTGTATGCGTTTC	No Blast hit	No Blast hit
miRNA_plate12_ES-T72_94-1	1	0	1	0	0	0	0	0	1	CATCAGCATATGAGGTCACGGT	No Blast hit	No Blast hit
miRNA_plate12_ES-T72_94-2	1	0	1	0	0	0	0	0	1	AAATAGGTTTCCATGTGCATGT	No Blast hit	No Blast hit
miRNA_plate13_ES-C0_07-2	1	1	0	1	0	0	0	0	0	TGCTGCGTGCTTTTTGTT	No Blast hit	No Blast hit
miRNA_plate13_ES-T24_20-1	1	0	1	0	0	0	1	0	0	CACCCGAAAAACCAACCTTGCG	No Blast hit	No Blast hit
miRNA_plate13_ES-T24_20-5	1	0	1	0	0	0	1	0	0	TAAAGGGCTTATAGGGAAGGAAG	No Blast hit	No Blast hit
miRNA_plate13_ES-T24_20-6	1	0	1	0	0	0	1	0	0	AAAGGGCTACTACTTTTGAGTCCG	No Blast hit	No Blast hit
miRNA_plate13_ES-T48_30-1	1	0	1	0	0	0	0	1	0	ACAAGAGGACCTTTCTGG	No Blast hit	No Blast hit
miRNA_plate14_ES-C24_03-1	1	1	0	0	1	0	0	0	0	CGAGTCCCACCAGAGTCGCCA	No Blast hit	No Blast hit
miRNA_plate14_ES-C24_04-3	1	1	0	0	1	0	0	0	0	CCCCTTGCCTCTCGGCGC	No Blast hit	No Blast hit

miRNA_plate15_ES_C0_09-1	1	1	0	1	0	0	0	0	0	AAAGTGCTTCCCTTTTGCGTGT	No Blast hit	No Blast hit
miRNA_plate15_ES_C48_19-1	1	1	0	0	0	1	0	0	0	AAAGTGCTTCCACTTTGTGTAA	No Blast hit	No Blast hit
miRNA_plate15_ES_C48_27-1	1	1	0	0	0	1	0	0	0	TGCCCACAAAGGGAAACCATTTT	No Blast hit	No Blast hit
miRNA_plate15_ES_C48_27-2	1	1	0	0	0	1	0	0	0	GGGCACACAAAGGGAAAACTCTTT	No Blast hit	No Blast hit
miRNA_plate15_ES_C48_27-4	1	1	0	0	0	1	0	0	0	TGTGCACGTGCACACATACTT	No Blast hit	No Blast hit
miRNA_plate2_ES-C0_02-4	1	1	0	1	0	0	0	0	0	TCCCGGCCAAAGCACCACTGGAA	No Blast hit	No Blast hit
miRNA_plate2_ES-C0_15-1	1	1	0	1	0	0	0	0	0	AGAATGACTACAGATTTTGAGTGT	No Blast hit	No Blast hit
miRNA_plate2_ES-C0_17-5	1	1	0	1	0	0	0	0	0	TGATTAAGAGGGACGGCCG	No Blast hit	No Blast hit
miRNA_plate2_ES-C0_22-10	1	1	0	1	0	0	0	0	0	ATCCCTTCGTGGTTGCCA	No Blast hit	No Blast hit
miRNA_plate2_ES-C0_22-3	1	1	0	1	0	0	0	0	0	TCGGAGAGCTGTCCCGGGGCGA	No Blast hit	No Blast hit
miRNA_plate2_ES-C0_29-3	1	1	0	1	0	0	0	0	0	TGAGTTCGGGGCCAGCCTGCTC	miRNA_plate2_ES-C0_29-3,1	No Folding
miRNA_plate2_ES-C0_37-1	1	1	0	1	0	0	0	0	0	AGTAATCCTAGCCCTAGCCGTA	No Blast hit	No Blast hit
miRNA_plate2_ES-C0_37-2	1	1	0	1	0	0	0	0	0	AAAGTGCTCCTACTTTTGAGTCT	No Blast hit	No Blast hit
miRNA_plate2_ES-C0_40-1	1	1	0	1	0	0	0	0	0	TGTGCAGATCCTTGCGAAACTGA	No Blast hit	No Blast hit
miRNA_plate2_ES-C0_40-2	1	1	0	1	0	0	0	0	0	CATCAGAGTGGATGCCCTCTCT	No Blast hit	No Blast hit
miRNA_plate2_ES-C0_43-1	1	1	0	1	0	0	0	0	0	TGACACCTGCCACCCAGCCCAAT	No Blast hit	No Blast hit
miRNA_plate2_ES-C0_49-9	1	1	0	1	0	0	0	0	0	AAGTGCCGCCAGGGTTTGAGTGT	No Blast hit	No Blast hit
miRNA_plate2_ES-C0_50-2	1	1	0	1	0	0	0	0	0	TCCCTGTCCTCCAGGAGCTCCTG	No Blast hit	No Blast hit
miRNA_plate2_ES-C0_53-1	1	1	0	1	0	0	0	0	0	TGAGTGCGTCCCATTTGTGTGT	No Blast hit	No Blast hit
miRNA_plate3_ES-C24_05-1	1	1	0	0	1	0	0	0	0	AAAGTGCCGCCAGGTTTTCAGTGT	No Blast hit	No Blast hit
miRNA_plate3_ES-C24_32-2	1	1	0	0	1	0	0	0	0	ACTCAAACTAAGGGGGCACTTTTT	No Blast hit	No Blast hit
miRNA_plate3_ES-C24_32-3	1	1	0	0	1	0	0	0	0	AAAGCGCTTCCACTTTGTGT	No Blast hit	No Blast hit
miRNA_plate3_ES-C24_32-6	1	1	0	0	1	0	0	0	0	TAAAGTGCATATACTGCACGTAG	No Blast hit	No Blast hit
miRNA_plate3_ES-C24_32-7	1	1	0	0	1	0	0	0	0	TGCTCGGCTAGGGCTAACACC	No Blast hit	No Blast hit
miRNA_plate3_ES-C24_32-8	1	1	0	0	1	0	0	0	0	AAAATGCTTCCACTTTGTGTGCT	No Blast hit	No Blast hit
miRNA_plate3_ES-C24_43-6	1	1	0	0	1	0	0	0	0	CCCCCTGACCCCCCTTTATTCAG	miRNA_plate3_ES-C24_43-6	No Folding
miRNA_plate3_ES-C24_52-2	1	1	0	0	1	0	0	0	0	ACAGTGAAACTGCGAATGGC	miRNA_plate3_ES-C24_52-2	No Folding
miRNA_plate3_ES-C24_56-6	1	1	0	0	1	0	0	0	0	CAAAGTGCTAACAGTGCAGGC	No Blast hit	No Blast hit
miRNA_plate3_ES-C24_71-1	1	1	0	0	1	0	0	0	0	GAGCAGCACGTAAATATTGGCG	No Blast hit	No Blast hit
miRNA_plate3_ES-C24_80-1	1	1	0	0	1	0	0	0	0	GCTATGCCGATCGGGTGTCCG	miRNA_plate3_ES-C24_80-1	No Folding
miRNA_plate3_ES-C24_80-6	1	1	0	0	1	0	0	0	0	CAAAGTGCTAACAGTGCATGTA	No Blast hit	No Blast hit
miRNA_plate3_ES-C24_81-1	1	1	0	0	1	0	0	0	0	TGCAAATCGGTCGTCCGA	No Blast hit	No Blast hit
miRNA_plate3_ES-C24_89-1	1	1	0	0	1	0	0	0	0	AAAGTGCTACTACTTTTAAGTCT	No Blast hit	No Blast hit
miRNA_plate3_ES-C24_89-3	1	1	0	0	1	0	0	0	0	TAAGTGCTTCCATGTTTGAGTCT	No Blast hit	No Blast hit
miRNA_plate3_ES-C24_90-3	1	1	0	0	1	0	0	0	0	AACAACTCACCTGCCGAAT	No Blast hit	No Blast hit
miRNA_plate3_ES-C24_90-6	1	1	0	0	1	0	0	0	0	TCTGGACACATGTGGCTT	No Blast hit	No Blast hit
miRNA_plate3_ES-C24_92-4	1	1	0	0	1	0	0	0	0	AAATGCCCCCAGGTTTTGAATGG	No Blast hit	No Blast hit
miRNA_plate3_ES-C24_95-2	1	1	0	0	1	0	0	0	0	TCACTCGGCTCGTGCGTCGATGT	No Blast hit	No Blast hit
miRNA_plate4_ES-C24_18-4	1	1	0	0	1	0	0	0	0	CGGCGATGATGACACTCC	No Blast hit	No Blast hit
miRNA_plate4_ES-C24_21-3	1	1	0	0	1	0	0	0	0	CTGTCAGGCCGACCTGTTCG	miRNA_plate4_ES-C24_21-3	No Folding
miRNA_plate4_ES-C24_30-3	1	1	0	0	1	0	0	0	0	CACTCTTGAGTTTGCCGCC	No Blast hit	No Blast hit
miRNA_plate4_ES-C24_46-1	1	1	0	0	1	0	0	0	0	TAAAGTACTTATAGTGCAGGTAGT	No Blast hit	No Blast hit
miRNA_plate4_ES-C24_46-2	1	1	0	0	1	0	0	0	0	ACTCAACCTATGGGGGCACTTTT	No Blast hit	No Blast hit
miRNA_plate4_ES-C24_46-4	1	1	0	0	1	0	0	0	0	GTACACACATGCACAAGCACACA	miRNA_plate4_ES-C24_46-4	No Folding
miRNA_plate4_ES-C24_67-3	1	1	0	0	1	0	0	0	0	CGGCGTCCGGTGAGCTCTC	No Blast hit	No Blast hit

miRNA_plate4_ES-C24_73-3	1	1	0	0	1	0	0	0	0	CCGGCGGCGGGCGCGGGG	No Blast hit	No Blast hit
miRNA_plate4_ES-C24_88-7	1	1	0	0	1	0	0	0	0	CAGCAGCAATTCATGTTTTAC	No Blast hit	No Blast hit
miRNA_plate5_ES-C48_07-7	1	1	0	0	0	1	0	0	0	TCCCCGGCACCTCCACCA	No Blast hit	No Blast hit
miRNA_plate5_ES-C48_22-1	1	1	0	0	0	1	0	0	0	TGTAACCATCCTACACTCTCAGCT	No Blast hit	No Blast hit
miRNA_plate5_ES-C48_22-2	1	1	0	0	0	1	0	0	0	AAAGTGCTATTACTTCTGAGTCT	No Blast hit	No Blast hit
miRNA_plate5_ES-C48_29-1	1	1	0	0	0	1	0	0	0	GCACGGGGTCGGCGGCGA	No Blast hit	No Blast hit
miRNA_plate5_ES-C48_45-8	1	1	0	0	0	1	0	0	0	AACCTGAACGTGCTGCGCGAGG	miRNA_plate5_ES-C48_45-8	No Folding
miRNA_plate5_ES-C48_45-9	1	1	0	0	0	1	0	0	0	CCCGCGGGGCCCCGTCGTCCCCC	No Blast hit	No Blast hit
miRNA_plate5_ES-C48_68-8	1	1	0	0	0	1	0	0	0	CACCGCCCGTCGCTACTACCG	miRNA_plate5_ES-C48_68-8	No Folding
miRNA_plate5_ES-C48_87-2	1	1	0	0	0	1	0	0	0	TGCGCAAATCCATGCAAAACTGA	No Blast hit	No Blast hit
miRNA_plate5_ES-C48_94-5	1	1	0	0	0	1	0	0	0	TACGCCTGTCTGAGCGTCGCTT	No Blast hit	No Blast hit
miRNA_plate5_ES-C48_95-1	1	1	0	0	0	1	0	0	0	AAGTGCTTCTTCTTTTGAGTTTT	No Blast hit	No Blast hit
miRNA_plate5_ES-C48_95-2	1	1	0	0	0	1	0	0	0	AAAGTGGTTCTTCTTTTGAGTTTT	No Blast hit	No Blast hit
miRNA_plate5_ES-C48_95-3	1	1	0	0	0	1	0	0	0	ACTCCAACTATGGGGGCCCTTTT	No Blast hit	No Blast hit
miRNA_plate6_ES-C48_05-1	1	1	0	0	0	1	0	0	0	CTCACCCGGCCCGGACAC	No Blast hit	No Blast hit
miRNA_plate6_ES-C48_20-10	1	1	0	0	0	1	0	0	0	CTGCGGAAGGATCATTAACGGGAG	miRNA_plate6_ES-C48_20-1	No Folding
miRNA_plate6_ES-C48_22-6	1	1	0	0	0	1	0	0	0	TCGACTCCCGGTGTGGGAACCA	No Blast hit	No Blast hit
miRNA_plate6_ES-C48_23-1	1	1	0	0	0	1	0	0	0	CGGGCGGTCCGCCGCGAG	No Blast hit	No Blast hit
miRNA_plate6_ES-C48_65-3	1	1	0	0	0	1	0	0	0	GAAAGCGGGTGCTTCACTTTT	No Blast hit	No Blast hit
miRNA_plate6_ES-C48_66-2	1	1	0	0	0	1	0	0	0	ACGCCGGCAGGCGCGGGTAACCCG	miRNA_plate6_ES-C48_66-2	No Folding
miRNA_plate6_ES-C48_75-8	1	1	0	0	0	1	0	0	0	AAAGTGCTACTACTTTTGAATCT	No Blast hit	No Blast hit
miRNA_plate6_ES-C48_88-11	1	1	0	0	0	1	0	0	0	AAAGTGCTACTACTTTTGTGTG	No Blast hit	No Blast hit
miRNA_plate6_ES-C48_88-9	1	1	0	0	0	1	0	0	0	CTAACCAGGATTCCCTCAG	No Blast hit	No Blast hit
miRNA_plate6_ES-C48_94-6	1	1	0	0	0	1	0	0	0	TAAGGTAGCCAAATGCCTCGTCA	miRNA_plate6_ES-C48_94-6	No Folding
miRNA_plate7_ES-T24_09-1	1	0	1	0	0	0	1	0	0	TGGCAGTGTCTTATCTGGTTGT	No Blast hit	No Blast hit
miRNA_plate7_ES-T24_09-2	1	0	1	0	0	0	1	0	0	TAAAAGTGCCTGCAGGTTTTGAGT	No Blast hit	No Blast hit
miRNA_plate7_ES-T24_09-3	1	0	1	0	0	0	1	0	0	AATTGCCACCAGGTTTTGAGTGT	No Blast hit	No Blast hit
miRNA_plate7_ES-T24_22-4	1	0	1	0	0	0	1	0	0	ATACTATGTCCCTTGTGTTTTGAG	miRNA_plate7_ES-T24_22-4	No Folding
miRNA_plate7_ES-T24_45-1	1	0	1	0	0	0	1	0	0	TATCAGCACGTAAATATTGTGTG	No Blast hit	No Blast hit
miRNA_plate7_ES-T24_66-5	1	0	1	0	0	0	1	0	0	GCTGGGACCTGTGGCTGTGTGTT	miRNA_plate7_ES-T24_66-5	No Folding
miRNA_plate7_ES-T24_74-1	1	0	1	0	0	0	1	0	0	AAAGTGCCTCCAGGTTTTGAGTGT	No Blast hit	No Blast hit
miRNA_plate7_ES-T24_74-2	1	0	1	0	0	0	1	0	0	GAACGCGCACTCTGGCGTGG	No Blast hit	No Blast hit
miRNA_plate7_ES-T24_78-2	1	0	1	0	0	0	1	0	0	GGACAACGGGGCCTGCTCCTCC	miRNA_plate7_ES-T24_78-2	No Folding
miRNA_plate7_ES-T24_80-1	1	0	1	0	0	0	1	0	0	TGAGTGCTACTACTTTTGAGTCT	No Blast hit	No Blast hit
miRNA_plate7_ES-T24_80-3	1	0	1	0	0	0	1	0	0	TGTTGTCACGGCCTCAATG	No Blast hit	No Blast hit
miRNA_plate7_ES-T24_80-6	1	0	1	0	0	0	1	0	0	AAATGGCTTCCCTTTTGTGTGT	No Blast hit	No Blast hit
miRNA_plate7_ES-T24_90-9	1	0	1	0	0	0	1	0	0	CCCTGTGGAGCAGTAGAA	No Blast hit	No Blast hit
miRNA_plate7_ES-T24_93-3	1	0	1	0	0	0	1	0	0	CGCGACCTCAAATCAGACGTG	No Blast hit	No Blast hit
miRNA_plate8_ES-T24_04-3	1	0	1	0	0	0	1	0	0	AAGGTGGGAAGCCTAGATGTCA	miRNA_plate8_ES-T24_04-3	No Folding
miRNA_plate8_ES-T24_06-4	1	0	1	0	0	0	1	0	0	GATCGATGTGGTGCTCCGGAC	No Blast hit	No Blast hit
miRNA_plate8_ES-T24_07-2	1	0	1	0	0	0	1	0	0	GTGGCTTGCTGCTTGTGACTGGGC	miRNA_plate8_ES-T24_07-2	No Folding
miRNA_plate8_ES-T24_09-4	1	0	1	0	0	0	1	0	0	AAAGTGCTCCCCTTTTGTGTGT	No Blast hit	No Blast hit
miRNA_plate8_ES-T24_19-5	1	0	1	0	0	0	1	0	0	TGACCAATGAACACTCTGACCA	No Blast hit	No Blast hit
miRNA_plate8_ES-T24_40-4	1	0	1	0	0	0	1	0	0	TCCCCGGCATCTCCACCA	No Blast hit	No Blast hit
miRNA_plate8_ES-T24_63-4	1	0	1	0	0	0	1	0	0	AAGGCAGAGGTAGATGGACTTC	No Blast hit	No Blast hit

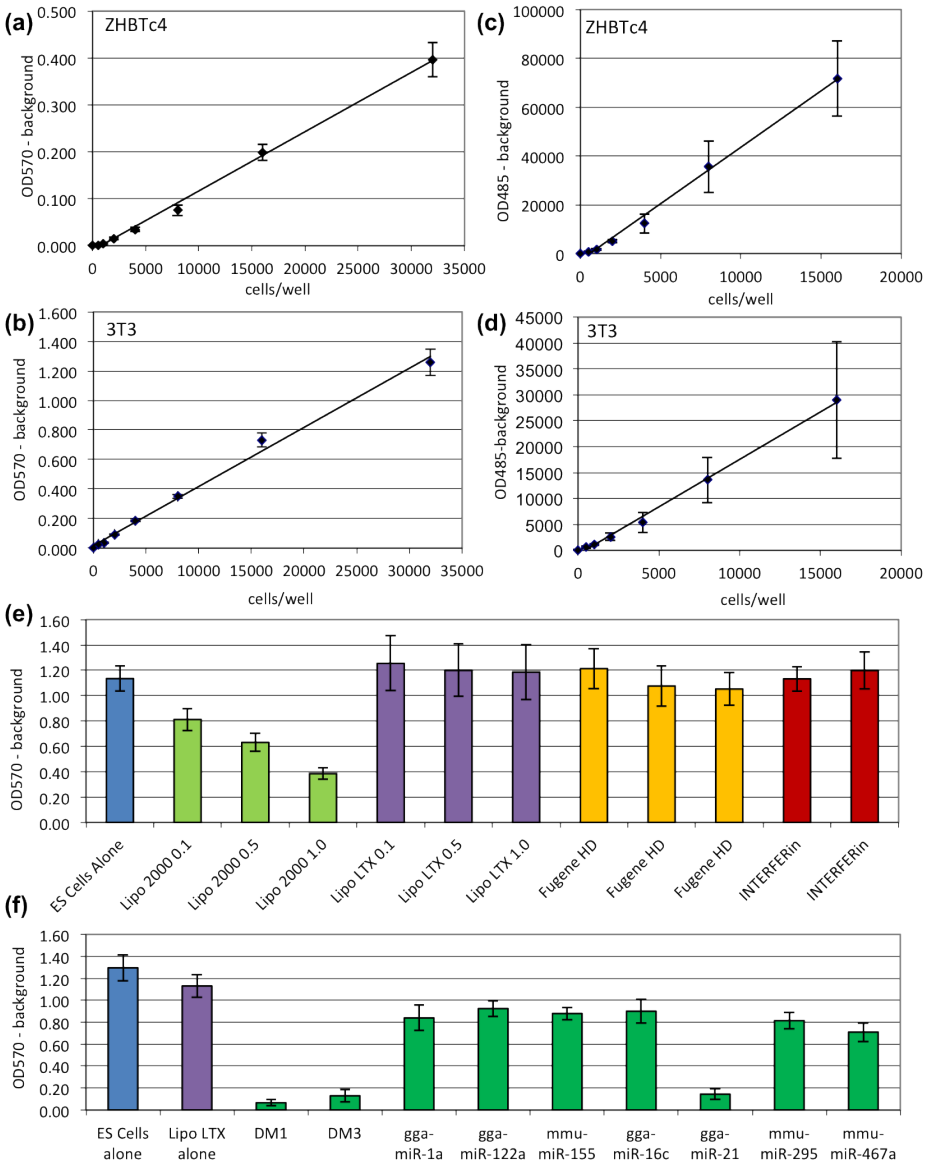
miRNA_plate8_ES-T24_65-11	1	0	1	0	0	0	1	0	0	GGCCTTTGATTTCAGTACTTGG	No Blast hit	No Blast hit
miRNA_plate8_ES-T24_82-5	1	0	1	0	0	0	1	0	0	AACGCTCATGTCTGACTG	No Blast hit	No Blast hit
miRNA_plate8_ES-T24_84-2	1	0	1	0	0	0	1	0	0	ACTCAAACTGTGAGCTCTTTTG	No Blast hit	No Blast hit
miRNA_plate8_ES-T24_94-1	1	0	1	0	0	0	1	0	0	ATCCCACCAGAGTCGCCA	No Blast hit	No Blast hit
miRNA_plate9_ES-T48_05-2	1	0	1	0	0	0	0	1	0	GTGGTGTGTTAGTTAATTTT	No Blast hit	No Blast hit
miRNA_plate9_ES-T48_23-4	1	0	1	0	0	0	0	1	0	CGGGAGCGGGGGACCACC	No Blast hit	No Blast hit
miRNA_plate9_ES-T48_26-3	1	0	1	0	0	0	0	1	0	AAAGTGCTTCCTTTTTGTGTGT	No Blast hit	No Blast hit
miRNA_plate9_ES-T48_33-3	1	0	1	0	0	0	0	1	0	TAAAGTGCTTATAGAGCAGGTAT	No Blast hit	No Blast hit
miRNA_plate9_ES-T48_47-5	1	0	1	0	0	0	0	1	0	TCGAATCCCACTCCTGACACCA	No Blast hit	No Blast hit
miRNA_plate9_ES-T48_54-5	1	0	1	0	0	0	0	1	0	CCCCTGGGCCTATCCTACAAT	No Blast hit	No Blast hit
miRNA_plate9_ES-T48_54-6	1	0	1	0	0	0	0	1	0	AAACTGCCGCCAGGTTTTGAGTGT	No Blast hit	No Blast hit
miRNA_plate9_ES-T48_54-7	1	0	1	0	0	0	0	1	0	AAATTGCTTCCCTTTTGTGTGT	No Blast hit	No Blast hit
miRNA_plate9_ES-T48_60-1	1	0	1	0	0	0	0	1	0	AGCTGCAAAGCCAGCTGACATTTT	miRNA_plate9_ES-T48_60-1,	No Folding
miRNA_plate9_ES-T48_68-12	1	0	1	0	0	0	0	1	0	CCGCGCGGCGCCTCGCCTCGGCC	No Blast hit	No Blast hit
miRNA_plate9_ES-T48_70-2	1	0	1	0	0	0	0	1	0	TACTGGTGAGTATCCCCGC	No Blast hit	No Blast hit
miRNA_plate9_ES-T48_82-6	1	0	1	0	0	0	0	1	0	CCCGGCGGGTGCCGGCGCG	No Blast hit	No Blast hit
miRNA_plate9_ES-T48_87-6	1	0	1	0	0	0	0	1	0	CGCGACCTCAAATCAAACG	No Blast hit	No Blast hit
miRNA_plate9_ES-T48_91-2	1	0	1	0	0	0	0	1	0	CATTGCCATAGTATGGTCGAACC	No Blast hit	No Blast hit
miRNA_plate9_ES-T48_94-1	1	0	1	0	0	0	0	1	0	TCACCTGGAGCATGTTTTTATT	No Blast hit	No Blast hit
miRNA_plate9_ES-T48_94-5	1	0	1	0	0	0	0	1	0	GGGATAAGGATTGGCTCTA	No Blast hit	No Blast hit
miRNA_plate12_ES-T72_61-5	32	22	10	4	6	12	3	5	2	AAAGTGCTTCCACTTTGTGTGT	No Blast hit	No Blast hit
miRNA_plate1_ES-C0_60-4	1	1	0	1	0	0	0	0	0	CACTCAAACCATGCGGCCACTTG	No Blast hit	No Blast hit
miRNA_plate11_ES-T72_15-4	6	1	5	1	0	0	0	1	4	GTGTGCATGTGCATGTGTGTATAT	miRNA_plate11_ES-T72_15-,	miRNA_plate11_E
miRNA_plate11_ES-T72_81-2	1	0	1	0	0	0	0	0	1	ATAAGTGCGTGCATGTATATGA	No Blast hit	No Blast hit
miRNA_plate12_ES-T72_53-6	1	0	1	0	0	0	0	0	1	ATGGTGACCACGGGTGGACGG	No Blast hit	No Blast hit
miRNA_plate15_ES_T24_35-2	1	0	1	0	0	0	1	0	0	ACTCAAACTGGGAGCTCTTTTG	No Blast hit	No Blast hit
miRNA_plate15_ES_T24_35-3	1	0	1	0	0	0	1	0	0	CAAATTGCTCATAGTGCAGGTAG	No Blast hit	No Blast hit
miRNA_plate16_ES_T72_02-3	1	0	1	0	0	0	0	0	1	CAAAGTGCTTGCAGTGCAGGTA	No Blast hit	No Blast hit
miRNA_plate16_ES_T72_03-2	1	0	1	0	0	0	0	0	1	TTGCGCCGCTAGAGGTGAAA	miRNA_plate16_ES_T72_03-	No Folding
miRNA_plate4_ES-C24_47-3	1	1	0	0	1	0	0	0	0	AAGCCGGGCCGTAGTGGCGCC	No Blast hit	No Blast hit
miRNA_plate5_ES-C48_23-4	1	1	0	0	0	1	0	0	0	AAAGTGCTTCCACTTTGTGGTGC	No Blast hit	No Blast hit
miRNA_plate6_ES-C48_30-7	1	1	0	0	0	1	0	0	0	TGTAGTGTTTCCTACTTTATGGT	No Blast hit	No Blast hit
miRNA_plate7_ES-T24_14-2	1	0	1	0	0	0	1	0	0	AAGTGTTATTACTTGTGAGTCTCG	No Blast hit	No Blast hit
miRNA_plate7_ES-T24_75-3	3	1	2	0	1	0	1	0	1	CAAAGTGCTCATAGTGCAGGTAA	No Blast hit	No Blast hit
miRNA_plate8_ES-T24_52-3	3	0	3	0	0	0	1	1	1	TCGATTCCCCGACGGGGAGCCA	No Blast hit	No Blast hit
miRNA_plate9_ES-T48_17-1	2	1	1	0	1	0	0	1	0	CCAGCCTGGGCTACACATTTTT	miRNA_plate9_ES-T48_17-1,	miRNA_plate9_ES
miRNA_plate9_ES-T48_96-8	1	0	1	0	0	0	0	1	0	TGCTATGATGAAGGCTATGTTG	miRNA_plate9_ES-T48_96-8,	No Folding

Appendix C - Optimisation of MTT and Cyquant assays

Optimisation of cell proliferation assays The MTT and Cyquant assays were tested to confirm that they gave a linear absorbance output from increasing numbers of ES cells and 3T3 cells in the conditions used (Figure a, b, c & d). Between 0 and 32,000 ZHBTc4 ES cells or 3T3 cells were plated per well of 96-well plates. Following ~24 hours of cell growth cultures were subjected to a 1, 2 or 4 hour incubation with MTT reagent. Analysis of MTT assay results showed that a 1 hour MTT incubation gave a linear absorbance output for between 2000 and 32,000 ZHBTc4 ES cells and a 2 hour MTT incubation gave a linear absorbance output for between 2000 and 32,000 3T3 cells. Scatter plots of these data are shown in Figure a and b. Scatter plots of other incubation times are not shown. In order to optimise the Cyquant assay between 0 and 32,000 ZHBTc4 ES cells or 3T3 cells were plated per well of 96-well plates. Following ~24 hours of cell growth cultures were subjected to incubation with Cyquant reagent. Analysis of Cyquant assay results showed that assays were linear between 1000 and 16,000 ZHBTc4 ES cells, and 1000 and 16,000 3T3 cells. In both cases the assay was not linear at 32,000 plated cells. The 32,000 plated cells point has not been included in the ZHBTc4 or the 3T3 optimisation graphs. Scatter plots of these data are shown in Figure c and d.

The MTT assay was used to assess the effects of different transfection reagents on cell viability. ES cells were plated and grown in the presence of different concentrations of Lipofectamine 2000, Lipofectamine LTX, Fugene HD or Interferin transfection reagents before being assayed for cell proliferation by MTT assay (Figure e). As shown in Figure e, addition of Lipofectamine LTX, Fugene HD and Interferin at the concentrations used had no effect on cell numbers. In contrast, addition of all concentrations of Lipofectamine 2000 caused a reduction in MTT assay absorbance. The reduction in absorbance was greater for higher concentrations of Lipofectamine 2000 suggesting the effect is concentration dependent. This result suggests that Lipofectamine 2000 is toxic to ES cells. In light of these results, Lipofectamine 2000 was replaced with Lipofectamine LTX as the transfection reagent for future experiments.

LNAs complementary to a range of different miRNAs were transfected into ES cells in order to select an appropriate LNA for use as a control for proliferation experiments (Figure f). The LNAs were transfected at a concentration of 225nM, which was deliberately high in order to correspond to the total concentration of LNA present if multiple LNAs were transfected into a single culture. As expected, transfection of any LNA caused a slight decrease in cell numbers relative to transfection reagent alone. Of the 9 LNAs tested, only 3 LNAs designed to target chicken miRNAs had a significant effect on cell numbers. Of the remainder, an LNA designed to inhibit expression of mmu-miR-155 was chosen for use as a control because in addition to it having little effect on cell numbers, this miRNA was not sequenced in the miRNA profiling experiment and was not detected in mouse ES cells by miRNA northern analysis (data not shown).



Optimisation of MTT and Cyquant Assays. The linear range of the MTT and Cyquant assays was determined for ZHBTc4 ES cells (a & c) and 3T3 cells (b & d). Various numbers of cells were plated/well of 96-well plates and grown for 24 hours then MTT (a & b) or CYQUANT (c & d) analysis was carried out. Scatter charts show mean absorbance of 12 replicate wells of cells plotted against number of cells/well. Error bars show SD. (e) 1500 ZHBTc4 ES cells were plated/well of a 96-well plate. After 24 hours of growth cells were incubated O/N with various volumes of different transfection reagents (0.1=0.1µl/well, 0.5=0.5µl/well, 1.0=1.0µl/well) before MTT analysis. Bar chart shows mean absorbance of 8 replicate wells of cells. Error bars show SD. (f) 1000 ZHBTc4 ES cells were plated/well of a 96-well plate and grown for 24 hours. Cells were transfected O/N with 225nM of different LNA oligonucleotides before MTT analysis. Bar chart shows mean absorbance of 8 replicate wells. Error bars show SD.

Appendix D - SILAC analysis data

D2

Protein IDs	Protein Names	R6K6/R0K0	R0K0/R6K6	Uniprot Accession No.	Peptides (seq)	Sequence Coverage [%]	Mol. Weight [kDa]	Sequence Length	Slice 1	PEP
IPI00131584.1	Mitochondrial carnitine/acylcarnitine carrier prot	5.581	0.179	Q9Z2Z6;Q7TPW6	5	16.3	33.026	301	5	6.93E-15
IPI00131887.3	Prostaglandin reductase 1;NADP-dependent leuk	5.308	0.188	Q91YR9;A2ALW3;Q4	9	38.3	35.56	329	13	3.51E-56
IPI00114733.1;IF	Serpin H1;Collagen-binding protein;47 kDa heat s	4.912	0.204	P19324;Q3TJK3;Q3T	5	18.9	46.589	417	6	1.50E-22
IPI00134191.3	Solute carrier family 2, facilitated glucose transp	4.869	0.205	P32037;Q3TPL8;Q4F	2	6.1	53.478	493	3	1.22E-13
IPI00229510.5;IF	L-lactate dehydrogenase B chain;LDH heart subu	3.578	0.280	P16125;Q5M8R7	10	33.8	36.572	334	16	3.49E-83
IPI00136703.1	Creatine kinase B-type;Creatine kinase B chain;B-	3.455	0.289	Q04447	13	55.1	42.713	381	28	4.90E-106
IPI00127691.1	Glutathione synthetase;Glutathione synthase;GS	2.513	0.398	P51855;Q3TWM2;Q3	1	3.6	52.246	474	2	0.11875
IPI00776276.1;IF	Four and a half LIM domains 1;Four and a half LIM	2.413	0.414	A2AEY2;A2AEX6;A2A	2	7.7	36.261	323	2	3.69E-09
IPI00653423.2;IF	Branched chain aminotransferase 1, cytosolic;16	2.382	0.420	B2KFA0;Q3TKJ1;Q8C	9	22.4	57.306	519	12	6.93E-49
IPI00131577.1;IF	Heme oxygenase 1;P32 protein;Putative unchara	2.215	0.452	P14901;Q3U5H8;Q3I	2	10.7	32.928	289	2	3.20E-13
IPI00653636.1;IF	Annexin A7;Annexin-7;Annexin VII;Synexin;Putat	2.002	0.499	Q07076;Q3TJ49;Q3U	6	12.8	52.251	485	6	8.88E-09
IPI00134704.1;IF	Quinone oxidoreductase;NADPH:quinone reduct	1.993	0.502	P47199;Q3UYY0;Q8C	4	18.1	35.268	331	4	2.42E-20
IPI00123202.1	Ataxin-10;Spinocerebellar ataxia type 10 protein	1.850	0.541	P28658;Q3TCF4;Q3T	15	36.6	53.706	475	17	2.11E-50
IPI00317309.5	Annexin A5;Annexin-5;Annexin V;Lipocortin V;En	1.836	0.545	P48036	2	7.8	35.752	319	2	6.90E-09
IPI00816884.1;IP	IPI00885560.1	1.827	0.547		2	5.2	36.869	346	2	3.20E-06
IPI00128904.1	Poly(rC)-binding protein 1;Alpha-CP1;hnRNP-E1	1.589	0.629	P60335	2	7.6	37.497	356	2	5.35E-07
IPI00130344.3;IF	Chloride intracellular channel protein 1;Nuclear c	1.579	0.634	Q9Z1Q5;Q3TIP8;Q54	5	25.7	27.013	241	7	1.94E-31
IPI00115085.1;IF	Inositol monophosphatase;Inositol-1(or 4)-mono	1.540	0.649	O55023;Q3TME6;Q3	1	4	30.436	277	2	1.41E-08
IPI00114945.1;IF	Septin-2;Neural precursor cell expressed develop	1.530	0.654	P42208;B2RRZ2	3	10.2	41.525	361	3	7.29E-08
IPI00113895.1;IF	Alpha-centractin;Centrosome-associated actin ho	1.484	0.674	P61164;Q8R5C5;Q8F	1	4.3	42.613	376	2	2.92E-11
IPI00353563.4;IF	Fascin;Singed-like protein	1.469	0.681	Q61553	3	8.9	54.507	493	4	8.62E-12
IPI00120457.1	Farnesyl pyrophosphate synthetase;Farnesyl dipl	1.436	0.696	Q920E5;Q3TMB3;Q3	4	16.1	40.581	353	7	4.64E-14
IPI00331394.3;IF	Aspartyl aminopeptidase;Dnpep protein (Mamm	1.417	0.706	Q9Z2W0;Q3TVK3;Q3	5	20.4	52.466	475	6	2.61E-25
IPI00163011.2	Thioredoxin domain-containing protein 5;Thiored	1.408	0.710	Q91W90;Q3TEE8;Q3	2	7	46.415	417	2	0.22334
IPI00119111.2	Calponin-3;Calponin, acidic isoform	1.395	0.717	Q9DAW9;Q99JN5	3	10.6	36.428	330	4	1.51E-17
IPI00875110.1;IF	Isocitrate dehydrogenase [NADP], mitochondrial;	1.392	0.719	P54071	12	37.1	52.399	469	14	4.87E-62
IPI00125091.1;IF	LIM and SH3 domain protein 1;MLN 50;LIM and S	1.373	0.728	Q61792;Q543N3;A2A	3	9.1	29.994	263	3	0.0018875
IPI00128904.1	Poly(rC)-binding protein 1;Alpha-CP1;hnRNP-E1	1.351	0.740	P60335	7	38.2	37.497	356	8	2.37E-43
IPI00112346.1;IF	Mitogen-activated protein kinase 14;Mitogen-act	1.345	0.743	P47811-1;P47811;B2	2	12.2	41.287	360	2	3.44E-15
IPI00265025.5	Galactokinase;Galactose kinase	1.322	0.757	Q9R0N0;Q80UL3;Q9	4	15.1	42.295	392	4	8.86E-14
IPI00120716.3	Guanine nucleotide-binding protein G(I)/G(S)/G(T	1.280	0.781	P62874;Q3TQ70;Q3U	5	19.4	37.377	340	5	8.43E-28
IPI00124287.1;IF	Polyadenylate-binding protein 1;Pabpc1 protein;	1.255	0.797	P29341;Q05DT1;Q3U	3	6.1	70.642	636	3	5.37E-18
IPI00130530.1;IF	Glyoxylate reductase/hydroxypyruvate reductase	1.247	0.802	Q91Z53;Q3T9Z2;B1A	4	14	35.328	328	5	2.49E-11
IPI00230204.7;IF	Aspartate aminotransferase, cytoplasmic;Transar	1.245	0.803	P05201;Q3UJH8	12	41.9	46.231	413	18	1.24E-100
IPI00116498.1;IF	14-3-3 protein zeta/delta;Protein kinase C inhibit	1.244	0.804	P63101;Q3UA58	5	22	27.771	245	3	1.20E-16
IPI00112053.1;IF	Sorting nexin-5	1.241	0.806	Q9D8U8;A2ANA4;Q3	3	8.4	46.797	404	5	2.07E-13
IPI00122698.1;IF	Histone-binding protein RBBP7;Retinoblastoma-b	1.241	0.806	Q60973;A2AFJ0;Q3U	6	23.1	47.789	425	8	5.76E-33
IPI00109142.4	S-formylglutathione hydrolase;Esterase D;Esterase	1.227	0.815	Q9R0P3	1	6.4	31.319	282	2	0.036882
IPI00135231.2;IF	Isocitrate dehydrogenase [NADP] cytoplasmic;Cy	1.209	0.827	O88844;Q3TJ51;Q5H	10	28.2	47.545	422	11	1.99E-47

D3

IPI00121887.5;IF	Transcription elongation factor A protein 1;Trans	1.194	0.838	P10711-1;P10711;Q3	3	11.6	33.88	301	3	5.62E-20
IPI00230707.6;IF	14-3-3 protein gamma	1.192	0.839	P61982;A8IP69	11	53.4	28.302	247	11	3.98E-69
IPI00313817.1;IF	Hepatoma-derived growth factor	1.190	0.840	P51859	2	11.8	26.268	237	2	3.20E-07
IPI00119680.3;IF	AP-1 complex subunit mu-1;Adaptor-related prot	1.172	0.853	P35585;Q3UG16	2	5.9	48.542	423	2	2.60E-05
IPI00132208.1;IF	DnaJ homolog subfamily A member 1;Heat shock	1.171	0.854	P63037;Q3TK61;Q5N	4	15.9	44.868	397	4	3.27E-18
IPI00224729.1;IF	2 days neonate thymus thymic cells cDNA, RIKEN	1.147	0.872	Q8C2Q7;O35737;Q3	5	12.1	51.217	472	7	1.69E-19
IPI00226930.2	Histone acetyltransferase type B catalytic subunit	1.142	0.876	Q8BY71;A2ATU9	6	28.6	50.182	423	9	8.19E-77
IPI00323881.1;IF	Importin subunit beta-1;Karyopherin subunit bet	1.128	0.887	P70168;Q3TFE8;Q3U	1	1	97.151	876	2	0.097578
IPI00132799.4;IF	P32-RACK (Complement component 1, q subcom	1.127	0.887	Q8R5L1;O35658	2	15.1	31.025	279	3	3.02E-12
IPI00224070.2;IF	Rho GTPase activating protein 1;Rho GTPase-acti	1.125	0.889	A2AGT9;Q8BQW4;Q	3	13.6	54.437	479	4	5.05E-14
IPI00454142.5;IF	Septin-11	1.124	0.890	Q8C1B7-1;Q8C1B7;C	7	21.1	49.694	431	10	5.52E-57
IPI00224626.3;IF	Melanocyte cDNA, RIKEN full-length enriched libı	1.120	0.893	Q3UGE0;Q3UNN1;Q	2	8.5	50.648	437	3	3.40E-13
IPI00169707.2;IF	Tropomyosin 3, gamma;Tropomyosin alpha-3 cha	1.120	0.893	Q8K0Z5;P21107-1;P2	3	10.9	33.28	285	4	2.37E-10
IPI00119663.3;IF	Mitogen-activated protein kinase 1;Extracellular	1.106	0.904	P63085;Q3UF82;Q92	13	50	41.275	358	15	9.79E-73
IPI00131056.1	Insulin-like growth factor 2 mRNA-binding protei	1.105	0.905	O88477	1	2.8	63.45	577	2	1.70E-06
IPI00775863.1;IF	Guanine nucleotide-binding protein G(I)/G(S)/G(T	1.099	0.910	P62880;Q3U9V4;P29	4	12.3	41.408	382	3	9.57E-13
IPI00222496.3;IF	Protein disulfide-isomerase A6;Thioredoxin doma	1.099	0.910	Q922R8;Q3THH1;Q3	7	24	48.689	445	10	1.57E-45
IPI00319950.4	Diphosphomevalonate decarboxylase;Mevalonat	1.096	0.912	Q99JF5;Q3UYC1	2	5.7	44.072	401	2	0.015145
IPI00226430.2;IF	3-ketoacyl-CoA thiolase, mitochondrial;Beta-ketc	1.095	0.913	Q8BWT1;Q3TIT9;Q3I	2	6.5	41.857	397	3	0.00012698
IPI00115627.4	Actin-related protein 3;Actin-like protein 3	1.093	0.915	Q99JY9;Q3TGE1;Q3T	8	29.7	47.357	418	9	1.88E-46
IPI00117352.1;IF	Tubulin beta-5 chain;Tubulin beta-2C chain;Tubu	1.091	0.917	P99024;B1B178;Q80	8	19.8	49.67	444	10	2.75E-14
IPI00121833.3;IF	3-ketoacyl-CoA thiolase A, peroxisomal;Beta-ketc	1.090	0.917	Q921H8;Q8BLD7;Q8	3	9.7	43.953	424	4	1.12E-11
IPI00177038.1;IF	Actin-related protein 2;Actin-like protein 2;12 da	1.088	0.919	P61161;Q5SW83;Q8	4	12.7	44.76	394	7	5.10E-17
IPI00108271.1;IF	ELAV-like protein 1;Hu-antigen R;Elav-like generi	1.082	0.924	P70372;Q3TT05;Q3U	3	8.9	36.069	326	3	5.45E-09
IPI00122548.3;IF	Adult male testis cDNA, RIKEN full-length enriche	1.078	0.928	Q3TTN3;Q5EBQ0;Q6	5	21.5	30.884	284	8	6.01E-27
IPI00408957.3	Suppressor of G2 allele of SKP1 homolog	1.056	0.947	Q9CX34;Q9CRG3	3	14	38.158	336	3	0.00089087
IPI00134809.2;IF	Dihydrolipoyllysine-residue succinyltransferase c	1.054	0.949	Q9D2G2-1;Q9D2G2;(	2	6.6	48.994	454	2	1.38E-23
IPI00109932.1	Probable ATP-dependent RNA helicase DDX6;DEA	1.052	0.951	P54823	4	12.8	54.191	483	7	2.03E-24
IPI00116761.1;IF	Translation initiation factor eIF-2B subunit alpha;	1.051	0.952	Q99LC8;Q3UZR8	3	11.8	33.816	305	3	1.78E-08
IPI00153107.3;IF	Bleomycin hydrolase	1.048	0.955	Q8R016	3	9.5	52.511	455	3	4.99E-07
IPI00323357.3;IF	Heat shock cognate 71 kDa protein;Heat shock 7(	1.047	0.955	P63017;Q3KQJ4;Q3T	5	8.7	70.87	646	6	5.51E-13
IPI00321718.4;IF	Prohibitin-2;B-cell receptor-associated protein B/	1.043	0.959	O35129;Q3V235	4	14.7	33.296	299	6	9.48E-25
IPI00122547.1;IF	Voltage-dependent anion-selective channel prot(	1.041	0.960	Q60930	5	15.3	31.732	295	7	2.14E-16
IPI00466919.7;IF	6-phosphogluconate dehydrogenase, decarboxyl	1.030	0.971	Q9DCD0;A2AH73;Q3	11	28.6	53.247	483	13	1.37E-34
IPI00135969.2	Adenylosuccinate synthetase isozyme 2;Adenylos	1.030	0.971	P46664;Q8C909	6	17.8	50.02	456	6	2.07E-31
IPI00221528.1	Beta-actin-like protein 2;Kappa-actin	1.029	0.972	Q8BFZ3	5	17.3	42.004	376	3	4.46E-10
IPI00314736.3;IF	Acidic leucine-rich nuclear phosphoprotein 32 fa	1.027	0.974	O35381;Q64G17	4	15	28.537	247	6	8.38E-12
IPI00459725.2;IF	Isocitrate dehydrogenase [NAD] subunit alpha, m	1.025	0.976	Q9D6R2-1;Q9D6R2;(	4	14.2	39.638	366	8	1.01E-14
IPI00314709.1;IF	Splicing factor, arginine/serine-rich 5;Pre-mRNA-	1.024	0.976	O35326;Q9D8S5;Q5I	2	10.7	30.944	270	3	1.32E-11
IPI00404551.1;IF	B6-derived CD11 +ve dendritic cells cDNA, RIKEN	1.024	0.977	Q8C243;P18242;Q05	3	11.2	48.373	445	3	8.25E-15
IPI00113141.1;IF	Citrate synthase, mitochondrial;Citrate synthase	1.021	0.980	Q9CZU6;Q0QEL9;Q8	6	11.4	51.736	464	7	2.20E-08
IPI00117352.1;IF	Tubulin beta-5 chain;Tubulin beta-3 chain;Tubuli	1.020	0.980	P99024;B1B178;Q80	6	23.4	49.67	444	6	2.36E-43

D4

IPI00311873.5	Serine/threonine-protein phosphatase PP1-beta	1.017	0.984	P62141;Q3TBE5;Q3T	5	19	37.186	327	4	1.65E-19
IPI00114491.1;IF	Cell division control protein 2 homolog;p34 prote	1.013	0.987	P11440;Q8R4A4;Q9	5	21.2	34.106	297	7	6.25E-18
IPI00133440.1;IF	Prohibitin;B-cell receptor-associated protein 32;A	1.013	0.987	P67778;Q3V2K0	5	23.9	29.82	272	8	4.24E-17
IPI00118930.1	Alpha-soluble NSF attachment protein;N-ethylma	1.011	0.989	Q9DB05;Q9CXX1	4	19.3	33.189	295	5	1.08E-26
IPI00228253.2;IF	Acetyl-CoA acetyltransferase, cytosolic;Cytosolic	1.008	0.992	Q8CAY6;A8XUS5;Q8I	2	8.6	41.297	397	2	1.59E-15
IPI00123199.1	Nucleosome assembly protein 1-like 1;NAP-1-rela	1.005	0.995	P28656;Q3TF41;Q8B	2	7.2	45.345	391	3	1.82E-08
IPI00230440.6;IF	Adenosylhomocysteinase;S-adenosyl-L-homocyst	1.004	0.996	P50247;Q3TF14;Q3U	11	37.7	47.688	432	15	7.72E-99
IPI00124771.1;IF	Phosphate carrier protein, mitochondrial;Phosph	1.000	1.000	Q8VEM8;Q3THU8;Q	5	14.6	39.632	357	5	1.53E-12
IPI00111045.1	Mitochondrial import inner membrane transloca	0.998	1.002	Q9D880	1	4.5	39.776	353	2	8.03E-09
IPI00336324.11;	Malate dehydrogenase, cytoplasmic;Cytosolic ma	0.996	1.004	P14152	8	29.9	36.511	334	10	5.29E-42
IPI00117896.3	Microtubule-associated protein RP/EB family me	0.994	1.006	Q61166;Q3U4H0	4	25.4	30.016	268	5	2.92E-19
IPI00120495.1;IF	Protein arginine N-methyltransferase 1;Prmt1 pr	0.991	1.009	Q9JIF0-1;Q9JIF0;Q80	9	29.9	42.435	371	12	3.78E-65
IPI00126635.1	Tumor-related protein (Isocitrate dehydrogenase	0.989	1.011	Q91VA7	2	11.2	42.194	384	2	3.01E-18
IPI00224575.1;IF	Heterogeneous nuclear ribonucleoprotein K	0.988	1.012	P61979-2;P61979;Q3	2	5.8	51.028	464	2	3.55E-11
IPI00308217.1;IF	Adenylosuccinate lyase;Adenylosuccinase;Lung R	0.984	1.016	P54822;Q3TMB8;Q8	4	11.8	54.807	484	4	1.24E-15
IPI00116356.5;IF	AP-2 complex subunit mu-1;Mu2-adaptin;AP-2 m	0.983	1.017	P84091;Q3TH69;Q3U	5	12.2	49.654	435	5	3.64E-15
IPI00323592.2	Malate dehydrogenase, mitochondrial	0.982	1.018	P08249	13	47.9	35.611	338	18	2.73E-60
IPI00750596.1;IF	Bone marrow macrophage cDNA, RIKEN full-leng	0.982	1.019	Q3UA17;Q9D050;Q7	3	17.9	34.554	312	5	7.15E-12
IPI00475378.4;IF	Bone marrow macrophage cDNA, RIKEN full-leng	0.982	1.019	Q3U5I2;Q5RJV5;Q6N	4	10.6	59.321	555	5	9.28E-13
IPI00283862.6	Proteasome subunit alpha type-1;Proteasome co	0.979	1.021	Q9R1P4;Q3TS44;Q8E	3	11	29.546	263	3	5.36E-05
IPI00263313.1	Developmentally-regulated GTP-binding protein	0.977	1.023	P32233;Q5NBZ3;Q8E	3	9.3	40.512	367	4	6.91E-14
IPI00406442.2	Succinyl-CoA ligase [GDP-forming] subunit alpha,	0.977	1.023	Q9WUM5	3	12.3	34.994	333	3	4.23E-09
IPI00269661.1;IF	Heterogeneous nuclear ribonucleoprotein A3;EG	0.977	1.023	Q8BG05-1;Q8BG05;(	4	15.6	39.652	379	8	1.35E-24
IPI00853914.1;IF	Heterogeneous nuclear ribonucleoproteins A2/B	0.970	1.031	O88569-1;O88569;O	7	8.4	87.62	814	15	6.51E-45
IPI00468481.2	ATP synthase subunit beta, mitochondrial	0.967	1.034	P56480	10	28.2	56.3	529	19	1.93E-44
IPI00127841.3;IF	ADP/ATP translocase 2;Adenine nucleotide trans	0.966	1.036	P51881;Q545A2;B0L	11	30.9	32.931	298	16	3.90E-38
IPI00123281.1;IF	Leucine-rich repeat-containing protein 59;17 day	0.964	1.037	Q922Q8;Q3UBL0;Q3	4	13.4	34.877	307	6	0.00045554
IPI00130343.2;IF	Bone marrow macrophage cDNA, RIKEN full-leng	0.962	1.039	Q3U6P5;Q9Z204-1;Q	4	14.4	36.905	334	5	9.61E-29
IPI00126716.3;IF	Eukaryotic initiation factor 4A-III;Eukaryotic trans	0.962	1.040	Q91VC3;Q3TEZ8;Q3U	7	17.5	46.839	411	5	1.71E-16
IPI00466128.3;IF	Alcohol dehydrogenase [NADP+];Aldehyde reduc	0.960	1.041	Q9JII6;Q3UJW9;Q54	4	12.3	36.586	325	4	6.13E-11
IPI00880400.1;IF	40S ribosomal protein S2;S4;LLRep3 protein;Bon	0.959	1.043	P25444;O89072;Q3T	13	39.6	31.231	293	19	1.64E-49
IPI00124692.1	Transaldolase	0.959	1.043	Q93092	14	43.6	37.387	337	17	3.49E-55
IPI00673290.1		0.958	1.044		1	9.1	17.798	165	3	3.92E-15
IPI00122565.1;IF	Rab GDP dissociation inhibitor beta;Guanosine di	0.955	1.047	Q61598-1;Q61598;Q	18	53.3	50.537	445	24	2.89E-89
IPI00624863.1;IF	Putative uncharacterized protein;Multifunctional	0.952	1.050	Q9CQ38;Q9DCL9	11	28	47.006	425	14	4.75E-50
IPI00132388.3	Basic leucine zipper and W2 domain-containing p	0.950	1.052	Q9CQC6;Q9CVI6	3	17.2	48.043	419	3	7.07E-45
IPI00136483.1	Proteasome subunit beta type-7;Proteasome sub	0.949	1.054	P70195;A2AQL6;Q5D	2	14.8	29.891	277	3	2.33E-09
IPI00130280.1;IF	ATP synthase subunit alpha, mitochondrial	0.947	1.056	Q03265	13	26.8	59.752	553	15	2.20E-58
IPI00119057.1;IF	Eukaryotic translation initiation factor 4E;eIF-4F 7	0.947	1.057	P63073;Q3TK95;Q8C	3	27.6	25.053	217	3	2.24E-15
IPI00320217.9;IF	T-complex protein 1 subunit beta;CCT-beta;Brain	0.940	1.064	P80314;Q542X7;Q9J	20	38.5	57.477	535	27	8.03E-115
IPI00134599.1;IF	40S ribosomal protein S3	0.938	1.066	P62908;O89068;Q3U	12	53.9	26.674	243	18	1.12E-39
IPI00120914.1	Eukaryotic translation initiation factor 3 subunit I	0.937	1.067	Q9DCH4;B0LAD0;Q3	5	16.9	38	361	6	5.10E-44

IPI00117087.1	Hsp90 co-chaperone Cdc37;Hsp90 chaperone pr	0.937	1.068	Q61081	5	17.9	44.593	379	5	3.25E-35
IPI00874482.1;IF	Actin, cytoplasmic 2;Gamma-actin;In vitro fertiliz	0.934	1.071	P63260;Q3UD81;Q4	10	38.7	41.792	375	2	1.79E-140
IPI00135087.3	COP9 signalosome complex subunit 5;Jun activat	0.933	1.072	O35864;Q3V0K7	4	13.8	37.548	334	4	6.52E-10
IPI00111831.1;IF	Nascent polypeptide-associated complex subunit	0.932	1.073	P70670;Q60817;Q3U	3	2	220.6	2187	4	5.01E-36
IPI00474883.2;IF	Capping protein (Actin filament) muscle Z-line, b	0.932	1.073	A2AMV7;P97856;Q3	3	16.3	33.767	301	5	1.07E-19
IPI00122350.4;IF	U1 small nuclear ribonucleoprotein A	0.930	1.076	Q62189;Q9CXX7	3	10.1	31.835	287	3	7.36E-06
IPI00123572.1	ADP-sugar pyrophosphatase;Nucleoside diphosp	0.930	1.076	Q9JKX6;A2ATT6	1	6.4	23.984	218	2	0.19889
IPI00110827.1;IF	Actin, alpha skeletal muscle;Alpha-actin-1;Actin,	0.925	1.081	P68134;Q61264;Q61	6	19.4	42.051	377	4	3.31E-33
IPI00118384.1;IF	14-3-3 protein epsilon;Tyrosine 3-monooxygenas	0.925	1.081	P62259;Q3V453;Q5S	18	60	29.174	255	37	3.37E-119
IPI00153103.3;IF	Multisynthetase complex auxiliary component p3	0.925	1.081	Q8R010;Q8R3V2	3	11.6	35.396	320	4	3.55E-05
IPI00110684.1	Inorganic pyrophosphatase;Pyrophosphate phos	0.922	1.085	Q9D819;Q3UA53;Q4	7	27.3	32.667	289	8	2.63E-55
IPI00874964.3;IF	Glyceraldehyde-3-phosphate dehydrogenase;Gly	0.920	1.088	P16858;A6H6A8;B0C	21	73.3	36.249	337	187	1.85E-205
IPI00114667.1	26S proteasome non-ATPase regulatory subunit 1	0.919	1.088	P26516;A1L3B8	5	24	36.539	321	8	7.79E-31
IPI00132942.1	Nuclear migration protein nudC;Nuclear distribut	0.919	1.089	O35685;A2A9F5	7	20.5	38.358	332	10	2.48E-24
IPI00123604.4;IF	40S ribosomal protein SA;Laminin receptor 1;p4C	0.918	1.090	P14206;B2CY77	7	25.8	32.838	295	10	5.24E-45
IPI00129928.2;IF	Fumarate hydratase, mitochondrial;EF-3	0.916	1.092	P97807-1;P97807;Q3	5	15.6	54.37	507	5	3.46E-28
IPI00317740.5;IF	Guanine nucleotide-binding protein subunit beta	0.916	1.092	P68040;Q3THP0;Q3T	13	54.6	35.076	317	17	1.54E-64
IPI00153740.1	Activator of 90 kDa heat shock protein ATPase ho	0.914	1.094	Q8BK64;Q3TL79	3	14.5	38.117	338	4	3.07E-14
IPI00554929.2;IF	Heat shock protein HSP 90-beta;HSP 84;Tumor-s	0.914	1.094	P11499;Q71LX8;Q80	4	7.6	83.324	724	5	4.87E-14
IPI00121758.1;IF	TAR DNA-binding protein 43;TARDBP S9 (TAR DN	0.911	1.098	Q921F2;Q3U591;Q3	2	5.3	44.547	414	2	0.00024148
IPI00122549.1;IF	Voltage-dependent anion-selective channel prot	0.911	1.098	Q60932-1;Q60932;B	11	42.2	32.351	296	16	5.49E-87
IPI00136251.1	DnaJ homolog subfamily A member 2	0.909	1.100	Q9QYJ0;Q3TFF0	3	9.2	45.745	412	3	7.28E-19
IPI00480307.1;IF	COP9 signalosome complex subunit 2;JAB1-conta	0.909	1.100	P61202-2;P61202;Q3	3	9.1	52.404	450	4	7.65E-19
IPI00307837.6;IF	Elongation factor 1-alpha 1;eEF1A-1;Elongation f	0.908	1.101	P10126;Q3TII3;Q3UA	5	10.4	50.113	462	6	7.14E-09
IPI00270326.1	26S protease regulatory subunit 7;Proteasome 2	0.908	1.102	P46471;Q3U5V3;Q3U	5	12.6	52.866	475	5	1.13E-15
IPI00828513.1;IF	17 days pregnant adult female amnion cDNA, RIK	0.907	1.103	Q3TFA5;Q3TJ97;Q3T	9	32.5	47.408	418	12	1.40E-48
IPI00458949.2;IF	Proteasome assembly chaperone 1;Down syndro	0.906	1.103	Q9JK23	2	8.3	33.104	289	2	1.81E-05
IPI00125971.1	26S protease regulatory subunit S10B;Proteasom	0.906	1.104	P62334;Q14AQ1;Q8C	5	14.4	44.172	389	5	1.34E-16
IPI00116308.1	Hsc70-interacting protein;Protein ST13 homolog;	0.906	1.104	Q99L47;Q3TGW8;Q3	7	19.1	41.655	371	7	2.21E-30
IPI00111412.3	60S ribosomal protein L4	0.905	1.105	Q9D8E6;Q564E8;Q5Z	18	35.8	47.153	419	28	3.54E-77
IPI00110885.2;IF	Cytochrome b5 reductase 3;13 days embryo liver	0.904	1.106	B0QZG1;Q9CY59;Q9	8	35.5	34.929	313	10	1.04E-33
IPI00856379.1;IF	Fructose-bisphosphate aldolase;Fructose-bispho	0.903	1.108	A6ZI44;Q6NY00;P05	16	52.9	45.12	418	32	6.48E-93
IPI00379245.2;IF	Glucosamine-6-phosphate isomerase 1;Glucosam	0.901	1.110	O88958;Q3TKA0;Q6A	7	34.3	32.55	289	9	1.10E-25
IPI00114925.1	13 days embryo liver cDNA, RIKEN full-length enr	0.900	1.111	Q3TFB5;Q99J77;Q9J	2	5.6	40.024	359	2	0.000749
IPI00314709.1;IF	Splicing factor, arginine/serine-rich 5;Pre-mRNA-	0.900	1.111	O35326;Q9D8S5;Q5I	2	8.5	30.944	270	3	4.07E-07
IPI00466610.5	Dual specificity mitogen-activated protein kinase	0.900	1.111	P31938;Q3TMJ8;Q3T	2	5.9	43.474	393	2	1.71E-13
IPI00454016.5;IF	S-adenosylmethionine synthetase isoform type-2	0.898	1.114	Q3THS6;Q3TED1;Q3	8	25.8	43.688	395	8	1.11E-46
IPI00321170.8;IF	60S ribosomal protein L3;J1 protein;Putative unc	0.898	1.114	P27659;Q3T9U9;Q3U	20	40.4	46.123	403	22	4.51E-76
IPI00222515.5	26S proteasome non-ATPase regulatory subunit 1	0.897	1.115	Q8BG32;Q3THG2;Q3	9	28	47.436	422	10	7.73E-57
IPI00121105.2	Hydroxyacyl-coenzyme A dehydrogenase, mitoch	0.895	1.117	Q61425	3	14.3	34.463	314	3	2.35E-19
IPI00116302.3;IF	Eukaryotic translation initiation factor 2 subunit 2	0.895	1.117	Q99L45;Q3ULL5;Q9C	9	29.9	38.092	331	10	4.00E-40
IPI00134820.1	Developmentally-regulated GTP-binding protein	0.895	1.117	Q9QXB9;Q5SX94	2	7.4	40.718	364	2	2.44E-08

D6

IPI00113655.1;IF	40S ribosomal protein S6;Phosphoprotein NP33	0.895	1.117	P62754;Q3TL53;Q5B	7	25.3	28.68	249	9	2.28E-35
IPI00381291.1;IF	26S proteasome non-ATPase regulatory subunit 4	0.894	1.119	O35226-2;O35226;Q	3	9	41.046	379	3	1.05E-08
IPI00457661.5	26S proteasome non-ATPase regulatory subunit 5	0.893	1.120	Q8BJY1;A2AUB0;Q5(	2	7.9	55.971	504	3	1.32E-42
IPI00132250.1	Eukaryotic translation initiation factor 3 subunit I	0.892	1.121	P60229;Q3UIG0;Q6L	9	24.5	52.22	445	11	7.70E-50
IPI00315550.3	Glutaredoxin-3;Thioredoxin-like protein 2;PKC-in	0.891	1.122	Q9CQM9	6	18.1	37.778	337	8	3.06E-45
IPI00553777.2;IF	Heterogeneous nuclear ribonucleoprotein A1;hn	0.890	1.123	P49312-2;P49312;P7	5	19.3	38.803	373	7	2.45E-29
IPI00312468.5	Eukaryotic peptide chain release factor subunit 1	0.888	1.126	Q8BWY3;P70199;Q3	9	27.9	49.03	437	10	4.43E-56
IPI00331345.5;IF	40S ribosomal protein S3a	0.888	1.127	P97351;Q3U5P8;Q3(	10	31.8	29.885	264	12	4.34E-25
IPI00136912.1;IF	Spermidine synthase;Putrescine aminopropyltrar	0.887	1.128	Q64674;Q3UJL7;Q54	7	40.4	33.995	302	9	1.16E-23
IPI00134300.1;IF	Lupus La protein homolog;La ribonucleoprotein;I	0.884	1.131	P32067;Q3TG93;Q5(	12	31.1	47.756	415	12	5.41E-51
IPI00277066.4;IF	S1 protein C2 (Heterogeneous nuclear ribonuclec	0.883	1.132	Q20BD0;Q3TMZ8;Q8	5	17.8	36.21	332	4	7.60E-24
IPI00230415.5;IF	Eukaryotic translation initiation factor 2 subunit 3	0.882	1.134	Q9Z0N1;Q3TML6;Q3	13	36.2	51.065	472	19	9.41E-87
IPI00307837.6;IF	Elongation factor 1-alpha 1;eEF1A-1;Elongation f	0.881	1.135	P10126;Q3TII3;Q3U/	16	55	50.113	462	56	5.24E-115
IPI00109615.1;IF	Nucleoporin Nup37;Putative uncharacterized prc	0.879	1.138	Q9CWU9;Q9CZ80	1	5.2	36.725	326	2	5.27E-10
IPI00137787.3	60S ribosomal protein L8	0.877	1.141	P62918;Q3THC7;Q3(	3	12.5	28.024	257	5	1.26E-11
IPI00269613.6;IF	Eukaryotic translation initiation factor 3 subunit I	0.875	1.143	Q9QZD9;A2AE04;Q3	4	17.8	36.46	325	5	4.15E-09
IPI00133066.3;IF	Bone marrow macrophage cDNA, RIKEN full-leng	0.874	1.144	Q3U8F5;Q9CQT4;Q9	7	18.9	52.895	456	9	5.36E-59
IPI00115580.4;IF	Eukaryotic translation initiation factor 3 subunit I	0.874	1.144	Q99JX4;A2A702;A2A	7	25.9	42.516	374	11	8.02E-35
IPI00131871.1;IF	COP9 signalosome complex subunit 4;JAB1-conta	0.873	1.145	O88544;Q14AI7;Q3T	4	14.8	46.284	406	5	7.69E-27
IPI00113738.2;IF	Tryptophanyl-tRNA synthetase, cytoplasmic;Tryp	0.872	1.146	P32921-1;P32921;Q3	5	17	54.357	481	6	2.66E-54
IPI00111556.1;IF	Serine/threonine-protein phosphatase 2A catalyt	0.872	1.147	P62715;Q8BN07;P63	2	9.7	35.575	309	2	3.75E-22
IPI00126048.2;IF	26S proteasome non-ATPase regulatory subunit 1	0.871	1.148	Q9WVJ2-1;Q9WVJ2;	10	30.1	42.809	376	10	9.86E-21
IPI00318841.4;IF	Elongation factor 1-gamma;eEF-1B gamma	0.869	1.151	Q9D8N0;Q4FZK2;Q8	11	30.7	50.06	437	16	5.80E-65
IPI00313222.5;IF	60S ribosomal protein L6;TAX-responsive enhanc	0.868	1.152	P47911;Q3UCH0;Q3	9	27	33.509	296	11	1.79E-36
IPI00131870.3	COP9 signalosome complex subunit 3;JAB1-conta	0.867	1.154	O88543;Q3TEA5;Q3(	2	7.3	47.832	423	2	4.77E-20
IPI00277066.4;IF	S1 protein C2 (Heterogeneous nuclear ribonuclec	0.865	1.156	Q20BD0;Q3TMZ8;Q8	9	28.9	36.21	332	11	2.10E-41
IPI00133920.3	Protein SEC13 homolog;SEC13-related protein;SE	0.864	1.157	Q9D1M0	2	11.8	35.565	322	5	5.06E-32
IPI00128202.1;IF	Eukaryotic translation initiation factor 3 subunit I	0.864	1.158	Q91WK2;Q3THW7;Q	7	27.3	39.832	352	9	1.16E-54
IPI00323971.2;IF	Inosine-5'-monophosphate dehydrogenase 2;IMI	0.861	1.162	P24547;O89058;Q3(	10	26.3	55.814	514	12	6.94E-72
IPI00119305.3;IF	Proliferation-associated protein 2G4;Proliferatior	0.859	1.164	P50580;Q05BN2;Q3	16	42.6	43.698	394	23	8.44E-86
IPI00828741.1;IF	HnRNP-associated with lethal yellow;RNA-bindin	0.858	1.166	A2AU63;Q64012-1;Q	3	14.1	33.188	312	3	2.91E-20
IPI00755892.1;IF	Heterogeneous nuclear ribonucleoprotein D-like;	0.856	1.169	Q9Z130	5	16.9	46.27	420	5	2.08E-27
IPI00849786.1;IF	Nuclease-sensitive element-binding protein 1;Y-b	0.855	1.170	P62960;Q3UBT1;Q6(	6	22.2	36.637	325	8	2.88E-63
IPI00319965.3	26S proteasome non-ATPase regulatory subunit 6	0.854	1.170	Q99JI4;Q8C1T2	8	29	45.536	389	9	3.49E-25
IPI00132194.3	Multisynthetase complex auxiliary component p4	0.854	1.171	P31230;Q3TUZ1;Q3(	2	9.1	35.166	319	3	7.98E-13
IPI00165694.1;IF	Mitochondrial import receptor subunit TOM34;T	0.853	1.172	Q9CYG7-1;Q9CYG7;(	2	7.1	34.278	309	2	0.023605
IPI00227773.1;IF	Serine/threonine-protein phosphatase PP1-gamr	0.848	1.180	P63087-2;P63087;Q3	5	19	38.504	337	9	7.88E-30
IPI00130016.1;IF	Acidic leucine-rich nuclear phosphoprotein 32 fa	0.842	1.187	P97822-1;P97822;Q3	4	23.1	29.622	260	6	4.91E-47
IPI00330958.2;IF	Heterogeneous nuclear ribonucleoprotein D0;AU	0.833	1.201	Q60668-1;Q60668;Q	9	27.6	38.354	355	15	3.70E-36
IPI00119138.1	Cytochrome b-c1 complex subunit 2, mitochondr	0.832	1.202	Q9DB77	5	15.9	48.234	453	5	2.41E-44
IPI00331092.7;IF	40S ribosomal protein S4, X isoform	0.832	1.202	P62702;Q3UXQ6;Q5	13	38.4	29.597	263	20	2.42E-61
IPI00775948.1;IF	60S ribosomal protein L7	0.828	1.208	P14148;Q3TK73;Q3(	8	23.7	32.416	279	20	6.23E-21

IPI00462453.4;IF	60S ribosomal protein L7a;Surfeit locus protein 3	0.826	1.210	P12970;Q58ET1;Q5E	11	29.2	32.624	288	24	3.58E-89
IPI00308706.4;IF	60S ribosomal protein L5;Blastocyst blastocyst cD	0.821	1.218	P47962;Q3THE1;Q3U	11	46.8	34.4	297	26	8.81E-59
IPI00132966.3;IF	40 kDa peptidyl-prolyl cis-trans isomerase;Cyclop	0.815	1.227	Q9CR16;Q3UB60;Q8	13	42.4	40.742	370	15	1.61E-64
IPI00133801.3;IF	Nuclear protein Hcc-1	0.809	1.236	Q9D1J3;B2RXM7	2	9	23.532	210	3	0.012215
IPI00474446.4	Eukaryotic translation initiation factor 2 subunit 3	0.808	1.237	Q6ZWX6;Q9CV24	9	37.1	36.108	315	13	5.84E-67
IPI00116753.4;IF	Electron transfer flavoprotein subunit alpha, mit	0.806	1.241	Q99LC5;B1B1B4	7	32.1	35.009	333	7	5.20E-50
IPI00127415.1;IF	Nucleophosmin;Nucleolar phosphoprotein B23;N	0.805	1.242	Q61937;Q3U536;Q5	9	52.4	32.56	292	22	2.40E-92
IPI00130343.2;IF	Bone marrow macrophage cDNA, RIKEN full-leng	0.805	1.243	Q3U6P5;Q9Z204-1;C	11	35.6	36.905	334	14	3.44E-59
IPI00133206.1;IF	26S protease regulatory subunit 6A;Proteasome	0.801	1.249	O88685;Q3THI5;Q3T	9	29.9	49.492	442	11	6.24E-55
IPI00132481.3;IF	Replication factor C subunit 5;Activator 1 subunit	0.798	1.253	Q9D0F6;Q3UDK3;Q5	2	8.3	38.096	339	2	6.34E-12
IPI00313475.1;IF	ATP synthase subunit gamma, mitochondrial;ATF	0.797	1.254	Q91VR2;Q3UD06;Q9	4	15.8	32.886	298	5	1.73E-10
IPI00555004.3;IF	Alcohol dehydrogenase class-3;Alcohol dehydrog	0.797	1.255	P28474;Q6P5I3	12	39.6	39.547	374	15	8.17E-113
IPI00121758.1;IF	TAR DNA-binding protein 43;TARDBP S9 (TAR DN	0.796	1.257	Q921F2;Q3U591;Q3'	7	25.1	44.547	414	9	2.52E-29
IPI00320208.3	Elongation factor 1-beta	0.796	1.257	O70251	4	23.1	24.693	225	5	4.60E-30
IPI00124677.5	Notchless protein homolog 1	0.791	1.264	Q8VEJ4;B1ARD5;Q3T	2	6.6	53.147	485	2	6.79E-11
IPI00409918.1;IF	Eukaryotic initiation factor 4A-II;ATP-dependent	0.789	1.268	P10630-2;P10630;P1	6	16.9	46.489	408	2	4.20E-24
IPI00111981.1;IF	Obg-like ATPase 1;GTP-binding protein 9	0.786	1.272	Q9CZ30-1;Q9CZ30;B	4	12.4	44.729	396	4	2.17E-19
IPI00112335.1	Dual specificity mitogen-activated protein kinase	0.784	1.275	P47809;Q543X6;Q8K	1	4.3	44.113	397	2	0.0016648
IPI00118676.3;IF	Eukaryotic initiation factor 4A-I;ATP-dependent F	0.784	1.275	P60843;Q3TFG3;Q3T	11	29.3	46.153	406	19	9.31E-85
IPI00113536.3;IF	Acidic leucine-rich nuclear phosphoprotein 32 fa	0.782	1.279	Q9EST5-2;Q9EST5;Q	3	10.9	37.447	329	5	4.13E-11
IPI00124372.3	4-trimethylaminobutyraldehyde dehydrogenase;	0.781	1.281	Q9JLJ2;A7VL18;Q3TC	7	18.3	55.887	518	7	5.59E-32
IPI00318550.5;IF	Interleukin enhancer-binding factor 2;Nuclear fac	0.781	1.281	Q9CXY6;Q3UXI9	7	31.5	43.062	390	8	1.34E-33
IPI00117312.1	Aspartate aminotransferase, mitochondrial;Trans	0.781	1.281	P05202	6	15.3	47.411	430	6	2.36E-27
IPI00111885.1;IF	Cytochrome b-c1 complex subunit 1, mitochondr	0.779	1.284	Q9CZ13;Q3THM1;Q3	2	5.8	52.768	480	2	0.0045918
IPI00118875.4;IF	Elongation factor 1-delta;9.5 days embryo parthe	0.776	1.288	P57776-1;P57776;Q6	7	11.7	72.93	660	8	2.76E-45
IPI00311131.1;IF	Replication protein A 32 kDa subunit;Replication	0.774	1.291	Q62193;Q3TE40;Q9S	2	11.1	29.718	270	2	2.11E-06
IPI00119581.2;IF	rRNA 2'-O-methyltransferase fibrillarin;Nucleolar	0.766	1.305	P35550;Q3UJS2	4	17.4	34.306	327	6	2.81E-17
IPI00848775.1;IPI00462605.5		0.764	1.310		14	31.7	43.801	407	2	4.23E-74
IPI00123878.4;IF	ATP-dependent RNA helicase DDX39;DEAD box p	0.763	1.310	Q8VDW0-1;Q8VDW0	9	30	49.067	427	11	8.27E-45
IPI00122442.1;IF	Branched-chain-amino-acid aminotransferase, m	0.763	1.311	O35855;Q3ULU3;Q8	4	13.7	44.127	393	4	4.10E-05
IPI00317794.5	Nucleolin;Protein C23	0.759	1.317	P09405;Q3TGR3;Q3T	6	11.2	76.722	707	6	4.41E-35
IPI00853924.1;IF	14-3-3 protein theta;14-3-3 protein tau	0.758	1.319	P68254-1;P68254;A3	5	17.8	34.348	303	2	5.14E-12
IPI00124938.1;IF	Ubiquitin carboxyl-terminal hydrolase isozyme L5	0.757	1.322	Q9WUP7-1;Q9WUP7	3	10.9	37.616	329	4	1.77E-22
IPI00110721.5;IF	Glyoxalase domain-containing protein 4	0.756	1.324	Q9CPV4-1;Q9CPV4;C	5	21.8	33.316	298	7	6.69E-26
IPI00118196.1	28S ribosomal protein S5, mitochondrial;S5mt;M	0.755	1.324	Q99N87;A2AHS8;Q3	1	3.2	48.206	432	2	0.14038
IPI00856490.1;IF	16 days embryo kidney cDNA, RIKEN full-length e	0.754	1.326	Q3TG45;Q3TI95;Q3T	3	9.1	39.93	353	5	1.24E-09
IPI00310880.4;IF	Putative uncharacterized protein	0.752	1.329	Q3TWW8;Q3UJV5;Q	3	6.2	39.025	339	4	0.010481
IPI00409462.2;IF	Spliceosome RNA helicase Bat1;DEAD box protein	0.752	1.329	Q9Z1N5;Q3TB18	9	31.8	49.035	428	6	5.67E-40
IPI00228548.6	Beta-enolase;2-phospho-D-glycerate hydro-lyase	0.749	1.335	P21550;Q4FK59;Q5S	4	16.1	47.024	434	2	1.02E-84
IPI00420724.4;IF	Putative uncharacterized protein;Mitotic checkpo	0.748	1.338	Q6ZWM5;Q8BH42;C	6	19.3	36.954	326	7	5.56E-19
IPI00133648.2	BRCA2 and CDKN1A-interacting protein	0.745	1.342	Q9CWI3	2	9.8	35.942	316	3	9.22E-10
IPI00119945.1	Nitrilase homolog 2	0.745	1.343	Q9JHW2	2	10.5	30.501	276	2	1.61E-12

IPI00828412.1;IF	Retinoblastoma binding protein 4 (Retinoblaston	0.744	1.344	A2A875;Q60972	6	23.8	47.655	425	4	1.62E-29
IPI00331251.1;IF	Short-chain acyl-CoA dehydrogenase;Short-chain	0.742	1.347	Q6LCR2;Q91W85;Q0	3	16.7	44.889	412	5	3.03E-13
IPI00133589.1;IF	Porphobilinogen deaminase;Pre-uroporphyrinog	0.740	1.352	P22907-1;P22907;Q3	2	11.4	39.302	361	3	2.57E-15
IPI00269661.1;IF	Heterogeneous nuclear ribonucleoprotein A3;EG	0.740	1.352	Q8BG05-1;Q8BG05;(	6	21.1	39.652	379	12	3.24E-40
IPI00130436.5	Brix domain-containing protein 2;Ribosome biog	0.739	1.353	Q9DCA5;Q8CEX3	4	19.8	41.241	353	4	1.59E-08
IPI00462072.3;IF	Alpha-enolase;2-phospho-D-glycerate hydro-lyas	0.736	1.359	P17182;Q5FW97;Q6	5	22.1	47.14	434	7	2.73E-56
IPI00130535.1;IF	Lipoamide acyltransferase component of branch	0.727	1.376	P53395;Q3TMF5;Q6	4	14.1	53.16	482	5	2.45E-18
IPI00407130.4;IF	Pyruvate kinase isozymes M1/M2;Pyruvate kinas	0.723	1.383	P52480-1;P52480;P5	2	6.2	57.844	531	2	1.10E-08
IPI00649191.2;IF	Proteaseome (Prosome, macropain) 28 subunit,	0.722	1.384	A2A4J1;P61290;Q4F	3	12.8	30.895	265	4	3.31E-20
IPI00124248.1;IF	Chloride channel regulator;Methylosome subunit	0.714	1.402	P97506;Q923F1;Q61	3	19.9	26.522	241	5	2.47E-35
IPI00118059.1;IF	Serine hydroxymethyltransferase, cytosolic;Glyci	0.714	1.402	P50431;Q8R0X9;Q9(	7	18.2	52.584	478	7	1.49E-34
IPI00469392.2;IF	Reticulon-4;Neurite outgrowth inhibitor;RTN4 (R	0.712	1.404	Q99P72-2;Q99P72;Q	4	4.1	126.61	1162	5	1.35E-24
IPI00127707.1;IF	Poly(rC)-binding protein 2;Alpha-CP2;Putative he	0.711	1.406	Q61990-1;Q61990;B	3	11.9	38.221	362	5	5.78E-11
IPI00849165.1;IF	Protein SET;Phosphatase 2A inhibitor I2PP2A;Ter	0.706	1.416	Q9EQU5-1;Q9EQU5;	2	8.9	35.309	305	3	9.56E-06
IPI00134746.5;IF	Argininosuccinate synthase;Citrulline--aspartate	0.706	1.416	P16460;Q3UEJ7;Q3U	9	25.7	46.584	412	12	4.01E-45
IPI00226943.1;IF	NEDD8-activating enzyme E1 catalytic subunit;Ub	0.705	1.418	Q8C878-1;Q8C878;Q	2	8	51.719	462	2	7.92E-06
IPI00624501.2	Arsenical pump-driving ATPase;Arsenite-transloc	0.702	1.424	O54984;Q3TAQ4;Q8	5	30.7	38.822	348	6	1.17E-38
IPI00555069.3;IF	Phosphoglycerate kinase 1	0.701	1.426	P09411;B0LAA9	21	55.6	44.55	417	29	4.61E-94
IPI00131873.1;IF	COP9 signalosome complex subunit 6;JAB1-conta	0.700	1.429	O88545;Q3UIT2	2	6.5	35.88	324	3	1.18E-08
IPI00856593.1;IF	CRL-1722 L5178Y-R cDNA, RIKEN full-length enric	0.687	1.456	Q3UJN2;P60122;Q3U	13	42.5	50.213	456	14	1.21E-58
IPI00135640.1;IF	26S protease regulatory subunit 8;Proteasome 2(	0.687	1.456	P62196;B1ARK2;Q9S	6	20.2	45.626	406	7	2.08E-47
IPI00323660.1;IF	Actin-like protein 6A;53 kDa BRG1-associated fac	0.683	1.463	Q9Z2N8;Q3TF62;Q5(	4	19.1	47.447	429	5	6.57E-30
IPI00462072.3;IF	Alpha-enolase;2-phospho-D-glycerate hydro-lyas	0.680	1.471	P17182;Q5FW97;Q6	25	63.1	47.14	434	118	1.01E-257
IPI00226073.2;IF	Heterogeneous nuclear ribonucleoprotein F	0.677	1.478	Q9Z2X1-1;Q9Z2X1;Q	6	21.2	45.729	415	8	1.58E-52
IPI00123557.3;IF	RuvB-like 2;p47 protein	0.674	1.483	Q9WTM5;Q3TXT7;Q	9	22.5	51.112	463	11	7.50E-59
IPI00224152.5	DNA-(apurinic or apyrimidinic site) lyase;Apurinic	0.672	1.488	P28352;Q544Z7	3	14.2	35.49	317	4	1.01E-22
IPI00115620.1;IF	Phosphoserine aminotransferase;Phosphohydrox	0.669	1.494	Q99K85;Q3U6K9;Q3	10	31.6	40.472	370	12	2.17E-41
IPI00130734.1	Thymidylate synthase	0.668	1.498	P07607;Q544L2;Q8C	3	11.1	34.958	307	4	9.93E-08
IPI00114862.1	THUMP domain-containing protein 1	0.667	1.500	Q99J36	2	7.7	38.884	350	3	6.96E-08
IPI00471475.1;IF	Plasminogen activator inhibitor 1 RNA-binding pr	0.662	1.510	Q9CY58-1;Q9CY58;Q	5	17.9	44.714	407	5	9.53E-26
IPI00111560.3;IF	Protein SET;Phosphatase 2A inhibitor I2PP2A;Ter	0.661	1.512	Q9EQU5-1;Q9EQU5;	5	19.4	33.377	289	6	4.05E-24
IPI00109813.1	12 days embryo head cDNA, RIKEN full-length en	0.661	1.514	Q9CX86	3	11.5	30.53	305	3	4.83E-19
IPI00120162.1;IF	Casein kinase II subunit alpha;Casein kinase II alp	0.660	1.514	Q60737;Q61177;Q8(	3	11	45.161	391	4	1.18E-12
IPI00126072.2	Synaptic vesicle membrane protein VAT-1 homol	0.659	1.516	Q62465;Q3TXD3;Q3I	5	17	43.096	406	5	2.13E-13
IPI00127707.1;IF	Poly(rC)-binding protein 2;Alpha-CP2;Putative he	0.658	1.519	Q61990-1;Q61990;B	4	14.1	38.221	362	3	6.27E-16
IPI00124979.2;IF	Putative uncharacterized protein;RNA binding m	0.655	1.527	Q9WV02;Q91VM5;A	3	9	42.3	391	3	1.71E-17
IPI00129105.2;IF	SUMO-activating enzyme subunit 1;Ubiquitin-like	0.650	1.537	Q9R1T2-1;Q9R1T2;Q	11	33.7	38.62	350	14	5.74E-42
IPI00121552.2;IF	Uridine 5'-monophosphate synthase;Orotate pho	0.640	1.563	P13439;Q3TVV7;Q54	10	36.2	52.292	481	13	7.63E-55
IPI00113232.3;IF	Cell growth-regulating nucleolar protein	0.637	1.570	Q08288	6	15.5	43.735	388	6	5.76E-26
IPI00314844.2;IF	Twinfilin-1;Protein A6	0.630	1.589	Q91YR1	2	12.3	40.079	350	2	0.005907
IPI00116221.2	Beta-lactamase-like protein 2	0.628	1.593	Q99KR3	2	12.5	32.754	288	2	1.53E-14
IPI00274407.1;IF	Elongation factor Tu, mitochondrial	0.625	1.600	Q8BFR5-1;Q8BFR5;Q	5	11.9	49.508	452	6	1.52E-17

IPI00130670.1	Serine-threonine kinase receptor-associated prot	0.622	1.609	Q9Z1Z2;B2RUC7	2	7.4	38.513	351	2	0.14678
IPI00661338.1;IF	2-oxoisovalerate dehydrogenase subunit beta, m	0.621	1.610	Q6P3A8-1;Q6P3A8;C	1	4.9	42.88	390	2	0.23091
IPI00112645.1;IF	Ribonucleoside-diphosphate reductase subunit M	0.621	1.612	P11157;Q3THV8	2	5.9	45.095	390	2	3.73E-05
IPI00119219.2;IF	Estradiol 17-beta-dehydrogenase 12;17-beta-hyc	0.616	1.623	O70503-1;O70503;Q	2	7.7	34.741	312	2	0.000134
IPI00113870.1;IF	Proliferating cell nuclear antigen;Cyclin;10, 11 da	0.603	1.657	P17918;Q542J9;Q91	9	41.8	28.785	261	19	1.73E-46
IPI00337893.2	Pyruvate dehydrogenase E1 component subunit	0.602	1.661	P35486;Q3UFJ3	2	9.5	43.231	390	2	1.00E-13
IPI00653841.1;IF	CRL-1722 L5178Y-R cDNA, RIKEN full-length enric	0.601	1.664	Q3UJV2;P47753;Q3T	2	10.3	33.294	290	2	2.64E-06
IPI00751369.1;IF	L-lactate dehydrogenase;L-lactate dehydrogenas	0.587	1.705	Q3TCI7;Q3THB4;Q3T	18	54.8	39.758	361	48	5.29E-141
IPI00223713.5;IF	Histone H1.2;H1 VAR.1;H1c;Histone H1.3;H1 VAR	0.585	1.709	P15864;Q3UXH2;Q5	5	21.2	21.266	212	6	2.03E-22
IPI00132042.1	Pyruvate dehydrogenase E1 component subunit	0.583	1.716	Q9D051	6	25.1	38.937	359	8	3.41E-33
IPI00308885.6;IF	60 kDa heat shock protein, mitochondrial;Heat sh	0.582	1.717	P63038-1;P63038;P6	2	8.6	60.955	573	2	8.59E-16
IPI00115650.4	Calcyclin-binding protein;Siah-interacting protein	0.562	1.778	Q9CXW3	3	10.9	26.51	229	4	6.53E-05
IPI00111218.1	Aldehyde dehydrogenase, mitochondrial;ALDH cl	0.513	1.948	P47738;Q3TVM2;Q3	8	20.6	56.537	519	9	5.81E-43
IPI00111218.1;IF	Aldehyde dehydrogenase, mitochondrial;ALDH cl	0.508	1.968	P47738;Q3TVM2;Q3	3	9.4	56.537	519	4	6.35E-09
IPI00134017.1	Cysteine and histidine-rich domain-containing pr	0.504	1.984	Q9D1P4	3	12.4	37.35	331	4	2.64E-11
IPI00468203.3;IF	Annexin A2;Annexin-2;Annexin II;Lipocortin II;Ca	0.502	1.990	P07356;Q3UCD3;Q5	3	8.3	38.676	339	3	0.00010762
IPI00118523.1	Adult male testis cDNA, RIKEN full-length enriche	0.482	2.075	Q9DAE2	2	5.7	41.912	385	2	1.39E-08
IPI00315452.5;IF	Purine nucleoside phosphorylase;Inosine phosph	0.449	2.225	P23492;Q4FJT6;Q54	8	34.9	32.277	289	11	2.28E-27
IPI00123278.1;IF	Pyrroline-5-carboxylate reductase 2;Pyrroline-5-c	0.444	2.254	Q922Q4;Q3TJ21;Q3	2	7.2	33.659	320	2	1.47E-11
IPI00454008.1	Serine hydroxymethyltransferase	0.430	2.326	Q3TFD0;Q99K87;Q9	8	18.7	55.76	504	9	4.43E-33
IPI00131259.2;IF	Galectin-3;Galactose-specific lectin 3;Mac-2 antig	0.365	2.739	P16110;Q3V471;Q8	2	8.3	27.515	264	3	9.42E-08
IPI00624175.1;IF	Lysophosphatidic acid phosphatase type 6;Acid p	0.361	2.767	Q8BP40-1;Q8BP40	3	8.6	47.624	418	4	1.27E-09
IPI00223757.4;IF	Aldose reductase;Aldehyde reductase;1 cell emb	0.343	2.913	P45376;Q3TCL2;Q3L	13	44	35.732	316	19	6.63E-24
IPI00114237.1;IF	C-terminal-binding protein 2	0.337	2.966	P56546-1;P56546;Q3	10	33.3	48.956	445	12	3.07E-45
IPI00762179.2;IF	Putative uncharacterized protein;Putative transp	0.323	3.099	Q9QUR4;P11260;Q6	2	5.9	42.866	371	2	0.012376
IPI00309089.6;IF	Septin-1;Differentiation protein 6;Peanut-like pro	0.305	3.277	P42209;B2RU74	2	11.5	42.019	366	2	1.93E-16
IPI00331660.5	L-threonine 3-dehydrogenase, mitochondrial	0.250	3.993	Q8K3F7	6	26.3	41.461	373	6	3.11E-20
IPI00323353.3;IF	Uridine phosphorylase 1	0.243	4.112	P52624;Q5SUC8	9	42.8	34.086	311	10	1.10E-21
IPI00331208.1	Uncharacterized protein FLJ11171 homolog	0.045	22.248	Q8BWQ4	1	1.2	87.142	767	37	0.1223
IPI00108085.1;IF	Nucleolysin TIAR;TIA-1-related protein;Nucleolys			P70318;Q545C1;P52	1	3.3	43.388	392	1	7.29E-05
IPI00760011.1;IF	10, 11 days embryo whole body cDNA, RIKEN full			Q9CZ03;Q6P8X1	1	2.7	46.648	406	1	7.62E-05
IPI00114819.3	Methylosome protein 50;WD repeat-containing p			Q99J09;Q3V2T0	1	9.4	36.942	342	1	1.93E-16
IPI00115607.3;IF	Trifunctional enzyme subunit beta, mitochondria			Q99JY0	1	2.3	51.386	475	1	0.016895
IPI00116748.1	NADH dehydrogenase [ubiquinone] 1 alpha subc			Q99LC3	1	4.2	40.603	355	1	0.030498
IPI00117705.1;IF	Dolichyl-diphosphooligosaccharide--protein glyco			O54734;Q3UC51;Q3	1	2.5	49.013	441	1	0.0020203
IPI00118762.6	Bystin			O54825;Q543N4	1	1.8	49.783	436	1	0.23563
IPI00118899.1;IF	Alpha-actinin-4;Non-muscle alpha-actinin 4;F-act			P57780;Q1A602;Q3U	1	1.1	104.98	912	1	0.040894
IPI00118904.1;IF	Protein farnesyltransferase/geranylgeranyltransf			Q61239;Q541Z2	1	4.8	44.013	377	1	0.1068
IPI00120457.1	Farnesyl pyrophosphate synthetase;Farnesyl diph			Q920E5;Q3TMB3;Q3	1	4.2	40.581	353	1	0.0052743
IPI00120790.2	p21-activated protein kinase-interacting protein			Q9DCE5;Q3UX26	1	4.7	42.116	382	1	0.00015777
IPI00121387.1	Guanine nucleotide-binding protein alpha-11 sub			P21278;Q3UPA1;Q9	1	4.2	42.024	359	1	0.022792
IPI00121471.1;IF	Serpin B6;Placental thrombin inhibitor;Proteinase			Q60854;Q3U3L3;Q4I	1	4.5	42.598	378	1	9.20E-13

D10

IPI00121562.1;IF	WD repeat-containing protein 12;YTM1 homolog	Q9JJA4;Q4V9X1	1	5.9	47.346	423	1	1.11E-07
IPI00830799.2;IF	Potassium voltage-gated channel subfamily KQT	Q9JK45;Q9CTU2	1	0.8	104.62	952	1	0.2162
IPI00122392.1	UPF0160 protein MYG1;Protein Gamm1	Q9JK81;Q05CL9;Q8R	1	2.6	42.722	380	1	0.016721
IPI00124073.1;IF	Nucleoredoxin;Protein Red-1	P97346-1;P97346;P9	1	6.4	48.343	435	1	9.76E-05
IPI00125939.2;IF	Cytosolic acyl coenzyme A thioester hydrolase;Lc	Q91V12-1;Q91V12;C	1	2.9	42.536	381	1	0.13816
IPI00776162.3;IF	Eukaryotic translation initiation factor 4 gamma 2	Q62448-2;Q62448;A	1	1.1	102.13	906	1	0.036479
IPI00126069.7;IF	Eif2b3 protein;Eukaryotic translation initiation fa	A0AUM9;A4FUS0;B1	1	3.3	50.488	452	1	0.15939
IPI00126940.1;IF	Adenosine kinase;Adenosine 5'-phosphotransfer	P55264-1;P55264;Q3	1	4.4	40.148	361	1	0.0019875
IPI00128334.3	Uromodulin;Tamm-Horsfall urinary glycoprotein	Q91X17;Q8CJA0	1	1.1	70.844	642	1	0.12302
IPI00129178.1	Ornithine aminotransferase, mitochondrial;Ornit	P29758;Q3TG75;Q3U	1	3.6	48.354	439	1	1.40E-05
IPI00130186.4	Vacuolar proton pump subunit C 1;V-ATPase sub	Q9Z1G3;Q3TG21;Q9	1	4.7	43.887	382	1	1.19E-08
IPI00130460.1	NADH dehydrogenase [ubiquinone] flavoprotein	Q91YT0	2	6.5	50.834	464	2	4.10E-12
IPI00130885.1	Heterogeneous nuclear ribonucleoprotein G;RNA	O35479	3	7.5	42.233	388	1	1.41E-08
IPI00133880.2	E3 ubiquitin-protein ligase RING2;RING finger pro	Q9CQJ4	1	3.3	37.623	336	1	0.00036688
IPI00136169.1;IF	Polyadenylate-binding protein 2;Poly(A)-binding	Q8CCS6-1;Q8CCS6;Q	1	4.3	32.296	302	1	0.0025318
IPI00136252.3	WD repeat-containing protein 18	Q4VBE8;Q8K265	1	3.2	47.211	431	1	3.54E-12
IPI00136655.1	2-amino-3-ketobutyrate coenzyme A ligase, mito	O88986;Q9CZ08	1	4.3	44.93	416	1	1.53E-14
IPI00137409.3	Transketolase;P68	P40142;O55021;Q9Z	1	2.7	67.63	623	1	0.0009987
IPI00139259.1	Splicing factor, arginine/serine-rich 10;Transform	P62996-1;P62996;Q5	1	5.6	33.665	288	2	6.90E-11
IPI00153950.1	Phosphoribosyl pyrophosphate synthetase-assoc	Q8R574;Q05BD4;Q5	1	3.8	40.88	369	1	0.016259
IPI00678657.3;IF	Splicing factor 3B subunit 4	Q8QZY9	1	5.4	44.625	425	1	1.16E-15
IPI00169463.1;IF	Tubulin beta-2C chain;Tubulin beta-2B chain;Tub	P68372;Q9CVR0;Q9I	6	24	49.83	445	1	6.70E-42
IPI00187240.1;IF	Probable ATP-dependent RNA helicase DDX47;DI	Q9CWX9;Q3TEK1;Q4	1	4.8	50.638	455	1	0.016062
IPI00187402.1;IF	tRNA-nucleotidyltransferase 1, mitochondrial;mi	Q8K1J6-1;Q8K1J6;Q8	1	3.5	49.895	434	1	0.10667
IPI00227720.3;IF	Protein DEK;9 days embryo whole body cDNA, RI	Q7TNV0;Q3U4W4;Q	2	6.6	43.158	380	2	5.39E-17
IPI00229274.1	Aminomethyltransferase, mitochondrial;Glycine	Q8CFA2;A2RSW6	1	3	44.008	403	1	0.019988
IPI00323644.4;IF	Trifunctional purine biosynthetic protein adenosi	Q64737-1;Q64737;Q	1	1.4	107.39	1010	1	6.36E-07
IPI00268927.1;IF	NudC domain-containing protein 3	Q8R1N4-1;Q8R1N4;E	2	8.8	40.89	363	2	0.003131
IPI00281011.7	MARCKS-related protein;MARCKS-like protein 1;I	P28667;B2KGE6	1	7.5	20.165	200	1	2.45E-09
IPI00313296.3	Ribonuclease inhibitor;Ribonuclease/angiogenin	Q91VI7;Q3UM23	1	4.2	49.816	456	1	6.50E-12
IPI00314788.5	Argininosuccinate lyase;Arginosuccinase	Q91YI0;Q3UUH0	1	2.4	51.739	464	1	9.39E-05
IPI00318725.4	Ribosome biogenesis regulatory protein homolog	Q9CYH6;Q80U76	1	3.6	41.551	365	1	0.15672
IPI00319956.1	Protein IMPACT;Imprinted and ancient gene prot	O55091	1	8.2	36.276	318	1	0.013042
IPI00323144.4;IF	Rpo1-1 protein;DNA-directed RNA polymerases I	A1L3C2;Q3UJZ9;P52	1	3.2	39.237	347	1	0.085451
IPI00331155.4;IF	Protein ADRM1;Adhesion-regulating molecule 1;	Q9JKV1	1	3.9	42.06	407	1	0.0015255
IPI00346834.1	Adult female vagina cDNA, RIKEN full-length enri	Q3UV17	1	1.5	62.844	594	1	0.021013
IPI00387232.1;IF	NSFL1 cofactor p47;p97 cofactor p47	Q9CZ44-3;Q9CZ44;Q	1	4.3	40.953	372	1	7.78E-07
IPI00395100.1;IF	Try10-like trypsinogen;Trypsinogen 10 (Trypsin 1	Q7M754;Q792Z1	1	4.1	26.531	246	1	0.033218
IPI00403909.2;IF	Armadillo repeat-containing protein 6	Q8BNU0-1;Q8BNU0;	1	6	50.683	468	1	0.0020285
IPI00458711.2	13 days embryo forelimb cDNA, RIKEN full-length	Q8BSQ7;Q8K0C4;Q9	1	3.2	56.775	503	1	0.00055958
IPI00461381.4;IF	Retrotransposon-derived protein PEG10;Paterna	Q7TN75-1;Q7TN75;C	1	1.1	109.85	958	1	0.081418
IPI00461969.2;IF	Cell division cycle protein 123 homolog	Q8CII2-1;Q8CII2;A2A	1	3	38.816	336	2	0.14086

IPI00474157.4;IF	Mitochondrial import receptor subunit TOM40 h	Q9QYA2;Q3UQD0	1	4.4	37.895	361	1	3.20E-06
IPI00475299.3	SET and MYND domain-containing protein 5;Reti	Q3TYX3;Q61159	2	4.8	47.095	416	3	8.75E-06
IPI00622371.3	Eukaryotic translation initiation factor 3 subunit	Q9Z1D1;Q3THA0;Q5	1	2.8	35.638	320	1	0.11842
IPI00626790.3;IF	Glutamine synthetase;Glutamate--ammonia ligas	P15105	1	4.3	42.119	373	1	2.78E-10
IPI00874774.1;IF	Estrogen related receptor, beta (13 days embryo	A2RTQ7;B2RRU9	1	2.1	48.327	433	1	0.072371
IPI00885509.1	Translation initiation factor eIF-2B subunit alpha/	Q9CQT1	1	7.6	39.41	369	1	0.17823
IPI00310323.1;IF	Retinoid-inducible serine carboxypeptidase;Serin	Q920A5;Q9D625;Q9	1	3.1	50.966	452	1	2.55E-06
IPI00111265.3	F-actin-capping protein subunit alpha-2;CapZ alp	P47754;Q3U7G3;Q3	1	5.2	32.967	286	1	0.0012692
IPI00111416.1	Syntaxin-12	Q9ER00;A2ADS0;Q3	1	5.5	31.195	274	1	0.015634
IPI00229539.2;IF	Histone H2B type 3-B;Histone H2B type 1-P;Histo	Q8CGP0;Q8CGP2-2;(	1	5.8	17.145	154	1	0.16786
IPI00112674.1;IF	Putative uncharacterized protein	Q9D9G7	1	2.8	28.406	251	1	0.17866
IPI00113052.1	Elongation factor Ts, mitochondrial	Q9CZR8;Q3TA37;Q8I	1	4	35.334	324	1	0.012535
IPI00113262.1	26S proteasome non-ATPase regulatory subunit	O35593;A2AR77;Q3U	1	4.2	34.577	310	1	6.52E-06
IPI00113517.1	Cathepsin B;Cathepsin B1;Cathepsin B light chain	P10605;Q3TC17;Q3T	1	5.3	37.279	339	1	6.14E-14
IPI00114329.1	Glutamate--cysteine ligase regulatory subunit;Ga	O09172;Q3T9P2;Q4F	1	5.5	30.534	274	1	1.42E-05
IPI00114401.1;IF	Emerin;TIB-55 BB88 cDNA, RIKEN full-length enri	O08579;Q3TIH6;Q3U	1	5.8	29.435	259	1	4.62E-08
IPI00114407.2;IF	THO complex subunit 4;Ally of AML-1 and LEF-1;T	O08583-1;O08583;Q	1	5.1	26.94	255	1	1.30E-06
IPI00115564.5;IF	ADP/ATP translocase 1;Adenine nucleotide trans	P48962;Q8BVI9	7	20.1	32.904	298	1	9.75E-13
IPI00116222.1	3-hydroxyisobutyrate dehydrogenase, mitochon	Q99L13;A0ZNJ2	1	4.8	35.44	335	1	0.10206
IPI00403468.3;IF	Transcription factor SOX-3;Transcription factor S	P53784;Q4VBD8;Q5I	1	1.9	50.786	470	2	0.21843
IPI00118723.3	Serine/threonine-protein phosphatase 2A regula	P58389;Q543N6;Q8(	1	9	36.71	323	1	0.066466
IPI00283671.4;IF	Uncharacterized protein C1orf77 homolog	Q9CY57-1;Q9CY57;Q	1	5.2	26.585	249	1	5.88E-11
IPI00121159.3;IF	Brix domain-containing protein 1	Q9JJ80	1	4.9	35.363	306	1	6.60E-08
IPI00122015.1	Protein FAM49B	Q921M7	1	4	36.776	324	1	2.71E-08
IPI00761873.1;IF	Mcm3ap protein;80 kDa MCM3-associated prote	Q6NWV4;Q7TS87;Q8	1	0.3	217.36	1971	1	0.13011
IPI00126255.1	Inositol monophosphatase 2;Inositol-1(or 4)-mor	Q91UZ5;Q3TAU6;Q3	1	3.4	31.716	290	1	0.0026404
IPI00678029.2;IF	Cell division protein kinase 4;Cyclin-dependent ki	P30285;Q545C3;Q8E	1	3.7	45.85	406	1	0.10373
IPI00128692.1	Sterol-4-alpha-carboxylate 3-dehydrogenase, de	Q9R1J0;Q3US15;Q7T	1	3.6	40.685	362	1	0.0063169
IPI00130185.1	Serine/threonine-protein phosphatase PP1-alpha	P62137;Q3U8W0	5	17.9	37.54	330	1	3.34E-18
IPI00130840.7;IF	Coatomer subunit epsilon;Epsilon-coat protein	O89079;Q9D1J2	1	7.1	34.567	308	1	0.20616
IPI00131909.1	RNA-binding protein PNO1	Q9CPS7	1	6.5	27.453	248	1	0.0010518
IPI00132096.3	S-methyl-5'-thioadenosine phosphorylase;5'-met	Q9CQ65;A2ANM4	1	6.7	31.062	283	1	1.07E-12
IPI00132578.1;IF	mRNA turnover protein 4 homolog;MRT4, mRNA	Q9D0I8;A2AMU9;A2	1	3.3	27.545	239	1	0.13248
IPI00132723.3;IF	Exosome complex exonuclease RRP42;Ribosoma	Q9D0M0;Q4VBW5;C	1	3.4	31.665	291	1	0.019455
IPI00135977.3	Chloride intracellular channel protein 4	Q9QYB1;Q543N5	2	11.5	28.729	253	1	0.00042516
IPI00749740.2;IF	WD repeat domain 5;WD repeat-containing prot	A2AKB2;Q3UNQ3;P6	1	3.4	38.427	350	1	0.0067437
IPI00153874.1	Transcription initiation factor IIB;General transcr	P62915;Q3ULN2;Q3U	1	3.5	34.819	316	1	0.16274
IPI00154004.1	Ubiquitin thioesterase OTUB1;Otubain-1;OTU do	Q7TQI3	1	5.5	31.27	271	1	6.75E-18
IPI00221414.1	WD repeat-containing protein 82	Q8BFQ4;B2RXQ8	1	6.4	35.079	313	1	2.86E-07
IPI00318548.3;IF	Splicing factor U2AF 35 kDa subunit;U2 auxiliary	Q9D883;Q14C24;Q8	1	5	27.815	239	1	0.15936
IPI00223216.5;IF	Thiosulfate sulfurtransferase;Rhodanese	P52196;Q545S0	1	3	33.466	297	1	0.024835
IPI00228548.6	Beta-enolase;2-phospho-D-glycerate hydro-lyase	P21550;Q4FK59;Q5S	2	7.6	47.024	434	1	4.74E-29

IPI00228616.5	Histone H1.1;H1 VAR.3	P43275;Q5SZ98	1	9.4	21.785	213	1	0.00065266
IPI00230133.5	Histone H1.5;H1 VAR.5;H1b	P43276;Q1WWK3;Q	1	5.4	22.576	223	1	0.0039886
IPI00265025.5	Galactokinase;Galactose kinase	Q9R0N0;Q80UL3;Q9	1	3.1	42.295	392	1	0.004276
IPI00276926.3;IF	NOD-derived CD11c +ve dendritic cells cDNA, RIK	Q3TDH6;Q8JZU2	1	3.5	33.931	311	1	0.0085787
IPI00277001.4	Proteasome subunit alpha type-4;Proteasome co	Q9R1P0;Q3TL95;Q3U	1	3.8	29.47	261	1	0.049901
IPI00308498.1	Caspase-3;Apopain;Cysteine protease CPP32;Yan	P70677;Q8BNT4	1	4.3	31.474	277	1	0.08632
IPI00308609.1;IF	Putative uncharacterized protein;Vesicular integ	Q8BJL4;Q9DBH5	1	3.4	40.429	358	1	0.0923
IPI00317794.5	Nucleolin;Protein C23	P09405;Q3TGR3;Q3T	1	1.8	76.722	707	1	7.58E-08
IPI00900411.1;IF	Ribose-phosphate pyrophosphokinase;Ribose-ph	Q32M04;Q9CS42;A2	1	2.5	35.093	320	1	0.14153
IPI00653266.1;IF	17 days embryo kidney cDNA, RIKEN full-length e	Q3UI84;Q99J62	1	2.4	40.918	370	1	0.05988
IPI00323357.3;IF	Heat shock cognate 71 kDa protein;Heat shock 7	P63017;Q3KQJ4;Q3T	1	2.2	70.87	646	1	0.042094
IPI00329942.4;IF	Vacuolar protein sorting-associated protein 26A;	P40336-2;P40336;P4	1	4.7	41.547	359	1	0.0029079
IPI00330057.4	Density-regulated protein	Q9CQJ6	1	8.6	22.166	198	1	0.00289
IPI00420261.5;IF	High mobility group protein B1;High mobility gro	P63158;Q3UBK2;Q3U	1	3.7	24.893	215	1	0.08522
IPI00331068.1	DNA repair protein RAD51 homolog 2;RAD51-like	O35719-1;O35719;Q	1	1.7	38.152	350	1	0.23446
IPI00346834.1	Adult female vagina cDNA, RIKEN full-length enri	Q3UV17	1	1.5	62.844	594	1	0.098054
IPI00347376.1	Diamine acetyltransferase 2;Spermidine/spermir	Q6P8J2;Q3TTS0;Q5F	1	4.1	19.305	170	1	0.21455
IPI00357056.5;IPI00755590.1			1	2.1	48.29	424	2	0.22565
IPI00648709.2;IF	E3 ubiquitin-protein ligase UBR4;N-recognin-4;Zi	A2AN08-5;A2AN08;A	1	0.7	575.17	5208	1	0.0013686
IPI00378933.4			3	20.3	13.595	118	1	8.07E-05
IPI00380309.3;IF	Putative RNA-binding protein Luc7-like 2;CGI-74	Q7TNC4-1;Q7TNC4;(	1	2.6	46.582	392	1	0.15871
IPI00761402.2;IF	Olfactomedin-4	Q3UZZ4;B2RUF9	1	1.1	60.967	538	1	0.14295
IPI00420807.3;IF	Splicing factor, arginine/serine-rich 1	Q6PDM2-1;Q6PDM2	1	4	27.744	248	1	0.0089592
IPI00555004.3	Alcohol dehydrogenase class-3;Alcohol dehydrog	P28474;Q6P5I3	1	4	39.547	374	1	5.73E-12
IPI00762185.2;IF	3-mercaptopyruvate sulfurtransferase;Adult mal	Q99J99;Q3UW66;Q5	1	3.3	33.267	299	2	0.0077238
IPI00652902.1;IF	Putative uncharacterized protein	Q3TXK7;Q8JZT4;Q92	1	4.5	40.489	355	1	8.62E-09

Appendix E - Alignment of human and mouse DGCR8 protein sequences

```
CLUSTAL 2.1 multiple sequence alignment


DGCR8human      METDESPSPLPCGPAGEAVMESRARPFQALPREQSPPPPLQTSSGAEVMDVGSGGDGQSE 60
DGCR8mouse      METYESPSPLPREPAGEAMMENRACPFQVLPHEQSPPPPLQTSSDAEVMDVGSGGDGQSE 60
                *** *******  *****:**.** ***.**:************.***************

DGCR8human      LPAEDPFNFYGASLLSKGSFSKGRLLIDPNCSGHSPRTARHAPAVRKFSPDLKLLKDVKI 120
DGCR8mouse      PPADDPFNFYGASLLSKGSFSKGRLLIDPNCSGHSPRTARHAPAVRKFSPDLKLLKDVKI 120
                 **:************************************************************

DGCR8human      SVSFTESCRSKDRKVLYTGAERDVRAECGLLLSPVSGDVHACPFGGSVGDGVGIGGESAD 180
DGCR8mouse      SVSFTESCRSKDRKVLYTGVERSTRPECGQLLSPVSGDVHACPFGGSVGNGVGLGGESAD 180
                ******************.**..*.*** ******************:***:*******

DGCR8human      KKDEENELDQEKRVEYAVLDELEDFTDNLELDEEGAGGFTAKAIVQRDRVDEEALNFPYE 240
DGCR8mouse      KKDEENELDQEKRVEYAVLDELEDFTDNLELDEEGTGGFTAKAIVQRDRVDEEALNFSYE 240
                ************************************:**********************.**

DGCR8human      DDFDNDVDALLEEGLCAPKKRRTEEKYGGDSDHPSDGETSVQPMMTKIKTVLKSRGRPPT 300
DGCR8mouse      DDFDNDVDALLEEGLCAPKKRRMEEKYGGDSDHPSDGETSVQPMMTKIKTVLKSRGRPPT 300
                ********************** ************************************

DGCR8human      EPLPDGWIMTFHNSGVPVYLHRESRVVTWSRPYFLGTGSIRKHDPPLSSIPCLHYKKMKD 360
DGCR8mouse      EPLPDGWIMTFHNSGVPVYLHRESRVVTWSRPYFLGTGSIRKHDPPLSSIPCLHYKKMKD 360
                ************************************************************

DGCR8human      NEEREQSSDLTPSGDVSPVKPLSRSAELEFPLDEPDSMGADPGPPDEKDPLGAEAAPGAL 420
DGCR8mouse      NEEREQNCDLAPSGEVSPVKPLGRSAELDFPLEEPDSMGGDSGSMDEKDPLGAEAAAGAL 420
                ******..**:***:********.*****:***:******.*.*. ***********.***

DGCR8human      GQVKAKVEVCKDESVDLEEFRSYLEKRFDFEQVTVKKFRTWAERRQFNREMKRKQAESER 480
DGCR8mouse      GQVKAKVEVCKDESVDLEEFRNYLEKRFDFEQVTVKKFRTWAERRQFNREMKRKQAESER 480
                **********************.****************************************

DGCR8human      PILPANQKLITLSVQDAPTKKEFVINPNGKSEVCILHEYMQRVLKVRPVYNFFECENPSE 540
DGCR8mouse      PILPANQKLITLSVQDAPTKKEFVINPNGKSEVCILHEYMQRVLKVRPVYNFFECENPSE 540
                ************************************************************

DGCR8human      PFGASVTIDGVTYGSGTASSKKLAKNKAARATLEILIPDFVKQTSEEKPKDSEELEYFNH 600
DGCR8mouse      PFGASVTIDGVTYGSGTASSKKLAKNKAARATLEILIPDFVKQTSEEKPKDSEELEYFNH 600
                ************************************************************

DGCR8human      ISIEDSRVYELTSKAGLLSPYQILHECLKRNHGMGDTSIKFEVVPGKNQKSEYVMACGKH 660
DGCR8mouse      ISIEDSRVYELTSKAGLLSPYQILHECLKRNHGMGDTSIKFEVVPGKNQKSEYVMACGKH 660
                ************************************************************

DGCR8human      TVRGWCKNKRVGKQLASQKILQLLHPHVKNWGSLLRMYGRESSKMVKQETSDKSVIELQQ 720
DGCR8mouse      TVRGWCKNKRVGKQLASQKILQLLHPHVKNWGSLLRMYGRESSKMVKQETSDKSVIELQQ 720
                ************************************************************

DGCR8human      YAKKNKPNLHILSKLQEEMKRLAEEREETRKKPKMSIVASAQPGGEPLCTVDV 773
DGCR8mouse      YAKKNRPNLHILSKLQEEMKRLAAEREETRKKPKMSIVASAQPGGEPLCTVDV 773
                *****:***************** ******************************
```