Induced Pluripotent Stem Cells as an Alternative to Embryonic Stem Cells for the Treatment of Type 1 Diabetes

A thesis submitted by

Natalie Francis

For the degree of Doctor of Philosophy from Imperial College London

Beta Cell Physiology Group

Department of Medicine

Imperial College London

Acknowledgements: Firstly, huge thanks to my NIBSC supervisor Chris, for all your help and support throughout my PhD. Thank you for staying positive even when experiments weren't working! Thanks to all my colleagues at NIBSC, especially to Mel for all your practical help, and to my colleagues at the UK Stem Cell Bank for passing on their knowledge (and cells!) when I needed them. Thanks also to my supervisor Guy at Imperial College for your helpful advice. Finally, thank you to my husband Rob, who proved that getting married during a PhD is not always a bad idea, and for your continuous love and support.

Declaration of originality: I declare that all work presented in this thesis is my own unless otherwise referenced.

Abstract

Type 1 diabetes mellitus (T1DM) results from auto-immune destruction of the insulin-secreting β-cells of the pancreas. The most common treatment is injection of exogenous insulin, but this allows only partial control over blood glucose levels, so other therapies are needed. Pancreatic islet transplantation has shown proof of principle for cell replacement therapy to treat T1DM. There are several sources of cells which could be used, but much of the focus has been on pluripotent stem cells, which are able to self-renew indefinitely in culture and give rise to any cell in the body. Insulin-expressing cells have successfully been produced from embryonic stem cells (ESCs) by recapitulating embryonic development *in vitro*. However, problems associated with the use of ESCs mean that an alternative cell source is needed. In 2006, it was discovered that 4 transcription factors can reprogram somatic cells into induced pluripotent stem cells (iPSCs). iPSCs provide an alternative source of pluripotent stem cells and can be derived in a patient-specific manner. iPSCs have been shown to differentiate *in vitro* into insulin-expressing cells, but it is unknown whether iPSCs are truly equivalent to ESCs. Important differences have been shown to exist between iPSCs and ESCs which may affect the ability of iPSCs to give rise to cells of a pancreatic lineage and therefore limit their usefulness for the treatment of T1DM. The aim of this project is to identify whether iPSCs are a viable alternative to ESCs for generating β-cells *in vitro* for cell replacement therapy to treat type 1 diabetes. The differentiation potential of iPSCs and ESCs to give rise to first definitive endoderm (the first stage in differentiation towards a pancreatic lineage) *in vitro* will be compared, and the involvement of miRNAs in differentiation of ESCs and iPSCs to definitive endoderm will be investigated.

Table Of Contents

List of Figures

List of Tables

Chapter 1
Introduction

Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Diabetes Mellitus

Diabetes mellitus is defined as "a group of metabolic diseases characterized by hyperglycaemia resulting from defects in insulin secretion, insulin action, or both." In 2000, there were around 171 million people with diabetes worldwide, and this is expected to rise to 366 million by 2030 (World Health Organisation, 2006).

1.1.1 Type 1 Diabetes Mellitus (T1DM)

Type 1 diabetes mellitus (T1DM) accounts for around 10% of cases of adult-onset diabetes, and results from T cell-mediated auto-immune destruction of the insulin-secreting β-cells of the pancreas, leading to insufficient insulin production and causing chronic hyperglycaemia (Anderson & Bluestone, 2005). Long-term complications associated with the microvascular damage caused by chronic hyperglycaemia include retinopathy, nephropathy, neuropathy, and cardiovascular disease (Report of the Expert Committee on the Diagnosis & Classification of Diabetes Mellitus, 2003).

Susceptibility to T1DM is largely inherited, with genetic susceptibility residing mainly in the HLA complexes, as well as in several other genetic loci termed insulin-dependent diabetes mellitus (*IDDM)* susceptibility genes, which are thought to be important regulators of the immune response. The HLA locus is thought to confer about 50% of the genetic susceptibility, with roughly 15% contribution from two other genes: *insulin-VNTR* (*IDDM2*) and *CTLA-4* (*IDDM12*), and minor contributions from the other *IDDM* genes (reviewed in Bluestone *et al*., 2010). However, although a number of susceptibility alleles have been identified through genome-wide association studies, these have not identified monogenic defects relating directly to disease onset: instead, it seems more likely that a global defect in immune regulation may underlie disease susceptibility. Onset of the disease may also require exposure to environmental factors, many of which are still poorly understood, but which are

thought to include certain viruses e.g. congenital rubella; environmental toxins; or certain foods (reviewed in Daneman, 2006).

The abnormal activation of the immune system in susceptible individuals leads to an inflammatory response within the pancreatic islets, termed insulitis. Histology of the pancreatic islets shows infiltration by macrophages and T cells. Islet antigen-reactive T cells have been identified both in the NOD mouse model and in human patients (Bluestone *et al*, 2010). There is also a humoral response with production of antibodies by B cells to β-cell antigens such as insulin (Atkinson *et al*., 1986), glutamic acid decarboxylase (Kaufman *et al*., 1992; Schott *et al*., 1994), and the protein tyrosine phosphatase IA2 (Lan *et al*., 1996; Lu *et al*., 1996). The presence of one or more of these antibodies is present in 85-90% of individuals diagnosed with T1DM, and can precede the clinical onset of T1DM by years (Expert Committee on the Diagnosis and Classification of Diabetes Mellitus, 2003). Other immune cell types are also involved in the pathogenesis of T1DM, including natural killer cells and monocytes (Bluestone *et al*., 2010).

1.1.2 Type 2 Diabetes Mellitus (T2DM)

T2DM is a complex disease, usually of late onset, which comprises around 90% of all diabetics. It is more heterogeneous than T1DM, as its aetiology involves not only a loss of responsiveness of β-cells to glucose, resulting in insufficient insulin production, but also end-organ insulin resistance. This results in a compensatory demand for increased β-cell insulin production, eventually leading to β-cell death (Expert Committee on the Diagnosis and Classification of Diabetes Mellitus, 2003).

T2DM is a complex polygenic disease, and is linked to polymorphisms in several genes (McCarthy & Zeggini, 2009). Environmental factors, especially obesity and other associated lifestyle factors such as physical inactivity and poor diet, are also known to contribute to T2DM pathogenesis. Obesity itself causes insulin resistance through increased release of non-esterified fatty acids, glycerol, hormones (including leptin and adiponectin), and pro-inflammatory cytokines from adipose tissue (Kahn *et al*., 2006).

Treatment for T2DM includes diet control and exercise, sometimes supplemented with oral hypoglycaemic agents. Exercise decreases hyperglycaemia as skeletal muscle activity does not require insulin for glucose uptake. Some oral hypoglycaemic agents used in T2DM treatment act to stimulate insulin secretion from the β-cells, while others act by slowing degradation of carbohydrates in the intestine, inhibiting hepatic glucose output, or increasing target-tissue responsiveness to insulin. Insulin therapy may also be needed in cases where first-line treatment fails to adequately control blood glucose levels (reviewed in Derosa & Sibilla, 2007).

1.1.3 Other Types of Diabetes

Gestational diabetes mellitus (GDM) is any degree of glucose intolerance which first occurs during pregnancy, and is associated with an increased risk of perinatal morbidity and mortality. It is caused by several hormones, including oestrogen, progesterone, and human placental lactogen, which cause insulin resistance, keeping blood glucose levels high so that adequate amounts can be taken up by the foetus. In order to compensate, β-cell numbers in the pregnant woman increase, resulting in higher insulin release, but in some cases this response is inadequate, resulting in hyperglycaemia. In the majority of cases, glucose regulation will return to normal after delivery (Expert Committee on the Diagnosis and Classification of Diabetes Mellitus, 2003).

Several forms of diabetes are associated with monogenic defects in β-cell function, which are inherited in an autosomal dominant pattern. Several common sites of mutation have been identified, and include HNF-1α, HNF-4α and glucokinase. These forms of diabetes result in the onset of hyperglycaemia at an early age, and are referred as maturity-onset diabetes of the young (MODY). They are characterised by impaired insulin secretion with no defects in insulin action. Conversely, some rare cases of diabetes are caused by genetically determined abnormalities of insulin action e.g. mutations of the insulin receptor, or inability to convert proinsulin to insulin. Other uncommon causes of diabetes include disease of the pancreas, endocrinopathies, and certain drugs or chemicals (Expert Committee on the Diagnosis and Classification of Diabetes Mellitus, 2003).

1.2 Structure of the Pancreas

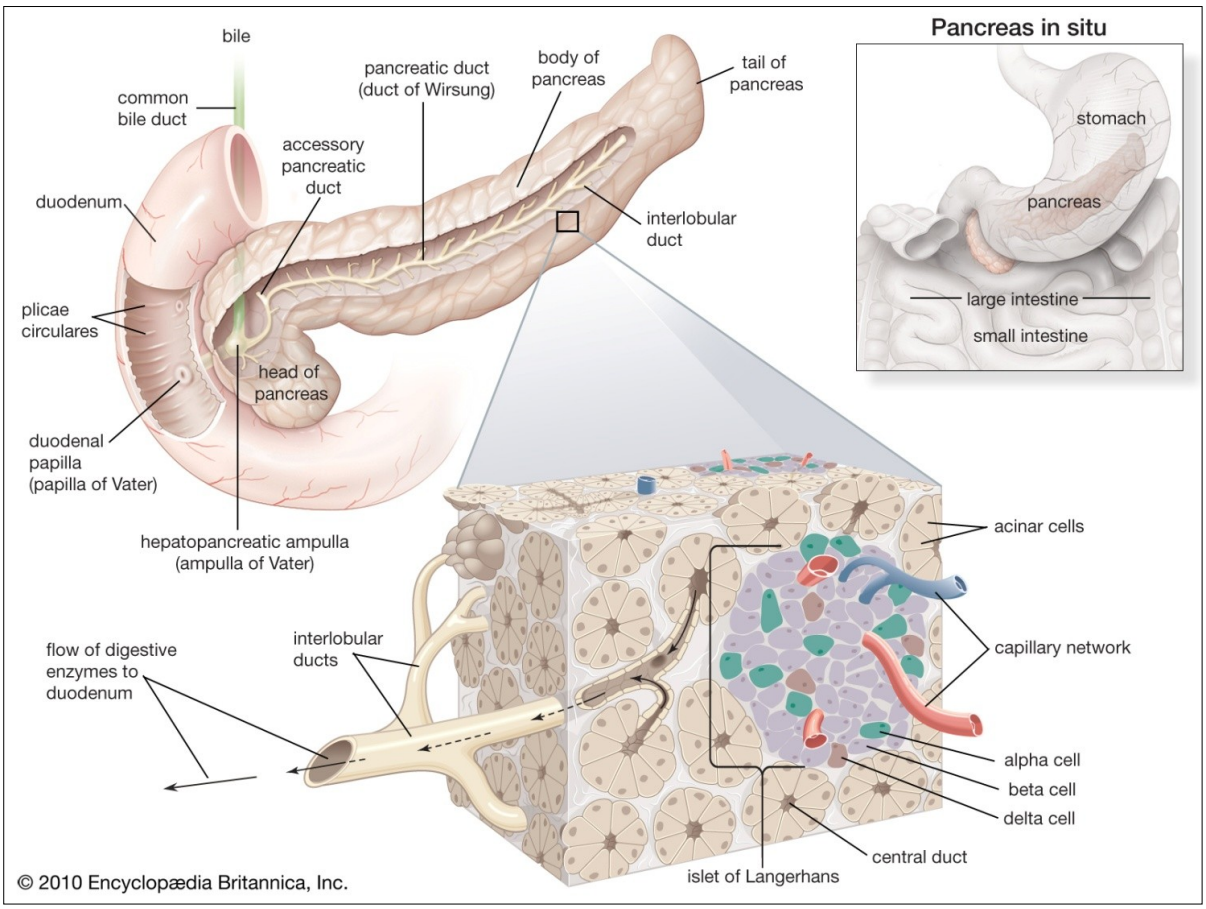

***Figure 1.1. Structure of the Pancreas (Encyclopaedia Britannica, 2014)**. Available from www.britannica.com/EBchecked/media/68636/Structures-of-the-pancreas-Acinar-cells-produce-digestive-enzymes-which.*

The mature pancreas (figure 1.1) has two distinct homeostatic functions. The larger exocrine part consists of acinar cells which produce and secrete digestive enzymes into the intestine via a highly branched ductal tree, while the endocrine part consists of the highly vascularised islets of Langerhans which contain the hormone-producing cells that control blood glucose levels. These are scattered throughout the pancreatic tissue and make up less than 2% of the pancreas's mass. The islets contain five cell types: insulin-secreting β-cells are the most abundant, making up an estimated 60-80% of the islet cells; the glucagon-secreting α-cells make up approximately 15-20% of the islet cells; followed by small numbers of somatostatin-secreting δ-cells, pancreatic polypeptide-secreting PP-cells and ghrelin-secreting ε-cells (Martini, 2006).

1.3 Developmental Biology of the Pancreas

Pancreatic development is a complicated process of morphological events accompanied by an elaborate pattern of cellular differentiation and lineage selection. These events are mediated by tissue interactions, signalling pathways and cascades of gene expression that determine cell fate. The expression of transcription factors throughout gastrulation and the formation of definitive endoderm (DE) are complex and overlapping due to the proximity of the three germ layers. Much of our knowledge of pancreatic development comes from studies in model organisms such as the mouse, chicken, zebrafish, and *Xenopus* and, despite differences in scale, pancreas development is highly conserved among these vertebrates.

1.3.1 Formation of Definitive Endoderm

Around 3-5 days post-fertilisation, the mouse embryo develops into a blastocyst which is comprised of an outer shell of trophoblast cells (the trophoectoderm) encasing the inner cell mass (ICM) and a hollow cavity known as the blastocoele. The ICM cells closer to the blastocoele are known as the visceral endoderm and contribute only to the formation of extra-embryonic tissue, while the cells closer to the trophoectoderm are known as the epiblast and give rise to the entire embryo (Rossant & Tam, 2004). The ICM is characterised by the expression of a number of transcription factors, such as OCT4, SOX2 and NANOG, each of which is required for its establishment and maintenance (Niwa *et al*., 2000; Avilion *et al*., 2003; Mitsui *et al*., 2003).

The process of gastrulation, which begins at approximately embryonic day (E) 6.5, subdivides the mouse embryo into the three germ layers: the ectoderm, which gives rise to the skin and central nervous system; the mesoderm, which gives rise to blood, bone, and muscle; and the endoderm, which gives rise to the respiratory and digestive tracts and their associated organs. The first sign of gastrulation is the formation of the primitive streak. The epiblast, at this stage, is a cup-shaped epithelial layer covered on its basal surface by the visceral endoderm. Cells that aggregate at the anterior primitive streak near the node have the potential to generate both endoderm and mesoderm, suggesting that they represent mesendoderm progenitors (Kimelman & Griffin, 2000; Rodaway & Patient, 2001). The transcription factor BRACHYURY is expressed in precursor cells throughout the primitive

streak, making this a marker of the putative mesendoderm population, but its expression is subsequently restricted to mesoderm and not endoderm cells (Kubo *et al*., 2004). These cells migrate through the primitive streak and then move laterally and anteriorly to develop into mesoderm and endoderm respectively. The endoderm cells emerge from the anterior segment of the primitive streak and are incorporated into the visceral endoderm, progressively displacing it to extra-embryonic sites (Lawson *et al*., 1991). Cells that migrate early give rise to anterior endoderm, and cells that migrate later give rise to posterior endoderm, suggesting that different regions of the primitive streak constitute different signalling environments that are responsible for the induction of specific lineages (Murry & Keller, 2008).

While the later stages of pancreas development have been extensively studied, relatively little is known about how endoderm is initially formed and specified in higher vertebrates: for example, which signals and genes direct the totipotent cells of the early mouse embryo to form a sheet of endoderm by the end of gastrulation, and how endoderm cells obtain positional identity and form a primitive gut tube. However, it is likely that other germ layers provide positional identity through soluble signalling factors. The primitive streak and node produce numerous growth factors, including FGF4, TGF-β family members and WNT3A, which have been implicated in patterning of the endoderm, although a direct role has not been demonstrated. FGF4 is expressed by the primitive streak (Wells & Melton, 2000), and the pancreas appears to arise from cells that receive intermediate levels of FGF4. Manipulations of FGF4 levels in the early embryo can expand or contract the pre-pancreatic domain (Dessimoz *et al*., 2006). Cells lacking FGFR1 fail to migrate through the PS and form endoderm/mesoderm. WNT3A is essential for primitive streak induction and germ layer formation. It inhibits the formation of ectoderm and induces posterior primitive streak. The TGF-β family of signalling molecules includes Nodal and BMP4. Nodal signalling is essential for primitive streak induction and germ layer formation (Murry & Keller, 2008), while BMP4 signalling is not essential but increases the efficiency of primitive streak formation through the induction of Nodal and WNT3A (Wells & Melton, 2000).

The two earliest markers of DE are SOX17 and CER1 (Lewis & Tam, 2006). CER1 expression in DE is transient and restricted to anterior DE (Belo *et al*., 1997; Biben *et al*., 1998). SOX17 is essential for appropriate formation of DE, as mice lacking SOX17 do not develop gut endoderm (Kanai-Azuma *et al*., 2002). FOXA2 also plays an important role in DE formation, as homozygous knockout in mice is lethal due to a lack of endoderm (Sasaki

& Hogan, 1993; Monaghan *et al*., 1993). The expression of CXCR4 is also observed on the cell surface of DE cells (McGrath *et al*., 1999; Nair & Schilling, 2008). Visceral endoderm also expresses BRACHYURY, FOXA2 and SOX17 (Wilkinson *et al*., 1990) but can be distinguished from DE by the expression of CXCR4 (Yasunaga *et al.*, 2005).

1.3.2 Pancreatic Specification

By the end of gastrulation, at approximately E7.5, the DE consists of a flat sheet of cells with anterior-posterior patterning which covers the mesoderm and ectoderm of the embryo. This flat sheet begins to roll up to form the primitive gut tube, while regionally distinct gene expression patterns emerge to specify the various endoderm organs. The DE is in contact with different structures along its anterior-posterior axis, including ectoderm, notochord plate, node, lateral mesoderm and primitive streak, which are likely to be a source of signals for patterning (Wells & Melton, 1999)

By E8.5, DE cells in the region of the foregut/midgut junction are already committed to a pancreatic fate, as demonstrated by their ability to differentiate into all cells of the pancreatic lineage (Wessels & Cohen, 1967). However, the first morphological evidence of pancreas development is not distinguishable until approximately E9.5, with the formation of two independent thickenings of the primitive gut tube, first dorsal, and then ventral (Pictet *et al*., 1972).

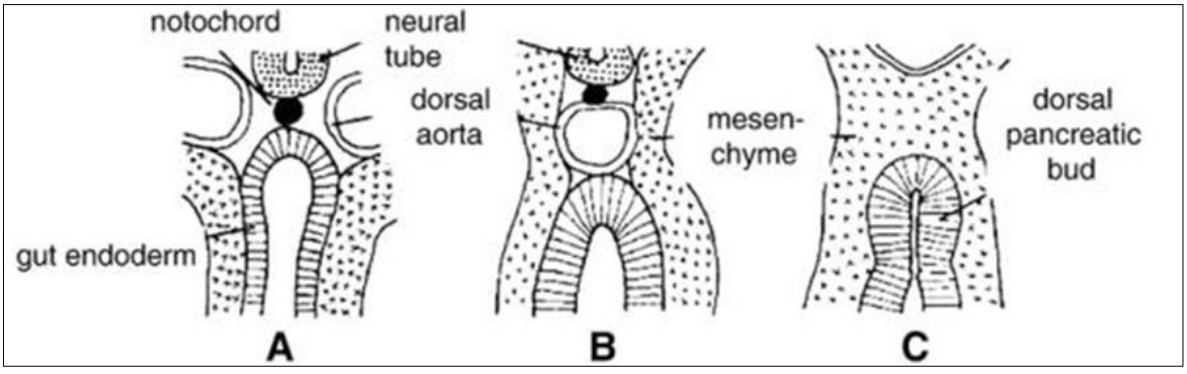

Figure 1.2. Early development of the dorsal pancreas. *(A) At E9, the gut endoderm is in contact with the notochord; (B) At E 9.5, the dorsal aortae fuse between the notochord and gut endoderm; (C) At E10, the mesenchyme surrounds the gut endoderm and the dorsal pancreatic bud has formed. (Slack, 1985).*

The dorsal pancreatic bud develops in close proximity to the notochord (figure 1.2) and this proximity is essential for dorsal bud development. Kim *et al*. (1997) showed that notochord removal from early chicken embryos *in vitro* prevented proper dorsal pancreas formation and inhibited the expression of pancreas-specific genes. Further studies showed that the morphogen sonic hedgehog (SHH) was absent from the pre-pancreatic endoderm, and grafting experiments with chicken endoderm showed that notochord proximity to the endoderm could suppress SHH expression (Hebrok *et al*., 1998). Analysis of the notochord-produced morphogens FGF2 and activin βB, a member of the TGFβ family, found that they were able to replace the notochord effect by inhibiting SHH expression (Hebrok *et al*., 1998; Kim & Melton, 1998). By E8.5-9.0, the notochord is no longer in contact with the dorsal pancreas due to the fusion of the two dorsal aortae between the notochord and gut endoderm (figure 1.2). Interaction between blood vessels and pancreatic endoderm has been shown to be critical for endocrine differentiation in the pancreas. *In vitro* studies of E8.5 dorsal endoderm revealed that it requires co-culture with adjacent dorsal aorta in order to differentiate into endocrine cells and express insulin (Lammert *et al*., 2001).

After this early dependence on the dorsal aorta, pancreatic epithelium becomes enveloped in pancreatic mesenchyme, separating it from the dorsal aorta. Condensation of the pancreatic mesenchyme and evagination of the underlying endoderm forms the pancreatic buds at around E9.5 and signals from the surrounding mesenchyme, including FGF10, TGF-β family members, and retinoic acid, are essential for the further proliferation and branching of the pancreatic buds. FGF10 enhances Notch signalling, which restricts expression of the transcription factor NGN3 to pancreatic progenitors (Apelqvist *et al*., 1999). Mice lacking FGF10 display impaired pancreatic epithelial growth, confirming the importance of signals from the surrounding mesenchyme (Bhushan *et al.*, 2001). TGF-β signals are also important for organising and regulating the growth and relative proportions of the endocrine and exocrine pancreas. Inactivation of the activin receptors has severe effects on endocrine development resulting in islet hypoplasia (Jorgensen *et al*., 2007). Retinoic acid (RA) signalling is also essential for pancreas development, as RA-deficient zebrafish fail to develop a pancreas or liver. In the mouse and *Xenopus*, RA signalling appears to be essential for dorsal pancreas formation (Stafford *et al*., 2004).

In contrast, the notochord plays no role in the development of the ventral pancreas, as they are not in contact with one another. Instead, ventral pancreatic development is closely linked

to liver development, with liver development being promoted by instructive signalling, such as FGF signalling, from the adjacent cardiac mesoderm. At the same time, in the absence of FGF signalling and cardiac mesoderm, the ventral pancreatic program is initiated. Thus, it appears that the default fate of the ventral endoderm is pancreatic and that signals from the cardiac mesoderm actively induce a hepatic state (Deutsch *et al*., 2001).

The gut endoderm widely expresses a number of transcription factors such as HNF1β (Ott *et al*., 1991), FOXA1 and FOXA2 (previously known as HNF3α and β; Monaghan *et al*., 1993) and HNF4α (Duncan *et al*., 1994; Taraviras *et al*. 1994), which delineate the region of the foregut endoderm which gives rise to the pancreas. The pancreatic domain is further refined by the expression of markers specifying a pancreatic fate, including HLXB9, PDX1 and ISL1.

One of the key components of pancreatic specification is the expression of PDX1, a transcription factor which is first expressed at E8.5 in the pre-pancreatic endoderm and throughout the developing ductal trees during pancreatic bud formation (Offield *et al*., 1996) until E13, when it becomes mainly restricted to the β-cells and some δ cells, although low expression of PDX1 is maintained in some duct and exocrine cells (Wilson *et al*., 2003). In the adult pancreas, PDX1 plays a critical role in the regulation of the transcription of genes associated with β-cell identity, such as insulin, GLUT2, islet amyloid polypeptide (IAPP) and glucokinase (Leonard *et al*., 1993; Ohlsson *et al*., 1993; Miller *et al*., 1994; Watada *et al*., 1996a; Watada *et al*., 1996b; MacFarlane *et al*., 2000; Chakrabarti *et al*., 2002). Lineage tracing experiments have demonstrated that mature pancreatic cells derive from PDX1-expressing progenitor cells (Gu *et al*., 2003). It has been suggested that members of the HNF1 and FOXA families of transcription factors may play a role in the control of PDX1 expression. Mice null for FOXA2 or HNF1β die before pancreas formation (Ang & Rossant, 1994; Weinstein *et al*., 1994; Coffinier *et al*., 1999), but embryoid bodies lacking FOXA2 fail to activate the *Pdx1* gene, which suggests that FOXA2 may be an upstream activator of *Pdx1* expression (Gerrish *et al*., 2000). PDX1 plays a critical role in early pancreas development: removal of PDX1 by gene targeting arrests pancreatic development after initial bud formation (Jonsson *et al*., 1994; Ahlgren *et al*., 1996; Offield *et al*., 1996), indicating that PDX1 is necessary for growth of the pancreatic buds but not for the initial induction of bud formation. Conversely, ectopic expression of PDX1 in non-pancreatic chick gut endoderm was also shown to induce the formation of pancreatic bud-like structures (Grapin-Botton *et al*., 2001).

The expression of HLXB9 is initiated around E8.0 in the gut endoderm and the notochord, and maintained in mature β-cells. Mice deficient for HLXB9 fail to express PDX1 in the dorsal pancreas, resulting in dorsal pancreatic agenesis (Harrison *et al*., 1999; Li *et al*., 1999). In the remaining ventral pancreas of HLXB9-deficient animals, islet malformations were observed, suggesting essential functions for HLXB9 in islet cell development. These studies demonstrate that HLXB9 is necessary for bud formation and its expression precedes that of PDX1 in the dorsal pre-pancreatic endoderm.

ISL1 is expressed in the developing pancreas and the central nervous system (Karlsson *et al*., 1990; Thor *et al*., 1991). Targeted disruption of the *Isl1* gene results in failure of development of the dorsal pancreatic mesenchyme and the complete absence of endocrine cells, while the ventral pancreatic epithelium and associated mesenchyme is unaffected. ISL1 has also been shown to play a role in the differentiation of islet cells and is expressed in mature β-cells (Ahlgren *et al*., 1997).

The specification of endocrine cells in the developing pancreatic endoderm is dependent on appropriate Notch signalling and expression of the pro-endocrine transcription factor NGN3 (figure 1.3). Although NGN3 is exclusively expressed in scattered cells in the pancreatic epithelium and not in differentiated endocrine cells (Apelqvist *et al*., 1999), lineage tracing has demonstrated that NGN3-positive cells function as endocrine precursors from which all four endocrine subtypes differentiate (Gu *et al*., 2002). NGN3 expression occurs in 2 distinct temporal waves, thereby generating "early" and "late" endocrine cells which have different developmental potential. Early NGN3-expressing cells exclusively give rise to *α*-cells, while NGN expression at later stages gives rise to β- and PP-cells after E11.5, and δ-cells after E14.5.

In the pancreas, the downstream target of Notch signalling is the *Hes1* gene, which in turn inhibits the expression of *Ngn3* by binding to several sites on the *Ngn3* promoter (Jensen *et al*., 2000; Lee *et al*., 2001). Animals lacking NGN3 failed to develop any endocrine cells while exocrine tissue and pancreatic ducts were nearly normal (Gradwohl *et al*., 2000). In addition, ectopic expression of NGN3 in the pancreatic epithelium caused premature differentiation of the pancreas into endocrine cells, resulting in nearly exclusive production of glucagon-expressing α-cells (Apelqvist *et al*., 1999; Schwitzgebel *et al*., 2000). These studies demonstrated that, in the absence of Notch signalling, NGN3 is essential for the initiation of endocrine differentiation in the pancreas. As Notch signalling restricts NGN3 expression to a

few scattered cells, other mechanisms must play a role in the upregulation of NGN3 expression. The NGN3 promoter has been shown to contain binding sites for a number of transcription factors, including HNF1, FOXA and HNF6 (Jacquemin *et al*., 2000; Lee *et al*., 2001). HNF6 is initially expressed throughout the pancreatic buds but is subsequently excluded from the mature pancreas (Rausa *et al*., 1997). Mice lacking HNF6 had severely reduced NGN3 expression in the developing pancreas leading to a severely reduced number of endocrine cells (Jacquemin *et al*., 2000), supporting a role for HNF6 as an upstream activator of NGN3 expression.

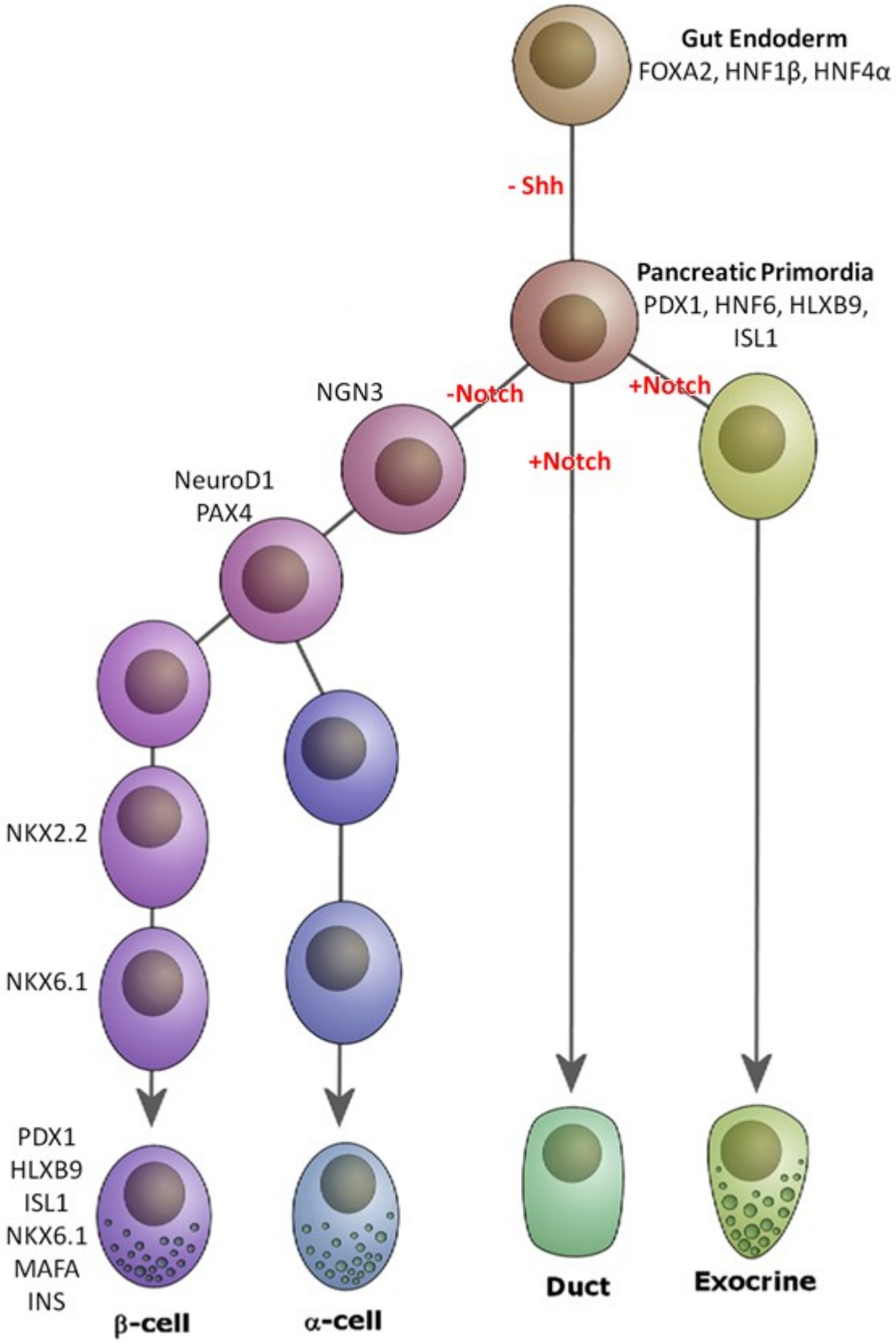

Figure 1.3.Endocrine specification of the developing pancreas. *The transcription factors expressed at each stage are labelled in black, and the instructive signals that specify certain cell types are shown in red. Adapted from Van Hoof et al,. 2009.*

NGN3 has also been shown to activate NEUROD1 (figure 1.3), which is expressed slightly later than NGN3 during pancreatic development. Unlike NGN3, its expression is maintained in mature islet cells and plays a role in the expression of a number of endocrine products including insulin (Naya *et al*., 1995; Qiu *et al*., 2002). The expression of NEUROD1 is dependent on NGN3 expression as NGN3-null mutant mice do not express NEUROD1 (Gradwohl *et al*., 2000), whilst NGN3 expression is unchanged in NEUROD1 null embryos. Similarly to NGN3, ectopic expression of NEUROD1 induces premature differentiation of endocrine cells in the pancreas, demonstrating that NEUROD1 is also a pro-endocrine gene (Schwitzgebel *et al*., 2000). However, unlike NGN3-null mutant mice, mice lacking NEUROD1 are still able to form all pancreatic endocrine cell types, but the number of cells is drastically reduced due to increased apoptosis (Huang *et al*., 2002).

Following initiation of the endocrine program, a set of transcription factors is necessary to convert NGN3-positive cells into the four mature endocrine cell types (figure 1.3). These factors can be divided into early factors such as PAX4, NKX2.2 and NKX6.1, which are co-expressed with NGN3 in endocrine precursors, and late factors such as PAX6, ISL1, MAFA and PDX1, which are found in more mature cells.

NKX2.2 expression is initiated at E9.5 throughout the dorsal pancreatic epithelium, and is expressed until E13 when it becomes localised to NGN3-positive cells. NKX2.2 expression persists in mature α-, β- and PP-cell types but not δ-cells. Mice lacking NKX2.2 display a complete absence of β-cells as well as a reduction in α- and PP-cells, whereas the number of δ-cells remained unaffected. Interestingly, NKX2.2 mutant islets contain a number of immature endocrine cells expressing IAPP and PDX1 but lacking other β-cell markers such as GLUT2 and glucokinase. Thus, it appears that in the absence of NKX2.2, β-cells are specified but are unable to differentiate to fully mature, insulin-producing cells. PAX4/NKX2.2 double deficient mice show a similar phenotype to NKX2.2 deficient mice, suggesting that PAX4 acts downstream of NKX2.2. This is further supported by the lack of alteration of NKX2.2 expression levels in PAX4-deficient mice (Sussel *et al*., 1998).

Expression of PAX4 is initiated at around E9.5 in the pancreatic buds and becomes progressively restricted to β-cells until approximately E15. However, little or no PAX4 can be detected in the mature pancreas (Smith *et al*., 1999). ARX and PAX4 play antagonistic roles in the specification of the endocrine cell subtypes, with ARX promoting specification of α- and PP-cells and PAX4 promoting specification of β- and δ-cells. Knockout of one of the

factors leads to a change in the proportion of cells of each lineage, although the total number of endocrine cells is unchanged. PAX4 inhibits the α/PP-cell fate through inhibition of ARX.

NKX6.1 is first detected at E9.5 in both pancreatic buds, where it persists until E13. After this, NKX6.1 expression becomes specifically restricted to β-cells where it is required for insulin secretion (Oster *et al*., 1998). A lack of NKX6.1 results in reduced β-cell numbers, while other endocrine cell types are unaffected. The islets of mice deficient in NKX2.2 expression also fail to express NKX6.1, suggesting that NKX2.2 is upstream from NKX6.1 in the β-cell differentiation pathway. In support of this, NKX2.2 expression is unaffected in NKX6.1-null mice and the pancreatic phenotype of mice lacking both NKX2.2 and NKX6.1 is identical to the phenotype of mice solely lacking NKX2.2, demonstrating that the lack of NKX6.1 expression has no effect on β-cell differentiation in the absence of NKX2.2 (Sander *et al*., 2000). NKX6.2, a paralogue of NKX6.1 shows a similar expression pattern but is not expressed in mature β-cells. NKX6.2-deficient mice do not display any obvious phenotype, while NKX6.1/NKX6.2 double mutants show a similar phenotype to the NKX6.1 knockout, apart from a reduction in glucagon-expressing cells, suggesting an additional role for NKX6.2 in α-cell formation. The misexpression of NKX6.1 in PDX1-producing cells in NKX6.2 deficient mice was found to rescue β-cell development, indicating that these factors have redundant roles in β-cell development (Binot *et al*., 2010).

The late factors are involved in the final steps of islet differentiation, after NGN3 expression and in conjunction with hormone expression. PAX6 is expressed in all hormone-producing cells throughout pancreatic morphogenesis, and plays an important role in islet cell differentiation. Conventional knockout experiments revealed its importance in the development of all islet cell types, as well as in the formation of well-structured islets of Langerhans (St-Onge *et al*., 1997). An additional role of this gene in maintaining postnatal β-cell phenotype was revealed as conditional ablation of PAX6 resulted in a diabetic condition with hypoinsulinaemia.

MAFA, MAFB and cMAF have all been shown to act in the terminal differentiation of α and β-cells. MAFA has been identified as specific to pancreatic β-cells, where it was found to directly interact with the promoter of the insulin gene and activate insulin expression (Matsuoka *et al*., 2003; Zhao *et al*., 2005). MAFA-deficient insulin-expressing cells are unable to release insulin in response to glucose. In contrast, ectopic expression of MAFA in embryonic chicken endoderm, as well as in non-pancreatic cell cultures, was found to be

sufficient to promote insulin secretion. Expression of MAFA is initiated at E13.5 and is restricted to insulin-positive cells throughout embryogenesis and adulthood (Matsuoka *et al*., 2004). Mice deficient for MAFA develop diabetes accompanied by a significant decrease in circulating insulin levels and abnormal islet architecture (Zhang *et al*., 2005).

As previously mentioned, PDX1 plays a role in the final differentiation of mature islets in addition to its early role in bud formation. After its early expression in the pre-pancreatic endoderm, PDX1 is downregulated as it cannot be detected in the NGN3 expressing endocrine progenitors (Schwitzgebel *et al*., 2000). Its expression is reactivated in most mature β-cells and in some δ-cells, where it plays a role in maintaining the differentiated phenotype and in the regulation of expression of a number of β-cell genes.

1.3.3 Human Pancreatic Development

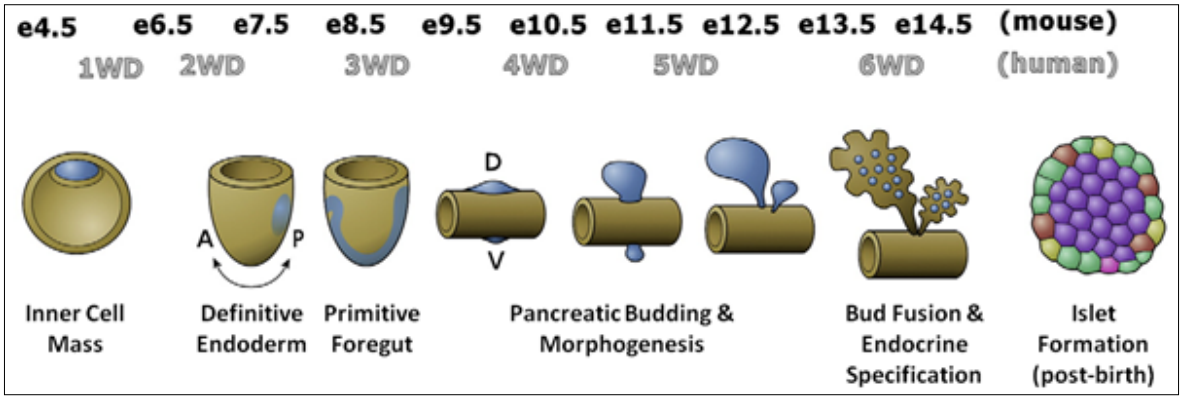

Figure 1.4. Approximate timescale of mouse and human pancreatic development. *Developmental stages are indicated as embryonic days of development (E) for mice and weeks of development (WD) for humans. Adapted from Van Hoof et al., 2009.*

Much of our current understanding of pancreatic development originates from mouse studies, as studies of human pancreatic development are limited by ethical constraints. However, development of the human pancreas may differ considerably from that of the murine model. In the mouse, the end of gastrulation and the formation of DE are completed by approximately E7.5 (figure 1.4). Pancreatic budding occurs by E9.5 and gut rotation brings the two buds in closer proximity around E12.5 (Van Hoof *et al*., 2009). In human

development, the first appearance of the pancreatic buds occurs just before 4 weeks of development and the two buds fuse at around week 6 of development. The first insulin-expressing cells appear very early in mouse development at approximately E10.5 and a second wave of β-cell differentiation occurs at birth which results in properly formed islets. In contrast, in human development, the first insulin-expressing cells do not appear until week 8 of development and the islets are formed by week 12-13 of development (Slack, 1995). The morphology of the pancreas also varies between mice and humans, as the murine pancreas consists of multiple lobes, whereas the human pancreas in a single solid organ. Arrangement of the endocrine islets also differs: in rodents, the β-cells tend to aggregate in the centre of the islets with the other cell types around the edge. However, in humans the β-cells are scattered throughout the islet (Brissova *et al*., 2005; Cabrera *et al*., 2006). Although the basic developmental program may be similar in humans and mice, the timing and/or activation of factors may differ which would be fundamental for the development *in vitro* differentiation protocols.

1.4 Insulin and Its Mechanisms of Action

1.4.1 Insulin Biosynthesis

Insulin is a 6kDa polypeptide hormone which is synthesised and released from the β-cells of the islets of Langerhans in the pancreas. The initial product of gene transcription is a pre-proinsulin molecule which is rapidly cleaved in the rough endoplasmic reticulum to proinsulin, a 9kDa precursor peptide consisting of 2 short peptide chains (A and B chains) and a connecting (C) peptide. The A and B chains are linked by two disulphide bonds between cysteine amino acid residues. The proinsulin molecules are then transported to the Golgi apparatus where they are packaged into secretory granules in which they are cleaved by converting proteases, removing the C-peptide and leaving the bound A and B chains to form the active insulin molecule. Insulin and C-peptide are stored in secretory vesicles and co-secreted in equal proportions by exocytosis in response to secretory stimuli (Steiner *et al*., 1985).

1.4.2 Insulin Secretion

The primary stimulus for insulin secretion is plasma glucose concentration. Glucose absorbed from the small intestine reaches the β-cells through the circulatory system and is transported into the cell by the GLUT2 transporter, where it is metabolised to glucose-6-phosphate by the enzyme glucokinase. This causes a rise in intracellular ATP concentration which in turn promotes the closure of the ATP-sensitive potassium (K_{ATP}) channels. Retention of potassium ions (K^+) within the cell depolarises the cell membrane causing the opening of voltage-sensitive calcium ion (Ca^{2+}) channels (VDCCs) leading to an influx of Ca^{2+}. The resulting rise in intracellular Ca^{2+} levels causes insulin secretion by exocytosis, where the insulin-containing secretory granules fuse with the plasma membrane and release insulin into the circulation (figure 1.5). High concentrations of extracellular Ca^{2+} have been shown to be a substitute for stimulatory levels of glucose demonstrating the critical role of Ca^{2+} ion influx in insulin release (Hedeskov, 1980).

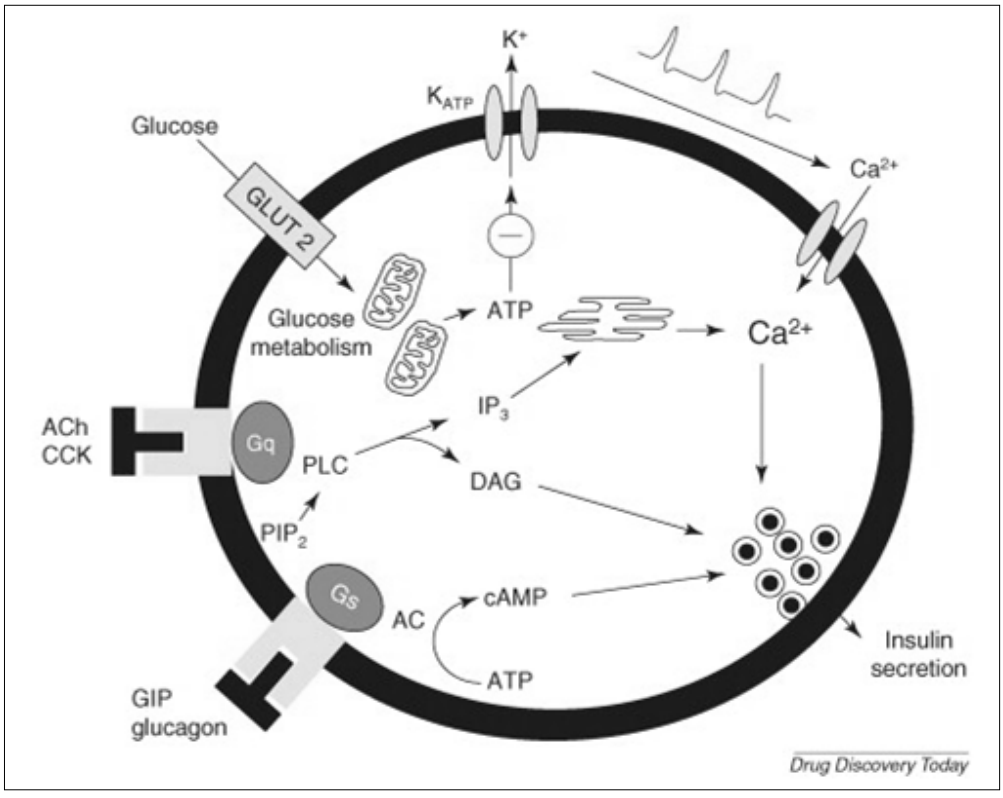

***Figure 1.5. Signal transduction in β-cells**. Insulin secretion is activated primarily by glucose but also by other stimuli, including amino acids, neurotransmitters and hormones. Jones et al., 2008.*

Insulin secretion is not solely dependent on a rise in plasma glucose levels; it can also be triggered by a number of other substances, such as amino acids, neurotransmitters and hormones, most of which require the presence of non-stimulatory blood glucose concentrations in order to be able to stimulate insulin secretion (figure 1.5). The autonomic neurons innervating the islets of Langerhans also influence insulin secretion: activation of the parasympathetic neurons, which occurs during a meal, stimulates insulin secretion; while activation of the sympathetic neurons, or an increase in the plasma concentration of adrenaline, inhibits insulin secretion. Neurotransmitters stimulate insulin secretion from the β-cell by activating the G protein Gq, which in turn activates the membrane-bound enzyme phospholipase C (PLC). PLC acts on phosphatidylinositol 4,5-biphosphate (PIP_2), which is present in small amounts in the inner half of the plasma membrane lipid bilayer. Activated PLC cleaves PIP_2 to generate inositol 1,4,5-triphosphate (IP_3) and diacylglycerol (DAG). IP_3 leaves the plasma membrane and diffuses through the cytosol to the endoplasmic reticulum (ER) where it binds to and opens IP_3-gated Ca^{2+}-release channels in the ER membrane. Ca^{2+} stored in the ER is released through the open channels, quickly raising intracellular Ca^{2+} which in turn triggers insulin release (Jones *et al.*, 2008). Insulin release is also partly under hormonal control. Glucose-independent insulinotropic peptide (GIP) and glucagon-like peptide (GLP) are secreted by the gastrointestinal tract in response to eating, and stimulate the secretion of insulin. This results in an earlier and greater release of insulin during a meal than would occur if plasma glucose concentration were the only stimulus (Vander *et al*., 1998, p596-600).

1.4.3 Insulin Action

Insulin induces its effects by binding to specific receptors in the plasma membrane of its target cells: muscle (both cardiac and skeletal), adipose tissue and liver. In these cells, insulin activates glycolysis, glycogenesis and lipogenesis, whilst simultaneously inhibiting glycogenolysis, gluconeogenesis, and lipolysis. Long-term regulation of insulin expression is mediated by glucose both through stimulation of insulin gene expression and inhibition of insulin mRNA degradation. Glucose also stimulates proinsulin biosynthesis preferentially over other islet proteins (Vander *et al*, 1998, p596-600).

Insulin binding to its tyrosine kinase receptor on the surface of its target cells triggers a number of signal transduction pathways (figure 1.6). Insulin binding first induces autophosphorylation at several tyrosine residues on the intracellular part of the receptor. This results in the recruitment of PI_3-kinase to the plasma membrane of the cell, bringing it within reach of its substrate, PtdIns(4.5)P_2, which is then phosphorylated. These events occur within the first minute of insulin binding to its receptor, and result in a significant increase in the concentration of PtdIns(3,4,5)P_3. Protein kinase B (PKB) binds to PtdIns(3,4,5)P_3, resulting in the recruitment of PKB from the cytosol to the plasma membrane, where it is phosphorylated by two other protein kinases, PDK1 and PDK2, resulting in its activation. Once activated, PKB dissociates from the plasma membrane and phosphorylates numerous substrates in both the cytoplasm and the nucleus which are important for regulating insulin-dependent processes.

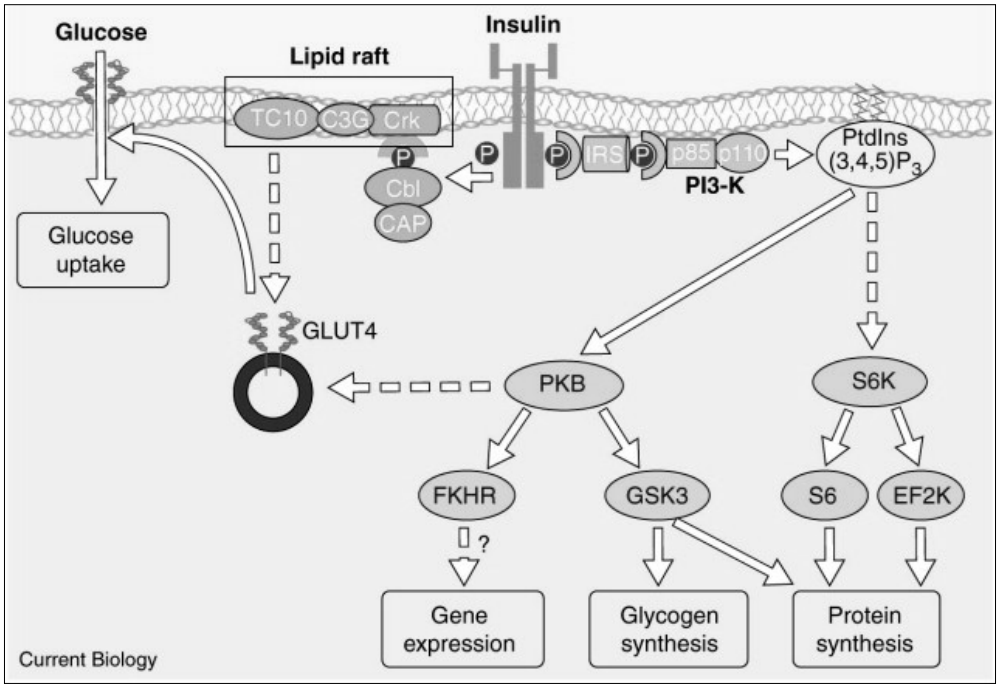

Figure 1.6. The Insulin Signalling Pathway. *Insulin binding to its receptor triggers signalling cascades that activates glycolysis, glycogenesis and lipogenesis, whilst simultaneously inhibiting glycogenolysis, gluconeogenesis, and lipolysis. Adapted from (Lizcano & Alessi, 2002).*

One of these substrates is glycogen synthase kinase 3 (GSK3), which is inactivated by PKB. This results in dephosphorylation of glycogen synthase, leading to the activation of glycogenesis. Inactivation of GSK3 also results in the dephosphorylation of eIF2B, which controls the initiation stage of protein synthesis, stimulating the synthesis of protein from

amino acids. Insulin also inhibits lipid metabolism in adipocytes by decreasing intracellular concentrations of cAMP, a process which appears to also be mediated by PKB through the phosphorylation and activation of PDE3B.

Another key action of insulin is to stimulate glucose uptake into its target cells by inducing the translocation of the glucose transporter GLUT4 to the plasma membrane, allowing uptake of glucose from the interstitial fluid by facilitated diffusion. Glucose transport into the liver is not directly insulin-dependent as hepatocytes have GLUT2 transporters in their cell membranes. Insulin activates hexokinase which phosphorylates glucose into glucose-6-phosphate. The lowered intracellular glucose concentration relative to the plasma allows the diffusion of glucose into hepatocytes via the GLUT2 transporters (Lizcano & Alessi, 2002).

In T1DM, the lack of insulin affects nutrient uptake and metabolism by the liver, skeletal muscle, and adipose tissue, resulting in hyperglycaemia. Without glucose to metabolise, these cells enter a fasting state. The liver initiates gluconeogenesis, resulting in a further increase in blood glucose; the adipose tissue breaks down its fat stores causing abnormal fat metabolism; the skeletal muscle breaks down its proteins to provide a substrate for ATP production. Abnormal fat metabolism results in ketone production from the liver which is poorly secreted by the kidneys resulting in diabetic ketoacidosis. High blood glucose concentrations can also saturate reabsorption from the kidney tubules resulting in a loss of glucose in the urine, causing less water to be reabsorbed and more to be excreted (known as osmotic diuresis) (Vander *et al*., 1998, p604-606).

1.5 Treatment of Type 1 Diabetes

1.5.1 Insulin Replacement

The discovery of insulin was the result of Nobel-prize winning experiments by Banting and Best in the 1920s. They found that removing the pancreas from a dog resulted in the development of diabetes, and that treatment with a pancreatic extract ameliorated the symptoms. Bertram Collip then joined the team and began the work of trying to purify the insulin from the pancreatic extracts. In 1922, the first patient, a fourteen-year-old boy, was

successfully treated with the purified insulin (Banting *et al*., 1922). More volunteers were recruited and as soon as 1923, insulin was being produced on a commercial scale, providing a life-saving treatment to many patients.

Since then, the most common treatment for T1DM has been injection of exogenous insulin, which, although lifesaving, is an invasive treatment regime that involves repeated testing of blood sugar levels and several injections daily. Up until the 1980s, animal insulins, obtained from pancreatic extracts of cows and pigs, were the only sources of insulin available. These were associated with complications such as variable pharmacokinetics, skin reactions at the sites of injection and immunogenicity. More recently, recombinant human insulin produced in genetically modified micro-organisms, has been more commonly used.

Different insulin regimens exist for the treatment of T1DM to permit individualisation of insulin therapy, and different biosynthetic insulin analogues are available which enable fine-tuning of the treatment: for example, "rapid-acting" insulin have an onset of 5-10 minutes, while "long-acting" insulin provides continuous activity for up to 36 hours (Crasto *et al*., 2009). The twice-daily regimen consists of twice-daily injections of short-acting and intermediate-acting insulin the doses of which can be adjusted according to blood glucose levels, meal times and exercise. However, this regimen requires a relatively rigid timing of meals and exercise. In an attempt to offer a more physiological replacement of insulin, the basal bolus regimen involves basal ("long-acting") insulin, used as background insulin to cover fasting, and short-acting insulin used at each meal. This therapy offers more flexibility for the patients but requires frequent blood glucose monitoring. Continuous subcutaneous insulin infusion (CSII) involves an insulin pump worn by the patient which delivers a continuous infusion of short-action insulin. The continuous infusion of insulin provides for background insulin requirement and pre-programmed infusion rates cover periods of activity, exercise or inactivity. Although reports indicate an improvement in quality of life with the use of CSII (Weinzimer *et al*., 2006), most randomised controlled studies did not demonstrate improvement in glycaemic control (DiMeglio *et al*., 2004; Fox *et al*., 2005).

When a major clinical trial was undertaken to compare intensive and conventional insulin therapies with regard to their effects on the development and progression of long-term complications associated with T1DM, it was found that keeping glucose levels as close to the normal range as possible slows the onset and progression of complications (Rohlfing *et al*., 2002). Although treatment with a mixture of long- and rapid-acting synthetic insulin

analogues, or with CSII, has improved control over blood glucose levels, the National Diabetes Audit for 2008-2009 found that only 28.6% of patients were achieving the NICE-recommended target level of glycated haemoglobin of 7.5% or less. In addition, these patients are more likely to suffer from episodes of hypoglycaemia. For these reasons, other therapies are needed.

1.5.2 Pancreas Transplantation

Even before the discovery of insulin, physicians were exploring the possibility of treating T1DM via pancreatic replacement strategies. The first successful whole pancreas transplant was carried out in 1966 (Kelly *et al*., 1967), and resulted in insulin independence. Although initial success rates were low, with methodological improvements and the use of immunosuppression the procedure started to become more widespread in the mid-1980s (Gaglia *et al*., 2005). However, there are problems associated with it: the shortage of suitable donors, the highly invasive nature of the procedure, and the requirement for lifelong immunosuppression afterwards. For these reasons, it is currently only carried out in diabetic patients with end-stage renal disease as a simultaneous pancreas and kidney transplant.

1.5.3 Islet Transplantation

Islet transplantation is a less invasive alternative to whole pancreas transplantation. It was first shown to be a viable approach for treating T1DM by reversing chemically induced diabetes in a rodent model (Ballinger & Lacy, 1972). However, technical limitations in islet isolation techniques, immunosuppression regimens, and low success rates, precluded its widespread use until protocols were optimised in a landmark study, where seven patients were treated with islet transplants, and all maintained insulin independence for up to one year (Shapiro *et al*., 2000). This approach was dubbed the 'Edmonton protocol' and included the selection of patients that suffered from severe hypoglycaemia, a steroid-free immunosuppressive protocol, islet preparation in the absence of xenogenic proteins, and an increase in the transplanted islet mass with islets from up to four donors used per recipient. More than 100 islet transplants are now carried out every year, with 44% of patients achieving insulin independence (Barton *et al*., 2012). However, although a clinical review in

the British Medical Journal in 2001 (Serup *et al*., 2001) anticipated that by 2010, islet transplantation would be the treatment of choice for most patients with type 1 diabetes, islet transplantation is still regarded as an experimental procedure, and is therefore currently only an option for type 1 diabetics with severe glycaemic lability, recurrent hypoglycaemia, or hypoglycaemia unawareness (de Kort *et al*., 2011), or those undergoing a simultaneous transplant, as these are the only patients for whom the significant side effects resulting from immunosuppression and islet perfusion are considered to outweigh the poor quality of life caused by diabetes.

One of the reasons why islet transplantation is still not widely used is because of the shortage of suitable donors. Currently, the only source of islets is heart-beating brain-dead donors. There have been suggestions that the use of non-heart-beating donors could widen the donor pool, but there would still be a shortfall. Pancreata that do become available are first offered for whole organ transplant, as this is an established procedure with recognised long-term benefit. In addition, islets from up to four donors are needed for each islet transplantation (Langer, 2010) due to the fact that as many as 50-60% of the transplanted islets may be lost immediately after transplantation (Harlan *et al*., 2009).

Another problem facing islet transplantation is graft deterioration, with only 10% of patients achieving insulin independence after 5 years due to loss of function of the transplanted islets. This failure has been attributed to both anti-donor antibodies and increased alloantibody titres, suggesting that some form of immunomodulatory therapy to control the auto-immune aspect of the disease will be needed. There is a need for better assays for monitoring these autoimmune and alloimmune responses (Harlan *et al*., 2009). It has been suggested that selecting other sites for transplantation, such as under the gastric or intestinal mucosa, or immunologically privileged sites such as the brain, thymus or testis, might better preserve the function of the transplanted islets. Currently, islets are infused into the liver via the portal vein, as this is the only site that has routinely shown success in large animals, and the pancreas would normally secrete insulin into the portal vein (Langer, 2010). Complications resulting from islet transplantation have been more common than originally thought, largely due to side effects of immunosuppression regimes, as well as peritoneal haemorrhage and partial portal vein thrombosis resulting from islet infusion into the portal vein (Gaglia *et al*., 2005).

There have been improvements in islet transplantation techniques since the Edmonton protocol, but the necessity of life-long immunosuppression, the scarcity of cadaveric donors and technical inefficiencies in islet isolation preclude widespread use of this therapy today as a replacement for insulin. Therefore there is a need for research into alternative sources of functionally competent, insulin-secreting β-cells as substitutes for donor islets to meet the clinical need for transplantation therapy.

1.5.4 Xenotransplantation

One solution to the problem of donor shortages is to use islets from other species. Much of the focus has been on pig islets, as they have several advantages. Firstly, high yields can be obtained using similar techniques to human islet isolation. Secondly, porcine insulin is very similar to human insulin, with only one amino acid difference (Trucco *et al*., 2007; Jones *et al*., 2008). Finally, pigs are amenable to genetic modification in order to promote cryoprotective or immunomodulatory properties, or to produce human insulin (Hering & Walawalkar, 2009).

However, two main obstacles prevent the widespread use of porcine islets in humans. Firstly, the hyper-immune response to xenografts due to the presence of the α1,3-galactosyl antigen on porcine cells has proved difficult to avoid, although in contrast to whole organ xenotransplants, this may not play a major role in the rejection of pig islet transplants in non-human primates (Hering & Walawalkar, 2009). The currently used intensive immunosuppressive regimen in pig islet transplantation may have severe side effects in humans, and is unlikely to be considered for routine clinical practice. However, advances have now been made in the breeding of genetically modified pigs which are free of this antigen: islets isolated from these pigs show decreased antigenicity and increased survival in non-human primate models of islet transplantation (Trucco *et al*., 2007). A study by Klymiuk *et al*. (2012) reported the transplantation of islets from transgenic pigs which had been engineered to express LEA29Y, a high-affinity variant of the T-cell co-stimulation inhibitor CTL4-4Ig, under the control of the insulin gene promoter. When these islets were transplanted into humanised mice, they were protected from immune rejection and were able to normalise hyperglycaemia, suggesting that this may represent a new approach to controlling the immune response to xenografts.

Secondly, the demonstration that porcine endogenous retroviral (PERV) sequences in the porcine DNA may become activated on xenotransplantation (van der Laan *et al*., 2000) has raised concerns over viral infections in humans receiving porcine islet transplants. However, another study has suggested that the potential threat of PERVs to xenotransplant recipients may have been over-estimated, because prolonged immunosuppression and exposure to porcine islet xenografts caused no detectable transmission of PERVs in *in vitro* or *in vivo* studies (Denner *et al*., 2008).

The interest in xenotransplantation has led to some controversial clinical trials. An Australian company, Living Cell Technologies Ltd., reported a clinical trial in Moscow in 1996, in which encapsulated porcine islet cells (DIABECELL®) were transplanted into two patients with diabetes, leading to reduced insulin requirements. Follow up work on this trial revealed evidence of residual, viable, encapsulated pig β-cells being retrieved from a transplant recipient almost a decade after xenotransplantation (Elliott *et al*., 2007). The approach used in this trial was subsequently strongly criticised by the International Xenotransplantation Association for being premature and potentially risky (Grose, 2007). Similarly, earlier clinical trials of porcine islet transplantation in Mexico (Valdes-Gonzalez *et al*., 2007) have been the subject of extensive criticism about the preclinical data used in justification of the trials, and about the reported efficacy of the trials (Sykes & Cozzi, 2006).

However, Living Cell Technologies Ltd. reported in a press release dated 5 May 2009 (www.lctglobal.com/latest-news.php), that in a clinical trial in Russia, insulin independence was observed for 2 of the 7 patients who received DIABECELL® and no adverse effects were reported. Patients had a reduction in insulin requirements of 6-25% and a reduction in both the severity and number of hypoglycaemic events. A subsequent phase II trial carried out in Argentina showed that patients treated with DIABECELL® also had reduced insulin requirements, and a reduced risk of hypoglycaemic incidents (www.lctglobal.com/latest-news.php, press release dated 22 November 2012). Another trial is due to start in New Zealand within the next few years. The commercial and clinical potential of pig islets as a source of transplant material for the treatment of T1DM is driving larger and more controlled clinical trials and the outcome of these will determine whether a future exists for the use of porcine islets as a substitute for human islets.

1.6 Cell Replacement Therapy

Islet transplantation has shown proof of principle for cell replacement therapy to treat type 1 diabetes. There are numerous different sources of tissue that could potentially be used, but there are certain requirements in order for them to be suitable for transplantation. First, they must be able to synthesise and store insulin, and release sufficient amounts in response to changes in circulating nutrients and a variety of hormones and neurotransmitters. Second, any source of replacement β-cells will also have to be able to generate large numbers of cells *in vitro* to meet the requirement of transplantation therapy, but this proliferative capacity must be tightly regulated to avoid post-transplantation expansion. Thirdly, they must show persistent engraftment and survival.

1.6.1 Human β-cells

An obvious starting point in the generation of replacement tissue is the expansion of mature islet β-cells, as studies have shown that, even many years after diagnosis, patients with T1DM still retain some β-cell mass and function (Wang *et al*., 2012). However, expansion of these cells *in vivo* or *in vitro* has proved to be a difficult task as, although β-cells are the predominant cell within pancreatic islets, they still only comprise around 2% of the total pancreatic cells, so that isolating a homogeneous population of β-cells has proven difficult. In addition, β-cells have a very limited proliferative capacity, and human β-cells are even less proliferative than rodent β-cells (Bonner-Weir et al, 2000b).

Although β-cells can be stimulated to proliferate *in vitro* by certain growth factors (Hayek *et al*., 1995; Beattie *et al*., 2002), their proliferation was accompanied by dedifferentiation and loss of function. A number of attempts have been made to restore insulin production in these expanded cells, but insulin expression was limited or inconsistent (Gershengorn *et al*., 2004; Lechner *et al*., 2005; Ouziel-Yahalom *et al*., 2006). A study by Bar *et al*. (2012) showed that these dedifferentiated cells can be redifferentiated to insulin-producing cells using HES1 shRNA to inhibit the Notch pathway, raising the possibility of producing functional β-cells *in vitro*. Similarly, Aly *et al*. (2013) showed increased proliferation of adult β-cells *in vitro* through manipulation of the Wnt signalling pathway. Another recent study identified a small molecule capable of inducing β-cell proliferation both in a rodent cell line, as well as in a

mouse model of diabetes (Shen *et al*., 2013). Despite this, it seems unlikely that expansion of human adult β-cells by *in vitro* manipulation will be able to generate the quantity of replacement β-cells required for cell replacement therapies.

During the past 30 years, a number of rodent β-cell lines have been established from x-ray induced insulinomas in adult rats (Gazdar *et al*., 1980; Asfari *et al*., 1992), or derived by transformation of adult hamster islet cells (Santerre *et al*., 1981). Other rodent β-cell lines have been generated by targeted oncogenesis in transgenic mice that have expressed the SV40LT under the control of the insulin promoter (Efrat *et al*., 1988; Miyazaki *et al*., 1990). These cell lines have been extremely useful for the study of rodent β-cells, but since there are many differences between rodent and human β-cells, attempts have been made to generate human β-cell lines.

Human β-cell lines have been generated from many different sources, including adult islets, foetal pancreases, or insulinomas (Levine *et al*., 1995; de la Tour *et al*., 2001; Demeterco *et al*., 2002) but insulin production by these cells was either extremely low or could only be maintained over a few passages. In 2005, a reversibly immortalised human β-cell line was established through the introduction and subsequent removal by Cre-recombinase-mediated excision of immortalising genes (Narushima *et al*., 2005). These cells expressed a mature β-cell phenotype, secreted insulin in response to glucose stimulation and rectified streptozotocin-induced hyperglycaemia in mice. Disappointingly, no new reports on the use of this line have since been published. In 2011, (Ravassard *et al*., 2011) reported establishment of a stable, functional β-cell line with glucose-inducible insulin secretion using targeted oncogenesis in human foetal pancreatic tissue. These cells were capable of proliferation and were stable over at least 80 passages. They expressed many specific β-cell markers without expression of markers of other pancreatic cell types, and were also capable of expressing insulin in response to glucose and other insulin secretagogues. In addition, transplantation of these cells in mice reversed chemically induced diabetes. However, after transplantation of these cells, the mice eventually became hypoglycaemic due to continued proliferation of these cells *in vivo*. Transformed β-cell lines derived from human β-cells offer the potential of generating *in vitro* the huge numbers of cells that would be required for transplantation therapy. However, while the establishment of a human β-cell line will be extremely useful for studying human β-cells, the use of genetically manipulated cells is unlikely to be acceptable for cell replacement therapy.

1.6.2 Pancreatic Stem Cells

The adult β-cell mass is not static but fluctuates in response to changing physiological conditions such as pregnancy (Sorenson & Brelje, 1997) and obesity (Butler *et al*., 2003). In addition, several approaches, including surgical ablation of pancreatic tissue and pancreatic duct ligation (PDL), were found to initiate endogenous regenerative programmes acting to replace the missing cells. β-cell regeneration could potentially occur through one of several mechanisms: replication of existing β-cells, neogenesis of β-cells from multipotent duct progenitor cells, or conversion of α-cells or non-endocrine cells into β-cells. However, the existence of stem cells in the adult pancreas remains controversial. Although stem cells have been proposed to exist in the adult exocrine pancreas, pancreatic ducts and in the islets themselves, they remain poorly characterised and their proliferative capacity *in vitro* is quite limited. In the case of exocrine and duct cells, it is not clear whether the mature cells can transdifferentiate directly into insulin-expressing cells or whether a step of dedifferentiation is required first.

Lineage tracing studies in mice suggest that, under most circumstances including acute pancreatic regeneration, new β-cells arise by a slow regeneration of existing β-cells. Dor *et al*. (2004) used pulse-chase labelling of β-cells to show that throughout adult life, or following partial pancreatectomy, new β-cells arise from pre-existing β-cells. Another study showed that, several weeks after toxin application, treated mice recovered from hyperglycaemia through the genesis of new β-cells that mainly originated from surviving β-cells (Nir *et al*., 2007). Teta *et al*. (2007) provided more insight into the mitotic behaviour of β-cells under normal physiological conditions, after partial pancreatectomy, and during pregnancy, and showed that new β-cells arose from a pre-existing population of β-cells, and not from repeatedly dividing pancreatic progenitors. This was also seen by another study, which demonstrated β-cell replication following PDL and its subsequent reversal, suggesting that once the disease state is removed, the β-cells can replicate and restore endocrine function (Hao *et al*., 2013).

However, several other studies have suggested alternative origins of β-cells during pancreas regeneration, most notably neogenesis of cells from multipotent duct progenitor cells. However, the evidence for this is conflicting, with several studies using lineage tracing techniques to support this hypothesis, and other studies using similar techniques to demonstrate that this is not the case. An early study by Wang *et al*. (1995) noted significant

islet hyperplasia following PDL. Pulse chase labelling indicated that proliferation of existing β-cells could not fully account for the hyperplasia, and it was therefore concluded that multipotent precursors and/or stem cells may be involved in these regenerative events. The concomitant increase in the numbers of β-cells and pancreatic duct cells suggested that newly formed β-cells may have a ductal origin. Since then, several studies have shown that pancreatic duct cells are able to give rise to insulin-expressing cells *in vitro* (Bonner-Weir *et al*., 2000a; Yatoh *et al*., 2007). Lineage tracing experiments on newly formed β-cells following PDL determined that duct cells could indeed give rise to islet and acinar cells, and that this conversion was accompanied by several rounds of replication followed by dedifferentiation and redifferentiation. In non-injured neonates, the naturally occurring postnatal increase in islet cell mass was also partially attributed to duct cell conversion (Inada *et al*., 2008). A study by Xu *et al*. (2008) provided evidence of β-cell renewal from existing β-cells and cells in the ductal lining of adult pancreas, which re-expressed the pro-endocrine gene NGN3 following PDL, as when these cells were isolated, they were found to be capable of giving rise to all 4 endocrine subtypes. Following partial pancreatectomy or PDL, PDX1 was found to be ectopically expressed in mature ductal structures, suggesting that it may play a crucial role in regenerative processes and direct the conversion of different pancreatic cell types (Li *et al*., 2010). More recently, a mechanism for this neogenesis has been proposed, through the activation of the Wnt target gene *Lgr5* following PDL, which marks progenitor cells in self-renewing tissues (Huch *et al*., 2013).

Several studies have demonstrated using lineage-tracing techniques that exocrine duct and acinar cells do give rise to endocrine cells, including β-cells, during embryonic development, but that after birth there is no detectable contribution of exocrine cells to β-cell growth in either the uninjured mouse (Kopinke & Murtaugh, 2010) or following PDL (Solar *et al*., 2009; Rankin *et al*., 2013; Xiao *et al*., 2013).Similarly, other studies showed that both endocrine and exocrine cells arise from SOX9-expressing cells during embryogenesis, but that this does not occur during adulthood, or following injury (Furuyama *et al*., 2011; Kopp *et al*., 2011). Studies also showed that although PDL did result in ectopic expression of NGN3 in ductal cells, this did not result in subsequent β-cell neogenesis (Kopp *et al*., 2011; Xiao *et al*., 2013). A more recent study which used PDL in combination with STZ treatment to remove all pre-existing β-cells found no evidence for β-cell regeneration, even after 10 months, but that the exocrine compartment of the pancreas did undergo regeneration (Cavelti-Weder *et al*., 2013). Some of these discrepancies in results may be due to the methods used to

measure β-cell mass (Chintinne *et al*., 2012; Rankin *et al*., 2013), misinterpretation of existing β-cells as new β-cells, or species-specific differences between mouse and rat.

Other studies have demonstrated the plasticity of pancreatic cells to transdifferentiate from one cell type to another. Thorel *et al.* (2010) generated a transgenic mouse model in which more than 99% of endogenous β-cells were chemically destroyed. Then, using a β-cell tracking system, the authors observed increased β-cell mass over time, and demonstrated that these new β-cells arose from healthy α-cells through transdifferentiation, showing a previously unappreciated flexibility of cell fate conversion of endocrine cells. Although the cells generated frequently secreted both insulin and glucagon, indicating an incomplete conversion, this study significantly extended knowledge of pancreatic cell plasticity and regeneration.

PDX1 is thought to play a crucial role in regenerative processes in the pancreas, and its ectopic expression has been shown to result in transdifferentiation of pancreatic cells (Taniguchi *et al.*, 2003; Miyatsuka *et al.*, 2006). Other transcription factors have also been shown to be important. Ectopic expression of ARX in adult β-cells forces them to adopt α- or PP-cell phenotypes, suggesting that the opposite conversion might be achieved in order to generate β-cells from other pancreatic cell types by the misexpression of PAX4 (Collombat *et al.*, 2003). Indeed, the same group later showed that inactivation of ARX, which has an opposing role to PAX4, in α-cells leads to their conversion into functional β-cells which were capable of normalising STZ-induced hyperglycaemia. In addition, this conversion also triggered islet cell neogenesis from pancreatic duct cells, with a reactivation of *Ngn3* expression (Courtney *et al*, 2013). Zhou *et al.* (2008) demonstrated 9 transcription factors that, when misexpressed, caused alterations in β-cell number or function. Overexpression of combinations of some of these factors in the exocrine pancreas of mice resulted in the conversion of acinar cells to insulin-secreting cells. The combination with the highest efficiency was PDX1, NGN3 and MAFA. This reprogramming did not involve any dedifferentiation or proliferating steps, suggesting direct transdifferentiation. The insulin-secreting cells produced were morphologically indistinguishable from endogenous β-cells, expressed all known β-cell markers, and were able to cure diabetes in streptozotocin-treated mice. PDX1 and MAFA were needed only transiently as they activated their endogenous homologues, while NGN3 expression was down-regulated one month post-infection, mimicking its transient expression pattern as normally seen during embryonic development.

It was not possible to use the same factors to force other cell types to acquire β-cell identity, suggesting that the close relation and epigenetic similarities of acinar and β-cells might be a prerequisite for a successful conversion. This combination of transcription factors was also used to convert pancreatic exocrine cells to β-cells *in vitro* (Akinci *et al.*, 2012). Another study showed that transient treatment of acinar cells with EGF and CNTF resulted in their *in vivo* conversion to functional β-cells, providing further evidence for the plasticity of these cells (Baeyens *et al*, 2014).

1.6.3 Other Adult Stem Cells

Many adult tissues have been shown to contain progenitor/stem cells, which are responsible for the normal maintenance and renewal of tissue. The proliferation capacity of these cells is limited and although this could restrict their utility as a source of replacement β-cells, it could also represent a safety advantage. In addition, these cells provide a source of autologous cells which may reduce the risk of graft rejection compared with an allograft. Tissue stem cells have been found in tissues characterised by high turnover such as bone marrow (Jiang *et al.*, 2002; Grove *et al.*, 2004), intestinal epithelium (Potten *et al.*, 1998) and liver (Theise *et al.*, 1999), as well as in tissues with low regenerative potential, such as brain (Galli *et al.*, 2003), kidney (Oliver *et al.*, 2004) and the retina (Tropepe *et al.*, 2000).

Liver Progenitor Cells: As the pancreas is derived from endoderm tissue, a logical place to start looking for stem cells able to give rise to insulin-expressing cells is in other endoderm-derived organs. The liver and the pancreas not only have a common developmental origin in the foregut (Pictet, 1972) but also share the expression of a number of transcription factors, such as the hepatic nuclear factors (Odom *et al.*, 2004). Furthermore, both liver and β-cells express the GLUT2 high-capacity glucose transporter and high specificity/low affinity glucokinase. In addition, the liver is one of the primary target sites for insulin, which is delivered from the pancreas through the portal vein. Finally, the liver is relatively accessible for biopsy and can regenerate after partial hepatectomy. Taken together, these properties make the liver a clinically attractive option for autologous cell grafting of liver-derived β-cells in patients with T1DM, thus avoiding the requirement for immunosuppression. Early

studies of the liver uncovered multipotent cells, termed hepatic oval cells, which have been shown to differentiate *in vivo* into a number of hepatic, pancreatic and intestinal cell types (Suzuki *et al.*, 2002) and *in vitro* into cells expressing pancreatic β-cell markers (Yang *et al.*, 2002; Nakajima-Nagata *et al.*, 2004; Yechoor *et al.*, 2009). For many of the same reasons, biliary tree stem cells may prove a promising source of adult stem cells for differentiation into endodermal lineages. The promise of these cells has been demonstrated for improving liver function (reviewed in Lanzoni *et al*, 2013).

Rodent and human hepatocytes can be induced to express pancreatic markers and adopt some functional aspects of β-cells (Yang *et al.*, 2002; Zalzman *et al.*, 2003; Cao *et al.*, 2004; Kaneto *et al.*, 2005; Kim *et al.*, 2007) via multiple approaches, including adenovirus-mediated gene transfer (Ferber *et al.*, 2000; Ber *et al.*, 2003; Kojima *et al.*, 2003; Sapir *et al.*, 2005). Some of these genetically modified cells had the capacity to improve hyperglycaemia in diabetic mice and expressed significant amounts of insulin. In many cases, the pancreatic transcription factor PDX1 was used to induce a pancreatic phenotype (Ferber *et al.*, 2000; Horb *et al.*, 2003; Ber *et al.*, 2003; Kaneto *et al.*, 2005; Yang *et al.*, 2006; Fodor *et al.*, 2007; Shternhall-Ron *et al.*, 2007). Similarly, the overexpression of another β-cell transcription factor, NEUROD-betacellulin, in mouse liver cells *in vivo* was reported to cause reversal of hyperglycaemia (Kojima *et al.*, 2003; Yatoh *et al.*, 2007). Another study showed that the forced expression of NGN3 alone was insufficient to transdifferentiate mature hepatocytes but was competent to convert hepatic progenitor cells into a functional β-cell lineage (Yechoor *et al.*, 2009). More recently, improved glucose control was shown, without transdifferentiation, by lentiviral delivery of an insulin gene construct to rat liver *in vivo* (Elsner *et al.*, 2008). Taken together, these studies demonstrate that liver and pancreatic cells may possess the plasticity of converting into each other under certain circumstances. However, there is so far not enough evidence that modified liver cells are able to undergo the *in vitro* expansion required to create the amounts of cells needed for cell replacement therapy.

Bone Marrow-Derived Stem Cells: Bone marrow contains two different types of stem cells: haematopoietic stem cells (HSCs), which can be isolated from bone marrow, peripheral blood or umbilical cord blood; and mesenchymal stem cells (MSCs), which can be found in bone marrow and umbilical cord blood. HSCs are responsible for the production of all blood cell types, while MSCs are involved in the formation of mesenchymal tissues e.g. bone, muscle,

and fat. It has been shown that bone marrow-derived stem cells possess a high plasticity and are able to transdifferentiate into a number of different cell lineages (Jiang *et al.*, 2002; Grove *et al.*, 2004). Bone marrow cells are already routinely collected for current medical procedures, so clinical harvesting procedures are already in place. Bone marrow could also potentially serve as an autologous source of stem cells, minimising the problem of immune rejection. All of these factors make bone marrow stem cells an attractive source of stem cells for cell replacement therapy.

An initial study reported the ability of transplanted bone marrow stem cells to migrate to the pancreas and to differentiate *in vivo* into endocrine pancreatic cells (Ianus *et al.*, 2003) but this was later contested by several studies (Choi *et al.*, 2003; Hess *et al.*, 2003; Lechner *et al.*, 2004; Taneera *et al.*, 2006) that failed to show evidence of transdifferentiation but suggested that bone marrow stem cells instead contributed to pancreatic regeneration by enhancing existing β-cell replication. Similarly, studies in the non-obese diabetic (NOD) mouse, a model of auto-immune diabetes, showed that bone marrow stem cell transplantation resulted in a proportion of mice regaining endogenous β-cell function but could not find any evidence that the bone marrow stem cells themselves replenished the β-cell mass (Nishio *et al.*, 2006; Suri *et al.*, 2006). Hasegawa *et al.* (2006) examined the role of bone marrow transplantation in the stimulation of insulin production in more depth, showing that, consistent with previous results, acute pancreatic injury was required to improve hyperglycaemia and restore islet cell numbers. In addition, although no bone marrow-derived insulin-producing cells were seen, regenerated islets were surrounded by bone marrow stem cell-derived haematopoietic cells, suggesting that they may have a supportive role in β-cell regeneration after injury.

Some advances have also been made in the differentiation of bone marrow stem cells into insulin-producing cells *in vitro*, either by manipulating the culture environment (Tang *et al.*, 2004; Hisanaga *et al.*, 2008) or by forced reprogramming of bone marrow stem cells using pancreatic transcription factors such as PDX1 (Li *et al.*, 2007). There have also been reports of MSCs in human umbilical cord blood which can differentiate *in vitro* into cells expressing markers of the endocrine pancreas (Pessina *et al.*, 2004) and into insulin-expressing cells (Sun *et al.*, 2007). *In vivo* studies have shown that transplantation of MSCs improved insulitis in a type 1 diabetic mouse model (Ende *et al.*, 2004). Although there appears to be promise in the generation of insulin-producing cells from bone marrow MSCs, further work is needed to

determine the efficiency of this differentiation process and the functional properties of the bone marrow stem cell-derived cells.

The immunosuppressive qualities of MSCs cells may also be advantageous in treating type 1 diabetes. A study by Urban *et al.* (2008) showed that transplantation of bone marrow-derived MSCs resulted in significant suppression of β-cell-specific T cell proliferation in the pancreas, which may allow β-cell regeneration. The first clinical attempt to use autologous stem cells to treat T1DM showed that, following transplantation of autologous HSCs in newly diagnosed patients undergoing immunosuppression, patients showed increased levels of c-peptide and achieved insulin independence. The rationale was to preserve residual β-cell mass and facilitate endogenous mechanisms of β-cell regeneration. However, this study used very aggressive immunosuppression which is unlikely to be acceptable for routinely treating patients with T1DM (Voltarelli *et al.*, 2007; Couri *et al.*, 2009). Bone marrow stem cells probably do not have the capacity to differentiate *in vivo* into reasonable numbers of β-cells and therefore may re-establish β-cell tolerance through immunosuppression and the regeneration of regulatory T cells.

Other Tissue Stem Cells: There have been sporadic reports of differentiation of insulin-expressing cells from a number of other tissues, including intestine, which shares an endodermal origin with the pancreas. The enteroendocrine K-cells in the intestinal epithelium possess a regulated secretory pathway, and secrete the insulinotropic hormone GIP in a glucose-regulated manner. In addition, some of the transcription factors involved in the development of intestinal endocrine cells are also involved in islet development. These properties make them attractive candidates for engineering insulin expression and regulated secretion. On the other hand, intestinal enteroendocrine cells are quite rare and the intestinal epithelium is not readily accessible for biopsy.

Early studies showed that primary mouse intestinal cells could be induced to produce insulin through treatment with GLP-1, which activated the expression of NGN3. Cells contained insulin secretory granules and could normalise hyperglycaemia in diabetic mice (Suzuki *et al.*, 2003; Yoshida *et al.*; 2002) but insulin production was constitutive rather than in response to glucose. Ectopic expression of PDX1 in combination with treatment with betacellulin or co-expression of ISL1 in the rat enterocyte cell line IEC-6 also resulted in the

activation of insulin expression (Kojima *et al.*, 2002). More recently, a study by Talchai *et al.* (2012) showed that *Foxo1* ablation in the gut epithelium of both adult and neonatal mice results in the formation of insulin-expressing cells, as well as other pancreatic hormone-expressing cells. These insulin-expressing cells expressed markers highly enriched in pancreatic β-cells but which are normally absent in the gut e.g. C-peptide, MAFA, PDX1 and NKX6.1. Immunocytochemistry revealed that insulin-expressing cells were distinct from cells producing PP, GCG and SST, indicating that they are not mixed lineage pancreatic progenitors. Unlike pancreatic endocrine cells, enteroendocrine cells arise from NGN3-expresing progenitors throughout life. Therefore, insulin-expressing cells in the gut should have greater regenerative capacity than islet β-cells. In this study, following ablation of insulin-expressing cells using streptozotocin, mice initially developed hyperglycaemia, but after 9 days, hyperglycaemia began to decrease, and most were able to survive without exogenous insulin. This improvement could not be attributed to β-cell regeneration in the pancreas but was instead due to regeneration of insulin-secreting cells in the gut. Another study using transient overexpression of the pancreatic transcription factors PDX1, MAFA and MGN3 showed conversion of intestinal crypt cells into glucose-responsive β-like cells (Chen et al, 2014).

Neural stem cells have been shown to differentiate into cells of all three germ layers (Clarke *et al.*, 2000) and subpopulations of neurons express functional elements that are characteristic of β-cells, including the KIR6.2 and SUR1 subunits of the K_{ATP} channel, voltage-operated Ca^{2+} channels, GLUT2 glucose transporters and glucokinase (Yang *et al.*, 1999). In separate studies, rat primary neural stem cells (Burns *et al.*, 2005) and a transformed human neural cell line (Hori *et al.*, 2005) were differentiated towards an insulin-expressing phenotype and demonstrated expression of a number of pancreatic markers including PDX1, NKX2.2, NKX6.2, NGN3, NEUROD/BETA2, PAX4, PAX6 and ISL1. Although the number of reports on the use of neural stem cells as β-cell precursors is limited, neural stem cells are amenable to expansion *in vitro* (Minger *et al.*, 1996) and the biology of human neural stem cells is becoming increasingly well understood, suggesting a potential future source of replacement β-cells. However, the inaccessibility of these cells remains a problem.

Other reports have suggested the production of insulin-producing cells from diverse tissue sources such as dermis (Shi *et al.*, 2004), blood monocytes (Ruhnke *et al.*, 2005) and mouse

salivary glands (Hisatomi *et al.*, 2004). However, these studies have yet to be shown to be reproducible and further work is required to determine the functional capacity of these cells.

1.7 Embryonic Stem Cells

Embryonic stem cells (ESCs) have two key characteristics which make them appealing for use in cell replacement therapy. Firstly, they can self-renew in culture indefinitely, providing a potentially unlimited source of cells, and secondly, they are pluripotent i.e. they are able to give rise to any cell in the body given the appropriate signals.

ESCs are derived from blastocyst-stage embryos, 3-5 days post-fertilisation (figure 1.7). The blastocyst is comprised of an outer layer of cells enclosing the inner cell mass and a hollow cavity called the blastocoele. The cells of the inner cell mass have the potential to develop into any cell type of the body, but after implantation, rapidly differentiate to other cell types with limited developmental potential. However, if these cells are removed from the blastocyst and cultured *in vitro* under appropriate conditions, they continue to proliferate indefinitely and maintain the ability to differentiate into any cell type.

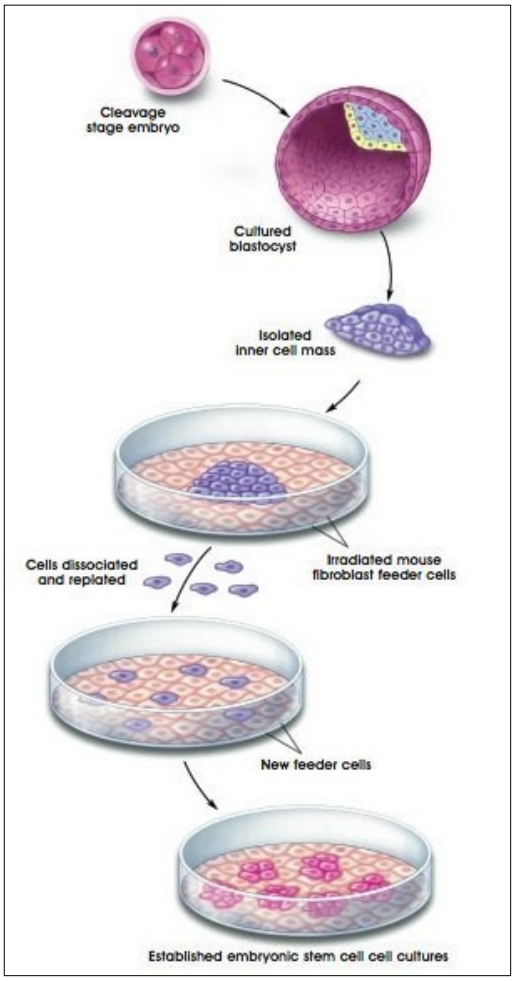

Figure 1.7. Derivation of ESC lines from cleavage-stage embryos. *The National Institutes of Health resource for stem cell research. Available online from: stemcells.nih.gov/info/2001report/appendixC.asp*

The derivation of mouse ESCs (mESCs) was first reported in 1981 (Evans & Kaufman, 1981; Martin, 1981). The term "embryonic stem cell" was introduced to distinguish these embryo-derived pluripotent cells from embryonal carcinoma (EC) cells, which were the first pluripotent cell lines to be established, derived from the undifferentiated compartment of both murine and human germ cell tumours (Finch & Ephrussi, 1967). Mouse ESCs express genes

characteristic of the early blastocyst such as OCT4 (Rosner *et al.*, 1990; Scholer *et al.*, 1990), SOX2 (Yuan *et al.*, 1995), REX1 (Rogers *et al.*, 1991) and UTF1 (Okuda *et al.*, 1998). These cells also express the cell surface antigen SSEA-1 (Krupnick *et al.*, 1994).

It took a further 17 years for the first derivation of human ESC (hESCs) lines to be reported (Thomson *et al.*, 1998). The cells displayed cell surface markers including SSEA-3, SSEA-4, the keratan sulphate antigens TRA-1-60 and TRA-1-81, and alkaline phosphatase (Thomson *et al.*, 1995; Thomson *et al.*, 1996). These cells also expressed high telomerase activity, showing that they are capable of extensive replication. These ESC lines maintained the capability of forming derivatives of all three germ layers when injected into severe combined immunodeficient (SCID) mice.

However, significant differences have been shown between mouse and human ESCs in terms of their culture conditions and gene expression. While the growth of mESCs in leukaemia inhibitory factor (LIF) and serum-containing media is sufficient to maintain them in an undifferentiated state in the absence of a feeder cell layer, this is not the case for hESCs (Thomson *et al.*, 1998; Reubinoff *et al.*, 2000). In addition, mouse and human ESCs differ in the expression of several cell surface antigens: SSEA-1 is expressed in undifferentiated mESCs where it is down-regulated upon differentiation, but it is not expressed in hESCs until these cells are differentiated. In addition SSEA-3, SSEA-4, TRA-1-60 and TRA-1-81 are expressed in undifferentiated hESCs but not in mESCs. Investigation into the gene expression of undifferentiated human and mouse ESCs revealed that hESCs expressed the transcription factors OCT4 and SOX2 in common with their mouse counterparts (Reubinoff *et al.*, 2000), though the expression levels of these genes, as well as of REX1 and UTF1 were found to be much higher in hESCs than in mESCs. hESCs share many characteristics with mouse epiblast stem cells (EpiSCs) which are derived from post-implantation epiblasts (Brons *et al.*, 2007), which are considered ‘primed’ pluripotent cells (unlike mESCs, which are considered ‘naive’ pluripotent cells), suggesting that they may not have the full developmental potential of mESCs. Despite these differences, the signalling pathways which control differentiation are very similar (Murray & Keller, 2008).

1.7.1 Differentiation of ESCs into Pancreatic β-cells

The first studies describing differentiation of mESCs into insulin-expressing cells used an antibiotic resistance gene driven by the insulin promoter to allow selection of cells that spontaneously differentiated into insulin expressing cells (Soria *et al.*, 2000). Soon afterwards, it was shown that hESCs could also spontaneously differentiate into insulin-expressing cells (Assady *et al.*, 2001), although the efficiency was very low. An early study described generation of insulin-producing cells from differentiated mESCs using a protocol modified from neural differentiation protocols (Lumelsky *et al.* , 2001), but there is now evidence that the insulin-positive cells produced were in fact of a neuronal lineage and that the insulin immunoreactivity observed resulted from insulin uptake from the medium (Rajagopal *et al.*, 2003). Naujok *et al.* (2009) extensively characterised ESCs differentiated using the protocols developed by Lumelsky *et al.* (2001) and D'Amour *et al.* (2006), showing that the cells produced by the first protocol have the characteristics of a neuronal lineage, while the second protocol produced cells with endodermal characteristics.

One strategy used to differentiate ESCs into insulin-expressing cells has been the over-expression of pancreas-associated transcription factors such as PDX1 (Miyazaki *et al*, 2004; Lavon *et al*, 2006) and PAX4 (Lin *et al.*, 2007; Liew *et al.*, 2008). Although these cells released marginal amounts of C-peptide in response to the insulin secretagogue tolbutamide, they did not respond to glucose (Liew *et al.*, 2008). More recently, Chen *et al.* (2011) showed that over-expression of SOX17, NGN3 and MAFA in mESCs resulted in increased expression of pancreatic genes, and, in the case of MAFA, enhanced glucose-stimulated insulin release. This was confirmed by another study in mESCs using over-expression of PDX1 and MAFA, together with either NGN3 or NEUROD (Xu *et al*, 2013). Other studies used over-expression of SOX17 to increase expression of DE markers (Qu *et al.*, 2008) or to establish stable endoderm progenitors which showed increased differentiation into DE and pancreatic lineages (Seguin *et al.*, 2008). However, the therapeutic scope of such genetically modified cells may be very limited. Li et al (2014) developed an approach that combined transient expression of reprogramming factors with small molecules to induce direct reprogramming of mouse fibroblasts into definitive endoderm-like cells, which could subsequently differentiate into pancreatic progenitor cells and insulin-expressing cells. The cells did not express pluripotency genes at any point, suggesting that they may be safer in terms of tumorigenicity than traditionally differentiated iPSCs.

Differentiation into Definitive Endoderm

More recently, the most successful protocols for generating insulin-expressing cells *in vitro* have been those which have recapitulated the signalling pathways important during *in vivo* pancreatic development, including TGF-β, Wnt, Hedgehog and Notch signalling (see section 1.3.2). The first step in the differentiation of ESCs into β-cells is the production of definitive endoderm (DE), a stage which has been extensively investigated and characterised in order to identify signals capable of inducing endoderm differentiation. Early studies found that DE could be induced in *Xenopus* embryos by activin A (Smith *et al*, 1990; Gurdon *et al*, 1994), which is often used as a replacement for Nodal in differentiation protocols as they both belong to the TGF-β family of signalling molecules, and this was further demonstrated in experiments in zebrafish and chicken (Grapin-Botton *et al*, 2007; Tam *et al*, 2003). The ability of Activin A to induce DE formation in ESCs was first demonstrated by Kubo *et al* (2004), who used a GFP reporter in the locus of the *Brachyury* gene to show that DE arises from a *Brachyury*$^{+}$ population capable of giving rise to both endoderm and mesoderm. Subsequent studies similarly used reporter ESC lines containing the *GFP* and *IL2Rα* reporter genes in the *Goosecoid* and *Sox17* loci respectively, to allow identification of culture methods capable of generating DE, and to distinguish DE from visceral endoderm (Yasunaga *et al.*, 2005; Tada *et al.*, 2005). Another study used an ESC reporter line that expressed CD4 in the *Foxa2* gene locus to show that both Wnt and TGF-β signalling are also important for DE formation (Gadue *et al.*, 2006).

In 2005, D'Amour *et al* published an influential protocol for the differentiation of DE that involved treatment with high concentrations of Activin A and low serum levels, resulting in a yield of *Sox17*-expressing cells of over 80%. Other markers of DE, including *Foxa2*, *Cxcr4*, and *Cer1* were also upregulated, while *Sox7*, a marker of visceral endoderm, was not. This protocol was further optimised in a later study (Kroon *et al*, 2008). It has now been established that Activin A has different effects depending on the concentrations used, with low concentrations (5-20ng/ml) able to maintain stem cells in an undifferentiated state under feeder-free conditions, intermediate concentrations (20-50ng/ml) resulting in the induction of mesoderm, and high concentrations (50-100ng/ml) resulting in the induction of DE. In addition, the timing of Activin A treatment also has an impact on the cells produced, with immediate inhibition of TGF-β signalling after endoderm formation resulting in induction of anterior foregut, and sustained TGF-β signalling resulting in posterior foregut and hepatic

development (Vallier *et al*, 2005). This may reflect the *in vivo* situation, where cells which migrate through the primitive streak earlier give rise to anterior DE, while cells which migrate later give rise to posterior DE. It has also been shown that inhibition of Activin A/Nodal signalling is required for further pancreatic specification of DE cells, further elucidating the timing needed for optimal differentiation (Cho *et al*, 2012).

While the role of Activin A has been extensively characterised, studies have also identified the roles of other signalling pathways in endoderm formation, including PI3K (McLean *et al.*, 2007), BMP4 (Teo *et al*, 2011; Xu *et al*, 2011), FGF (Sui *et al*, 2012), and Wnt (D'Amour *et al*, 2006) signalling. In a follow-up to the study published by D'Amour *et al* (2005), Nostro *et al* (2011) further optimised differentiation conditions in order to enhance the efficiency of differentiation, both at the DE stage, and at the final insulin-producing cell stage. This study further elucidated the temporal requirements for TGF-β, Wnt and BMP4 signalling, and showed that BMP4, bFGF and VEGF could replace serum in differentiation. Being able to carry out differentiation in serum-free conditions is an advantage, both in terms of reproducibility and in being more clinically relevant. More recently, a comprehensive study by Loh *et al* (2014) further elucidated the developmental signals and timing required for high efficiency differentiation into DE.

Culture conditions also play an important role in differentiation to DE, as demonstrated by Kubo *et al* (2004), who showed that high serum concentrations resulted in a lower efficiency of differentiation into DE, due to the presence of undefined growth factors and hormones in serum. Reduction of serum concentration to 0.2-0.5% was shown to be the most successful for generating DE (D'Amour *et al*, 2005). Whether the cells are differentiated in monolayer culture or as embryoid bodies is also likely to affect the differentiation process, and the necessary dose and timing of signalling factors (Sulzbacher *et al*, 2009). In an attempt to recreate more physiological conditions, and improve the quality of the cells produced, Van Hoof *et al* (2011) developed a protocol for the differentiation of ESCs into DE in adherent, size-controlled, clusters, which were able to give rise to homogenous clusters of pancreatic progenitors. However, Nostro *et al* (2011) found that differentiation kinetics were faster in monolayer cultures vs. embryoid bodies, with endodermal gene expression arising by day 3 in monolayer but not until day in embryoid bodies. However, embryoid body culture has the advantage of being more easily scaled up. Chetty *et al* (2013) found that the differentiation efficiency varied depending on the density of the cells at the start of the protocol, with cells

plated at higher densities forming more DE than those at low densities, suggesting that contact-mediated growth inhibition might improve the cells' response to differentiation signals. Treating low density cells with DMSO prior to differentiation improved the efficiency of differentiation to a level comparable with high-density cells. This was true for differentiation into several lineages, including DE. Another study linked differences in propensity of stem cells to form DE to the cell cycle state of the cells, and showed that the activity of Activin signalling during cell cycle progression is controlled by cyclin D proteins. Using small molecules to manipulate CDK4/cyclin D could improve differentiation efficiency, even to the extent where Activin A was not required (Pauklin & Vallier, 2013). Taylor-Weiner *et al* (2013) showed the role of extracellular matrix components, such as fibronectin and lamimin, in DE formation.

Differentiation into Insulin-Expressing Cells

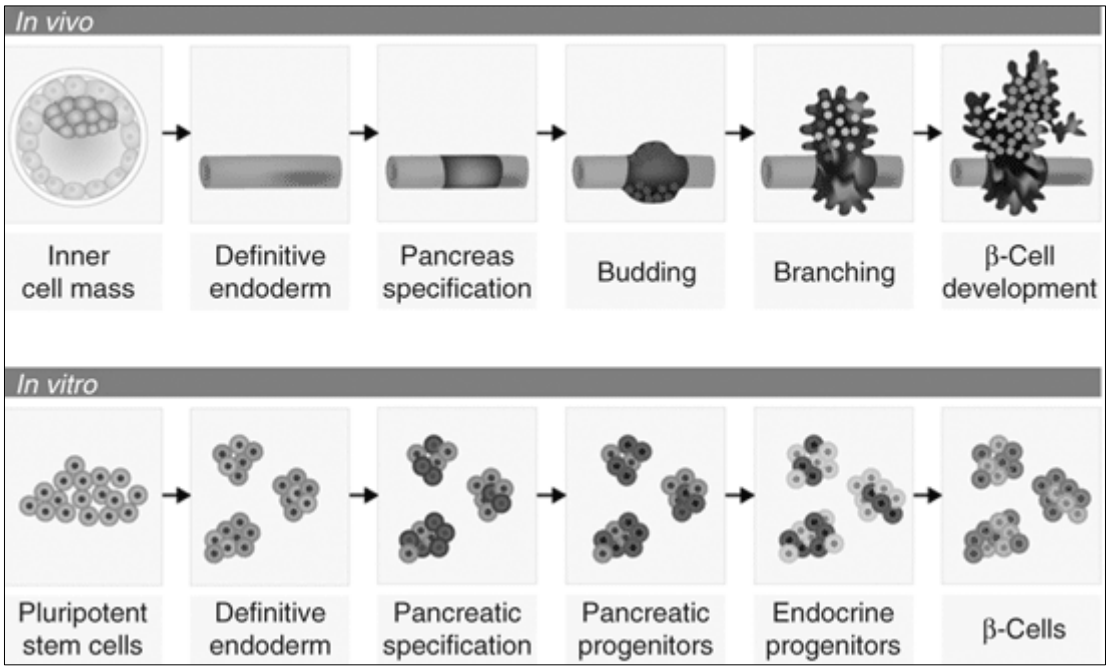

Figure 1.8. In vitro directed differentiation of pluripotent stem cells toward insulin-producing β-cells recapitulates in vivo development. *During embryonic development, pluripotent cells in the inner cell mass are progressively differentiated toward pancreatic endocrine cells, through sequential formation of definitive endoderm, pancreatic progenitors, and endocrine progenitors. Adapted from Maehr, 2011.*

In a follow-up of their previous study, and using knowledge of *in vivo* pancreas development, D'Amour *et al.* (2006) developed a 5-stage differentiation protocol that guided ESCs to differentiate into β-cells through the sequential formation of DE, primitive foregut, pancreatic endoderm, endocrine progenitors and finally pancreatic endoderm cells (figure 1.8). The cells produced using this protocol had an insulin content similar to that of human islets, and were positive for both insulin and C-peptide (although at very low efficiency, with only 7.3% of cells positive for insulin). The cells released insulin in response to various secretagogues, although not glucose. In addition, some cells were polyhormonal, suggesting that the cells produced had an immature phenotype.

Several other studies have adopted a similar sequential approach, but used different combinations of growth factors to induce differentiation. Jiang *et al* (2007) used a feeder- and serum-free protocol to differentiate ESCs into insulin-expressing cells using slightly different factors. 4% of the cells within the islet-like clusters were positive for c-peptide, and these were also capable of releasing insulin in response to glucose. These cells were further characterised in a subsequent study by the same group, which found that the islet-like clusters generated using this protocol contained cells which expressed insulin, glucagon and c-peptide, which were occasionally co-expressed in the same cells. However, when transplanted into diabetic mice, these islet-like clusters were unable to completely normalise hyperglycaemia, although they did extend survival and showed evidence of c-peptide release in response to glucose (Eshpeter *et al*, 2008). Shim *et al.* (2007) used a shorter protocol, treating ESCs with foetal calf serum, Activin A and retinoic acid, which produced a $PDX1^{+}/FOXA2^{+}$ population indicative of pancreatic endoderm, but did not produce significant numbers of insulin-expressing cells *in vitro*. However, when the precursor cells were transplanted into mice, they were able to reverse STZ-induced diabetes, and insulin-positive cells were observed when the graft was removed, suggesting further maturation *in vivo*. This was confirmed by another study, in which transplanted cells were able to correct hyperglycaemia in diabetic mice (Jiang *et al*, 2007). Cho *et al.* (2008) reproduced the protocol developed by D'Amour *et al* (2006), but added betacellulin and nicotinamide to the final stage, which allowed sustained expression of PDX1, but the efficiency of differentiation was lower, with very few insulin-expressing cells observed, and little or no insulin release in response to glucose. Shiraki *et al.* (2008) used a reporter ESC line expressing GFP in the *Pdx1* gene locus to investigate pancreatic differentiation, showing that co-culture of ESCs with a mesodermal cell line was able to induce differentiation into DE and PDX1-expressing

cells. They then elucidated the signalling molecules, concentrations and timings involved in differentiation to produce higher yields of DE and PDX1-expressing pancreatic progenitors which were able to give rise to both endocrine (including insulin-expressing) and exocrine cell types when transplanted into immunodeficient mice. An additional study that further elucidated the roles of signalling molecules in differentiation into pancreatic progenitors showed that retinoic acid plays a role in the differentiation of ESC-derived DE cells into $PDX1^+$ progenitors, and allows differentiation at a high efficiency by inhibiting hepatic differentiation (Cai *et al.*, 2010). Sui *et al* (2013) elucidated the role of BMP4 signalling in the later stages of pancreatic differentiation, while Nostro *et al* (2011) further elucidated the timing and doses of signalling molecules required at each stage of differentiation, and produced populations of cells containing more than 25% c-peptide-positive cells. A study using treatment with Activin A, retinoic acid, bFGF and nicotinamide resulted in the production of islet-like structures with distinct monohormonal cells, including insulin-expressing cells. These cells showed low levels of insulin release in response to glucose when cultured in monolayer, but this was increased when the cells were grown in suspension (Liu & Lee, 2012).

Despite these promising results demonstrating that ESCs can give rise to pancreatic endocrine cells *in vitro*, most of these studies produced cells that were functionally restricted, either having a polyhormonal phenotype, or lacking appropriate insulin release in response to glucose. Many of the cells had characteristics more reminiscent of foetal pancreatic cells. Perhaps this is not surprising, given that most protocols allow around 20 days for the differentiation of ESCs into insulin-expressing cells, even though insulin-expressing cells are not detected in the human foetal pancreas until around 8 weeks of development. A study by Patterson *et al* (2012) showed that this is a common feature of several *in vitro* differentiation protocols, with the differentiated progeny showing more similarity in terms of morphology and gene expression to cells normally found at very early stages of development *in vivo*. While increasing the length of time in culture did allow the cells to mature somewhat, they were still very different from the equivalent adult cells, suggesting that increased time in culture alone is not enough to drive maturation, and that additional signals are needed. However, Kroon *et al.* (2008) succeeded in producing a population of pancreatic progenitor cells which, several weeks after transplantation into immunodeficient mice, were shown to release C-peptide and insulin in response to glucose, demonstrating that unknown factors in the *in vivo* environment were capable of inducing these cells to complete differentiation into a

more mature β-cell phenotype. However, only a few mice were able to normalise hyperglycaemia. This result was confirmed by another group using a different differentiation protocol (Shim *et al.*, 2007).

Since then, there have been efforts to elucidate the factors needed for β-cell maturation. One study demonstrated that ESCs did not optimally differentiate *in vitro* until the cells were grown layered with other islet cells, showing a need for either growth factors or direct cell contact from α- or δ-cells for β-cell maturation *in vivo* (Kahan *et al.*, 2003). A more recent study elucidated the differences between immature and mature β-cells from postnatal mice, and found differences in glucose-stimulated insulin secretion (GSIS) between them. Comparison of gene expression between mature and immature β-cells showed that this is very similar between them, but that some genes, including *MafB* and *Ucn3*, showed differential expression. *Ucn3* expression was also different in immature β-cells derived from ESCs compared to those that had been allowed to mature through transplantation into mice, suggesting that this gene is a good marker to distinguish between immature and mature β-cells derived from ESCs *in vitro* (Blum *et al*, 2012). Another study identified aberrant chromatin remodelling in insulin-expressing cells derived *in vitro* compared to those matured in mice (Xie *et al*, 2013). However, the precise factors involved in β-cell maturation remain unknown, and, until they are elucidated, it is unlikely that fully functional mature β-cells will be generated *in vitro*.

Another problem facing the differentiation of β-cells from ESCs is the variability in the cells produced, even when using the same protocol. This clearly demonstrates that culture conditions and differentiation protocols still need to be optimised. The efficiency of differentiation also depends on the cell lines used, as cell lines have different propensities to differentiate into a pancreatic lineage (Osafune *et al.*, 2008), making it difficult to compare different studies. Nostro *et al* (2011) suggested that this was due to different levels of endogenous BMP4 signalling between different cell lines, and could be controlled by manipulating BMP4 levels.

In addition to the variability noticed between different cell lines, another problem facing the differentiation of pluripotent stem cells into insulin-expressing cells is the low efficiency of current differentiation protocols, with most studies reporting a final yield of around 4-7% insulin-expressing cells, and an optimised protocol still only resulting in a 25% yield (Nostro *et al*, 2011). If these cells are to be produced on a large scale for clinical use, the efficiency

will need to be improved. An appealing alternative may be to propagate intermediate progenitor cells, as was recently demonstrated by Cheng *et al.* (2012), who generated self-renewing endoderm progenitor (EP) cell lines from both human ESCs and iPSCs, which expressed key endodermal genes such as FOXA2 and SOX17, and were able to differentiate into both hepatic and pancreatic cells. The β-cells generated from these progenitors produced insulin at around 20% of levels seen in adult human islets in response to glucose, and in addition showed low expression of other endocrine hormones e.g. SST, GCG. Despite their extensive proliferative capacity *in vitro*, when injected into immunodeficient mice they did not form teratomas. A similar cell line has also been described in mice (Morrison *et al.*, 2008). Sneddon *et al* (2012) described the self-renewal *in vitro* of both endoderm and endocrine progenitors when co-cultured with mesenchymal cells. Although they were unable to identify which growth factors released by the mesenchymal cells were responsible for this effect, it is likely to be due to a combination of signalling pathways, including BMP, EGF and FGF signalling (Sui *et al*, 2013). Importantly, continued passage of endoderm or endocrine progenitors *in vitro* had no detrimental effect on later stages of differentiation, and insulin-expressing cells could still be detected after maturation *in vivo* (Sneddon *et al*, 2013). Greggio *et al* (2013) described a three-dimensional culture system allowing the expansion of pancreatic progenitors *in vitro*.

Something else that may contribute to the variability of the β-cells produced is that the protocols have typically used recombinant growth factors to induce differentiation, and these may vary in their biological activity. The use of small molecules to control differentiation is a promising alternative, as they can be chemically synthesised and purified, allowing standardisation. In addition, they are more easily produced and therefore less expensive. Studies have identified small molecules that can induce differentiation into DE (Borowiak *et al.*, 2009; Zhu *et al.*, 2009) and pancreatic progenitors (Chen *et al.*, 2009). More recently, Kunisada *et al.* (2012) replaced several growth factors used in differentiation protocols with small molecules, replacing Wnt3A with CHIR00921 (a GSK3β inhibitor) in DE formation, and Noggin with Dorsomorphin (an inhibitor of BMP type 1 receptors) in pancreatic specification. Addition of SB431542 (a TGF-β type 1 receptor inhibitor) was able to increase NGN3 expression, as well as that of other pancreatic markers.

The culture conditions may also have an impact on the quality of the cells produced. An early report showed that EB differentiation of ESCs resulted in higher endodermal and pancreatic

gene expression than differentiation of monolayer cultures (Xu *et al*, 2006). It has been suggested that the size of β-cell clusters may affect their insulin-releasing ability: clusters containing multiple β-cells release more insulin than single β-cells (Meda *et al.*, 1990), and *in vitro*-derived β-cells secrete more insulin when in three-dimensional clusters than when grown in monolayer (Brereton *et al.*, 2006). This was also shown when pancreatic progenitors derived from ESCs were transplanted into mice: a study by Van Hoof *et al.* (2011) developed a method for controlling the size of cell clusters and determined the optimum number of cells within each cluster for differentiation into pancreatic endoderm.

Before these cells can be considered for a clinical application, there are concerns that need to be addressed. Firstly, the potential tumorigenicity of the cells produced, as the many rounds of replication in culture that stem cells or cell lines undergo before transplantation may lead to accumulation of abnormalities that could potentially be oncogenic (Ben-David *et al*., 2011). In addition, the differentiation protocols are inefficient, and usually lead to a heterogeneous population of cells which may include pluripotent cells, which could give rise to teratomas if included in the transplant. It is therefore important to find strategies to avoid tumorigenesis, including the elimination of undifferentiated cells, since contamination could lead to the formation of teratomas. This is further discussed in section 1.8.2.

Secondly, there are concerns about the immune challenges faced by transplanted cells. There are two immune challenges facing the transplantation of β-cells derived from ESCs. The first is allograft rejection due to immunological mismatches between donor and recipient, and it seems likely that some form of immunomodulatory therapy will be needed to prevent graft failure. In solid organ and bone marrow transplants, it has been well established that the best strategy for reducing graft rejection is to minimise mismatching between donor and recipient. ABO blood typing is particularly important, as a mismatch will result in graft rejection, but in terms of HLA matching, the matching of HLA-A, HLA-B and HLA-DR is most important. However, the better the match, the greater the chance of graft survival, and the lower the requirement for immunosuppression (Grinnemo *et al*, 2008). Techniques developed in the field of whole organ transplantation to reduce graft rejection can also be applied to stem cell transplantation, including the use of immunomodulatory drugs and antibodies, co-stimulatory blockade to inhibit activation of T cells (Huber *et al*, 2013), or the induction of mixed haematopoietic chimerism by simultaneous transplantation of haematopoietic stem cells from

the same donor to induce tolerance to the transplanted cells (reviewed in Boyd & Fairchild, 2010.)

Some research has suggested that national or international banks of HLA-typed ESC lines could provide HLA- and ABO blood group-matched cells for the majority of individuals. Although a large number of ESC lines would be needed to provide every potential recipient with a perfect HLA match, one study has suggested that a bank of just 150 ESC lines could provide a beneficial match for the majority of the UK population (Taylor *et al*, 2005). For less genetically diverse populations, such as the Japanese population, as few as 50 ESC lines may provide a beneficial match. However, most ESC lines are derived from surplus embryos generated for IVF, so the MHC haplotypes are unknown at the time of ESC derivation and rare haplotypes may not become available. In addition, ethnic minorities are poorly represented in this pool of embryos (Fairchild, 2010). One way around this may be to generate ESC lines using somatic cell nuclear transfer, in which the nucleus from a somatic cell is transferred into an enucleated oocyte, which is then cultured to the blastocyst stage, from which ESCs of a known genetic background can be derived. This was only recently achieved in human cells (Tachibana *et al*, 2013).

The second immune challenge is the problem of auto-immunity in T1DM. This has been the subject of much research, and can be divided into antigen-specific immunomodulation, including vaccination with antigens associated with the onset of T1DM e.g. GAD65; and non-antigen-specific immunomodulation, which aim to more generally suppress the autoimmune response, including treatment with T cell-inhibiting drugs e.g. cyclosporine A; T cell-inhibiting antibodies; thymoglobulin, which inhibits not only T cells but also B cells, dendritic cells, NK cells and endothelial cells; B cell-specific antibodies; cytokines; and co-transplantation of autologous bone marrow stem cells (reviewed in Waldron-Lynch & Herold, 2011).

One method which could be used to protect transplanted cells from the immune system, potentially providing a solution to both immune challenges, is encapsulation of the transplanted cells. This method has been tested repeatedly in previous years, initially with poor cell survival (Lacy *et al.*, 1991; De Vos *et al.*, 2002), although more recently, improved methods of encapsulation have shown better results (Kirk *et al*., 2014; Yahknenko *et al.*, 2014). As described in section 1.5.4, encapsulation of pig islets has been suggested to prevent

their rejection by the immune system. Survival and function of stem cell-derived insulin-expressing cells in capsules has also been demonstrated (Bruin *et al*, 2013).

1.8 Induced Pluripotent Stem Cells

In 2006, a significant advance in the field of stem cell research was made with the discovery that 4 transcription factors could reprogram somatic cells into induced pluripotent stem cells (iPSCs). Reprogramming was initially carried out in mouse fibroblasts through screening of 24 candidate genes that had all been implicated in the establishment and maintenance of pluripotency in ESCs. This pool of candidates was subsequently narrowed down to 4 transcription factors: OCT4, KLF4, c-MYC and SOX2, which were sufficient to enable reprogramming (Takahashi & Yamanaka, 2006). These reprogrammed cells were shown to be similar to ESCs in terms of their morphology and gene expression, and were karyotypically normal. Although iPSCs from this first experiment were not germline-competent, they were able to give rise to all three germ layers *in vitro*. Germline-competent iPSCs have since been generated (Okita *et al.*, 2007). Reprogramming of human fibroblasts was achieved a year later (Takahashi *et al.*, 2007; Yu *et al.*, 2007), and since then, iPSCs have been generated from numerous different human cell types (Aasen *et al.*, 2008; Aoi *et al.*, 2008; Haase *et al.*, 2009; Giorgetti *et al.*, 2009; Loh *et al.*, 2010; Maherali *et al.*, 2008; Utikal *et al.*, 2009; Sun *et al.*, 2009), including pancreatic β-cells (Stadtfeld *et al.*, 2008*)*, as well as from several other species, including rhesus monkey (Liu *et al.*, 2008).

Reprogramming of somatic cells into iPSCs holds great promise for the treatment of a range of diseases, including T1DM, through cell replacement therapy (figure 1.9). As an alternative source of pluripotent cells, iPSCs have advantages over ESCs, as they are patient-specific, and do not require the use of human embryos or oocytes in their derivation. iPSCs also have potential for the treatment of genetic defects through gene therapy in combination with cell replacement therapy, as the cells can be extensively characterised to ensure that gene repair is specific, reducing the risks of random, viral-mediated gene therapy. Proof-of-principle for gene therapy in combination with cell replacement therapy was shown by a study in which iPSCs were generated from mice models of sickle cell anaemia and the human sickle haemoglobin allele was corrected using gene-specific targeting. Haematopoietic progenitors were obtained from these corrected iPSCs and were able to reconstitute the haematopoietic

system and cure sickle cell anaemia when transplanted into irradiated mice (Hanna *et al.*, 2007).

In addition to their promise in cell replacement therapy, iPSCs are also extremely useful for disease modelling and drug screening *in vitro* (figure 1.9). iPSCs can be differentiated *in vitro* into several types of cells which will be useful for drug screening, including hepatocytes, cardiomyocytes and neurons. These should help to predict whether drugs have toxic effects on the liver, heart and brain at an early stage of research and development.

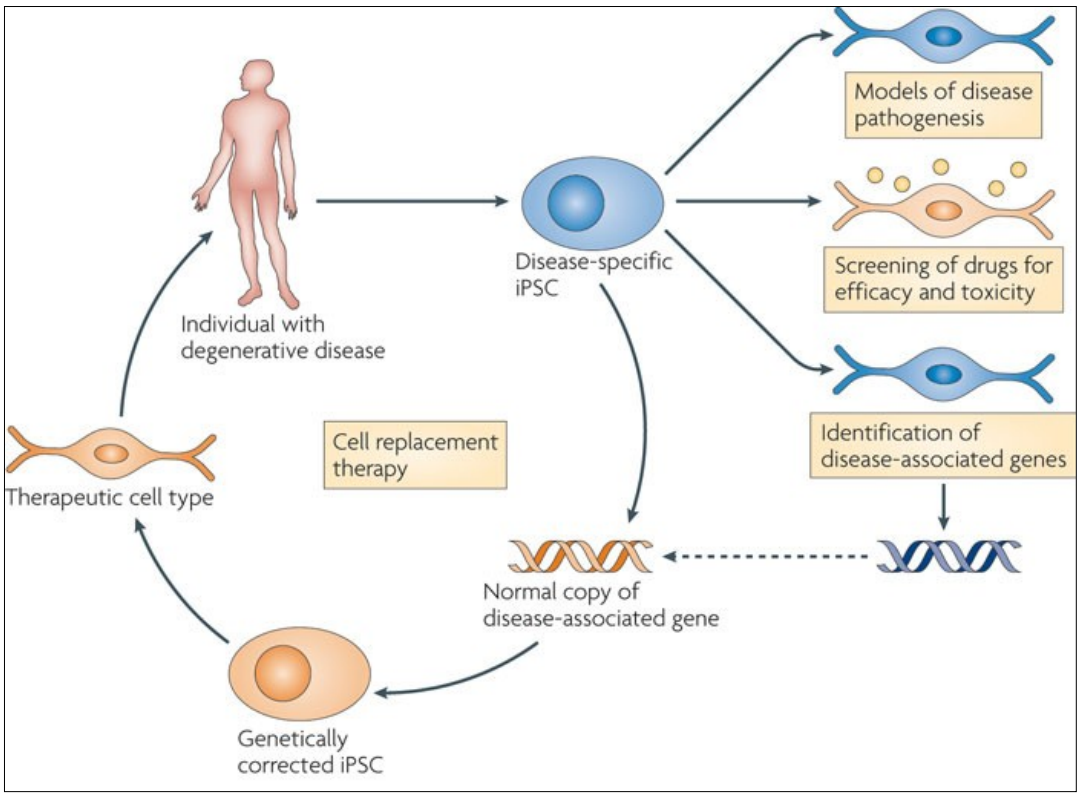

Figure 1.9. Derivation of iPSCs & their applications in cell replacement therapy, disease modelling and gene therapy. *Fairchild, 2010.*

Many cells affected by disease have defied attempts to culture them *in vitro*, or their phenotype has been changed by adaptation to tissue culture. Embryos shown to carry genetic disease using pre-implantation genetic diagnosis can yield ESC lines that model genetic disorders, but the vast majority of diseases that show more complex genetic patterns of

inheritance are not accessible in this way. The increased accessibility of disease-specific iPSCs compared to ESCs means that numerous iPSC lines derived from patients of different ages and with specific diseases are now available. These provide an extremely useful tool for disease modelling and drug screening. Although early studies focussed on diseases with monogenic defects, more recently iPSCs have been generated from patients with more genetically complex diseases, including T1DM (Maehr *et al.*, 2009). Currently, insights into T1DM come largely from rodent models such as the non-obese diabetic (NOD) mouse model, which although extremely useful, have limitations as they are rarely fully representative of the disease in humans. Given that T1DM seems to result from auto-immune interaction between T cells and β-cells, it may be useful to differentiate iPSCs from diabetic patients into β-cells and observe their interaction with the patient's T cells *in vitro* so that the auto-immune aspect of the disease can be recapitulated *in vitro*.

However, the usefulness of iPSCs for disease modelling is limited by current differentiation protocols. Although iPSCs can in theory give rise to any cell in the body, so far *in vitro* differentiation protocols have only been developed for a small subset of cell types. In addition, it is difficult to achieve a pure population using current differentiation protocols, and the cells produced frequently have an immature phenotype characteristic of foetal cells. This limits their usefulness for modelling adult diseases. In some cases, iPSCs may not be suitable for disease modelling: for example, iPSCs derived from patients with fragile X syndrome continued to silence the *Fmr1* gene after differentiation, in contrast to ESCs derived from pre-implantation genetic diagnosis, which reactivate *Fmr1* after differentiation (Urbach *et al*, 2010). Also, iPSCs could not be generated from patients with Fanconi anaemia unless the genetic defect was repaired first, so are unsuitable for modelling this disease *in vitro* (Muller *et al*, 2012).

1.8.1 Differentiation of iPSCs into Pancreatic β-cells

β-cells have been derived from iPSCs generated both from normal human fibroblasts (Tateishi *et al.*, 2008; Zhang *et al.*, 2009; Alipio *et al.*, 2010), as well as from fibroblasts from both a mouse model of T1DM (Jeon *et al.*, 2012) and human patients with T1DM (Maehr *et al.*, 2009). In addition, insulin-expressing cells have also been produced from rhesus monkey iPSCs (Zhu *et al.*, 2011) which may be useful for pre-clinical trials of iPSC-based therapies

for T1DM. The protocols used have been developed from those used to generate β-cells from ESCs, starting with the formation of DE. These studies have shown that not only do the cells produced demonstrate specific markers of the pancreatic lineage, such as expression of PDX1, HNF6, HNF4α and NKX6.1; they are also able to release insulin in response to glucose, suggesting that they may be functional. Saito *et al.* (2011) showed that both ESCs and iPSCs were able to form islet-like structures *in vivo*, which were able to release insulin in response to glucose and were able to normalise hyperglycaemia when transplanted into mice. However, the same problems associated with β-cells differentiated from ESCs also apply to those derived from iPSCs, such as the variability in the cells produced, their potential tumorigenicity and the problem of the auto-immune aspect of the disease. In addition, there are still issues specific to iPSCs that need to be resolved before they can be considered as an alternative to ESCs for a therapeutic application, which are discussed below.

Reprogramming Method		Cell Type	Reprogramming Factors	Efficiency (~ %)	Time (days)	Advantages	Disadvantages	References
Integrating	Retroviral	Fibroblasts, keratinocytes, β-cells, neural stem cells, peripheral blood cells, cord blood cells	OSKM OSNL OSK OKM OK OSKM + VPA OSK + VPA OS + VPA O O + NaB	0.01 - 1	8-21	High efficiency	Safety concerns over genomic integration and incomplete viral silencing. Slow kinetics. Safety concerns over oncogenicity of reprogramming factors.	Takahashi & Yamanaka, 2006; Yu *et al*, 2007 Aasen *et al*, 2008; Kim *et al*, 2008; Huangfu *et al*, 2008; Nakagawa *et al*, 2008; Wernig *et al*., 2008; Giorgetti *et al*, 2009; Kim *et al*, 2009; Loh *et al*, 2009; Zhu *et al*, 2010; Brown *et al*, 2010; Bar-Nur *et al*, 2011
	Lentiviral	Fibroblasts, cord blood cells, adipose stem cells	OSKM OSNL	0.01 – 1.1	20-30	High efficiency Can reprogram dividing and non-dividing cells	Safety concerns over genomic integration and incomplete viral silencing.	Blelloch *et al*, 2007; Haase *et al*, 2009; Sun *et al*, 2009; Hu *et al*, 2010
	Inducible lentiviral	Fibroblasts, keratinocytes, β-cells, peripheral blood cells, melanocytes	OKSM OKSM + C/EBPα OSK + n-Myc OSNL OKM	0.1 - 2	6 – 30	Reasonable efficiency Allows controlled expression of reprogramming factors	Genomic integration. Requires transactivator expression.	Maherali *et al*, 2008; Hockemeyer *et al*, 2008; Stadtfeld *et al*, 2008; Emlini *et al*, 2009; Utikal *et al*, 2009; Staerk *et al*, 2010
Excisable	Transposon	Fibroblasts	OSKM	0.1	8-10	Reasonable efficiency Fast kinetics Non-integrating	Labour-intensive screening of excised lines. Requires removal of multiple transposons.	Woltjen *et al*, 2009
	Cre-Lox cassette	Fibroblasts	OSK	0.1-1	21-28	Reasonable efficiency Non-integrating	Labour-intensive screening of excised lines LoxP sites retained in the genome.	Soldner *et al*, 2009

Non-integrating	Adenoviral	Fibroblasts, hepatocytes	OSKM	0.001	24-30	Non-integrating	Low efficiency. Slow kinetics. Small chance of genomic integration.	Stadtfeld *et al*., 2008; Zhou W*et al*., 2009
	Plasmid DNA	Fibroblasts, cord blood cells	OSKM OSNL OSKML + NaB	0.001	14	Non-integrating	Low efficiency. Requires multiple transfections.	Okita *et al*, 2008; Si-Tayeb *et al*, 2010; Chou *et al*. 2011
	Episomal vectors	Fibroblasts	OKSMNL	0.001	18	Non-integrating	Low efficiency. Labour intensive screening of lines.	Yu *et al*, 2009
DNA-free	Sendai virus	Peripheral blood cells	OSKM	1	21-25	Non-integrating	Difficulty in purging cells of replicating virus.	Seki *et al*, 2010
	Protein	Fibroblasts	OKSM + VPA	0.001	30-35	Non-integrating Direct delivery of reprogramming factors	Low efficiency. Slow kinetics. Requires multiple transfections. Requires large quantities of purified protein.	Zhou H *et al*., 2009
	mRNA	Fibroblasts	OSKM	1 - 4	20	High efficiency Fast kinetics Non-integrating Avoids innate immune response	Requires multiple rounds of transfection High toxicity	Warren *et al*, 2010
	Modified RNA	Fibroblasts	OKSM OSK + GLIS1	0.01 - 2	25-30	High efficiency Requires only one transfection	Requires modulation of the innate immune response Slow kinetics	Yoshioka *et al*, 2013
	miRNA	Fibroblasts, adipose cells	miR-302/367 cluster + VPA miR-200c/302/369	0.1	6-15	Non-integrating High efficiency Fast kinetics No exogenous transcription factors	Requires multiple rounds of transfection	Anokye-Danso *et al*., 2011; Miyoshi *et al*, 2011
	Chemical compounds	Fibroblasts, Adipose stem cells	DZNep, VPA, CHIR99021, 616452, Tranylcypromine, TTNPB,	0.2	28	More cost-effective More easily standardised Non-immunogenic	Low efficiency Slow kinetics	Hou *et al*, 2013

Table 1.1. Developments in the techniques for reprogramming somatic cells into iPSCs, and their respective advantages and disadvantages. *Abbreviations: O = OCT4, S = SOX2, K = KLF4, M = C-MYC, N = NANOG, L = LIN28, VPA = valproic acid, NaB = sodium butyrate. Adapted from Robinton & Daley, 2012.*

1.8.2 Considerations Prior to Clinical Application of iPSCs

Reprogramming Method

All four factors used in initial reprogramming experiments are overexpressed in at least some types of tumour, and *c-Myc* in particular is known to play a central role in both pluripotency and cancer pathways (Lee *et al*, 2013), raising concerns about a potential increase in tumorigenicity in these cells, and leading to efforts to remove or replace these factors with less oncogenic factors (table 1.1). For example, in mouse fibroblasts, SOX1 and SOX3 can replace SOX2; KLF2 can replace KLF4; and L-MYC and N-MYC can replace C-MYC (Blelloch *et al*., 2007; Nakagawa *et al*., 2008). A different combination of factors (OCT4, SOX2, NANOG and LIN28) has also been shown to reprogram neonatal human fibroblasts (Yu *et al.*, 2007). In addition, some cell types have endogenous expression of one or more of the reprogramming factors, removing the need for exogenous expression in reprogramming e.g. fibroblasts express C-MYC and KLF4, and it has been shown that C-MYC can be omitted from the reprogramming mix (Nakagawa *et al.*, 2008; Wernig *et al.*, 2008); while neural progenitor cells have high endogenous expression of SOX2 and C-MYC and have been reprogrammed initially in the absence of SOX2 (Eminli *et al.*, 2008; Kim *et al.*, 2008) and then using only OCT4 (Kim *et al.*, 2009), although the efficiency of reprogramming is reduced, and the inaccessibility of these cells make it unlikely that they will ever be widely used. In addition, iPSCs generated using one or two reprogramming factors showed a marked reduction in efficiency of differentiation into neural and haematopoietic lineages compared to ESCs and iPSCs generated using four factors (Löhle *et al*., 2012).

Another concern is the use of integrating viral vectors to deliver the reprogramming factors. The first reprogramming experiments used retroviral (Takahashi & Yamanaka, 2006; Takahashi *et al*., 2007) and lentiviral vectors (Blelloch *et al*., 2007; Yu *et al*., 2007; Brambrink *et al*., 2008; Hockemeyer *et al*., 2008; Maherali *et al*., 2008; Stadtfeld *et al*., 2008). However, both lentiviruses and retroviruses lead to integration of the viral transgenes into the genome which may not be completely silenced in the iPSCs generated (Lois *et al*., 2002). Genomic integration has been shown to alter gene function (Kustikova *et al*., 2005) and viral transgene re-activation in iPSC-derived chimeric mice has been implicated in tumorigenesis (Nakagawa *et al*., 2008). Also, random integration may influence the molecular signature of iPSCs by disrupting regulatory regions in the human genome. Transcriptional analysis showed that transgene expression from incompletely silenced viral vectors could perturb global gene expression in iPSCs (Soldner *et al*, 2009). A recent study demonstrated a propensity for

virally-reprogrammed iPSCs to revert to pluripotency following differentiation, leading to teratoma formation following transplantation into mice (Polanco et al., 2013). For these reasons, it is unlikely that these cells would be acceptable for clinical applications.

Attempts to make iPSCs more clinically applicable have led to the development of non-integrating reprogramming methods (summarised in table 1.1), such as reprogramming using chemical compounds in place of the reprogramming factors (Zhu *et al*., 2010), as well as the use of expression plasmids (Okita *et al.*, 2008), episomal vectors (Yu *et al.*, 2009), piggyBac transposition (Woltjen *et al*., 2009), Cre-recombinase-based excisable viruses (Soldner *et al.*, 2009), membrane-soluble protein-induced reprogramming (Zhou *et al*., 2009), transient transfection with plasmid DNA (Si-Tayeb *et al*., 2010), delivery of reprogramming factors conjugated with cell-penetrating peptide (Kim *et al.*, 2009) and reprogramming using synthetic modified mRNA (Warren *et al.*, 2010).

Reprogramming strategies using small interfering RNAs (siRNAs) or microRNAs (miRNAs) may offer an alternative reprogramming strategy to generate more clinically relevant iPSCs. Zhao *et al.* (2008) showed that siRNA targeting *p53* and *Utf1* dramatically increases reprogramming efficiency. More recently, a lentiviral vector encoding stem cell-specific miRNAs were used to reprogram somatic cells to pluripotency at a higher efficiency than the traditional reprogramming factors. The reasons for this increase in efficiency may include the fact that miRNA expression does not require protein translation, leading to a fast response, and also that miRNAs are able to target hundreds of mRNAs which can lead to a rapid phenotypic change (Anokye-Danso *et al.*, 2011). This was quickly followed by reprogramming by direct transfection of mature double-stranded miRNAs (Miyoshi *et al.*, 2011) which has the additional advantage of avoiding viral integration. iPSCs generated in this way have been reported to be less tumorigenic when compared with iPSCs generated by other approaches. This may be due to the fact that miRNA-mediated reprogramming does not require exogenous expression of c-MYC; that the miRNAs used are known to target several cell cycle regulators, resulting in the attenuation of the G1/S phase transition; and that these miRNAs result in the upregulation of tumour suppressor genes such as *p16Ink4a* and *p14/19Arf* (Lin *et al.*, 2010). However, there may also be disadvantages to using this approach: firstly, in order to provide the necessary amount of miRNAs over the course of reprogramming it was necessary to carry out repeated transfections; and secondly, the need for such large amounts of synthetic nucleotides may greatly increase the cost of large-scale reprogramming needed for the use of iPSCs in a clinical setting.

Choice of Donor Cell Type

The efficiency and kinetics of reprogramming vary depending on the starting cell type: for example, retroviral reprogramming of mouse embryonic fibroblasts can be achieved in 8-12 days, while human fibroblasts typically take around 20-25 days. Fibroblasts were used in early reprogramming experiments, and are still the most popular cell type due to their accessibility, commercial availability, and ease of maintenance in culture. However, reprogramming of fibroblasts is relatively slow and inefficient. Since then, many different cells have been reprogrammed into iPSCs, each with their own advantages and disadvantages (table 1.1).

Keratinocytes (Aasen *et al.*, 2008; Maherali *et al.*, 2008) represent an easily accessible source of cells (obtained from plucked hair follicles) that can be reprogrammed with ~100-fold increase in efficiency and ~2-fold increase in speed compared to fibroblasts, although this was only reported for juvenile cells, and the efficiency is likely to be less for older cells. In addition, they require extended culture *in vitro* to obtain enough cells for reprogramming, which increases their likelihood of accumulating genetic abnormalities. Furthermore, due to their physical location, keratinocytes have prolonged exposure to UV light, which further increases their risk of DNA damage. Melanocytes (Utikal *et al.*, 2009) have many of the same advantages and disadvantages of keratinocytes, in that they are accessible and can be reprogrammed at a higher efficiency and speed, but they may have accumulated genetic abnormalities.

Adipose stem cells from patients of a range of ages have also been reprogrammed at a higher efficiency (~ 20-fold increase) and speed (~2-fold faster) than fibroblasts. They are readily available through lipoaspiration, which is a more invasive procedure, but a small amount of tissue provides large numbers of cells for reprogramming, removing the need for their expansion *in vitro*. In addition, the derivation of iPSCs from adipose stem cells does not require the support of feeder cells (Sun *et al.*, 2009).

iPSCs derived from blood cells are an appealing prospect, as they can be obtained through a minimally invasive procedure, or through existing blood banks, and does not require the extended expansion in culture that fibroblasts do. The first study to show that haematopoietic cells can be successfully reprogrammed into iPSCs used $CD34^+$ cells from the peripheral blood of patients undergoing G-CSF mobilisation. While this increases the yield of $CD34^+$ progenitor cells, which have been suggested to be more amenable to reprogramming than terminally differentiated peripheral blood cells, it is also associated with side effects such as headache, nausea, and bone pain. This study

used retroviral reprogramming, which requires cells to be dividing, so cells were expanded in culture prior to reprogramming. However, this led to a decrease in the numbers of progenitor cells and an increase in numbers of differentiated cells over time. The reprogramming efficiency was still relatively low (0.01-0.02%) (Loh *et al.*, 2009). Subsequent studies reprogrammed peripheral terminally differentiated T cells, which are abundant in peripheral blood, so that only 1ml of blood was needed to provide sufficient cells for reprogramming, and which can be easily expanded in culture (Brown *et al.*, 2010; Seki *et al.*, 2010; Staerk *et al.*, 2010). A potential concern of using T cell-derived iPSCs is the persistence of T cell receptor gene rearrangements in the iPSC genome, and their potential effects on subsequent differentiation, but so far, no significant differences in differentiation potential have been observed between ESCs and T cell-derived iPSCs. A considerable advance was made by Seki *et al.* (2010), who used a non-integrating approach by using Sendai virus to deliver the reprogramming factors, which resulted in higher reprogramming efficiency. More recently, Chou *et al.* (2011) used an episomal vector to reprogram both peripheral blood and cord blood cells.

Another potential source of haematopoietic cells for reprogramming is cord blood (Haase *et al.*, 2009; Giorgetti *et al.*, 2009: Chou *et al.*, 2011), which is routinely available from commercial and public banks, covering a wide range of haplotypes. In addition, the isolation of progenitors is more efficient from cord blood than peripheral blood, and cord blood cells have higher proliferation potential than peripheral blood cells. The fact that the cells are younger also means that they are less likely to have acquired potentially harmful genetic abnormalities. However, the availability of autologous cells is restricted to those who had their cord blood banked at birth. It is also unclear how long these cells can be cryopreserved for and still generate iPSCs: one study used cord blood that had been cryopreserved for more than 5 years (Giorgetti *et al.*, 2009), but many patients who could potentially benefit from cell replacement therapy using cord blood-derived iPSCs would not need them until much later in life. In addition, cord blood cells have been reprogrammed using just OCT4 and SOX2 without the need for the oncogenes c-MYC and KLF4, which makes them more clinically relevant (Giorgetti *et al.*, 2009). The differentiation status of the starting cell type also has an impact on reprogramming: for example, haematopoietic progenitor cells can generate 300 times more iPSC colonies than terminally differentiated T or B cells (Kim *et al.*, 2010).

An additional consideration when choosing a starting cell type for reprogramming is that there is emerging evidence that iPSCs retain an epigenetic memory of their starting cell type (Polo et al, 2010). This may adversely affect their ability to differentiate into cells of other lineages, with some

studies showing that iPSCs have lower efficiencies of differentiation into haematopoietic, neuroepithelial, neuronal and cardiac lineages than ESCs (Feng *et al*, 2010; Hu *et al*, 2010). However, in some cases, epigenetic memory may be advantageous for generating cells that are difficult to generate through *in vitro* differentiation, such as pancreatic β-cells: a study in which iPSCs were generated from pancreatic β-cells showed that these iPSCs could differentiate into β-cells *in vitro* at considerably higher efficiencies than iPSCs derived from other sources (Bar-Nur *et al*, 2011).

Efficiency and Speed of Reprogramming

The reprogramming process is very slow (~ 4 weeks) and inefficient (~0.01%) making it difficult to scale up this technology to the level needed for use of autologous cells for cell replacement therapy. The use of small molecules and soluble factors to increase the efficiency of reprogramming is appealing given their ease of use, ready availability, potential for standardisation, and lack of permanent genome modification. Numerous small molecules (and other factors) which improve the efficiency of reprogramming or replace some of the reprogramming factors have now been identified (reviewed in Feng *et al*, 2009). It has recently been shown that a combination of several chemical compounds is sufficient to reprogram mouse fibroblasts and adipose stem cells into iPSCs without the need for any additional reprogramming factors (Hou *et al*, 2013), although it has yet to be shown whether this is possible for human cells. Certain culture conditions, such as carrying out reprogramming in hypoxic conditions (Yoshida *et al*., 2009), have also been shown to improve efficiency.

Since reprogramming depends on the resetting of the epigenetic state, numerous epigenetic modifiers have been identified which increase the efficiency of reprogramming (Mikkelsen *et al*, 2008; Shi *et al.*, 2008), including valproic acid (Huangfu *et al.*, 2008). However, the broad and non-specific effects of these epigenetic modifiers may result in global dysregulation of gene expression: for example, the demethylating agent 5-azacytidine (5-AZA) is mutagenic (Jackson-Grusby *et al.*, 1997), and mice with global alterations in DNA methylation develop tumours at a high frequency (Gaudet *et al*., 2003).

Culture Methods and Standardisation

In the past few years, standards for the identification of iPSCs and for the assessment of their functional equivalence to ESCs have become widely accepted. Assessing the success of reprogramming into iPSCs starts with the identification of compact colonies that have distinct borders and well-defined edges, and that consist of cells which have a large nucleus, large nucleoli and little cytoplasm. Reprogramming is a stochastic process that results in the production of a wide range of colony morphologies, and while some of them may appear similar to ESCs, further characterisation is needed to identify true, fully reprogrammed iPSCs, including the expression of certain molecular markers, as well as functional assays.

Molecular markers which allow the identification of fully reprogrammed iPSCs include the expression of a network of pluripotency genes, including *Oct4*, *Sox2* and *Nanog*, at comparable levels to ESCs; as well as the expression of a number of cell surface markers including SSEA-3/4, TRA-1-81 and TRA-1-60 (Maherali & Hochedlinger, 2008). Alkaline phosphatase has been widely used as a marker for pluripotency, but recent evidence shows this to be insufficient to identify iPSCs, as cells that have been only partially reprogrammed may also stain positive (Chan *et al*, 2009). iPSCs must be able to grow and maintain pluripotency independently of transgene expression. iPSCs must also be epigenetically similar to ESCs, with DNA demethylation at the promoters of pluripotency genes, X chromosome reactivation in female cells, and the presence of bivalent chromatin domains at developmental genes (Maherali & Hochedlinger, 2008).

Analysis of the pluripotency of mouse iPSCs can be rigorously assessed through the generation of chimaeric mice, which evaluates the potential of iPSCs to contribute to the normal development of adult tissues when injected into the blastocyst. Germline transmission is assessed by the ability of these chimaeras to give rise to offspring derived entirely from iPSCs (Okita *et al.*, 2007). The test with the highest stringency is tetraploid complementation, which involves the generation of extra-embryonic tissue such as the placenta by fusion of embryonic cells at the two-cell stage. The resulting tetraploid embryo can then be combined with normal diploid iPSCs, giving rise to a foetus in which all extra-embryonic tissue is derived from the tetraploid cells and the foetus is derived from the diploid iPSCs. This measures the ability of the iPSCs to direct the normal development of an entire organism. This test has only been accomplished for a small number of iPSC lines, and it appears that not all iPSC lines are able to contribute to all-iPSC mice (Stadtfeld *et al*, 2010). Due to obvious ethical constraints, these tests for pluripotency cannot be carried out for human iPSCs. The most commonly used method of assessing the pluripotency of human iPSCs is differentiation *in*

vitro, either as embryoid bodies (compact balls of cells that form loosely organised tissues which resemble the gastrulating embryo) or in monolayer cultures, in order to assess the differentiated cells for markers of all three embryonic germ layers. The teratoma formation assay involves measuring the ability of iPSCs to differentiate into all three germ layers *in vivo*, following injection of undifferentiated cells into immunodeficient mice. Although this is the most stringent assay available for testing the pluripotency of human iPSCs, it does not assess whether the iPSCs are able to give rise to every cell in the body, nor does it assess the contribution of the iPSCs to the germline (Robinton & Daley, 2012).

In order to be clinically relevant, iPSCs need to be generated and cultured using defined media components that are free of animal components. Many cell lines are grown on a feeder layer of mouse fibroblasts, which has the possibility of contaminating iPSCs with animal pathogens, as well as transferring potentially immunogenic antigens to the stem cells (Martin *et al*, 2005). Currently, the only ESC lines licensed for clinical use were not derived under xeno-free conditions (Crook *et al*, 2007), although they were subsequently maintained in culture under xeno-free conditions. In 2012, another group developed the first animal-component free and good manufacturing practice (GMP)-compliant ESCs (Tannenbaum *et al*, 2012). One possible alternative to mouse feeders is to use autologous skin fibroblasts from the same patient as a source of feeder cells (Takahashi *et al*, 2009). Another possibility is the use of synthetic basement membranes such as Matrigel™, which is a soluble basement membrane extract of the mouse Engelbreth-Holm-Swarm tumour that gels at room temperature, forming a reconstituted basement membrane which can be used to maintain pluripotent stem cells without a feeder cell layer. However, this is still not a completely defined formula, and is derived from a mouse cell line, so is not compatible with xeno-free culture. Extracellular matrix prepared from human fibroblasts in combination with xeno-free medium has also been reported (Meng *et al*, 2010), and several other groups have developed synthetic surfaces capable of maintaining pluripotent stem cells (reviewed in Villa-Doaz *et al*, 2012).

There has also been a move towards more defined cell culture medium, with Rajala *et al* (2010) reporting the development of a fully defined xeno-free medium, capable of supporting the expansion of ESCs and iPSCs. In addition, mTeSR-2 is a defined, xeno-free medium for the maintenance of pluripotent stem cells in culture (www.stemcelltechnolgies.com). A study by Bergstrom *et al* (2011) compared four xeno-free culture systems. Many of the advances in GMP culture of ESCs will also apply to iPSCs, and recent studies have reported reprogramming under feeder- and xeno-free conditions (MacArthur *et al*, 2012; Warren *et al*, 2012), which is a big step towards the derivation of

clinical grade iPSCs. However, the scale up of stem cell culture to the levels required for clinical application remains challenging.

Several studies have described the use of suspension culture for large-scale growth of pluripotent stem cells in culture (Andang *et al*, 2008; Amit *et al*, 2011; Zweigerdt *et al*, 2011), while more recently a study by Fluri *et al* (2012) reported the derivation and expansion of iPSCs in suspension culture, which would allow large-scale production using bioreactors. However, they used reprogrammable mouse embryonic fibroblasts, which just require the addition of doxycycline for the reactivation of pluripotency genes and subsequent reprogramming, so it is unclear whether the reprogramming efficiency would be high enough when using cells that have not been genetically modified in this way. A subsequent study carried out reprogramming of unmodified cells in stirred suspension bioreactors, but they used retroviral vectors for reprogramming, making the study less clinically relevant than if they had used a non-integrating method (Shafa *et al*, 2012). Again, it is unclear whether the reprogramming efficiency would be high enough.

In addition, new strategies will also need to be developed for the differentiation of pluripotent stem cells into insulin-expressing for cell replacement therapy, as currently many protocols rely on undefined products such as recombinant growth factors for the direction of cell fate. Commercial products for differentiation are now becoming available which have the advantages of being fully defined, serum- and animal product-free, and less variable in their biological activity, making them more clinically applicable.

Tumorigenicity

There are two concerns regarding the tumour-forming capacity of iPSCs: malignant transformation of the transplanted differentiated cells due to the use of integrating vectors or oncogenic reprogramming factors, and benign teratoma formation resulting from the transplantation of grafts contaminated with undifferentiated cells. Studies in both mouse (Miura *et al*, 2009) and human (Gutierrez-Aranda *et al*, 2010) cells found that transplantation of embryoid bodies derived from iPSCs gave rise to teratomas at higher efficiencies than those derived from ESCs, and that this was due to the presence of higher numbers of undifferentiated cells in the transplant. This makes finding strategies to remove any undifferentiated cells from the transplant even more important.

Early reports reported the removal of teratomas retrospectively through the use of suicide genes and chemotherapy (Schuldiner *et al.*, 2003; Cao *et al.*, 2007; Vazquez-Martin *et al*, 2012), but these methods have several disadvantages including adverse effects, drug resistance, and the fact that they were used retrospectively. Removal of undifferentiated cells prior to graft transplantation is preferable, as this reduces the risk to patients, as well as reducing the need for post-transplantation monitoring. Several techniques have been developed, including antibody-induced cytotoxicity (Choo *et al*, 2008; Tan *et al*, 2009), lectin binding for the selection of pluripotent cells (Wang *et al*, 2011), fluorescence-activated cell sorting (FACS), and magnetic-activated cell sorting (MACS) (Tang *et al*, 2011). The last two methods can be used either to eliminate pluripotent cells (Ben-David *et al*, 2013) or enrich differentiated cells based on their cell surface marker expression. Enrichment of differentiated cells has the advantage of purifying the graft to include only a single cell type, but this relies on knowledge of cell surface markers that can adequately distinguish the cells of interest from any other cells which may have arisen during directed differentiation, and usually requires a combination of markers. Recent studies identified novel surface markers for cells at different stages of the pancreatic differentiation process cells (Kelly *et al*, 2011; Fishman *et al*, 2012). However, these approaches have disadvantages including the loss of cells during sorting, and possible changes in cellular properties (Wang *et al*, 2012). The use of small molecules for selectively inducing cell death in pluripotent stem cells is also of interest: high-throughput screening identified two drugs capable of inducing death of undifferentiated ESCs (Conesa *et al*., 2012; Lee *et al*, 2013). Another study identified differences in adhesion strengths between pluripotent stem cells and differentiated cells which could be used to eliminate pluripotent cells from mixed cultures (Singh *et al*, 2013).

Another recent study raised another safety issue, by showing that virally-reprogrammed iPSCs may reactivate pluripotency genes after differentiation, leading to the formation of teratomas. However, this study only looked at early stages of differentiation, and iPSCs reprogrammed using episomal vectors did not have this problem (Polanco *et al*, 2013). However, this is another issue that will need to be considered before iPSCs can be considered for clinical use.

Immunogenicity

Although it has been generally assumed that iPSCs, because they can be derived from an autologous source, would be immune-tolerated by the patient, recent evidence has shown that in mice, iPSCs generated by either retroviral or episomal approaches are rejected by the recipient's immune system

more often than equivalent ESCs due to abnormal gene expression which contributes to increased immunogenicity of the cells (Zhao *et al*., 2011). However, this study only looked at the immune response to undifferentiated cells. A more clinically relevant study looked at terminally differentiated cells derived from ESCs and iPSCs, and found that cells derived from syngeneic iPSCs were no more immunogenic than those derived from syngeneic ESCs, either *in vivo* or *in vitro* (Guha *et al*, 2013). This study found no differences in the expression of the genes identified by Zhao *et al* (2011) between ESCs and iPSCs. These results were confirmed by a subsequent study which again found no evidence of immune rejection of syngeneic ESC- or iPSC-derived tissues (Araki *et al*, 2013).

The immunogenicity of iPSCs may depend on the cell type and the developmental stage of the starting material used for reprogramming, as iPSCs derived from some cell types (e.g. cord blood) appear to be innately less immunogenic than iPSCs derived from others (e.g. adult skin) (Liu *et al*, 2013), suggesting that epigenetic mechanisms may be involved in immunogenicity. In addition, iPSCs reprogrammed using retroviral reprogramming have been suggested to be more immunogenic than those generated using non-integrating methods (Zhao *et al*, 2011). This may be due to the fact that the retroviral transgenes can alter gene expression. However, due to the time it takes to differentiate a patient's own cells for therapy, and the expense of making these cells, it may be more appropriate to turn to banking of iPSCs. With the diversity of starting material available from donors, it is considerably easier to generate cell lines of a known genetic background that would provide an immunological match for most of the population compared to an ESC bank, although it is unlikely that even the best possible match from an iPSC bank will be able to entirely escape immune recognition following transplant.

1.9 Differences Between iPSCs and ESCs

Although early reports of iPSCs demonstrated that they were equivalent to ESCs in terms of their expression of pluripotency markers and ability to differentiate into cells from all three germ layers (Maherali *et al*, 2007; Wernig *et al*, 2007), recent in-depth analyses have revealed numerous subtle differences between the two cell types. Disparities were first observed in the differentiation abilities of ESCs and iPSCs, in both teratoma formation and *in vitro* differentiation assays. Just like ESC lines, iPSC lines have considerable heterogeneity and show different propensities for differentiation (Abeyta *et al*, 2004; Osafune *et al*, 2008; Martinez *et al*, 2012; Mills et al, 2013), and some studies

have shown that iPSCs have lower efficiencies of differentiation into haematopoietic, neuroepithelial, neuronal and cardiac lineages compared to ESCs (Feng *et al*, 2010; Hu *et al*, 2010; Narsinh *et al*, 2011). Studies in both mouse (Miura *et al*, 2009) and human (Gutierrez-Aranda *et al*, 2010) cells found that transplantation of embryoid bodies derived from iPSCs gave rise to teratomas at higher efficiencies than those derived from ESCs, and that this was due to the presence of higher numbers of undifferentiated cells in the transplant. iPSCs are generally less successful in generating high percentage chimaeras and live mice through tetraploid complementation than ESCs (Stadtfeld *et al*, 2010). While some researchers argued that these results show that iPSCs have an intrinsically lower capacity to differentiate into certain lineages *in vitro* than ESCs, others have shown otherwise. A more recent large-scale study compared the propensity of 49 iPSC lines and 10 ESC lines to differentiate into a neural lineage, and found that seven iPSC lines showed significantly reduced differentiation capacity, which corresponded to differences in the expression levels of several genes. However, the majority of iPSC lines could not be distinguished from ESC lines (Koyanagi-Aoi *et al*, 2013). In contrast, Kwon *et al* (2014) found that iPSCs differentiated into dopaminergic neurons at a higher efficiency than ESCs. These studies have led to questions about whether iPSCs are indeed identical to ESCs, and if they are not, what the functional implications of this might be for the use of iPSCs for *in vitro* applications such as disease modelling and drug screening, and *in vivo* use for cell replacement therapy.

1.9.1 Gene Expression

Early studies comparing the gene expression signatures of ESCs and iPSCs found that they were very similar, and distinct from other cell types (Muller *et al*, 2008). However, more recent studies have identified differences between the two cell types, although there has been disagreement over whether these are stochastic differences present between different cell lines, or indicative of a unique iPSC gene expression signature that separates them from ESCs. One study suggested that iPSCs have a unique gene expression signature that is shared among different iPSCs from different labs, generated using different reprogramming techniques (Chin *et al*, 2009). This study argued that the overlap in the gene expression signatures of iPSCs from different reprogramming experiments shows that the iPSC state is not stochastic. However, a subsequent study criticised these results, arguing that although Chin *et al* (2009) had identified 15 genes that were consistently differentially expressed between iPSCs and ESCs from several laboratories, these genes were not always consistently up- or downregulated across the iPSC lines. Meta-analysis showed that the transcriptomes of ESCs and

iPSCs do not cluster together; neither do they form two separate clusters separating iPSCs and ESCs. Instead, the transcriptomes generally form distinct clusters representing the lab of origin. This lab-specific clustering was not observed for the fibroblast lines from which the iPSCs were derived, which clustered together irrespective of the lab of origin. In addition, many of the genes that are differentially expressed between stem cells from different laboratories have also been previously reported to differ between ESCs and iPSCs. This study suggests that culture conditions may have a large impact on iPSCs and ESCs, resulting in lab-specific gene expression patterns that overshadow differences in gene expression between ESCs and iPSCs (Newman & Cooper, 2010). Interestingly, a large scale comparison of ESCs and iPSCs found that ESCs derived at Kyoto University could be distinguished from the other cell lines included in the analysis, providing further evidence for lab-specific differences between cell lines (Koyanagi-Aoi *et al*, 2013). Contradictory to this, another study argued that the lab-specific differences that were observed were due to analysis of array data without statistical correction of batch effects. When analysis was performed with statistical correction of batch effects, the lab of origin ceased to be the main difference between the cell lines, and differences between iPSCs and ESCs was seen to be the main difference (Wang *et al*, 2011). In addition, when genes that were differentially expressed were grouped into functional modules (groups of genes with similar functions), it was found that several of these were consistently different between ESCs and iPSCs, providing further evidence for an iPSC state that is distinct from ESCs, and suggesting that there may be a functional effect of these differences. However, a recent study carried out at the human pluripotent stem cell database showed that there were no genes that were consistently differentially expressed between ESCs and iPSCs, although there were genes that were differentially expressed between cell lines (Mallon *et al*, 2013), so there is still no consensus on whether iPSCs have a unique gene expression signature that can distinguish them from ESCs.

Investigating DNA methylation patterns of ESCs and iPSCs showed that a very small number of genes were differentially methylated between ESCs and iPSCs from different sources, and suggested that these perhaps represented a set of aberrantly methylated genes than can distinguish between ESCs and iPSCs (Ruiz *et al*, 2012). Analysis of data from previous studies confirmed that this subset of genes was also aberrantly methylated in these studies. These epigenetic aberrations were associated with changes in gene expression, and 20-50% of these aberrations persisted after differentiation. Differentiated samples segregated based on whether the starting cell type was an ESC or iPSC, which could explain functional differences observed between ESCs and iPSCs.

A study by Narsinh *et al* (2011) investigated expression of genes known to be involved in pluripotency and early lineage commitment in 3 ESC lines and 4 iPSC lines (derived from several cell types using different reprogramming methods), and found that, although the populations of both ESCs and iPSCs meet pluripotency criteria, when individual cells were analysed, iPSCs showed significantly higher variability than ESCs in expression of pluripotency-associated marker genes and surface antigens. This was still true for iPSC populations that were selected for using TRA-1-60/SSEA-4 surface antigen expression, suggesting that the variability is not due to the presence of higher numbers of differentiated cells in the population. This study did not report passage numbers, but if the cells were analysed at early passages, the higher variability observed in iPSC population could be due to the presence of incompletely reprogrammed cells in the iPSC population.

A handful of studies have analysed differences in gene expression between genetically identical mouse ESCs and iPSCs. A study by Stadtfeld *et al* (2010) assessed their pluripotency through their ability to contribute to all-iPSC mice through tetraploid complementation. This study found that, although all ESCs were able to give rise to tetraploid embryos, most iPSCs could not. Comparison of mRNA and miRNA expression levels showed that only one region was differentially expressed between the two cell types: the *Dlk1-Dio3* region on chromosome 12. This region, containing the genes *Gtl2* and *Rian*, as well as 26 miRNAs, was aberrantly silenced in most iPSCs compared to ESCs, and only iPSCs that express this region are able to support the development of all-iPSC mice. Silencing of this region is due to hypermethylation and hyperacetylation, but is not due to aberrant expression in the somatic cell of origin, and is maintained upon differentiation. However, treatment with histone deacetylase inhibitor VPA was able to rescue the defect, and the iPSCs were thereafter able to give rise to all-iPSC mice. This study suggests that a relatively small subset of transcripts distinguishes miPSCs from mESCs, and that differences seen in other studies are likely to be due to variation in genetic background or insertion of viral transgenes.

A study by Liu *et al* (2010) showed that expression of miRNAs encoded in the *Dlk1-Dio3* imprinted region, including miR-134, -323-5p, -409-5p, -495, and -543, in iPSCs that were not able to give rise to all-iPSC mice compared to fully pluripotent stem cell lines. Among the targets of these miRNAs were several components of the polycomb repressive complex 2 (PRC2). Decreased repression of these components in partially pluripotent cells may result in the maintenance of histone acetylation and failure to demethylate the *Dlk1-Dio3* region, which would be consistent with the reduced gene and miRNA expression levels. The authors of this study suggested that the expression levels of this region may provide a fast and effective method to characterise the pluripotency of iPSCs.

Stadtfeld *et al* (2012) further elucidated the timing of repression of maternal *Dlk1-Dio3* transcripts during reprogramming, with transcriptional downregulation preceding the acquisition of aberrant DNA methylation and gene silencing. This process is dependent on DNMT3A, and promoted by DNMT3l. They also investigated different culture conditions that were able to prevent aberrant silencing of this region, and showed that treatment with ascorbic acid can prevent but not reverse aberrant DNA methylation of the *Dlk1-Dio3* region. Cells treated with ascorbic acid maintained activating histone marks which are otherwise frequently lost during reprogramming. These cells then maintained normal DNA methylation levels, whereas cells that were not treated with ascorbic acid frequently lost activating histone marks, facilitating the recruitment of DNMT3A, and leading to progressive hypermethylation during reprogramming. Interestingly, ascorbic acid did not appear to affect methylation of any other genes other than in this region.

However, it is unknown whether this is also true for human cells: an evaluation of published data did not indicate aberrant regulation of the *Gtl2* homologue *Meg3* in hiPSCs (Yu *et al*, 2009). In addition, mouse and human iPSCs and ESCs are significantly different to one another, likely representing different developmental stages of the embryo, and for this reason any conclusions drawn from miPSCs are not necessarily applicable to hiPSCs. Very recently, a comparison of isogenic hESC and iPSC lines was carried out, differentiating an ESC line into neural progenitors, and then reprogramming these cells into iPSCs. This study found no differences in gene expression or DNA methylation between the ESC and iPSC lines tested, in either differentiated or undifferentiated cells, although the sample group was small (Mallon et al, 2014).

While there is still no definitive answer as to whether there are any consistent differences in gene expression between ESCs and iPSCs, it is clear that there is variability between cell lines which may have a functional impact on their pluripotency and ability to differentiate. The differences in gene expression observed between different cells lines could be due to several factors, including the presence of genomic and epigenomic aberrations, the presence of an "epigenetic memory" of the somatic cell of origin in iPSCs, and differences in microRNA expression.

1.9.2 Genomic Stability

It has long been established that ESCs are susceptible to acquiring chromosomal abnormalities with prolonged culture. The fact that the inner cell mass from which ESCs are derived is a transient state

existing only for a few days, may mean that genetic changes as a result of adaptation to culture cannot be avoided. The most common mutations in ESCs are aneuploidy of chromosomes 17, 12, 20 and X (Draper *et al*, 2004; Maitra *et al*, 2005; Baker *et al*, 2007; Spits *et al*, 2008), and these changes correspond with changes in gene expression (Narva *et al*, 2010; Mayshar *et al*, 2010), and have been suggested to be advantageous for growth. This is supported by the fact that many genes on these chromosomes are involved in processes such as self-renewal, differentiation and apoptosis (Draper *et al*, 2004), and also that aneuploid ESCs grow faster and are highly clonogenic (Baker *et al*, 2007). Despite these differences, aneuploid ESCs seem to maintain expression of pluripotency genes, and the ability to differentiate *in vitro*. However, chromosomal abnormalities are not observed in all ESC cultures, and it is unclear whether certain cell lines are more prone to developing abnormalities, or whether the instability is a result of certain culture methods: for example, it has been shown that the passage of ESCs by mechanical dissection preserves genomic integrity better than enzymatic passaging (Maitra *et al*, 2005), and one report only observed abnormalities in cells cultured in knockout serum replacement and not in those cultured in serum (Spits *et al*, 2008). A larger study carried out by Mayshar *et al* (2010) showed that cell lines prone to aberrations did not necessarily cluster together; instead, the main contribution seemed to be the laboratory of origin, suggesting that culture conditions have an effect. In addition, several studies have linked these abnormalities to an increased risk of oncogenesis, as several of these genes are also linked to cancer (Baker *et al*, 2007; Lefort *et al*, 2008; Narva *et al*, 2010), emphasising the importance of carefully characterising ESCs that are to be used for clinical applications, and keeping time in culture to a minimum.

While some chromosomal aberrations, such as gains at chromosomes 12, are common to both iPSCs and ESCs that are maintained in culture over long periods of time, some are specific to each cell type: for example, gains of chromosome 17 were observed only in ESCs and gains of chromosome 8 were seen only in iPSCs (Mayshar *et al*, 2010; Taapken *et al*, 2011). The question is whether iPSCs are more prone to genetic abnormalities than ESCs. The evidence seems to suggest that, in addition to chromosomal abnormalities acquired as a result of extended time in culture, iPSCs are also prone to develop aberrations as a result of the reprogramming process itself, or from mutations already present in the somatic cells of origin, as large numbers of genomic aberrations are seen in new iPSC lines shortly after reprogramming. These could result from a number of sources: mutations in the cell of origin that have been carried over through the reprogramming process; mutations acquired during reprogramming due to replication and metabolic stress; mutations resulting from the integration of the viral reprogramming transgenes; and mutations that arise as a result of adaptation to culture.

There is disagreement about which factor is the main contributor to genomic abnormalities in iPSCs, and also about whether iPSCs are more prone to genomic abnormalities than ESCs.

A study by Mayshar *et al* (2010) identified aberrations in early passage iPSCs that are likely to result from selection during the reprogramming process and initial adaptation to culture. These are unlikely to result from the starting population, as the aberrations could not be identified in the somatic cells, and multiple iPSC lines from the same starting population did not develop the same aberrations. It is also unlikely to be entirely due to the reprogramming method, as iPSCs derived without using viral vectors were also prone to aberrations. This was also observed in a study by Laurent *et al* (2011) who showed that in iPSCs, deletions appeared at the earliest passage numbers and often decreased over time in culture, suggesting that they are positively selected for during reprogramming and negatively selected against during culture. Several of the deletions involved tumour-suppressor genes. In addition, duplications arose over long-term passage, and these genes, which were often associated with tumorigenicity or cell proliferation, were positively selected for during culture. This was also observed by Hussein *et al* (2011), who found that the number of copy number variations (CNVs) in iPSCs was about twofold higher than in either ESCs or fibroblasts. Many of these CNVs fell outside 'normal' CNVs identified in large-scale screening of healthy individuals, and were observed at early passage numbers, suggesting that they arose as a result of reprogramming or early adaptation to culture. However, the number of CNVs decreased over extended passaging to a level similar to that seen in ESCs, presumably due to negative selection of these abnormalities during extended culture. A study which performed karyotype analysis on a large number of iPSCs and ESCs found no difference in the incidence of chromosomal aberrations between ESCs and iPSCs (Taapken *et al*, 2011), which is in contrast to other reports (Martins-Taylor *et al*, 2011).

iPSCs that are generated from adult somatic cells may have an increased mutational load due to pre-existing mutations. This is especially a concern with iPSCs generated from skin fibroblasts, as the skin is exposed to UV light which may increase DNA damage and subsequent mutations. For this reason, the age of the donor is also likely to affect the mutational load of the cells used to generate iPSCs. Several studies have showed that some mutations observed in iPSCs could also be observed in the fibroblast populations (Gore *et al*, 2011; Young *et al*, 2012). However, iPSCs had a significantly higher mutational load than fibroblasts, which indicates that genomic abnormalities observed in iPSCs are not entirely due to mutations in their cells of origin, but instead are likely to also arise during reprogramming or early adaptation to culture. Evidence for this includes the fact that early passage iPSCs contained new mutations not found in the fibroblast population (Gore *et al*,

2011). These mutations often became fixed in the iPSC line, as they could also be observed after extended passaging. New mutations were observed in the later passage cells, which are likely due to adaptation to culture. This is in contrast to previous studies (Laurent *et al*, 2011; Hussein *et al*, 2011) which showed that iPSCs initially had high levels of mutations that were likely to result from the reprogramming process, but that these were selected against over passaging until the mutational load was similar to that of ESCs.

Another possible reason for differences between ESCs and iPSCs may be due to differences in mitochondrial DNA. A study using a mouse model with an error-prone mitochondrial DNA polymerase demonstrated that iPSCs could be generated from somatic cells with a high mitochondrial mutational load, but that these cells had defects in both proliferation and differentiation (Wahlestedt et al., 2014).

While most of these studies looked only at genomic abnormalities in undifferentiated stem cells, a study by Laurent *et al* (2011) found that the most rapidly arising genomic aberrations were identified in cells undergoing directed differentiation. This shows that differentiation can be a highly selective process, and that genomically aberrant cells can rapidly take over a differentiating population. This stresses the importance of assessing the genomic integrity of differentiated cells as well as pluripotent cells, as this may have an impact on the function of these cells.

1.9.3 Epigenomic Stability

Differences between ESCs and iPSCs are not confined to genetic differences, but also include epigenetic differences. Detailed analysis of the epigenetic differences between ESCs and iPSCs was made possible by the development of high-throughput sequencing technologies and the generation of single nucleotide genome-wide maps of DNA methylation. The main features of the ESC epigenome are also seen in iPSCs, including genome-wide DNA methylation patterns and the establishment of bivalent histone marks at specific loci. However, despite these similarities, a core set of differentially methylated regions (DMRs) that seems to represent hot spots of failed epigenetic reprogramming has been identified (Lister *et al*, 2011). The DMRs are enriched for genes that are important for developmental processes. In the most exhaustive comparison so far, Kim *et al* (2010) reported that there were more DMRs in iPSCs than in ESCs. However, these DMRs were not linked to specific loci and therefore do not represent consistent differences between iPSCs and ESCs. The high

incidence of DMRs in iPSCs that are not present either in somatic cells or in ESCs suggests that they are stochastic and arise during reprogramming.

Several studies have looked at histone modifications in ESCs and iPSCs and showed that although reprograming results in a reduction of repressive chromatin marks, resulting in an epigenome that is more similar to that of ESCs than the starting cell type, there are some differences between the two cell types: some differences are due to residual histone modifications that are also seen in the cell of origin, while others are unique to iPSCs and are likely to result from aberrant reprogramming. These patterns correspond to differences in gene expression between ESCs and iPSCs (Pick *et al*, 2009; Hawkins *et al*, 2010). However, in contrast to these results, a study by Guenther *et al* (2010) showed that no more variation was observed between ESCs and iPSCs than within ESC lines or iPSCs lines, and also found no evidence that iPSCs contain histone marks reflective of their cell of origin. Although some variation in histone modifications was noticed between cell lines, these differences had little or no effect on gene expression. A study by Pick *et al* (2009) investigated differences in genomic imprinting in ESCs and iPSCs. In general, ESCs show a substantial degree of genomic imprinting stability but aberrant biallelic expression of some genes was observed in some iPSC lines. This biallelic expression was due to significant DNA demethylation in the promoter in the affected iPSC lines.

1.9.4 Epigenetic Memory

Another factor which has been proposed to result in differences in gene expression between ESCs and iPSCs is the presence of an "epigenetic memory" of the somatic cell of origin in iPSCs. This may lead to significant and persistent donor-cell gene expression, due to aberrant silencing of somatic genes during the reprogramming process, and may affect the differentiation potential of iPSCs. This phenomenon was first noticed when studies were carried out to investigate differences in gene expression between ESCs and iPSCs. Chin *et al* (2009) carried out genome-wide expression analysis in both ESCs and iPSCs, and found that genes expressed at a higher level in ESCs vs. iPSCs are also expressed at higher levels in ESCs vs. fibroblasts. Of genes that are expressed at higher levels in iPSCs than ESCs, many appear to be inefficiently silenced from the fibroblast state. This suggests that iPSCs have not fully silenced the expression pattern of the somatic cell from which they are derived, and failed to fully activate genes characteristic of ESCs. This was corroborated in another study looking at iPSCs generated from fibroblasts (Doi *et al*, 2009), and it was subsequently

shown that this effect is not limited to fibroblast-derived iPSCs (Marchetto *et al.* 2009; Ghosh *et al*, 2010). iPSCs were more similar to their donor cell type than to any other donor cell type, with significant residual gene expression. In addition, these iPSCs also showed incomplete induction of genes important for pluripotency compared to ESCs. These results suggest that differences in gene expression between ESCs and iPSCs may be due to incomplete reprogramming resulting in incomplete silencing of genes expressed in the donor cell. However, this study also pointed out that another possible source of variation could be due to a heterogeneous population of iPSCs, some of which are fully reprogrammed and some of which are not. This would skew the results towards an intermediate state of gene expression.

In order to remove any potentially confounding effects of donor-specific differences in gene expression, Polo *et al* (2010) compared the transcriptional and epigenetic patterns, as well as the differentiation potential, of genetically identical, transgene-free iPSCs derived from different somatic cell types. They found that iPSCs derived from different cell types exhibit different gene expression patterns, and in particular that iPSCs retained gene expression patterns indicative of their cell of origin. iPSCs derived from different cell types were also distinguishable based on their DNA methylation patterns, although the promoter regions of genes that had been shown to be differentially expressed between iPSCs from different origins were not differentially methylated. This suggests that methylation differences are more subtle than gene expression differences, and that other epigenetic mechanisms may be involved. The promoters of differentially expressed genes did show different patterns of histone modification. This has subsequently been shown by several other studies, using iPSCs derived from various different somatic cell types (Kim *et al*, 2011; Ruiz *et al*, 2012).

In contrast to previous studies, a study by Ohi *et al* (2011) found no starting cell type-specific differences between iPSCs; instead, they reported that while some of the differentially expressed genes were specific to one type of somatic cell, most were found to be expressed in all three somatic cell types analysed, suggesting that the epigenetic memory is general and not specific to the starting cell type. They also did not identify any differences in differentiation efficiency between the iPSCs derived from different somatic cell types. Unlike many of the previous studies which identified epigenetic memory in iPSCs, this study attempted to correlate the differences in gene expression between ESCs and iPSCs with epigenetic differences, and showed that differential gene expression between ESCs and iPSCs was due to differences in promoter DNA methylation, with genes that are overexpressed in iPSCs compared to ESCs showing hypomethylation in iPSCs, suggesting that

establishment of DNA methylation during reprogramming may be insufficient. This could not be attributed to differences in expression levels of *de novo* methyl transferases (DNMTs), but the authors did note that the affected genes tended to be isolated from other genes, suggesting that the recruitment of the DNA methylation machinery to these isolated genes may be inefficient.

Several studies have shown a reduced differentiation efficiency of iPSCs compared to ESCs, and have suggested that this may be due to epigenetic memory. A study by Polo *et al* (2010) showed that genetically identical iPSCs derived from different cell types exhibit different differentiation potential: for example, iPSCs derived from fibroblasts were less able to give rise to haematopoietic lineage cells than iPSCs derived from B cells. This was also shown in a study by Kim *et al* (2010), who demonstrated that iPSCs retain a memory of their cell of origin, which translates into differences in differentiation propensity, and that this is due to residual DNA methylation at loci required for haematopoietic differentiation. Importantly, they also showed that the epigenetic memory of these cells can be reset, through subsequent rounds of differentiation and reprogramming, or through treatment with epigenetic-modifying drugs. In some cases, epigenetic memory may be advantageous for generating cells that are difficult to generate through *in vitro* differentiation, such as pancreatic β-cells: for example, Bar-Nur *et al* (2011) showed that iPSCs generated from human pancreatic β-cells maintained open chromatin structure at key β-cell genes together with a unique DNA methylation signature that distinguished them from other PSCs, and that these iPSCs had an increased ability to differentiate into insulin-producing cells *in vivo* and *in vitro*.

Some studies have argued that the main factor influencing the differentiation propensity of iPSCs is not the cell type used for reprogramming, but instead is due to the genetic background of the donor (Kajiwara *et al*, 2012; Shao *et al.,* 2013; Mills et al., 2013). However, a study by Thatava *et al* (2012) found that the differences between iPSC clones from the same donor were greater than the differences between iPSCs from different donors.

Differences between ESCs and iPSCs may be attenuated after continued passaging (Chin *et al*, 2009; Polo *et al*, 2010) and not passed on to the differentiated progeny of the iPSCs (Patterson *et al*, 2012). This is likely to be due to early-passage iPSC lines gradually resolving transcriptional and epigenetic differences, and provides evidence that the reprogramming process may take longer than previously thought. This shows the importance of reporting the passage numbers of cells used in these comparisons. In contrast, other studies have shown that differentially methylated regions in iPSCs are transmitted to differentiated progeny, and cannot be erased through continued passaging (Chan *et al*, 2009; Mikkelsen *et al*, 2008; Hu *et al*, 2010; Lister *et al*, 2011).

It is becoming clear that the heterogeneity and behaviour of stem cells is more complex than had previously been thought, and that iPSCs and ESCs are neither identical nor distinct populations; instead, they seem to overlap, with greater variability within each population than is observed between the populations. This variability is likely to have functional implications for the cells. It may be more useful to consider each cell line in terms of its quality and utility, in order to choose the best cell line for the application. With a view to this end, some researchers have produced a 'scorecard' to evaluate the character of both iPSCs and ESCs, and to predict the quality and utility of any pluripotent cell in a high-throughput manner: a large-scale study by Bock *et al* (2011) used gene expression profiling, DNA methylation mapping, and a high-throughput quantitative differentiation assay to establish genome-wide reference maps for patterns of gene expression and DNA methylation in order to provide a baseline against which comparisons of new ES and iPS cell lines can be made in order to assess their quality and utility.

1.9.5 miRNA Expression

There have been several studies which have investigated differences in miRNA expression between ESCs and iPSCs, but although there has been some overlap in the miRNAs identified, so far there is little evidence for consistent differences between ESCs and iPSCs. Wilson *et al* (2009) found that, although ESCs and iPSCs had very similar miRNA expression profiles, there were subtle differences between the 2 cell types. Specifically, although some stem cell-specific miRNAs (e.g. the miR-302, miR-17 and miR-106 clusters) were highly expressed in both ESCs and iPSCs, others, including the miR-520 and miR-371 clusters, were more highly expressed in ESCs than in iPSCs. In addition, some miRNAs, including miR-886-5p and let-7a, which are associated with differentiated cells, were more highly expressed in iPSCs than ESCs, suggesting that they may be incompletely silenced during reprogramming. This study observed a downward trend in miRNA expression when transitioning from fibroblast to iPSC, but the expression levels of many miRNAs did not decrease in iPSCs to the levels observed in ESCs. However, this study did not report the passage numbers of the cells used, and if they were low, reprogramming may have been incomplete. In a meta-analysis focussing on differences in gene expression between ESCs and iPSCs, Chin *et al* (2009) also carried out miRNA profiling, and found little difference in miRNA expression between ESCs and iPSCs; however a few miRNAs were consistently differentially expressed between ESCs and iPSCs. These included some of the miRNAs identified by Wilson *et al* (2009), suggesting that iPSCs may have a miRNA expression signature unique from that of ESCs. Some of these differences were alleviated

over time in culture. However, another study, using a larger sample group, also carried out miRNA profiling, and found that pluripotent stem cells could be divided into two distinct groups based on their miRNA expression levels, but that this segregation was not dependent on cell type: one group consisted of mainly ESCs with some virally reprogrammed iPSCs, while the other consisted of iPSCs generated in a number of ways, as well as the ESC line H9. These groups differed in the expression of a group of miRNAs including miR-199a* and miR-302a, whose targets include pro-proliferative cyclins and oncogenes, and the p53-regulating miRNAs miR-92 and miR-141 (Neveu *et al*, 2010). This study did not report differences in expression of miR-302, miR-371 or miR-502 clusters between ESCs and iPSCs, as reported by other studies (Chin *et al*, 2009; Wilson *et al*, 2009), although there was some variability in their expression between cell lines. An analysis of miRNA expression in genetically identical miPSCs and mESCs showed reduced expression of the miRNAs encoded in the *Dlk1-Dio3* imprinted region in some iPSCs (Liu *et al*, 2008). In contrast to these results, a large-scale comparison of 49 iPSC lines and 10 ESC lines found no significant differences in miRNA expression between the two cell types (Koyanagi-Aoi *et al*, 2013), and another study found that hierarchical clustering could not segregate ESCs and iPSCs, although further investigation revealed there were some miRNAs that were differentially expressed between ESCs and iPSCs. However, these differences were not consistent with previous studies (Razak *et al*, 2013).

There have been fewer studies investigating differences in the differentiated progeny of ESCs and iPSCs. Wilson *et al* (2009) reported differences in miRNA expression in embryoid bodies derived from ESCs and iPSCs. Christodoulou *et al* (2011) reported differential expression of BMP4 during differentiation to DE, possibly linked to aberrant silencing of the *Dlk1-Dio3* region, but this did not appear to be associated with any reduction in differentiation propensity. The functional significance of differences between ESCs and iPSCs is still unclear, and it is worth noting that although differences in miRNA expression may be statistically significant, they may not be biologically relevant, and this is something that needs to be investigated further before iPSCs can be considered for a clinical application.

1.10 MicroRNA

MicroRNAs (miRNAs) are small (18-25 nucleotides), non-coding RNAs that regulate gene expression post-transcriptionally. Although miRNAs predominantly negatively regulate mRNA expression, a few examples of positive regulation of target genes by miRNAs through activation of

translation have been described (Place *et al.*, 2008; Vasudevan *et al.*, 2007). miRNAs were first identified in studies of *Caenorhabditis elegans*, where they play an important role in regulating the timing of larval development (Lee *et al.*, 1993; Wightman *et al.*, 1993; Reinhart *et al.*, 2000). Since then, over 21,000 miRNAs have been identified in numerous different species (www.mirbase.org), and they are involved in almost every biological process, including development, metabolism, and ageing, as well as in many human diseases, most notably cancer (Slack, 2010). Cells express a miRNA signature characteristic of their lineage and developmental state (Strauss *et al.*, 2006).

1.10.1 Biosynthesis and Action of miRNAs

miRNAs are present in the genome either as independent transcription units or as clusters. Their coding sequences can be either intragenic (i.e. located in the exons or introns of so-called "host" genes) or intergenic. Intragenic miRNAs may be co-transcribed with the host genes or regulated by their own specific promoters. However, the majority of characterised miRNAs are intergenic and orientated antisense to neighbouring genes, suggesting that they are transcribed independently (Bartel, 2004).

The biogenesis of miRNAs is shown in figure 1.10. Most miRNAs are transcribed by RNA polymerase II. Primary miRNA (pri-miRNA) transcripts in the nucleus are recognised and cleaved into precursor miRNA (pre-miRNA) by a complex containing RNase III enzyme Drosha and RNA binding protein DGCR8. Pre-miRNAs are transported into the cytoplasm by Exportin5, where they are processed by DICER into double stranded mature miRNA. Although it was initially thought that one of the strands of the mature duplex, the "guide" strand, was preferentially incorporated into the Argonaute-containing RISC complex to induce degradation or repress translation of target mRNA, while the other strand (designated with a *) was released and degraded, it is now recognised that either one of the strands of the miRNA duplex can have biological activity, and they are more commonly denoted as 3p and 5p. miRNAs recognise and bind specific sequences in target mRNAs, usually in the 3' untranslated region (3'UTR) (Bartel, 2004). Although the number of miRNAs is small compared to that of mRNAs (the human genome is thought to encode around 1,000 miRNAs, while the number of mRNAs is estimated to be around 30,000), a single miRNA is able to target hundreds of mRNAs, and one mRNA may be targeted by several miRNAs. As a result, the potential for fine tuning translation of mRNAs by miRNAs is enormous.

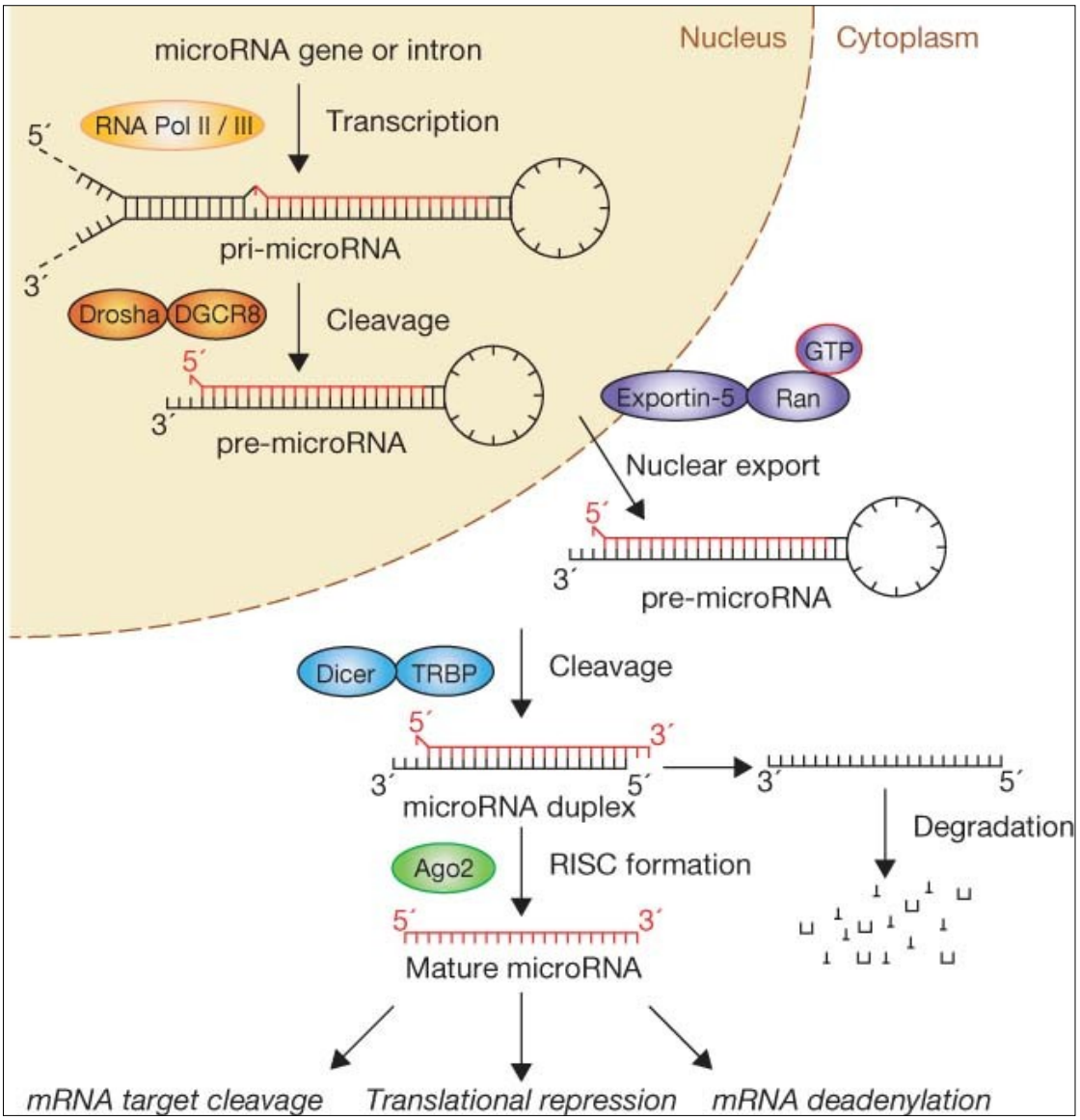

Figure 1.10 miRNA biogenesis. *pri-miRNA transcripts in the nucleus are cleaved into pre-miRNA, which are transported into the cytoplasm by Exportin5, where they are processed by DICER into double stranded mature miRNA. The guide strand is incorporated into the RISC complex to induce degradation or repress translation of target mRNA. Adapted from Winter et al, 2009.*

1.10.2 The Role of miRNAs in Pluripotency

Experiments disrupting the miRNA processing enzymes DICER and DGCR8 show the importance of miRNAs in ESC proliferation, cell cycle and differentiation. DICER-null ESCs show defects in proliferation and differentiation, including prolonged G0 and G1 phases in the cell cycle, failure to express differentiation markers, decreased levels of DNA methylation and DNMTs, and increased telomere recombination and elongation (Kanellopoulou *et al.*, 2005). DGCR8-null ESCs arrest in the

G1 phase, cannot silence self-renewal genes e.g. *Oct4*, *Rex1*, *Nanog* and *Sox2*, and show reduced expression of differentiation markers. They also do not form teratomas (Wang *et al.*, 2007). The differences in phenotype between DICER- and DGCR8-null ESCs suggest that other small RNAs may play a role in ESC proliferation, cell cycle and differentiation.

There is an overall trend of increasing miRNA expression complexity through differentiation, with PSCs expressing relatively few miRNAs, and more differentiated cells expressing more miRNAs. This is in contrast to the pattern of mRNA expression, which has been shown to be restricted with increasing development, with fewer genes expressed in more mature cells (Strauss *et al.*, 2006). Early experiments in human and mouse ESCs showed that a small number of miRNAs are specifically expressed in pluripotent stem cells, and are quickly downregulated upon differentiation. These miRNAs are highly conserved between mice and humans, and include the miR-290 family (miR-290, -291a, 291b, -292, -293, 294 &- 295) in mice (Houbaviy *et al.*, 2003), its human homologues the miR-371 family (miR-371, -372, -373), and the miR-302 cluster (miR-302a, -302b, -302c, -302d and -367) which is present in both in mice and humans (Suh *et al.*, 2004). Many ESC-specific miRNAs are present in the genome in clusters, which are co-transcribed as polycistronic transcripts, suggesting common upstream regulation and co-ordinated expression patterns. They often have similar seed sequences, suggesting common mRNA targets (Martinez & Gregory, 2010). Such a high degree of functional redundancy suggests that they play a critical role in maintaining the pluripotent state of ESCs. The targets of these miRNAs included many genes with promoters bound by the ESC-associated transcription factors NANOG, SOX2 and OCT4, suggesting an important role in pluripotency and self-renewal.

These results were confirmed by larger microarray profiling experiments (Laurent *et al.*, 2008), which also identified an additional cluster of miRNAs on chromosome 19, the miR-520 family, present in ESCs, as well as the miR-17 cluster and its paralogous miRNAs in the miR-106 cluster. Even larger scale experiments were made possible with the advent of next-generation sequencing methods, which allowed high-throughput analysis of both previously identified and novel miRNAs (Morin *et al.*, 2008; Bar *et al.*, 2008). In addition, they showed that miRNAs frequently showed variation from their 'reference' sequences, producing multiple mature miRNAs termed "isomiRs". Much of this variability can be explained by variability in DICER1 or Drosha cleavage positions in the pre-miRNA hairpin structure. The relative abundance of different isomiRs may vary across tissues; however the functional implications of these modifications are still unclear.

miRNAs and Gene Regulation

The core molecular circuitry of stem cells, composed of pluripotency-associated transcription factors, interacts with miRNAs to activate genes involved in pluripotency and self-renewal and suppress genes involved in differentiation. The transcription factors OCT4, SOX2 and NANOG act co-ordinately and are central to the establishment and maintenance of the pluripotent state. They form a positive auto-regulatory circuit in which each factor binds to and activates its own promoter, as well as the promoters of the other pluripotency factors, resulting in maintenance of high levels of expression. They also co-occupy the regulatory regions of hundreds of different genes which are divided into two groups: one group are transcriptionally active in ESCs and include transcription factors, chromatin modifiers and components of stem cell-specific signalling pathways, while the other group of genes are silenced in ESCs due to the co-occupation of their promoters by members of the Polycomb group proteins (transcriptional repressors that regulate lineage choices during development and differentiation), and include a number of transcription factors involved in differentiation and fate commitment (Lakshmipathy *et al*., 2010; Tiscornia & Belmonte, 2010).

miRNAs are an integral part of the gene networks regulated by the pluripotency factors OCT4, SOX2 and NANOG, and help to poise ESCs for rapid and efficient differentiation. OCT4, SOX2, and NANOG occupy the promoters of most ESC-specific miRNAs, as well as another subset of miRNAs that are normally silenced by Polycomb group proteins in ESCs but are upregulated during differentiation into certain lineages (Marson *et al*, 2008). In the undifferentiated state, both OCT4 and OCT4-induced miR-302 directly repress NR2F2 (a transcription factor expressed in differentiated cells) at the transcriptional and post-transcriptional level respectively. During differentiation, NR2F2 directly inhibits OCT4, triggering a positive feedback loop for its own expression (Rosa *et al.*, 2011). Exogenous expression of c-MYC induces the expression of miRNAs involved in negatively regulating genes involved in differentiation, resulting in the attenuation of differentiation, and therefore the maintenance of pluripotency, in these cells (Lin *et al*, 2009).

The ESC-specific miR-290 family of miRNAs (which includes miR-291-3p, miR-294 and miR-295) can also enhance the efficiency of somatic cell reprogramming in murine cells using 3 reprogramming factors, OCT4, SOX2 and KLF4, but not the efficiency of reprogramming with 4 factors, including c-MYC, suggesting that these miRNAs enhance reprogramming through acting downstream of c-MYC (Judson *et al*., 2009). Lentiviral transfection of the miR-302 cluster of miRNAs, which includes miR-302a-d and miR-367, can replace all the previously defined factors to reprogram somatic cells to iPSCs at a higher efficiency (Anokye-Danso *et al*., 2011).

Reprogramming by direct transfection of mature double-stranded miRNAs (Miyoshi *et al*., 2011) has the additional advantage of avoiding viral integration.

"Incoherent feed-forward" regulation, where a transcription factor both activates and represses its target, is important in fine-tuning expression in ESCs. For example, *Lefty1* and *Lefty2* are both actively expressed in ESCs, and are directly occupied at their promoters by OCT4, SOX2, and NANOG. However, these transcription factors also occupy the promoter of the miR-290 cluster, which repress *Lefty1* and *Lefty2*. Therefore ESC transcription factors activate expression of *Lefty1* and *Lefty2*, but also fine-tune its expression by activating a family of miRNAs that repress *Lefty1* and *Lefty2* (Marson *et al.*, 2008).

miRNAs and Cell Cycle Regulation

ESCs are able to undergo infinite and rapid self-renewal due to their unique cell cycle, in particular their shortened G1 phase. Small RNAs have been implicated in ESC proliferation based on the phenotype of DICER- and DGCR8-null ESCs, which show reduced proliferation ability, suggesting that miRNAs normally suppress inhibitors of the G1/S phase transition. Many ESC-specific miRNAs are involved in regulation of the cell cycle, and are often referred to as the ESC cell cycle-regulating (ESCC) miRNAs. Wang *et al*. (2008) used a screening strategy to investigate whether miRNAs introduced into DGCR8-null ESCs could rescue the defects in proliferation. Several miRNAs, including the miR-290 family, were able to do so. These miRNAs shared a similar seed sequence, suggesting that they target the same mRNAs, and using a combination of these miRNAs did not show any further improvement, suggesting that they may be functionally redundant. This study showed that ESC-specific cell cycle regulating miRNAs increase ESC proliferation through repression of cell cycle inhibitors, including p21, CDKN1A, RB1, RBL1, RBL2 and LATS2, resulting in increased G1/S transition.

Card *et al.* (2008) first showed a connection between the pluripotency genes and regulation of the cell cycle, through the activity of the miR-302 cluster of miRNAs. OCT4, SOX2 and NANOG bind to the promoter region of the miR-302 cluster and are required for its expression. Increased expression of miR-302 results in an increase in S-phase cells and a decrease in G-phase cells through regulation of cell cycle regulators including *CyclinD1* and *Cdk4*, while reduced expression has the opposite effect (Card *et al*, 2008). It is now known that many of the stem cell-specific miRNAs,

including the miR-21 cluster, the miR-17 cluster, and the miR-15b cluster, are involved in promoting the G1-S phase transition of the cell cycle through repression of cell cycle inhibitors including *p21cip*, *Rbl2*, and *Lats2*, which increases activity of *cyclinE/cdk2* activity (Gangaraju & Lin, 2009; Sengupta *et al.*, 2009; Tiscornia & Belmonte, 2010). The ESCC miRNAs, which include the miR-302 cluster, the miR-371-373 cluster, and the miR-106a~363 cluster, can also enhance the efficiency of reprogramming of human cells (Subramanyam *et al*., 2011; Liao *et al*., 2011).

Melton *et al.* (2010) showed that the cell cycle in pluripotent stem cells is regulated by two families of miRNAs that play opposing roles: the ESCC miRNAs are expressed in ESCs, and the let-7 family of miRNAs are expressed in somatic cells. In *Dicer*-null ESCs, introduction of ESCC miRNAs rescued the cell cycle defects associated with *Dicer* knockout, while introduction of let-7 miRNAs rescued the differentiation defects. However in wild-type ESCs, introduction of let-7 had no effect on pluripotency, suggesting that the ESCC miRNAs normally antagonise its effects, and the let-7 miRNAs are not able to induce differentiation until the ESCC miRNAs are downregulated. Ectopic expression of LIN28 is also able to enhance reprogramming efficiency through inhibition of let-7 biogenesis (Melton *et al*., 2010).

1.10.3 The Role of miRNAs in Differentiation

The switch from pluripotency to differentiation requires downregulation of pluripotency genes, decreased self-renewal and activation of lineage-specific gene expression. The fact that ESC-specific miRNAs involved in maintaining pluripotency do not rescue the differentiation defects of DICER1- or DGCR8-deficient cells indicates that different miRNAs are involved in differentiation.

One of the most important regulators of differentiation is the let-7 family of miRNAs, members of which are widely expressed in differentiated cells, but not in pluripotent cells (Chen *et al.*, 2007; Morin *et al.*, 2008). As ESCs differentiate, expression of the pluripotency genes are downregulated, resulting in decreased expression of the ESCC family of miRNAs, and therefore decreased expression of LIN28. This decrease in LIN28 expression allows de-repression of let-7 expression. This is enhanced by a positive feedback loop, where let-7 suppresses its own negative regulator LIN28. This interaction between LIN28 and let-7 provides robustness to the differentiation switch. The let-7 family directly target numerous genes involved in promoting the G1/S phase transition such as *Cdk6*, *Cdc25a* and *Ccnd2*, and indirectly target several oncogenes which normally promote

proliferation, such as *Nras*, *Kras*, *Hmga2* and *c-Myc*). Other miRNAs also play a role in cell cycle regulation during differentiation: the miR-15a/16 cluster also represses several genes involved in promoting the G1/S phase transition, including *Cdk6*, *Card10* and *Cdc27* (Ivey & Srivastava, 2010).

The let-7 family is also involved in silencing the self-renewal programme by directly targeting the expression of several pluripotency factors, including c-MYC and LIN28. Let-7 also inhibits the downstream targets of the pluripotency factors in order to stabilise the differentiated state (Melton *et al.*, 2010). Tissue-specific miRNAs often play a role in the switch from pluripotency to differentiation through the repression of pluripotency genes. In mice, miR-134, -296 and -470 repress the pluripotency factors NANOG, OCT4 and SOX2, and miR-200c, -203 and -182 repress SOX2 and KLF4 (Tay *et al.*, 2008). In humans, miR-145 directly represses the expression of pluripotency factors OCT4, SOX2 and KLF4 (Xu *et al.*, 2009).

Some miRNAs associated with pluripotency also play a role in differentiation. The miR-290 cluster of miRNAs, which plays a role in maintenance of self-renewal in stem cells, also plays a role in differentiation through repression of another target gene, *Rbl2*, leading to increased expression of DNMTs and increased methylation of the OCT4 promoter, preventing its expression (Benetti *et al.*, 2008). Ectopic expression of the miR-290 cluster in DICER-deficient ESCs can also rescue differentiation defects. A recent study has shown that the miR-290 cluster promotes mesendoderm formation through repression of *Pax6* and therefore ectoderm formation (Kaspi *et al*, 2013).

As pluripotent cells adopt certain fates, lineage-specific genes are transcriptionally activated and genes that would drive differentiation towards other fates are suppressed. Active repression of lineage-specific genes is important for pluripotency and is largely mediated by Polycomb group proteins. One Polycomb protein, EZH2, is highly expressed in pluripotent cells but is downregulated during skeletal muscle differentiation through the action of miR-214, allowing activation of lineage-specific genes (Juan *et al*, 2009). Similar mechanisms almost certainly exist in differentiation to other lineages.

A number of studies have demonstrated the importance of miRNAs in commitment to a specific embryonic germ layer during development. In mouse ESCs, elevated levels of miR-296 and miR-134 enhanced the expression of genes associated with differentiation toward ectoderm (Tay *et al.*, 2008). In human ESCs, members of the miR-30 family have been shown to regulate the embryonic ectoderm development protein which is important in the development of the neural tube (Song et al., 2011), and miR-125 isoforms have been shown to potentiate early neural specification (Boissart et

al., 2012). The miR-302 family is expressed in undifferentiated ESCs and has been shown to be critical in the maintenance of pluripotency. In addition, a study by Rosa & Brivanlou (2011) demonstrated that this family also plays a role in the earliest stages of differentiation, as in humans, repression of the *Lefty1* and *Lefty2* genes by miR-302 results in increased TGF-β signalling, leading to increased endoderm and mesoderm formation, and decreased neuroectoderm formation. The miR-17-92 cluster plays an important role in mesoderm differentiation and knockout in mouse embryos results in severe developmental defects affecting the heart and lungs (Ventura et al., 2008). Foshay et al. (2009) demonstrated the importance of one member of this cluster, miR-93, in the regulation of both endoderm and mesoderm differentiation through its regulation of *Stat3*. miR-124 is highly expressed in human ESCs and targets two regulators of cytoskeletal rearrangement important in cell migration during mesodermal differentiation(Berardi et al. 2012). Another cluster of pluripotency-associated miRNAs, the miR-290 cluster, regulate *Pax6* expression in early lineage commitment, and knockout of these miRNAs in mice resulted in an increased propensity to form ectoderm at the expense of mesoderm and endoderm (Kaspi et al. 2013).

One early event that is critical in the commitment of differentiating cells to an endodermal or mesodermal lineage is epithelial-to-mesenchymal transition (EMT). Members of the miR-200 family of miRNAs have been implicated in EMT and have been shown to be important in targeting Smad interacting protein-1 (SIP1) (Korpal *et al*, 2008; Chng *et al*, 2010). SIP1 promotes neuroectoderm formation at the expense of mesendoderm and so its inhibition has a positive effect on mesendoderm commitment (Chng *et al*, 2010). In addition, downregulation of miR-200a is important in definitive endoderm formation, as it results in reduced repression of ZEB2, leading to downregulation of E-cadherin and subsequently allowing EMT (Liao *et al*, 2013). Other miR-200 family members also regulate EMT through their targets ZEB1 and ZEB2 (Gregory *et al*, 2008).

miRNAs Involved in Pancreatic Development

Full knockout of the miRNA processing enzyme, DICER1, in murine embryos results in embryonic lethality, whereas partial knockdown allows the development of mice which are histologically normal apart from the pancreas (Morita et al. 2009). Conditional deletion of DICER1 in the developing pancreas results in defects in all pancreatic lineages, but it particularly affects the insulin-secreting β-cells (Lynn et al. 2007). These studies demonstrate the critical role that miRNAs play in pancreatic development.

Numerous miRNAs have been identified as playing a role in pancreatic development (figure 1.12). Several groups (Tzur et al. 2008; Hinton et al. 2010; Kim et al. 2011; Tsai et al. 2010; Porciuncula et al. 2013) have investigated miRNA expression at the DE stage of pancreatic development. Although there is relatively little overlap in the miRNAs identified in these studies, which is likely to be due to the wide range of cell lines and differentiation protocols used, some miRNAs, including miR-375, miR-708 and miR-744, were identified in several of the studies, providing strong evidence for their involvement in DE formation. However these studies have, so far, only identified a few targets for these miRNAs. An early study identified *Timm8a* as a target of miR-375 but the authors were unable to elucidate a function for this pathway in the formation of endoderm (Hinton, et al., 2010). More recently, a study by Liao *et al* (2013) identified a direct relationship between miR-200a and *Sox17*, but not *Foxa2* (another predicted target of miR-200a), and postulated that downregulation of miR-200a promotes endoderm formation through promotion of EMT and de-repression of *Sox17*.

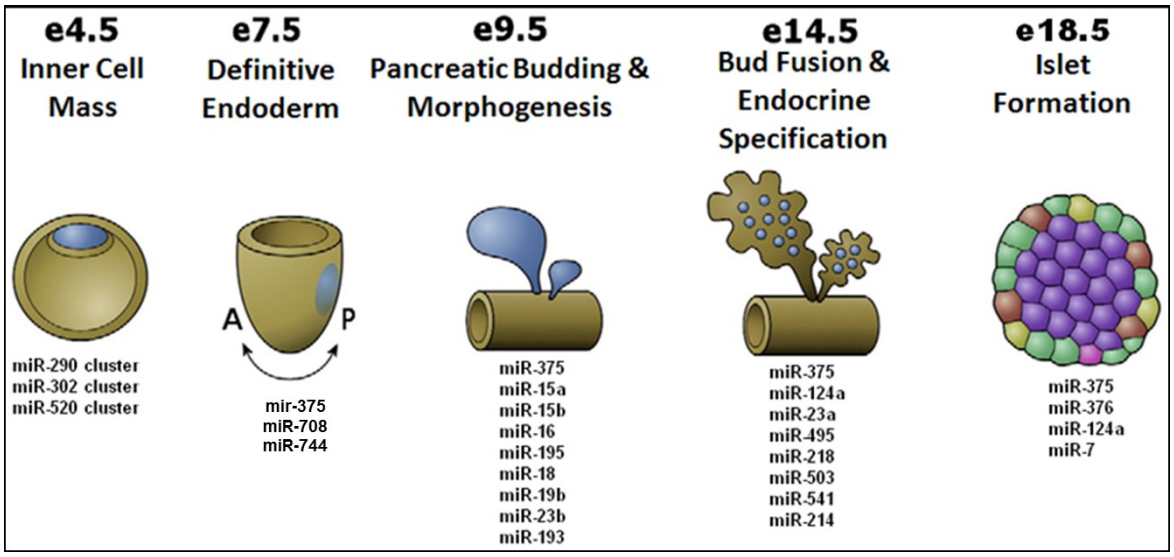

Figure 1.11. miRNAs expressed at the different stages of pancreatic development. *Adapted from Van Hoof et al., 2009.*

The role of miRNAs in the later stages of pancreatic development has been better elucidated. miR-375 was one of the first miRNAs identified in the pancreas (Poy *et al.*, 2004), and remains one of the best characterised. It is expressed in the pancreas and pituitary gland, organs linked by their role in hormone secretion, and expression levels increase throughout pancreas organogenesis. Loss-of-function studies showed that miR-375 is essential for β-cell formation in zebrafish (Kloosterman *et*

al., 2007), and that miR-375 knockout mice have decreased numbers of β-cells and increased numbers of α-cells. This shows the importance of miR-375 in the establishment of normal pancreatic cell mass through the targeting of a group of genes which control cellular growth and proliferation in the developing pancreas (Poy *et al.*, 2009). Further evidence for the involvement of miR-375 in pancreas development is provided by the fact that its expression is regulated by several transcription factors important in pancreatic development and function, including HNF6, INSM1, NGN3, NEUROD1, and PDX-1 (Keller *et al.*, 2007). A recent study showed that overexpression of miR-375 in differentiating embryoid bodies resulted in formation of islet-like clusters, upregulation of endocrine genes including insulin, glucagon, and somatostatin, and release of insulin in response to glucose. Interestingly, the cells produced using this protocol were reported not to be polyhormonal, and the fact that they produced insulin in response to glucose suggests that they are more mature than insulin-expressing cells produced using previously reported (Lahmy *et al*, 2013).

miR-124a shows a similar expression pattern to miR-375, increasing during pancreas development. Its targets include FOXA2, a transcription factor important for β-cell differentiation, pancreatic development, glucose metabolism and insulin secretion. Other predicted targets of miR-124a include NEUROD1 (Baroukh *et al.*, 2007). miR-124a and miR-23b have also been shown to be involved in regulation of HES1 during neuronal development (Wang *et al*, 2010), and might have a similar role in pancreatic development, as HES1 plays an important role in endocrine specification by controlling the numbers of $NGN3^+$ cells.

miR-7 is highly and specifically expressed in the endocrine pancreas where it is co-localised with glucagon and insulin in mature α- and β-cells, respectively. At E12.5, 13.5 and 14.5, miR-7 is co-localised with NGN3-expressing cells, suggesting that it is induced in newly specified endocrine cells. miR-7 expression is also abolished in NGN3 knockout mice (Kredo-Russo *et al.*, 2012). miR-127 and miR-382 showed a similar pattern of expression (Correa-Medina *et al.*, 2009). miR-7 regulates expression of PAX6, a transcription factor important in the development of the endocrine pancreas. This is important since under-expression of PAX6 results in glucose intolerance, as well as having effects on the eye, while over-expression causes eye abnormalities and induces apoptosis in the brain and the endocrine pancreas. Knockdown of miR-7 led to increased expression of PAX6, leading to increased expression of ARX, which is specifically expressed in α-cells, as well as PAX4 and MAFB, which are expressed in β cells. Insulin and glucagon levels were also increased, due to an increase in the numbers of α- and β-cells. There was a decrease in the number of ε-cells, suggesting that miR-7 knockdown acts upstream of PAX6 to promote differentiation of α- and β-

cells at the expense of ε-cells. Conversely, overexpression of miR-7 led to a decrease in PAX6 and downregulation of insulin and MAFB, both direct targets of PAX6, while ghrelin expression was increased (Kredo-Russo *et al.*, 2012).

Several other miRNAs are also expressed during pancreatic development: miR-495 and miR-218 are involved in early pancreas morphogenesis and endocrine specification (Simion *et al*. 2010); miR-23a and miR-23b are co-expressed in the developing pancreas, and repress HNF6 expression (Joglekar *et al.*, 2007); miR-19b is highly expressed in pancreatic progenitor cells and regulates *Neurod1*(Zhang *et al*. 2011); overexpression of miR-26a during pancreatic development increased islet numbers, expression of endocrine genes, and cell proliferation (Fu *et al*. 2013). miR-15a, miR-15b, miR-16 and miR-195a regulate translation of *Ngn3* during pancreas regeneration, and may play a similar role during development, although no direct interaction of these miRNAs with *Ngn3* mRNA could be found.

miRNAs Involved in Pancreas Function

The deletion of DiCER1 in mature β-cells has less effect on islet morphology than deletion in the developing pancreas, although both the number of β-cells and the release of insulin were reduced (Kalis et al. 2011; Melkman-Zehavi et al. 2011) demonstrating the importance of miRNAs for mature β-cell maintenance and function through regulation of both insulin biosynthesis and exocytosis.

miR-375 is required for normal glucose homeostasis through its control of insulin secretion in mature β-cells. It negatively regulates *Mtpn* expression (Poy et al., 2004), which may control insulin exocytosis either through the actin network, or through NFκB signalling. miR-375 also regulates *Pdk1* expression, which is involved in regulation of insulin gene expression, as well as cell signalling downstream of the insulin receptor (El Ouaamari et al., 2008), showing that miR-375 regulates β-cell function through at least two mechanisms. Knockout of miR-375 also results in reduced β-cell proliferation, and a number of predicted targets of miR-375 are genes regulating cellular growth (Poy et al., 2009). Other miRNAs involved in β-cell proliferation include miR-29a (Bagge et al., 2012) and miR-7, which inhibits components of the mTOR signalling pathway(Wang et al., 2013).

Similarly to miR-375, miR-124a also regulates *Mtpn* expression (Krek et al. 2005), but additionally regulates insulin secretion through its regulation of *Foxa2*, downstream targets of which include

components of the β-cell K_{ATP} channel, which is involved in insulin exocytosis. miR-124a is also predicted to directly target other components of the exocytotic machinery, including *Rab27a*, *Snap25*, *Rab3a*, *Syn1*, and *Noc2* (Lovis et al. 2008), as well as *Creb-1, Isl-1,* and *Vamp-3*, which are part of the transcriptional network downstream of the insulin receptor (Baroukh et al., 2007). miR-124, miR-29a and miR-29b have each been shown to be important regulators of the pyruvate/lactate (monocarboxylate) transporter MCT1 in β-cells. This transporter is a member of a group of housekeeping genes which are usually highly selectively suppressed in these cells (Sekine et al. 2012). However, levels of miR-124a have been reported to be low or undetectable in mouse and human islets, suggesting that miR-29a and miR-29b may be the most important in this context (Pullen et al. 2011), although the levels of miR-124a may increase during aging, at least in the rat (Tugay & Regazzi, 2013).

miR-9 regulates insulin secretion through inhibition of both *Oc-2*, leading to increased levels of granuphilin (Plaisance et al., 2006), and *Sirt1* (Ramachandran et al., 2011) . Similarly, mir-96 also regulates insulin secretion in β-cells through increased levels of granuphilin, but through a pathway that does not involve *Oc-2*. Overexpression of miR-96 also resulted in decreased levels of *Noc2*, a protein which plays a positive role in insulin exocytosis (Lovis et al., 2008). Several miRNAs, including miR-7 (Melkman-Zehavi et al., 2011), miR-30d (Tang et al., 2009; Zhao et al., 2012), miR-133a (Fred et al., 2010), and miR-15a (Sun et al., 2011), regulate β-cell function through the control of insulin gene expression and biosynthesis. The expression of many of these miRNAs is controlled by glucose concentration, demonstrating how these miRNAs form part of the regulatory network controlling β-cell function.

Recent studies have demonstrated the importance of miRNAs in maintaining β-cell function, as their dysregulation in T2DM contributes to a loss of pancreatic β-cell phenotype (Talchai et al., 2012). Locke et al (2014) recently demonstrated that miR-187 is strongly up-regulated in islets from T2DM human donors compared to healthy control donors, leading to reduced expression of the *Hipk3* gene, and a subsequent reduction in glucose-stimulated insulin secretion. Another study comparing miRNA expression between T2DM human donors and healthy control donors demonstrated aberrant repression of a cluster of miRNAs expressed in the *Dlk1-Meg3* gene locus, including miR-376a, miR-432 and miR-495, in T2DM, and subsequently demonstrated that this downregulation was linked to increased β-cell apoptosis through the gene targets *Iapp* and *Tp53inp1* (Kameswaran et al. 2014).

In addition to its importance during early lineage commitment, EMT is also important for maintaining mature pancreas function. Isolated human pancreatic islet cells undergo EMT in culture, a process which coincides with a loss of function including a lack of insulin secretion. Several proteins, including VIMENTIN, N-CADHERIN, β-CATENIN, SNAIL1, SNAIL2 and TWIST, are overexpressed during EMT and have been shown to play a role in this process (Ouziel-Yahalom et al., 2006). Interestingly, analysis of human pancreatic islets showed that, although mesenchymal proteins are not detected, mesenchymal gene transcripts are present at very high levels. This suggests that EMT is controlled by post-transcriptional regulation of these mesenchymal genes, and a study demonstrated that members of the miR-30 family negatively regulate transcription of mesenchymal genes to maintain the epithelial phenotype of islet endocrine cells. During EMT, expression of the miR-30 family is reduced, allowing increased transcription of mesenchymal genes and an alteration in cell phenotype (Joglekar et al, 2009). While this process is important for in vitro culture of isolated islets, it is also likely to contribute to the maintenance of the mature, islet cell phenotype. DiCER1 deletion in mature β-cells has a smaller effect on islet morphology than deletion in the developing pancreas, although both the number of β-cells and insulin release were reduced (Kalis et al., 2011; Melkman-Zehavi et al., 2011), demonstrating the importance of miRNAs for mature β-cell maintenance and function through regulation of both insulin biosynthesis and exocytosis, as well as cell signalling downstream of the insulin receptor .

1.11 Hypothesis & Aims

The aim of this project is to identify whether iPSCs are a viable alternative to ESCs for generating β-cells *in vitro* for cell replacement therapy to treat type 1 diabetes. The differentiation potential of iPSCs and ESCs to give rise to DE (the first stage in differentiation towards a pancreatic lineage) *in vitro* will be compared, and mechanisms important in the differentiation of ESCs and iPSCs to DE will be investigated, focussing on the involvement of miRNAs in this process.

Chapter 2
Materials & Methods

Contents

List Of Figures

List of Tables

Chapter 2: Materials & Methods

2.1 Generation of Induced Pluripotent Stem Cell (iPSC) Lines

Although there are a number of ways of generating iPSCs (as previously described in chapter 1), for the purposes of this project new iPSC lines were generated using retroviral transduction of four reprogramming factors (Oct4, Sox2, Klf4 and c-Myc) as described by Takahashi & Yamanaka (2007) (see figure 2.1 for an overview of the process). While this approach results in integration of the viral transgenes into the genome of the iPSCs generated, rendering the cells unsuitable for clinical applications, it also has one of the highest efficiencies of reprogramming.

2.1.1 Preparation of Cell Lines

Fibroblast Cell Culture: 2 human fibroblast cell lines, human foreskin fibroblasts (HFF) and neonatal human dermal fibroblasts (HDFn) were used for reprogramming into iPSCs. These cell lines were obtained from the UK Stem Cell Bank (www.ukstemcellbank.org.uk). Cryopreserved cells were thawed in a 37°C water bath and transferred into warmed fibroblast medium (table 2.1) in a cell culture flask and incubated at 37°C, 5% CO_2. When the cells reached 80-90% confluence, the cells were passaged. The medium was aspirated and the cells washed once with PBS. 0.25% trypsin/1mM EDTA was added to the flask, which was then incubated until the cells detached from the surface of the flask. The trypsin was inactivated by the addition of medium, and the cells were then transferred into another flask. When making frozen stocks, the cells were counted after trypsinisation, using 4% trypan blue solution to ensure that only live cells were included in the count. The cells were spun down and the medium removed. $1x10^6$ cells were cryopreserved in 1ml of cryopreservation solution (10% DMSO in foetal calf serum).

Fibroblast Medium		
Reagent	*Volume (ml)*	*Final Concentration*
DMEM	440	
Foetal Calf Serum	50	10%
200mM L-Glutamine	5	2mM
1% Penicillin-Streptomycin	5	1mg/ml
Total	**500**	

Table 2.1. Volumes and final concentrations of each component of fibroblast medium.

293FT Cell Culture: The 293FT cell line is supplied with the ViraPower™ Lentiviral Expression System, and is useful for generating high-titre lentivirus. It is a fast-growing, highly transfectable clone derived from human embryonal kidney cells transformed with the SV40 large T antigen from the pCMVSPORT6Tag.neo plasmid, controlled by the human cytomegalovirus promoter which is high-level and constitutive. Presence of the SV40 large T antigen allows very high levels of protein to be expressed from vectors containing the SV40 origin, allowing maximal production of a replication-incompetent HIV-1-based lentivirus that is used to deliver a mouse retrovirus receptor (Slc7a1) to the fibroblasts, which then allows the retroviruses carrying the pluripotency factors needed for reprogramming to infect the fibroblasts (figure 2.1). Cryopreserved cells were thawed in a 37°C water bath and transferred into warmed 293FT medium (table 2.2) in a T25 cell culture flask and incubated at 37°C, 5% CO_2. The next day, the medium was changed to 293FT medium containing Geneticin (50µg/ml), which is important for selecting the pCMVSPORT6Tag.neo plasmid-containing cells, as they contain a resistance gene. The cells were maintained in this medium until confluent, when they were passaged as described above. Frozen stocks were made as described above.

293FT Medium		
Reagent	*Volume (ml)*	*Final Concentration*
DMEM	430	
Foetal Calf Serum	50	10%
100x Non-Essential Amino Acids	5	0.1mM
200mM L-Glutamine	5	2mM
100mM Sodium Pyruvate	5	1mM
1% Penicillin-Streptomycin	5	1mg/ml
Total	**500**	

Table 2.2. Volumes and final concentrations of each component of 293FT medium.

Plat-E Cell Culture: The Plat-E cell line is a retrovirus packaging line based on the 293FT cell line. It allows high and stable expression of viral structural proteins (Morita *et al.*, 2000). In iPSC generation, it is used to package retroviruses containing pMX plasmid DNA (prepared as described in section 2.1.2) encoding the 4 transcription factors needed for reprogramming to pluripotency (OCT4, KLF4, SOX2 and c-MYC), as well as one encoding GFP to enable monitoring of transfection efficiency. These retroviruses are capable of infecting fibroblasts expressing the mouse retrovirus receptor Slc7a1. Cryopreserved cells were thawed in a 37°C water bath and transferred into warmed FP medium (table 2.3) in a T25 cell culture flask and incubated at 37°C, 5% CO_2. The next day, the medium was changed to FP medium containing Blasticidin (1µg/ml) and Puromycin (0.1µg/ml), as these cells include a resistance gene that allows antibiotic selection of the cells containing the packaging constructs. The cells were maintained in this medium until confluent, when they were passaged as described above. Frozen stocks were made as described above.

FP Medium		
Reagent	*Volume (ml)*	*Final Concentration*
DMEM	447.5	
Foetal Calf Serum	50	10%
1% Penicillin-Streptomycin	2.5	0.5mg/ml
Total	**500**	

Table 2.3. Volumes and final concentrations of each component of FP medium

2.1.2 Preparation of pMX Plasmids

Glycerol stocks of *E. coli* containing plasmids for each of the reprogramming factors and GFP (Takahashi & Yamanaka, 2006) were obtained from AddGene (www.addgene.org).

Plate Preparation: Luria broth (LB) was melted in the microwave using low power and then incubated in a 50°C water bath until cool. 40µg Ampicillin with sodium salt was added to 400ml LB to give a final concentration of 0.4µg/ml. 20ml of LB was poured into a 100mm dish, which was then left to cool and set. One plate was prepared for each plasmid.

Recovery of Single Colonies from Stored Cultures: A sterile wire loop was used to streak an inoculum of bacteria from a glycerol stock onto the LB plate. The plate was incubated upside down at 37°C for 12 hours.

Preparation of Bacteria for Plasmid Preparations: A starter colony was prepared by inoculating a single colony from the freshly streaked plate into 10ml LB medium containing 0.4µg/ml Ampicillin. The starter culture was grown up for 8 hours with vigorous shaking. The starter culture was then diluted 1:500 in LB medium containing 0.4µg/ml Ampicillin, and the culture was grown up overnight at 37°C with vigorous shaking.

Isolation of Plasmid DNA: The SNAP midiPrep Kit was used to lyse, precipitate and isolate the plasmid DNA. The bacterial culture was harvested by centrifugation at 6000 x g for 15min at 4°C and the medium was poured off. The cell pellet was resuspended in 4ml of resuspension buffer, then 4ml of lysis solution was added. The solution was mixed by inverting 5-6 times and incubated at room temperature for 3min. 4ml of precipitation salt was added, and the solution was mixed by inverting 6-8 times before being placed on ice for 5min. The tubes were inverted several times to ensure even formation of precipitate. A SNAP MidiPrep Column A (filtering) was placed inside a 50ml conical tube, and the solution was transferred onto this column and centrifuged for 5min at 3000 x g. The column was discarded, and 12ml of binding buffer was added to the filtrate, which was transferred to a SNAP MidiPrep Column B (binding) and centrifuged at 1000 x g for 1min. The flow-through was discarded. 5ml of wash buffer was added, and the tube was centrifuged at 2000 x g for 1min. The flow-through was discarded. 5ml of final wash buffer was added and the tube was centrifuged at 2000 x g for 2min. 10ml of final wash buffer was added to the column and the tube was centrifuged at 2000 x g for 2min. The flow-through was discarded. The tube was then centrifuged at 4000 x g for 5min to dry the column. The column was then transferred to a new 50ml conical tube and 750µl of sterile water was added directly to the column. The column was incubated at room temperature for 3min, and then centrifuged at 4000 x g for 5min. The column was discarded and the purified plasmid DNA was stored at -20°C until needed. The DNA was quantified using the Nanodrop spectrophotometer.

2.1.3 Reprogramming of Fibroblasts

Day 1: Seeding of 293FT Cells: Cells were seeded to three 100mm dishes at a density of $4x10^6$ cells per dish in 10ml 293FT medium and incubated overnight at 37°C, 5% CO_2.

Day 2: Transfection of 293FT Cells: 9µg ViraPower™ packaging mix (supplied as part of the ViraPower™ Lentiviral Expression System), containing plasmids that supply the cellular machinery needed to produce lentivirus, and 3µg of pLenti6-mSlc7a1 plasmid were added to 1.5ml Opti-MEM low serum medium and mixed. 36µl of Lipofectamine 2000, which improves transfection efficiency, were added to 1.5ml Opti-MEM and incubated at room temperature for 5 minutes. The DNA and lipofectamine were combined and incubated at room temperature for 20 minutes. The medium on each dish of the 293FT cells was replaced and 3ml of the DNA mixture was added drop-wise to the cells, which were incubated at 37°C, 5% CO_2 overnight. The medium was changed after 24 hours.

Day 3: Seeding Fibroblasts: Fibroblasts were seeded at a density of $1.5x10^5$ cells per well in four 6-well plates for each cell line.

Day 4: Lentiviral Transduction of Fibroblasts: The supernatant containing the lentivirus was aspirated from the 293FT cells and filtered through a 0.45µm cellulose acetate filter to remove any cells that may have been aspirated with the supernatant. The supernatant was diluted 1:2 with Opti-MEM, and 4µg/ml Polybrene (a cationic polymer used to increase the efficiency of infection) was added. The lentivirus-containing supernatant was added to 20 wells of each fibroblast line, and the remaining 4 wells were left untreated as a negative control. The cells were incubated at 37°C, 5% CO_2 and after 5-6 hours the medium was replaced with 2ml of FP media per well. The medium was replaced again the next day.

Day 9: Seeding of PlatE Cells: Cells were seeded to five 100mm dishes at a density of $3.6x10^6$ cells per dish in 10ml FP medium.

Day 10: Transfection of PlatE Cells: 27µl Fugene6 transfection reagent (a non-liposomal reagent that transfects DNA into cells with high efficiency and low toxicity) was added to 0.3ml Opti-MEM and incubated for 5 minutes at room temperature in a 1.5ml tube for each plasmid. 9µg of plasmid pMX DNA encoding OCT4, SOX2, KLF4, c-MYC and GFP was

added drop-wise to each tube containing the Fugene 6/Opti-MEM mix and incubated at room temperature for 15 minutes. Each plasmid was added drop-wise to the PlatE cells, which were incubated overnight at 37°C, 5% CO_2. The medium was replaced the next day.

Day 11: Seeding Transduced Fibroblasts: The fibroblasts transduced with the mSlc receptor (HDFn/mSlc and HFF/mSlc) and the negative controls (HDFn/neg and HFF/neg) were seeded to 4x 6-well plates at a density of $1.5x10^5$ cells per well.

Day 12: Retroviral Transduction of Fibroblasts: The supernatant containing the retrovirus was removed from each dish of PlatE cells and filtered through a 0.45µm cellulose acetate filter. 4µg/ml Polybrene was added, and the 4 retrovirus-containing mixtures were combined in equal parts. The supernatant from the GFP-infected cells was kept separate. The medium was removed from the fibroblasts and replaced with 1ml of retrovirus-containing supernatant per well. Some wells were left untreated as a negative control, while some were treated with GFP-containing supernatant. The cells were incubated at 37°C, 5% CO_2 and after 5-6 hours the medium was replaced with 2ml of FP medium per well. The medium was changed again the next day.

Day 15: The medium on the cells was replaced with FP media containing 5µM valproic acid (VPA), which is a histone deacetylase inhibitor that improves the efficiency of reprogramming (Huangfu *et al.*, 2008).

Day 16: Mitomycin C-inactivated iSNL feeder cells (see section 2.2.1) were seeded to 6 gelatine-coated 6-well plates at a density of $1.5x10^6$ cells per well.

Day 17: The potential iPS cells (piPSCs) and uninfected controls were seeded to iSNL feeder plates at a density of $1\text{-}2x10^4$ cells per well, and RNA was isolated. This was later used to check for exogenous transgene expression by qRT-PCR.

The medium on the piPSCs and uninfected controls was replaced every other day with knockout serum replacement medium (table 2.5) containing 5µM VPA until the appearance of colonies 2-3 weeks after transfection. When large enough, the colonies were manually passaged onto new feeder cells and maintained as described in section 2.2.

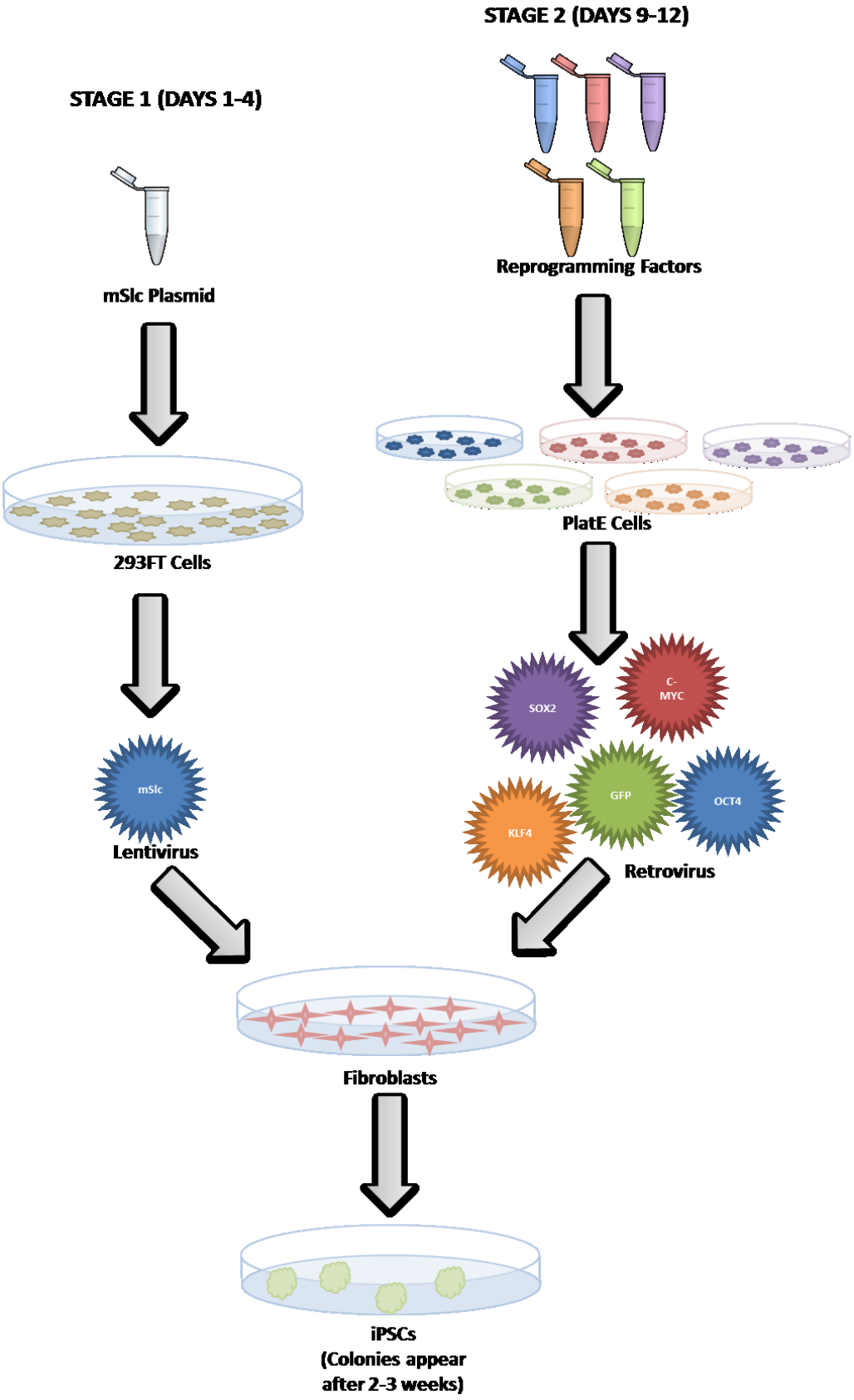

Figure 2 1.Overview of the process of reprogramming of fibroblasts to induced pluripotent stem cells.

2.2. Maintenance of Pluripotent Stem Cells in Culture

2.2.1 SNL Feeder Cell Culture & Inactivation

Feeder cell layers consist of adherent growth-arrested but viable cells. These cells act as a substratum on which stem cells can be grown in a co-culture system. Stem cells are dependent on contact with a feeder layer for survival and expansion, as the feeder layer provides an extracellular matrix and matrix-associated factors, and also secrete nutrients and signalling molecules into the medium. In this case, SNL cells, which are immortalised cells clonally derived from a mouse fibroblast cell line transformed with neomycin resistance and murine LIF genes, were used as feeders. These cells are frequently used as feeders when culturing mouse pluripotent stem cells, as they constitutively express LIF which these cells require for their maintenance in culture, but can also support the growth of human pluripotent stem cells (Pan *et al.*, 2010). As these cells are immortal, they are easy to culture and can be expanded over many passages, unlike human and mouse embryonic fibroblasts, which are an alternative source of feeder cells.

Cell culture dishes were coated with 0.1% gelatine, which improves adherence of cells, and incubated at 37°C for an hour. Excess gelatine was then removed. Cryopreserved cells were thawed in a 37°C water bath and transferred in 10ml warmed SNL medium (table 2.4) onto a gelatine-coated 100mm dish and incubated at 37°C, 5% CO_2 until confluent, when they were passaged. The medium was aspirated and the cells washed once with PBS. 0.25% trypsin/1mM EDTA was added to the dish, which was incubated until the cells detached from the surface of the flask. The trypsin was inactivated by the addition of medium. The volume was adjusted to 200ml by adding medium and 10ml of the cell suspension was transferred to each gelatine-coated 100mm dish and incubated at 37°C, 5% CO_2 until confluent.

SNL Medium		
Reagent	*Volume (ml)*	*Final Concentration*
DMEM	460	
Foetal Calf Serum	35	7%
200mM L-Glutamine	5	2mM
Total	**500**	

Table 2.4. Volumes and final concentrations of each component of SNL medium.

16 of these dishes were mitotically inactivated using Mitomycin-C (once inactivated, they are referred to as iSNL cells). 0.3ml of 0.4mg/ml Mitomycin-C solution was added to each dish, giving a final concentration of 12µg/ml. The dish was incubated for 2 hours 15 minutes at 37°C, 5% CO_2. The medium was removed and the dishes washed with PBS twice. The cells were trypsinised as described above and the cells were then resuspended in 2.5 ml media. The cells were counted, using 4% trypan blue solution to ensure that only live cells were included in the count, and then centrifuged and the medium aspirated. Cells were frozen at a concentration of $1.5x10^6$ cells in 1ml of freezing solution (10% DMSO in foetal calf serum) and stored at -80°C. The remaining 4 dishes were not inactivated and the SNL cells were cryopreserved as replacement stocks.

Cryopreserved iSNL cells were thawed into 10-12ml warmed SNL medium and mixed gently by pipetting up and down. 1ml of this cell suspension was added to a gelatine-coated IVF dish and the cells were incubated overnight at 37°C, 5% CO_2. The next day the medium was replaced with knockout serum replacement medium in order to optimise attachment of stem cell colonies.

2.2.2 Maintenance of Pluripotent Stem Cells on Feeder Cells

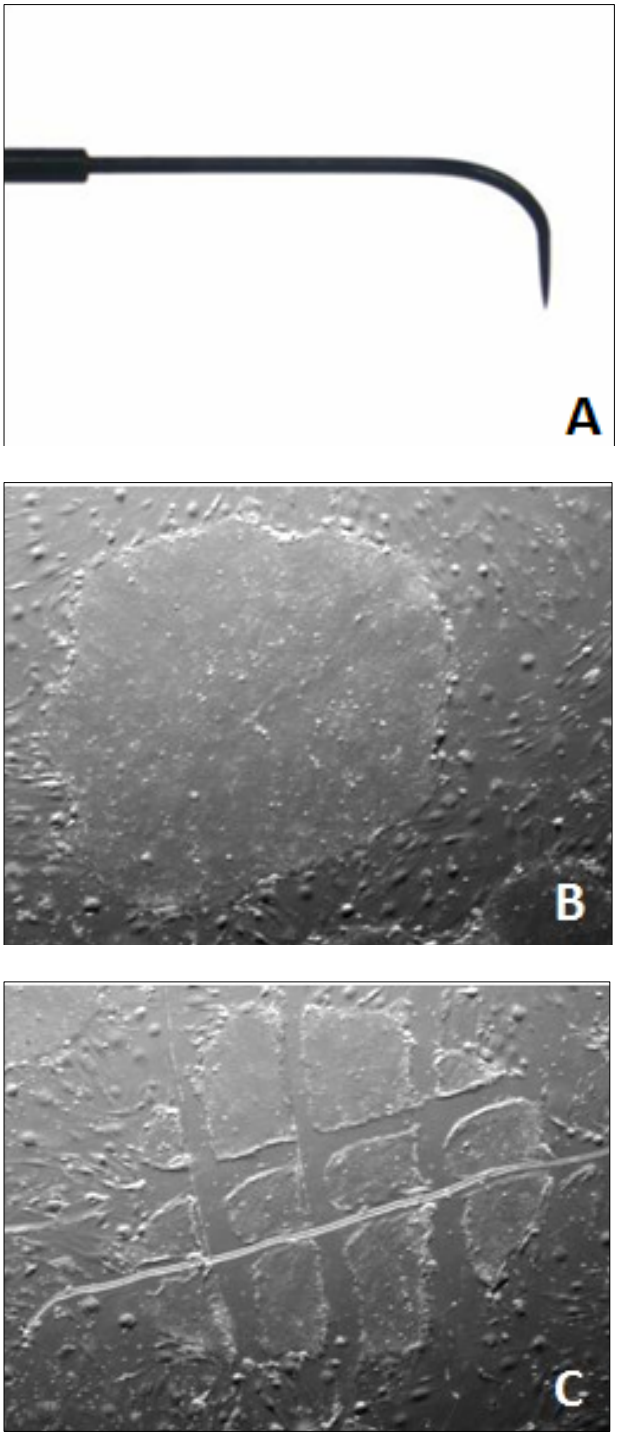

Figure 2.2. Passage of stem cell colonies. *(A) Hooked cutting tool used for manual dissection of colonies. (B) Undifferentiated stem cell colony grown on iSNL feeders. (C) Undifferentiated colony manually dissected into approximately equal pieces for passage.*

Stem cells were thawed into warmed knockout serum replacement (KOSR) medium (table 2.5) and transferred to an IVF dish containing iSNL feeder cells. Stem cells start to differentiate around the outside and from the centre of the colonies as they are maintained in culture. To preserve the stem cells in an undifferentiated state, undifferentiated sections of the colony were manually excised using a hooked cutting tool (figure 2.2) under a dissecting microscope and transferred to a new layer of iSNL cells, evenly spaced to avoid the colonies merging. The plate was transferred very carefully to the incubator (37°C, 5% CO_2). Once the colonies had adhered (after about 24 hours), the medium was changed every 2-3 days.

Knockout Serum Replacement (KOSR) Medium		
Reagent	*Volume (ml)*	*Final Concentration*
Knockout DMEM	383.8	
Knockout serum replacement	100	20%
200mM L-glutamine	5	2mM
0.1mM non-essential amino acids	5	1%
β-mercaptoethanol	1	0.1mM
bFGF	0.2	4ng/ml
1% Penicillin-streptomycin	5	1mg/ml
Total	**500**	

Table 2.5. Volumes and final concentrations of each component of KOSR medium.

2.2.3 Preparation of Matrigel™-Coated Plates

Matrigel™ is a soluble basement membrane extract of the mouse Engelbreth-Holm-Swarm tumour that gels at room temperature, forming a reconstituted basement membrane which can be used to maintain pluripotent stem cells without a feeder cell layer. The major components of Matrigel™ are laminin, collagen IV, entactin, and heparan sulfate proteoglycan. Growth factors, collagenases, plasminogen activators and other undefined components are also present in Matrigel™.

Matrigel™ was thawed at room temperature or overnight at 4°C, and then aliquoted into the dilution factor specified by the manufacturer. An aliquot was thawed on ice and added to 25ml of cold DMEM medium to coat four 6-well plates (1ml/well). All pipette tips, tubes and plates were kept at 4°C prior to use in order to prevent the Matrigel™ from adhering to them.

Diluted Matrigel™ was added immediately to the plates, ensuring that the surface of the wells was fully coated, and left at room temperature for one hour prior to use. Coated plates could be kept for 7 days at 4°C. Matrigel™ was removed immediately before use and 2ml mTeSR-1™ medium were added to each well.

2.2.4 Maintenance of Pluripotent Stem Cells on Matrigel™

Undifferentiated hESCs were maintained in mTeSR-1™ medium, which is a defined, serum-free medium developed for stem cell research, supplemented with 100ml mTeSR™1 5X Supplement per 400ml. Cells were incubated at 37°C, 5% CO_2 and the media was changed every 2-3 days. The hESC colonies were passaged every 5-7 days by manual dissection using a hooked tool (figure 2.2) under a dissecting microscope and transferred onto a new Matrigel™-coated 6-well plate. Undifferentiated colonies were cut and fragments were re-plated into 2-3 wells for each colony.

2.2.5 Cryopreservation of Stem Cells

Pluripotent stem cells were stored at -196°C in liquid nitrogen as vitrified cells in open-pulled straws. Vitrification accompanied by rapid freezing and thawing avoids the formation of ice crystals that could be potentially harmful to the cells. The vitrification protocol involves the stepwise exposure of colony fragments to two vitrification solutions with increasing cryoprotectant concentration; in this case DMSO and ethylene glycol. Vitrification solution 2 contains sucrose as an osmotic buffer to protect the colony fragments from the cryopreservant for long-term storage. Direct immersion of open-pulled straws containing the colony fragments in small droplets of vitrification solution allows rapid cooling rates, which is required to avoid the formation of ice crystals.

The holding medium (table 2.6) was prepared aseptically and used to prepare VS1 and VS2 (tables 2.7 & 2.8, respectively) which were sterile filtered through a 0.2µM filter, aliquoted and stored at -20°C. Each thawed aliquot was used once and then discarded. VS1 and VS2 were warmed to 37°C and a 200µl drop of VS1 and two 100µl drops of VS2 were placed on a sterile culture dish. Under the dissecting microscope, small colony fragments were

mechanically dissected as described previously and approximately 4-6 fragments were transferred to the VS1 drop in a minimum volume of medium (approximately 4-5µl) for 1min, before transferring the colony fragments to the first drop of VS2, then to the second drop after 25 seconds. The cells were recovered in a minimum volume (4-5µl) as a drop on the culture dish. The open-pulled straw was touched to the drop, which was aspirated by capillary action. The straw containing the cells was then plunged into a 5ml cryovial which had been previously pierced and immersed into the liquid nitrogen. The cryovial was then transferred to a liquid nitrogen container for long-term storage.

Holding Medium		
Reagent	*Volume (ml)*	*Final Concentration*
DMEM	86.5	
Foetal calf serum	10	10%
L-glutamine	1	2mM
1M HEPES Buffer	2.5	25mM
Total	**100**	

Table 2.6. Volumes and final concentrations of each component of holding medium.

Vitrification Solution 1 (VS1)		
Reagent	*Volume (ml)*	*Concentration*
Holding medium	32	80%
DMSO	4	10%
Ethylene glycol	4	10%
Total	**40**	

Table 2.7. Volumes and final concentrations of each component of vitrification solution 1

Vitrification Solution (VS2)		
Reagent	*Volume (ml)*	*Concentration*
Holding medium	To 40ml	80%
DMSO	4	10%
Ethylene glycol	4	10%
Sucrose	6.84g	
Total	**40**	

Table 2.8. Volumes and final concentrations of each component of vitrification solution 2.

2.3. *In Vitro* Differentiation of Stem Cells

2.3.1 Differentiation to Definitive Endoderm

The formation of definitive endoderm (DE) is a key step in the differentiation of pluripotent stem cells towards a pancreatic lineage (Courtney et al., 2010). Given that the successful generation of pancreatic progenitors and subsequently the development of more mature hormone-expressing cells is critically dependent on this stage, we have focussed our experiments on the generation of definitive endoderm. In order to optimise the yield of definitive endoderm, I have tested three different protocols.

<table>
<tr><td colspan="8">Stage 1
Formation of definitive endoderm</td></tr>
<tr><td>Day</td><td colspan="2">1</td><td colspan="2">2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td></tr>
<tr><td>D’Amour</td><td colspan="2">RPMI
L-glutamine (2mM)
Activin A (100ng/ml)
Wnt3A (25ng/ml)</td><td colspan="3">RPMI
L-glutamine (2mM)
Activin A (100ng/ml)
Foetal Calf Serum (0.2%)</td><td colspan="4"></td></tr>
<tr><td>Borowiak</td><td colspan="7">RPMI
L-glutamine (2mM)
FCS (0.2%)
IDE1 (400nM)</td><td></td></tr>
<tr><td>Nostro</td><td colspan="6">RPMI
Wnt3A (3ng/ml)
Activin A (100ng/ml)
BMP4 (5ng/ml)
VEGF (20ng/ml)
bFGF (2.5ng/ml)</td><td colspan="4"></td></tr>
</table>

Table 2.9. Reagents and concentrations used to direct differentiation of stem cells to definitive endoderm using protocols developed by D’Amour et al, 2005; Nostro et al, 2011; or Borowiak et al, 2009.

The use of small molecules in directed differentiation protocols is of great interest, as they can be chemically synthesised and purified, allowing a greater degree of standardisation than recombinant growth factors. In addition, they are often cheaper to produce. A study by Borowiak *et al.* (2009) described the use of the small molecules IDE1 and IDE2 to improve the efficiency of DE formation in comparison to an Activin A-based protocol (D’Amour et al, 2005). As a result, I have investigated the effect of IDE1 on the efficiency of DE formation in iPSC and ESC lines. In 2011, Nostro *et al.* reported an optimised protocol for the differentiation of pluripotent stem cells into DE, and I have also investigated the efficiency of this protocol for DE formation in iPSC and ESC lines.

Stem cells were maintained in culture as previously described in section 2.2. Before the start of the differentiation protocol on day 1, RNA from colonies containing undifferentiated cells was taken. Colonies undergoing differentiation were washed with PBS before being exposed to either the D'Amour protocol, the Nostro protocol, or the Borowial protocol according to the time scale shown in table 2.9. The medium was prepared aseptically as described in table 2.9 and sterile filtered through a 0.2µm filter. Cells that were allowed to spontaneously differentiate over the same timescale were also included in the comparison as a control, and were treated with RPMI only. Cells were harvested at the end of stage 1 and analysed as described in section 2.4.

2.4. Characterisation of iPS Cells

2.4.1 Reverse Transcription-Polymerase Chain Reaction (RT-PCR)

RT-PCR for the analysis of mRNA levels is a combination of 3 steps: i) reverse transcription of RNA into cDNA, ii) amplification of cDNA using the polymerase chain reaction, and iii) detection and quantification of amplification products. Traditional methods relied on end-point analysis, where the cDNA is amplified for a specified number of cycles, and then the products visualised through gel electrophoresis. However, there are several disadvantages to this method, including poor precision and low sensitivity. More recently, the advent of real-time PCR (qPCR) allows quantitative analysis of RNA in real time, using fluorescent reporter dyes to combine the amplification and detection stages of the PCR reaction. In this case, SYBRGreen, a double-stranded DNA-binding dye which fluoresces when it intercalates into a double stranded DNA molecule, was used. The increase in fluorescent signal after each PCR cycle is proportional to the amount of DNA produced. Intercalating dyes are inexpensive and simple to use, and, because they are not sequence specific, they can be used in any reaction. However, since SYBRGreen binds indiscriminately to double-stranded DNA, primer-dimer artefacts and amplification errors can contribute to the detected fluorescence. In addition, multiplexing (analysis of several different genes in one PCR reaction) is not possible using this system.

Isolation of Total RNA: Total RNA was isolated using the RNeasy Mini Kit. Stem cell colonies were isolated by mechanical dissection into 350µl lysis buffer containing Buffer

RLT and 1% (v/v) β-mercaptoethanol. This is a highly denaturing guanidine thiocyanate-containing buffer, which inactivates RNases to ensure purification of intact RNA. The cells were homogenised by vortexing before an equal volume of 70% ethanol was added to the lysate to optimise RNA binding to the RNeasy membrane. The lysate was applied to the RNeasy spin column and spun at 8000 x g for 1min to bind the RNA to the membrane and 350µl Buffer RW1 was then added to the column. The column was centrifuged at 8000 x g for 1min and the flow-through was discarded.

On-column DNase digestion was performed using the RNase-free DNase kit. The lyophilised DNase I was dissolved in 550µl PCR-grade water and mixed gently by inverting the vial. For long-term storage of DNase I, the stock solution was divided into 50µl aliquots and stored at -20°C. A working solution of DNase I was freshly prepared by adding 10µl DNase I stock solution to 70µl Buffer RDD, mixed by gently inverting the tube. 80µl DNase I incubation mix were then added directly to the RNeasy spin column membrane and incubated for 15min at room temperature.

The RNeasy spin column was washed with 350µl Buffer RW1, centrifuged at 8000 x g for 1min and the flow through was discarded. 500µl Buffer RPE was added to the RNeasy spin column and centrifuged at 8000 x g for 1min and the flow-through was discarded. A further 500µl Buffer RPE was added to the spin column, centrifuged at 8000 x g for 2min and the flow through was again discarded. The column was spun dry at 8000 x g for 1min and then placed in an RNase, DNase-free 1.5ml tube. 30µl of RNase-, DNase-free water were added directly to the membrane. The column was incubated for 1min at room temperature before being centrifuged at 8000 x g for 1min to elute the RNA. The eluate was reapplied to the spin column membrane and centrifuged at 8000 x g for 1min to maximise RNA concentration. The RNA was analysed using the Nanodrop spectrophotometer to check the yield and quality of RNA isolated. Isolated RNA was stored at -80°C.

Reverse Transcription of Total RNA into cDNA: 9.5µl total RNA, 1µl Oligo$(dT)_{15}$ primers (0.5µg/µl) and 1µl random primers (0.5µg/µl) were added to a sterile 0.5ml microcentrifuge tube and the mixture was incubated at 70°C for 5min to melt secondary structure within the template then immediately cooled on ice for 5min to prevent secondary structure from reforming. Oligo$(dT)_{15}$ primers bind to the polyA tail of the mRNA, while random primers

bind to random sequences in the mRNA. A master mix was prepared according to table 2.11 and 8.5µl was added to each sample. The samples were then incubated for 50 min at 42°C to allow the activation of the MMLV-RT enzyme and the synthesis of cDNA, followed by 15min at 70°C to inactivate the enzyme. Newly synthesised cDNA was stored at -20°C.

Reverse Transcription Mix	
Reagent	*Volume Needed For 1 Reaction (µl)*
5x Reaction Buffer	4
10mM dNTPs	1
40U/µl RNAsin	2
400mM DTT	0.5
MMLV-RT	1
Total	**8.5**

Table 2.10. Reagents and volume needed for the reverse transcription reaction.

Amplification of cDNA by PCR: Forward and reverse oligonucleotide primers, complementary to the sequence to be amplified, were chemically synthesised by Integrated DNA Technologies and are detailed in table 2.12. The primers were supplied lyophilised, and a 100µM stock solution was made and diluted to a working concentration of 10µM.

RT-PCR Reaction Mix		
Reagent	*Volume Needed For 1 Reaction (µl)*	*Concentration*
5x Reaction Buffer	4	1x
10mM dNTPs	0.5	*250µM of each nucleotide*
25mM $MgCl_2$	1.2	1.5mM
10µM Forward Primer	1	0.5µM
10µM Reverse Primer	1	0.5µM
GoTaq Polymerase	0.2	1U
cDNA	2	-
PCR-grade Water	10.1	-
Total	**20**	

Table 2.11. RT-PCR reaction mix.

The cDNA samples were diluted 1:3 to remove any potential inhibitors of the PCR from the reverse transcription reaction. A PCR mix was prepared as described in table 2.11 in a sterile

microcentrifuge tube and briefly vortexed and centrifuged. As a negative control, the cDNA was replaced with water. The tubes were placed into a PCR thermo-cycler and the PCR cycles were performed using the following parameters: incubation at 94°C for 2min to activate the GoTaq® polymerase; 40 cycles of 94°C for 45sec (denaturation), 58°C for 30sec (annealing) and 72°C for 1min (elongation); and a final 10 minutes at 72°C to allow PCR products to anneal into double-stranded products. PCR products were visualised by gel electrophoresis.

	Gene	Forward primer (5' to 3')	Reverse primer (5' to 3')	Amplicon Size (bp)
Reference Gene	CYCG	CTTGTCAATGGCCAACAGAGG	GCCCATCTAAATGAGGAGTTGGT	83
Pluripotency Genes	OCT4	GACAGGGGGAGGGGAGGAGCTAGG	CTTCCCTCCAACCAGTTGCCCCAAAC	78
	SOX2	GGGAAATGGGAGGGGTGCAAAAGAGG	TTGCGTGAGTGTGGATGGGATTGGTG	151
	C-MYC	GCGTCCTGGGAAGGGAGATCCGGAGC	TTGAGGGGCATCGTCGCGGGAGGCTG	328
	KLF4	ACGATCGTGGCCCCGGAAAAGGACC	TGATTGTAGTGCTTTCTGGCTGGGCTCC	397
	REX1	CAGATCCTAAACAGCTCGCAGAAT	GCGTACGCAAATTAAAGTCCAGA	306
	TDFG1	AGATGGCCCGCTTCTCTTA	GAGATGGACGAGCAAATTCC	108
	NANOG	ATGCCTCACACGGAGACTGT	AGGGCTGTCCTGAATAAGCA	66
Exogenous Transgenes Used For Reprogramming	OCT4	CCCCAGGGCCCCATTTTGGTACC	TTATCGTCGACCACTGTGCTGCTG	-
	SOX2	GGCACCCCTGGCATGGCTCTTGGCTC		-
	C-MYC	CAACAACCGAAAATGCACCAGCCCCAG		-
Ectoderm Markers	NESTIN	TGCGGGCTACTGAAAAGTTC	TGTAGGCCCTGTTTCTCCTG	64
	GFAP	AGAGGGACAATCTGGCACA	CAGCCTCAGGTTGGTTTCAT	71
	TUBBIII	GCAACTACGTGGGCGACT	CGAGGCACGTACTTGTGAGA	85
Mesoderm Markers	BRACH	GCTGTGACAGGTACCCAACC	CATGCAGGTGAGTTGTCAGAA	106
Definitive Endoderm Markers	SOX17	GGCGCAGCAGAATCCAGA	CCACGACTTGCCCAGCAT	61
	FOXA2	GGGAGCGGTGAAGATGGA	TCATGTTGCTCACGGAGGAGTA	89
	CER	ACAGTGCCCTTCAGCCAG	ACTACAACTACTTTTTCACAGCCTTCGT	52
	CXCR4	CACCGCATCTGGAGAACCA	GCCCATTTCCTCGGTGTAGTT	79

Table 2.12. Oligonucleotide primers used in PCR reactions.

2.4.2 Gel Electrophoresis

Preparation of Agarose Gels: 2% agarose gels were made by dissolving 1g electrophoresis grade agarose in 50ml of 1X TBE buffer (table 2.13) and the mixture was heated in a microwave for approximately 2min until the agarose had dissolved. The solution was cooled and 5µl of 10mg/ml ethidium bromide (to give a final concentration of 1µg/ml) was added. The mixture was poured into a gel tray and allowed to set.

10x Tris Borate EDTA (TBE) Buffer (1L)	
Reagent	*Amount (g/L)*
Tris-(hydroxymethyl) methylamine	109.3
Boric Acid	55.65
EDTA	9.31

Table 2.13. Mass of each component of TBE buffer. *TBE buffer was used to make the gels and as a running buffer in electrophoresis. The volume was made up to 1L with water and the buffer was stored at room temperature. For use, the 10X buffer was diluted 10-fold to obtain a 1X TBE buffer.*

Gel Electrophoresis: 6X blue/orange loading dye was added to the PCR products in a ratio of 1:6. The samples were then loaded onto an agarose gel in an electrophoresis tank filled with 1x TBE buffer. Electrophoresis was carried out at 70V for around 15 minutes. Afterwards, DNA was visualised under UV light using a transilluminator. Images were captured using a gel imaging system (InGenius LHR-Syngene). An estimation of DNA band size was obtained by comparison with DNA markers (e.g. 100bp DNA ladder) run in parallel with the samples.

Purification of DNA from Agarose Gels: DNA was purified from agarose gels using the QIAquick gel extraction kit. The DNA fragment was excised from the gel using a clean scalpel blade and transferred to a DNase-, RNase-free tube. The mass of the DNA-containing gel slice was determined and 5 volumes of buffer PB, containing a pH indicator, were added for every 1 volume of gel (500µl/0.1g gel) and the gel slice was incubated in the buffer at 55°C for approximately 5-10min to allow the gel to melt. The mixture was transferred to a spin column which was centrifuged for 1min at 13,000rpm to bind the DNA to the silica-based membrane and the flow-through was discarded. The DNA was then washed twice with

750µl wash buffer PE and after centrifugation (1min at 13,000rpm) the flow-through was again discarded, and the column further centrifuged for 4min at 13,000rpm to dry the membrane. The column was then placed into a 1.5ml microcentrifuge tube and the bound DNA was eluted with 30µl of PCR-grade water, which was added directly to the silica membrane and incubated for 1min at room temperature, followed by centrifugation for 1min at 13,000rpm. Purified DNA was stored at -20°C.

Preparation of Standard Curves for qRT-PCR: qRT-PR relies on the generation of standard curves for each gene of interest. To generate a template from which serial dilutions could be made to construct a standard cure, a PCR reaction was then performed, as described above, to create a standard curve for each gene of interest in subsequent qRT-PCR experiments. A standard curve for a reference gene was carried out in parallel in order to enable normalisation of any results. *Cyclophilin G* (*CYCG*) was chosen as a reference gene for these experiments as it has previously been shown that expression levels are not altered upon differentiation of stem cells (D'Amour et al., 2006), and this had previously been internally validated in our lab (Courtney, 2010).

2.4.3 Quantitative PCR (qRT-PCR) for Detection of mRNA Expression

PCR reactions were set up in 20µl reactions as described in table 2.14, using either target cDNA or standard cDNA. Standard dilutions were prepared as in section 2.4.2 and run together with the samples of interest. A negative control in which the template cDNA was replaced with water was also included in each reaction. PCR reactions were carried out on the Qiagen Rotor-Gene™ 6000 using the following PCR cycling conditions: 10min at 95°C; followed by 40 repeated cycles of 5sec at 95°C, 15sec at 58°C and 10sec at 72°C. The fluorescence was acquired at the end of the elongation step. A melt curve analysis was carried out to ensure amplification of a single product, and the products were run on a gel to confirm the amplicon was the correct size (table 2.12).

qRT-PCR Reaction Mix	
Reagent	*Volume Needed For 1 Reaction (µl)*
2x Sensimix Plus *(containing reaction buffer, heat-activated Taq Polymerase, dNTPs, 6mM $MgCl_2$, internal reference and stabilisers, and SYBR green)*	10
10µM Forward Primer	1
10µM Reverse Primer	1
cDNA	2
PCR-grade H_2O	6
Total	**20**

Table 2.14. qRT-PCR reaction mix.

Data Analysis: Individual reactions are characterised by the PCR cycle at which fluorescence first rises above a defined threshold background fluorescence (figure 2.3), known as the threshold cycle (C_T). The more target cDNA there is in the starting material, the lower the C_T. Plotting the concentration for a set of known standards versus the C_T gives the standard curve (figure 2.4). In these experiments, the standards were assigned arbitrary concentrations and the amount of target in the unknown samples of interest is determined by measuring the C_T of the samples and using the standard curve to determine their starting concentration. To normalise gene expression for variation in the amount of RNA between different samples, the concentration for the gene of interest is divided by the concentration from a reference gene in the same sample.

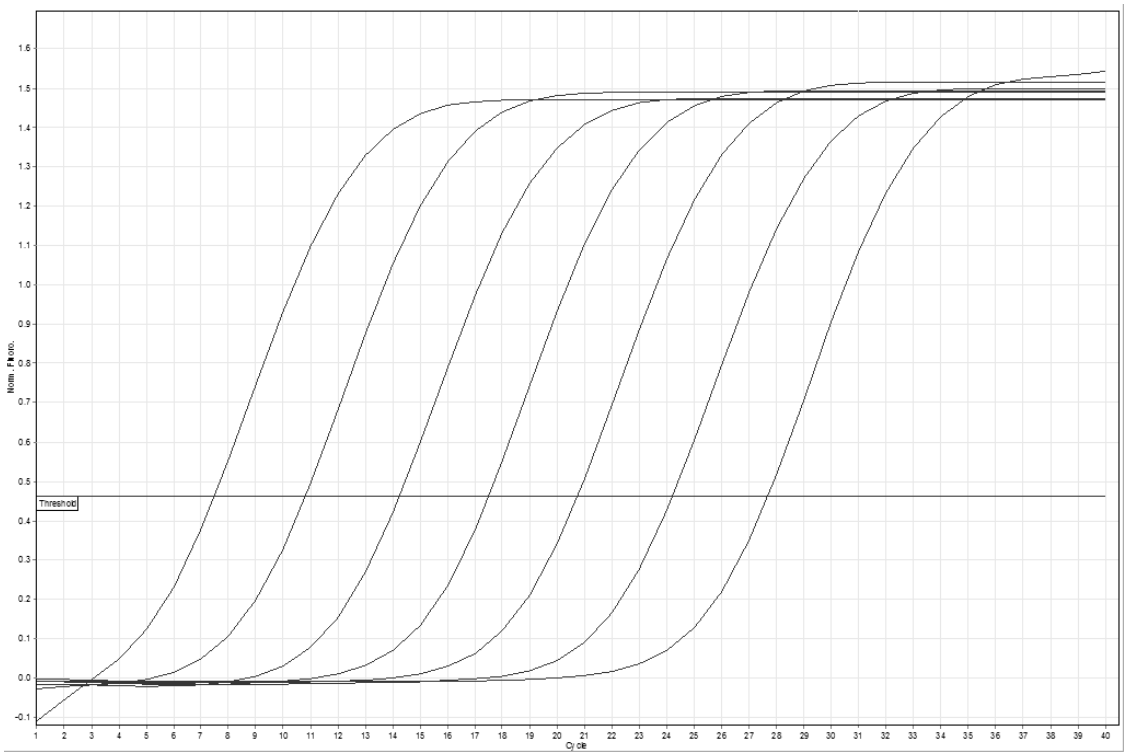

Figure 2.3. Amplification plot *of fluorescence versus cycle number for seven serially-diluted standards, showing the cycle at which fluorescence first rises above a defined threshold background fluorescence (C_T).*

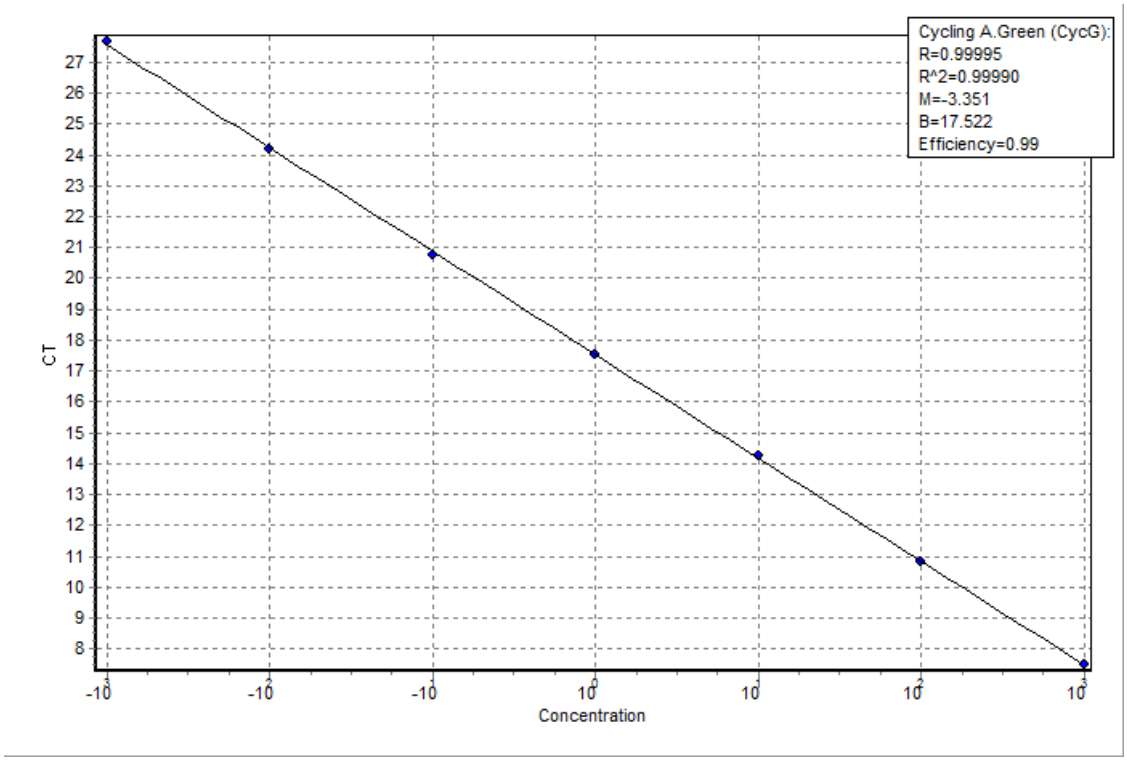

Figure 2.4. Standard curve *generated by plotting the assigned concentration of the standards against the C_T. This can then be used to determine the concentration of the unknown samples relative to the standards.*

2.4.4 Immunocytochemistry

Immunocytochemistry is used to visualise the cellular location and abundance of a specific protein within a cell using antibodies that specifically bind to the protein of interest. These antibodies are labelled with a fluorescent tag, such as fluorescein isothiocyanate (FITC), that will absorb light and emit at a different wavelength. Direct immunofluorescence uses fluorescently-labelled primary antibodies whilst indirect immunofluorescence uses fluorescently-labelled secondary antibodies specific to the primary antibody. A list of the primary and secondary antibodies used is shown in table 2.15.

Primary Antibody	Dilution	Specificity
Goat anti-human Oct4 IgG	1:500	Primary
Goat anti-human Sox17 IgG	1:500	Primary
Mouse anti-human SSEA-3 IgM	1:100	Primary
Mouse anti-human TRA-1-81 IgM	1:100	Primary
Rabbit anti-human Nanog IgG	1:500	Primary
Rabbit anti-goat AlexaFluor® 488	1:100	Secondary
Goat anti-mouse IgG + IgM (FITC)	1:100	Secondary
Goat anti-rabbit IgG (FITC)	1:100	Secondary
Goat IgG	1:500	Isotype Control
Mouse IgG	1:500	Isotype Control
Rabbit IgG	1:500	Isotype control

Table 2.15. List of antibodies used.

Immunocytochemistry on Formalin-Fixed Cells: Immunofluorescence measurements of intracellular proteins require the cells to be fixed and permeabalised to allow the antibody inside the cell. The medium was aspirated and the cell colonies were washed twice with PBS. The cells were fixed with 500µl of 4% paraformaldehyde and incubated for 20min at room temperature before being washed 3 times with PBST (PBS containing 0.01% TritonX 100), and then incubated in 1ml PBST at room temperature for 15min. 500µl of primary antibody (table 2.15), diluted appropriately in PBST, was added to the cells which were then incubated in the dark overnight at 4°C. The cells were washed 3 times with PBST and 500µl secondary antibody at the appropriate dilution (table 2.15) was then added to the cells and left to incubate at room temperature for 1hr. 200µl of 5µg/ml Hoescht DNA stain solution was

added to the cells and incubated for 1min at room temperature. The cells were then washed twice for 5min in PBST. Control dishes were also prepared, by omitting the primary antibody to ascertain the specificity of the antibody, and also using an isotype control together with the secondary antibody to ascertain the degree of non-specific binding. The cells were visualised using a Nikon Eclipse TS100 inverted microscope.

Live Cell Immunocytochemistry: If the proteins of interest are on the surface of the cell, as is the case with the pluripotent stem cell-specific keratan sulphate antigen TRA-1-60, the cells do not need to be fixed or permeabalised, allowing live cell staining. Cells were cultured as described in section 2.2 for at least 5 days before staining. 2µl of 0.5ng/ml TRA-1-60 StainAlive dye was diluted in 198µl of medium to give a final concentration of 5pg/ml. The medium was aspirated and 200µl of antibody solution was added. After 30mins incubation at 37°C, 5% CO_2, the cells were washed twice with medium and left covered with 1ml of medium. The cells were visualised using a Nikon Eclipse TS100 inverted microscope.

Alkaline Phosphatase Staining: Undifferentiated pluripotent stem cells are characterised by high levels of expression of alkaline phosphatase, which can be visualised using a Napthol/Fast Red Violet staining solution. The cells were cultured for at least 5 days before staining. The media was aspirated and the cells were fixed with 4% formaldehyde for 2 minutes. The cells were washed with TBST (20mM Tris-HCl, pH7.4; 0.15M NaCl, 0.05% Tween-20). The staining solution was prepared by mixing Fast Red Violet, Napthol AS-BI phosphate and water in a ratio of 2:1:1. This was added to the cells which were then incubated at room temperature in the dark for 15min. The cells were washed with TBST and covered with PBS to prevent drying. The cells were visualised using a Nikon Eclipse TS100 inverted microscope.

2.5 Investigation of microRNA Expression

Several characteristics of miRNAs make investigating their expression challenging. Unlike mRNAs, which all have a poly(A) tail, miRNAs do not have a common sequence which can be used to selectively enrich miRNAs in the sample. This is important as miRNAs make up only about 0.01% of the total RNA in a sample, and profiling needs to be able to distinguish mature miRNAs from other RNA species, as well as from pre- and pri-miRNA sequences, which will contain the mature miRNA sequence. Different mature miRNAs can be produced from the same pre-miRNA: distinct miRNAs are generated from the 5' and 3' arms of the pre-miRNA, and mature miRNAs can vary in length due to post-transcriptional modifications, giving rise to different "isomiRs" which may have differences in stability of function. In addition, miRNAs within a family may vary from each other by as little as one nucleotide in the sequence, so profiling methods need to be highly sensitive to be able to detect these differences. The copy number of different miRNA species within a cell varies over several orders of magnitude, and while the average copy number of a miRNA has been estimated to be around 500 per cells, some cell type-specific miRNAs are present at over 10,000 copies per cells. Finally, due to their short length, variance in the CG content of miRNAs leads to a wide variance in melting temperatures for annealing reactions, creating miRNA-specific bias. There are three main techniques used for profiling miRNA expression, each with their own advantages and disadvantages: microarray, qRT-PCR, and high-throughput RNA sequencing.

2.5.1 Microarray Analysis

A microarray works on the basis that RNA can hybridise specifically to complementary oligonucleotide target sequences which are printed onto a glass slide. Sample RNA is labelled with a fluorescent tag and the labelled RNA hybridises to complementary probes on the microarray slide. When the slide is scanned with a laser, a measurement of the fluorescence allows a measurement of the amount of RNA bound to that target (figure 2.5).

Microarray analysis has the advantage of being a high-throughput method for profiling miRNA expression, allowing investigation of thousands of miRNAs in parallel: in this case, the microarray contains capture probes for all the microRNAs annotated in miRBase version

18. It is also relatively inexpensive. However, limitations of this approach include the fact that the labelling reaction can result in the addition of a large and variable number of nucleotides, and other RNA species may also be labelled, contributing to background noise, and may cross-hybridise with miRNA probes. The variation in the CG content of miRNAs can lead to different annealing temperatures for different miRNAs. To avoid this problem, LNA (locked nucleic acids) were incorporated into the array, as they have an increased affinity for complementary miRNA sequences. In addition, a limited range of quantification, imperfect specificity for some closely related miRNAs, and the inability to absolutely quantify miRNA means that this technique is most suited to carrying out comparisons between two states e.g. differentiated and undifferentiated stem cells, rather than absolute quantification of miRNA expression levels.

In a traditional two-colour array, two different samples to be compared are labelled with two different colour fluorescent dyes, usually Cy3 (green) and Cy5 (red). The two labelled samples are mixed together and hybridised to a single microarray. Relative intensities of each fluorophore bound to each capture probe may then be used to identify up-regulated and down-regulated genes. In the experiments described here, a common reference approach was used, allowing comparison of all samples against each other. Each sample is hybridized on the same slide as a common reference sample, which is a mixture of all the samples used in the array, therefore reflecting their complexity. One colour dye is used for the sample, and the other for the common reference. As the reference sample is the same in all the arrays carried out, all samples can be directly compared with each other. This is similar in principle to single colour experiments, but allows for optimal inter-array performance monitoring and normalization, since the common reference gives a common baseline. It also minimizes the influence of dye specific differences on data analysis.

Microarray analysis was used to analyse microRNA expression in samples of undifferentiated ESCs and iPSCs, as well as in ESCs and iPSCs that had undergone differentiation to definitive endoderm, as described in section 2.3. Samples were taken in triplicate, and total RNA, including the small RNA fraction, was isolated as described in section 2.5.2. Microarray hybridisation and image scanning was carried out by Exiqon Services, Denmark, who then supplied the raw data for analysis at NIBSC.

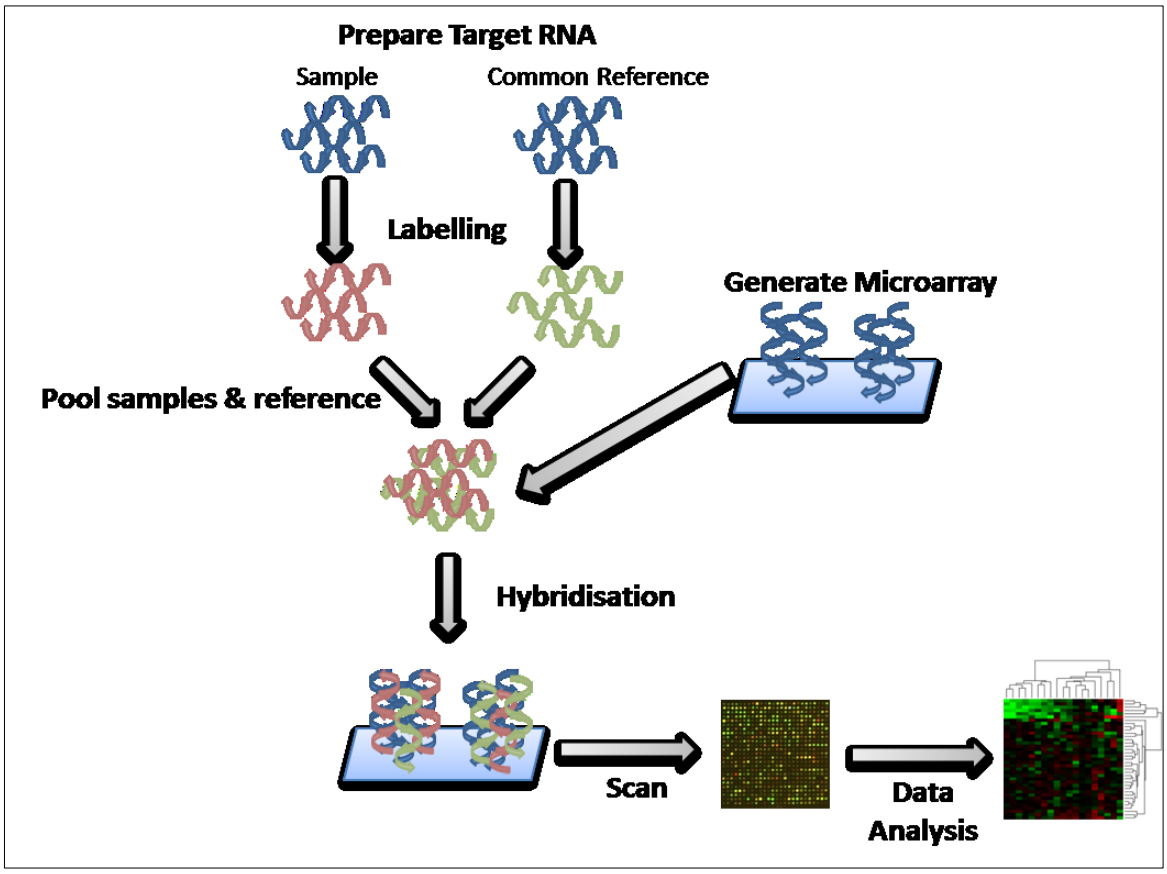

Figure 2.5. Overview of microarray analysis of microRNA expression.

Prior to microarray analysis, RNA yield and quality was analysed using RNA 6000 Nano chips on the Agilent 2100 Bioanalyzer, which is a microfluidics-based platform for sizing, quantification and quality control of DNA, RNA, proteins and cells. The chips contain sample wells, gel wells and a well for an external standard (ladder). Micro-channels create interconnected networks among these wells. During chip preparation, the micro-channels are filled with a sieving polymer ("gel") and fluorescence dye. Then, samples and ladder with marker are loaded in each well. Once the wells and channels are filled, the chip becomes an integrated electrical circuit. Charged molecules like DNA or RNA are electrophoretically driven by a voltage gradient, similar to gel electrophoresis. Dye molecules intercalate into RNA strands and these complexes are detected by laser-induced fluorescence. Data is translated into gel-like images (bands) and electropherograms (peaks). The software automatically compares the unknown samples to the ladder fragments to determine the concentration of the unknown samples and to identify the ribosomal RNA peaks. The software also assigns an RNA integrity number (RIN) which is a measure of the quality of the

sample. A RIN of higher than 8 (out of maximum possible score of 10) is required for use in downstream applications such as microarray analysis, as a RIN lower than this indicates that the RNA is at least partially degraded, and this can compromise results. A typical electropherogram from an RNA 6000 Nano chip is shown in figure 2.6, which shows clearly defined peaks for the ladder, and the 18S and 28S ribosomal subunits. In a more degraded RNA sample, the 18S and 28S peaks would be lower, and other small peaks would also be present throughout the electropherogram.

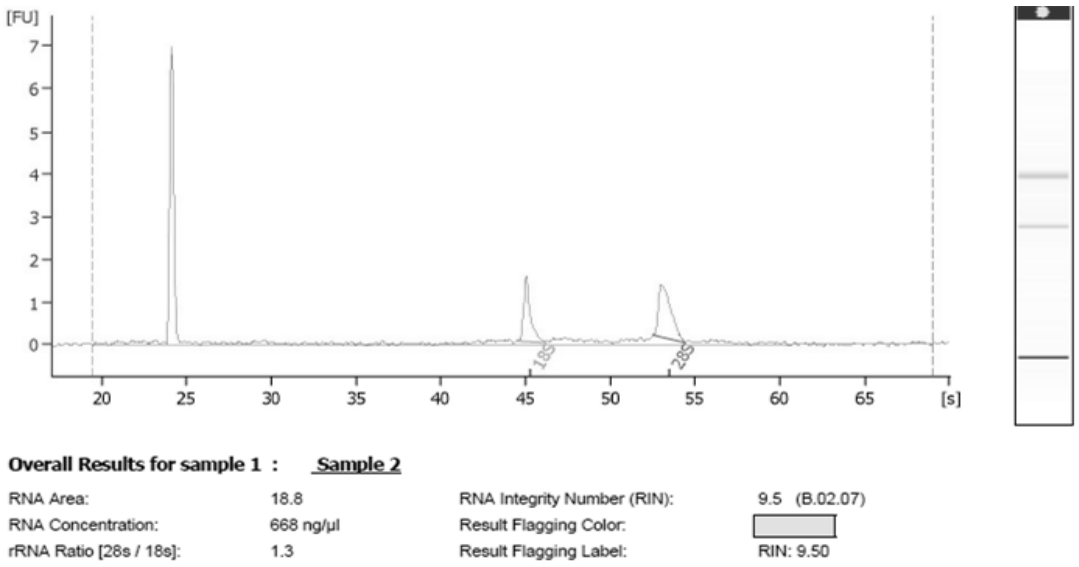

Figure 2.6. Typical electropherogram from an RNA 6000 Nano Bioanalyzer chip used to analyse RNA quantity and quality. *The first peak shows the RNA ladder used as an external standard, while the two smaller peaks show the ribosomal subunits. There is little degradation of the sample, and it has a RIN of 9.5, which is adequate for use in further analysis.*

Data Analysis: The microarray slides were analysed using the NimbleGen MS200 scanner (Roche). Signal intensity data were extracted from the scanned images of each array using ImaGene9 software. Background correction, normalisation and miRNA expression analysis was carried out using Nexus Expression software.

Background correction was carried out for each array slide using the Normexp background correction algorithm (Ritchie et al, 2007). This is based on signal intensity calculations that do not use subtraction to obtain a background corrected value. Using this algorithm, the background threshold was calculated for each array slide as 1.2x the 25th percentile of the

overall signal intensity of the slide. miRNAs with intensities above background level in less than 4 of the 24 samples were removed from subsequent analysis, as when selecting candidates for validation, it is important to ensure that the signal intensity of the probes is well above the background. If microRNA capture probe signals are close to the background, this indicates that the microRNAs are present at a concentration that is close to or below the detection level of the array. For these probes, calculation of expression levels becomes unreliable.

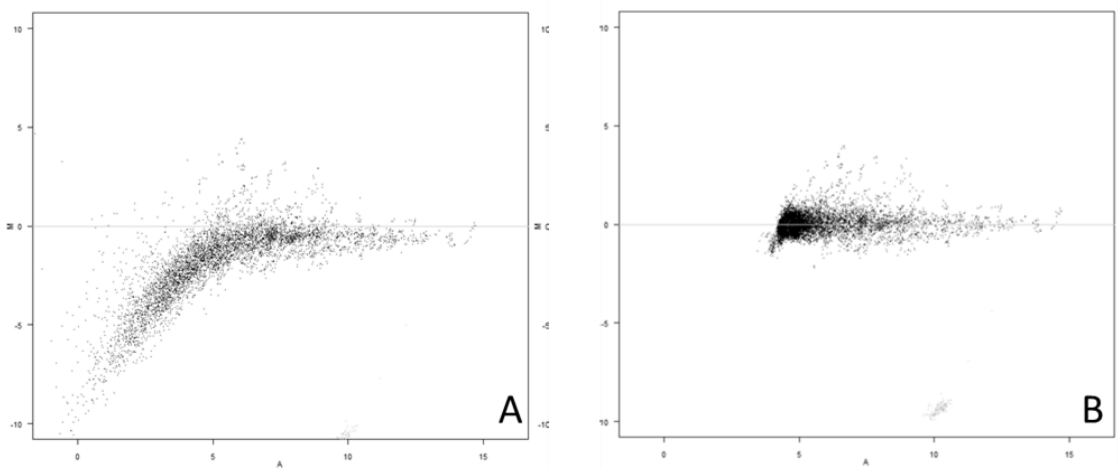

Figure 2.7. Before (left panel) and after (right panel) normalisation of data using Lowess normalisation.

Normalisation is a mathematical adjustment of data to eliminate systemic errors such as dye bias and differences in labelling, hybridisation and scanning which can occur when performing a microarray experiment. The aim of normalisation is to enable comparison of data from different samples. When performing dual-colour arrays, it must be taken into account that the two dyes used have a different dynamic range for their concentration-dependent light emission: for example, for low intensity signals, green fluorescent dyes show higher signal intensity compared to red fluorescent dues. In these experiments, normalisation was performed using Lowess normalisation, which smooths the log ratio/log mean-intensity scatterplot to a linear distribution (figure 2.7). It works under the assumption that the majority of the signals between samples do not differ and enforces an equal overall mean on all signal intensities.

Comparison of miRNA Expression: For comparisons of miRNA expression between samples, the log fold change was calculated and compared. For this comparison, the calculated p-values were based on Students' T-test. Due to the high numbers of miRNAs being tested in parallel, microarrays are prone to give false results. This can be controlled by using a multiple testing correction (Benjamini-Hochberg method) which adjusts p-values derived from multiple statistical tests.

2.5.2 qRT-PCR for Detection of miRNA Expression

Due to the high chance of false positives occurring during microarray analysis, the results need to be validated by qRT-PCR, as this is a more reliable method of analysing miRNA expression. Several features of miRNAs make profiling their expression levels using qRT-PCR challenging. They are short (~22 nucleotides), which makes primer design difficult; the GC content of different miRNAs is heterogeneous, leading to variability in melting temperatures of nucleic acid duplexes between different miRNAs; mature miRNAs lack a common sequence feature allowing their selective purification from other RNA species e.g. a poly(A) tail; and the target sequence of a mature miRNA is also present in the primary and precursor miRNAs (for an overview of miRNA biosynthesis and processing see figure 1.10). In order to solve these problems, a poly(A) tail is added to miRNAs during the reverse transcription of RNA to cDNA, which allows selective amplification. Addition of a universal tag sequence during reverse transcription allows the use of a universal primer for all miRNA PCR assays, together with a primer specific for the miRNA of interest. This makes primer design easier, and also increases the length of the amplicon so that variability in melting temperature can be reduced.

miRNA Isolation: Total RNA, including the miRNA-containing fraction, was isolated from samples using the miRNeasy™ Mini Kit. Stem cell colonies were isolated by mechanical dissection (see section 2.2) into 700µl QIAzol lysis reagent, which contains phenol and guanidine thiocyanate, and is designed to facilitate cell lysis, inhibit RNases and remove DNA and proteins from the lysate. Lysates were incubated at room temperature for 5min. 140µl chloroform was added to each tube, which was shaken vigorously for 15sec and then

incubated at room temperature for 2-3 minutes. The samples were centrifuged at 4°C for 15 min at 12,000 x g, allowing them to separate into phases: a clear colourless aqueous phase (containing RNA), a white interphase (containing DNA) and a red organic phase (containing proteins). The upper aqueous phase (around 350µl) was transferred to a new collection tube and 1.5 volumes (usually 525µl) of 100% ethanol were added and mixed thoroughly by pipetting up and down. The sample was transferred into an RNeasy mini spin column and centrifuged at 8000 x g for 1 minute. The flow-through was discarded. On-column DNase digestion was performed as previously described.

500µl Buffer RPE was added to the RNeasy spin column, which was centrifuged at 8000 x g for 1 minute to wash the column. Another 500µl was added to the RNeasy spin column, which was centrifuged at 8000 x g for 2 minutes to dry the column. The column was transferred to a new 2ml collection tube and centrifuged at 13000rpm for 1 minute. The column was transferred to a 1.5ml collection tube and 30µl of RNase-free, DNase-free PCR-grade water was pipetted directly onto the spin column membrane. The tube was centrifuged for 1 minute at 8,000 x g to elute the RNA. The eluate was reapplied to the spin column membrane and centrifuged at 8000 x g for 1min to maximise RNA concentration. Isolated total RNA was stored at -80°C.

Reverse Transcription Of microRNA into cDNA: The miScript™ reverse transcription kit was used to reverse transcribe miRNA into cDNA. The reverse transcription kit includes miScript reverse transcriptase mix, which is a blend of enzymes comprising a poly(A) polymerase and a reverse transcriptase, and miScript RT buffer, which contains Mg^{2+}, dNTPs, Oligo-dT primers and random primers. The poly(A) polymerase adds a poly(A) tail to the miRNA, while the Oligo-dT primers add a universal tag sequence on the 5' end, which allows binding of the universal primer during the PCR step. Polyadenylation of miRNAs and reverse transcription are carried out in parallel in the same step.

Total RNA and reverse transcriptase were thawed on ice. MiScript™ RT buffer and RNase-free water were thawed at room temperature. The reverse transcription master mix was prepared on ice according to table 2.16 and added to template RNA. The reverse transcription reaction was carried out at 37°C for 60minutes followed by 95°C for 5 minutes to inactivate the reverse transcriptase. The cDNA samples were stored at -20°C.

Reverse Transcription Mix	
Reagent	*Volume Needed For One Reaction (µl)*
5x miScript™ RT Buffer	4
RNase-free water	5
miScript™ Reverse Transcriptase Mix	1
RNA	10
Total	**20**

Table 2.16. Reagents and volumes needed for the reverse transcription reaction.

qRT-PCR For Detection Of miRNA Expression: PCR analysis of mature miRNA was carried out using a forward primer specific for the miRNA of interest in combination with the miScript™ SYBR Green PCR kit, which contains a universal primer used in all miRNA PCRs, using the reaction mix shown in table 2.17. 18µl of master mix was added to 2µl of cDNA. Standard dilutions were prepared as described in section 2.4.2 and run together with the samples of interest. A standard curve for a reference gene was also set up following the same procedure in order to enable normalisation of any results. A negative control in which the template cDNA was replaced with water was also included in each reaction. PCR reactions were carried out on the Qiagen Rotor-Gene™ 6000 using the reaction conditions as follows: 15min at 95°C; 40 repeated cycles of 15sec at 94°C, 30sec at 55°C and 30sec at 70°C. The fluorescence was acquired at the end of the extension step. A melt curve analysis was carried out to ensure amplification of a single product.

PCR Master Mix	
Reagent	*Volume Needed For One Reaction (µl)*
2x QuantiTect SYBR Green PCR Master Mix	10
10x miScript Universal Primer	2
10 miScript Primer Assay	2
PCR-grade Water	4
cDNA	2
Total	**20**

Table 2.17. qRT-PCR reaction mix for detection of miRNA

2.5.3 Identification of Gene Targets of miRNAs

Identifying miRNA targets in mammals is challenging, largely due the limited complementarity between miRNAs and their target mRNAs, and the fact that the interactions between miRNAs and their target genes are still not entirely understood. The lack of high throughput experimental methods for miRNA target identification means relying on target prediction algorithms. Several independent groups have established computational algorithms designed to predict target genes of miRNAs (reviewed in Rajewsky, 2006)

miRNA targets usually have at least one region that has complementary base pairing to the "seed" sequence at the 5' end of the miRNA. Most miRNA recognition sites can be found within the 3' untranslated region (UTR) of the target mRNA, although in theory miRNAs could bind to any part of the mRNA (Saito & Saetrom, 2010). Complete complementarity between a miRNA and a target mRNA is rare, but as little as a 6-base pair match is sufficient to suppress gene expression, because RISC can tolerate small mismatches within the seed region (Thomson *et al.*, 2011).

No existing target prediction software has been able to include all the currently known features of miRNA-target interactions, which include: 1) Complementarity to the miRNA seed sequence: when predicting targets of miRNAs, considering only perfectly complementary seed types increases specificity but might miss many potential targets, while considering moderately complementary seed types as well increases sensitivity but is also likely to increase the number of false positives; 2) Evolutionary conservation of the target site among different species by comparative sequence analysis: in order to reduce the number of false positives, several target prediction algorithms perform conservation analysis across species. However, even conserved miRNAs have a large number of non-conserved targets. There are also numerous miRNAs that are not conserved between species (Witkos *et al.*, 2011); 3) Free energy of the miRNA-mRNA complex: this can help determine the likelihood of binding between a miRNA and an mRNA, although since data sets of identified miRNA-mRNA duplexes are very limited, and a low free energy of hybridisation does not guarantee accurate prediction of miRNA target genes, it is difficult to resolve appropriate thresholds of free energy; 4) mRNA sequence features outside the target site: some target prediction algorithms take into account target site accessibility i.e. the free energy cost of unfolding the mRNA secondary structure surrounding the target site to assess the likelihood of binding

between a miRNA and an mRNA; 5) The presence of multiple target sites: strong miRNA targets tend to have multiple target sites within their mRNA sequence (Saito & Saetrom, 2010), as many miRNAs are co-expressed and are likely to regulate the same mRNA co-ordinately. Multiple target sites in the same 3' UTR can potentially increase translation suppression and enhance specificity of gene regulation (Min & Yoon, 2010). Although the effect of multiple target sites generally seems to be additive, miRNA targeting can also be synergistic (Saito & Saetrom, 2010).

Different target prediction programs use different combinations of these features in order to predict miRNA targets. TargetScanS requires perfect complementarity with 7 or more bases of the miRNA seed region, and conservation between at least 5 species for the portion of the target site that binds to the miRNA seed. The seed match must occur at exactly corresponding positions in a cross-species UTR alignment. In addition, it adds a context score which takes mRNA sequence features outside of the target site into account. The estimated false positive rate is 22%, and all known miRNA-target interactions were successfully predicted (with no false negatives) (Min & Yoon, 2010).

PicTar combines base-pairing and thermodynamic stability by allowing targets with imperfect seed matching providing they pass a strict free energy cut-off for imperfect matches. It also requires that the seed match occur at overlapping positions in a cross-species UTR alignment for at least 5 species. In addition, it considers clustering of co-expressed miRNAs, matching miRNAs with targets that are expressed in the same context and requires binding sites that are co-regulated by multiple miRNAs across species. It has an estimated precision level of >60% (Min & Yoon, 2010).

miRanda selects target genes on the basis of three properties: 1) sequence complementarity, using a modified dynamic programming approach that recognises the importance of seed binding but does not require perfect seed complementarity; 2) the free energy of miRNA-mRNA complexes; and 3) the conservation of target sites. The entire target site must occur with at least 90% identity at exactly corresponding positions in a cross-species UTR alignment. It only requires conservation between human and rodent genomes, but also has the additional option of including more extensive species conservation. miRanda has an estimated false-positive rate of ~24% and was able to correctly identify 9/10 previously validated targets (Min & Yoon, 2010).

DIANA-microT uses a sequence of 38 nucleotides to progressively go through a 3'UTR of a potential target in order to identify possible miRNA recognition elements (MREs). It calculates the minimum binding energy at each step, which is compared with the outcomes obtained from scrambled sequences with the same dinucleotide content as real 3' UTRs. DIANA-microT allows a weak binding at the 5' seed region, unlike TargetScan and PicTar, if there is additional base pairing between the 3' end of the miRNA and the target gene. In addition, it only requires conservation between human and rodent. DIANA-microT is reported to have the highest precision level of all prediction software at 66% (Min & Yoon, 2010), and its web server contains links to UniProt for protein information; iHOP for functional and bibliographic information; miRBase for miRNA information; and KEGG pathway for pathway analysis.

PITA combines the free energy cost of unfolding the secondary structure of the mRNA surrounding the target site with the free energy cost of miRNA-mRNA pairing in order to predict miRNA targets, and does not rely on conservation between species.

Algorithms used for predicting targets typically predict hundreds if not thousands of target genes for each miRNA, most of which are not true targets (Thomas *et al.*, 2010). The false positive rate of prediction algorithms has been variously reported as between 24-70% (Thomson *et al.*, 2011). This underscores the need for experimental data to demonstrate genuine miRNA targets and functions. Combining results from multiple prediction programs is encouraged in order to reduce the probability of false positives/negatives. TargetScanS, miRanda and PicTar are the most commonly used (Min & Yoon, 2010). In a comparison of target prediction programs, TargetScanS, PicTar and miRanda used either alone or in combination had the best trade-off between sensitivity and specificity (Witkos *et al.*, 2011). Although the union of all target prediction programs achieves the highest sensitivity, it also achieves the lowest specificity. Conversely, although the intersection of all programmes achieves the highest specificity, it also achieves the lowest sensitivity (Sethupathy *et al.*, 2006). Target prediction algorithms also sometimes fail to predict biologically relevant targets. Many genes identified by overexpression or silencing of miRNAs in combination with microarray analysis were not predicted by target prediction software, which can largely be explained by the application of stringent criteria regarding the evolutionary conservation of putative binding sites (Krutzfeldt *et al.*, 2006).

In addition, many resources are now available to extend the usefulness of target prediction programs. TarBase is a database containing miRNA targets that have been experimentally validated, and is linked to other databases to extend information about miRNAs and their target genes. miRecords curates predicted targets generated by 11 different prediction programs, and also includes validated targets together with their prediction results. miRGator integrates target prediction (combining TargetScan, MiRanda and PicTar), functional analysis, gene expression data and genome annotation. GoMir combines JTarget and TAGGO. JTarget integrates target prediction from 4 programs (TargetScan, miRanda, RNAhybrid, PicTar) and an experimental database (TarBase) while TAGGO performs GO clustering with the common genes obtained from JTarget and analyses how many target genes share a common GO category. MMIA finds target mRNAs by combining computational prediction results and expression data analysis, based on the fact that the expression of a miRNA must be reciprocal to that of its target mRNA.

2.5.4 Luciferase Assay in 293FT Cells

In order to determine whether miRNAs were acting directly by binding to genes which had been identified as putative targets, a luciferase assay was carried out in which a miRNA mimic or inhibitor were co-transfected into 293FT cells together with a reporter plasmid carrying a *Gaussia luciferase* gene and the 3'UTR of the gene of interest, which is where the majority of miRNA-target interactions occur. The 3'UTR plasmid is shown in figure 2.8. The plasmids encoded in the lentiviral vectors used to knockdown or overexpress miRNAs are shown in figures 2.9 and 2.10, respectively. Both vectors contain a reporter gene (eGFP for the mimic, and mCherry for the miRNA inhibitor), allowing monitoring of transfection efficiency. The effects of manipulating miRNA expression were analysed using qRT-PCR, as described in section 2.4.3.

An overview of the luciferase assay is shown in figure 2.11. If the gene is a direct target of the miRNA of interest, the miRNA (either transfected or present endogenously in the cell) binds to the 3'UTR and degrades it. The luciferase gene is not translated and no luminescence is observed. However, when an inhibitor of the miRNA is transfected into the cells, it prevents miRNA binding to the 3'UTR. The luciferase gene is translated, resulting in luminescence which can be detected using a luminometer. In this case, a plasmid was used

which also contained a reporter gene, secreted alkaline phosphatase (seAP), which allows normalisation of luminescence measurements between different samples. In addition to the 3'UTR plasmid, a control plasmid was also used which has a scrambled sequence in place of the 3'UTR which does not correspond to any known miRNA binding sites. Scrambled miRNA mimic and inhibitor lentiviral vectors were also included as negative controls.

293FT cells were maintained in culture as described in section 2.1.1. 24hrs prior to infection, 293FT cells were trypsinised and plated at a density of $3x10^4$ cells per well of a 24-well plate in 293FT medium, giving a confluency of 60-70%. After 24hrs, the medium was removed from the cells, and they were washed once with Opti-MEM reduced serum medium. 0.5ml of Opti-MEM supplemented with 4µg/ml Polybrene was added to each well, and 10µl of lentivirus or 1µl of control lentivirus (giving a MOI of approximately 10) was added to the cells according to the plate map shown in table 2.20. After 5-6 hours, the medium on the cells was replaced with 293FT medium. The medium was changed again the next day. 72hrs after infection with lentivirus, cells were transfected with the 3'UTR-containing plasmid. Cells were washed once with Opti-MEM and 0.5ml of Opti-MEM was added to the cells. 2.5µl of Lipofectamine 2000 transfection reagent was diluted in 25µl of Opti-MEM. 500ng of plasmid DNA was diluted in 50µl of Opti-MEM and the DNA and Lipofectamine were combined and incubated at room temperature for 5mins. The DNA-Lipofectamine mixture was then added to the cells according to the plate map shown in table 2.19. The cells were incubated for 5-6hrs at 37C, 5% CO2 before the medium on the cells was replaced with 293FT medium. The medium was replaced again the next day.

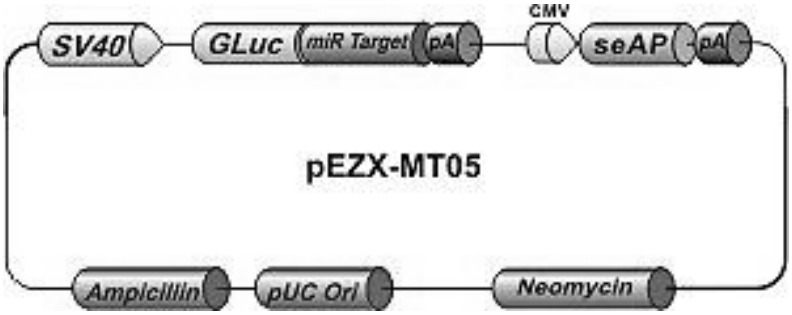

Figure 2.8. Plasmid used in luciferase assays to determine miRNA-target relationships. *The plasmid contains the 3'UTR of the gene of interest coupled to a luciferase gene which allows monitoring of miRNA binding to this region. This plasmid also contains a constitutively expressed reporter gene (alkaline phosphatase) which allows normalisation of luminescence measurements between different samples.*

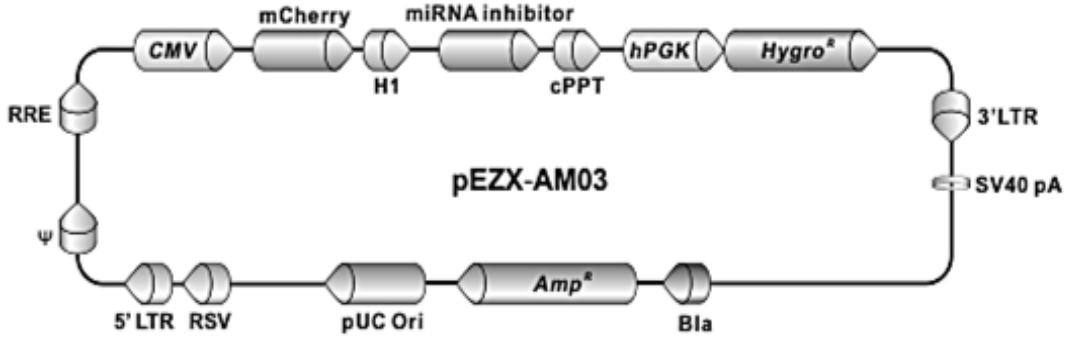

Figure 2.9 miArrest™ lentiviral vector used to repress expression of miRNAs of interest. *The vector contains the miRNA inhibitor sequence, as well as a mCherry reporter gene and an antibiotic resistance gene (hygromycin) which allows stable selection of transfected cells.*

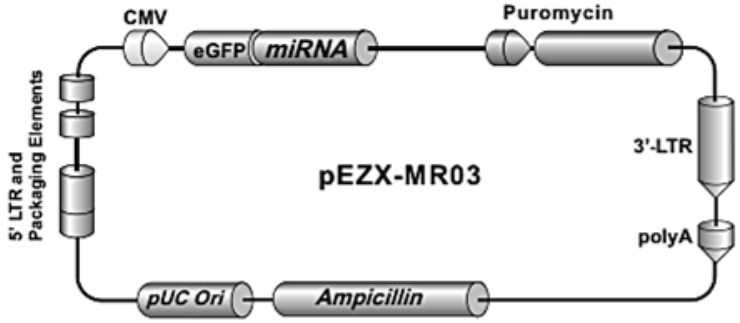

Figure 2.10 miExpress™ lentiviral vector used to overexpress miRNAs of interest. *The vector contains the precursor miRNA sequence, as well as an eGFP reporter gene and an antibiotic resistance gene (puromycin) which allows stable selection of transfected cells.*

Plate Map					
Vector only	Vector only	Vector only	Control vector only	Control vector only	Control vector only
Vector + mimic	Vector + mimic	Vector + mimic	Control vector + mimic	Control vector + mimic	Control vector + mimic
Vector + inhibitor	Vector + inhibitor	Vector + inhibitor	Control vector + inhibitor	Control vector + inhibitor	Control vector + inhibitor
Vector + mimic + inhibitor	Vector + mimic + inhibitor	Vector + mimic + inhibitor	Untransfected Cells	Untransfected Cells	Untransfected Cells

Table 2.18. Plate map used in luciferase assays.

72hrs after plasmid transfection, the culture medium was collected and analysed for luminescence using the SecretePair™ Dual Luminescence Assay Kit. Two luminescence measurements were taken, one for *Gaussia luciferase* (GLuc), which is the reporter for miRNA binding, and one for secreted alkaline phosphatase (seAP), which is constitutively expressed and can be used to normalise GLuc measurements between samples.

For measurement of GLuc, the GL-S buffer was used. The 10x buffer was thawed at room temperature and then vortexed to mix. 100µl of 1x buffer was prepared per sample by diluting 1:10 in distilled water. The GLuc assay working solution was prepared by adding 10µl of substrate GL to 1ml of buffer GL-S. This was mixed well and incubated at room temperature in the dark for 25mins. 10µl of each sample was added to a well of a 96-well plate. 100µl of GLuc assay working solution was added to each well and gently mixed. The plate was incubated at room temperature for 1min before measuring luminescence.

For measurement of seAP, 50ul of medium from each sample was heated at 65ºC for 10-15min and then cooled on ice. 10x buffer AP was thawed at room temperature and then vortexed to mix. 100µl of 1x buffer was prepared per sample by diluting 1:10 in distilled water. The seAP assay working solution was prepared by adding 10µl of Substrate AP to 1 ml of 1x buffer AP. This was mixed well and incubated at room temperature in the dark for 5-10mins. 10µl of each sample was added to a well of a 96-well plate. 100µl of seAP assay working solution was added to each well and gently mixed. The plate was incubated at room temperature for 5-10mins before measuring luminescence with the luminometer. To obtain normalised luminescence measurements, the GLuc readings were divided by the seAP readings. ANOVA statistical analysis was performed to determine differences between samples.

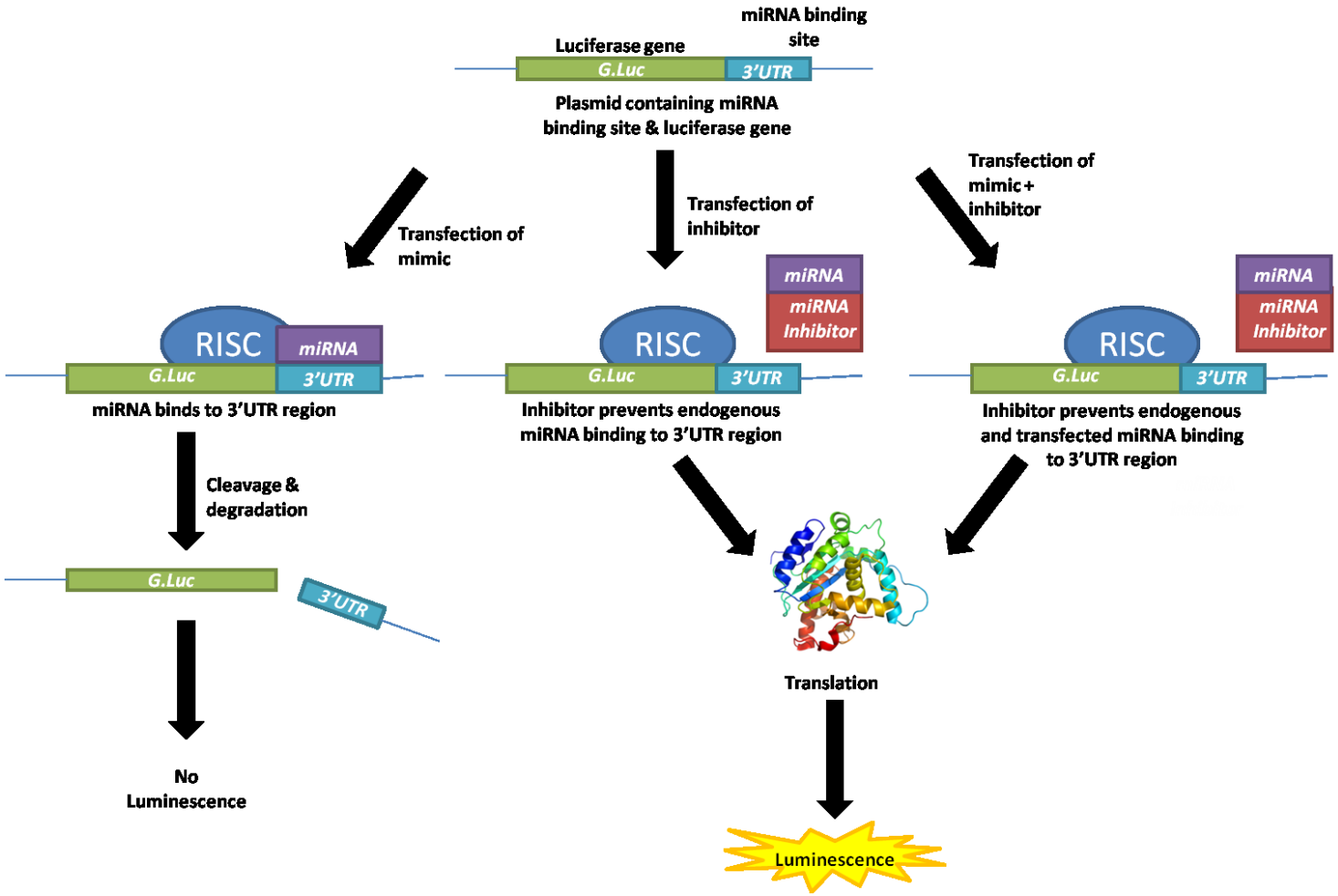

Figure 2.11. Overview of luciferase assay to detect miRNA-target interactions. *Cells are transfected with a plasmid containing a luciferase gene and the 3'UTR of the gene of interest. The cells are then transfeted with i) a miRNA mimic; ii) a miRNA inhibitor; iii) both a miRNA mimic and inhibitor. When the miRNA of interest (either transfected or present endogenously in the cell) binds to the 3'UTR, it is degraded and there is no luciferase expression. However when the miRNA of interest is prevented from binding by an inhibitor, the luciferase gene can be translated, resulting in luminescence.*

Chapter 3
Generation & Characterisation of Induced Pluripotent Stem Cell Lines

Contents

List of Tables

List of Figures

Chapter 3: Generation & Characterisation of Induced Pluripotent Stem Cell Lines

3.1 Introduction

Embryonic stem cells (ESCs) were first derived from blastocyst-stage mouse embryos in 1981 (Evans & Kaufman, 1981; Martin, 1981), although it took a further 17 years for the first derivation of human ESC (hESC) lines to be reported (Thomson *et al.*, 1998). These cells have two defining characteristics: firstly, they can self-renew in culture indefinitely, and secondly, they are pluripotent i.e. they can give rise to any cell in the body. They also display cell surface markers including SSEA-3, SSEA-4, the keratan sulphate antigens TRA-1-60 and TRA-1-81, and alkaline phosphatase, and also have high telomerase activity, showing that these cells are capable of extensive replication (Thomson *et al.*, 1995; Thomson *et al.*, 1996).

In 2006, a significant advance in the field of stem cell research was made with the discovery that 4 pluripotency-associated transcription factors (OCT4, KLF4, c-MYC and SOX2) could reprogram somatic cells into induced pluripotent stem cells (iPSCs). Reprogramming was initially carried out in mouse fibroblasts (Takahashi and Yamanaka, 2006), and reprogramming of human fibroblasts was achieved a year later (Takahashi *et al.*, 2007; Yu *et al.*, 2007). Since then, iPSCs have been generated from numerous different human cell types, as well as from several other species, including rhesus monkey. These cells were shown to be similar to ESCs in terms of their morphology and gene expression, and were karyotypically normal. Although iPSCs from this first experiment were not germline-competent, they were able to give rise to all three germ layers *in vitro*. Germline-competent iPSCs have since been generated (Okita *et al.*, 2007).

In the past few years, standards for the characterisation of ESCs and iPSCs have become widely accepted (Brivanlou et al, 2003; Maherali & Hochedlinger, 2008). Assessing the success of reprogramming into iPSCs starts with the identification of compact colonies that have distinct borders and well-defined edges, which consist of cells which have a large nucleus, large nucleoli and little cytoplasm. Pluripotent stem cells also have a short cell cycle with a truncated G1 phase, resulting in rapid growth (Smith *et al*, 2009). Molecular markers

which allow the identification of fully reprogrammed iPSCs include the expression of a network of pluripotency genes, including *Oct4*, *Sox2* and *Nanog*, at comparable levels to ESCs; as well as the expression of a number of extracellular markers including SSEA-3/4, TRA-1-81 and TRA-1-60 (Maherali & Hochedlinger, 2008). Alkaline phosphatase has been widely used as a marker for pluripotency, but evidence shows this to be insufficient to identify fully reprogrammed iPSCs, as cells that have been only partially reprogrammed may also stain positive (Chan et al, 2009). iPSCs must also be able to maintain their pluripotency independently of viral transgene expression, and be epigenetically similar to ESCs, with DNA demethylation at the promoters of pluripotency genes, X chromosome reactivation in female cells, and the presence of bivalent chromatin domains at developmental genes (Maherali & Hochedlinger, 2008). In addition, profiling of microRNA expression in pluripotent stem cells revealed a set of miRNAs that are uniquely expressed in pluripotent stem cells (Suh *et al*, 2004) which could potentially be used as an additional test for pluripotency.

Analysis of the pluripotency of mouse iPSCs can be rigorously assessed through the generation of chimaeric mice, which evaluates the potential of iPSCs to contribute to the normal development of adult tissues when injected into the blastocyst. Germline transmission is assessed by the ability of these chimaeras to give rise to offspring derived entirely from iPSCs which have the genomic integrity of the injected iPSC line and the ability to form functional germ cells (Okita *et al.*, 2007). The test with the highest stringency is tetraploid complementation, which involves the injection of iPSCs into blastocysts to measure the ability of iPSCs to direct the normal development of an entire organism. This test has only been accomplished for a small number of iPSC lines, and it appears that not all iPSC lines are able to contribute to all-iPSC mice (Stadtfeld *et al*, 2010). Due to ethical constraints, these tests for pluripotency cannot be carried out for human iPSCs. The most common method of assessing the pluripotency of hiPSCs is differentiation *in vitro*, either as embryoid bodies (compact balls of cells that form loosely organised tissues which resemble the gastrulating embryo) or in monolayer cultures, in order to assess the differentiated cells for markers of all three embryonic germ layers. The teratoma formation assay involves measuring the ability of iPSCs to differentiate into all three germ layers *in vivo* following injection of undifferentiated cells into immunodeficient mice. Although this is the most stringent assay available for testing the pluripotency of human iPSCs, it does not assess whether the iPSCs are able to give rise to every cell in the body, nor does it assess the contribution of the iPSCs to the germline

(Robinton & Daley, 2012). In addition, the ethical considerations of using lab animals for a routine pluripotency test, particularly where numerous iPSC lines need to be assessed, means that there has been a move away from the teratoma formation assay in favour of *in vitro* tests.

Reprogramming to pluripotency is a stochastic process that results in the production of a wide range of colony morphologies, and while some of them may initially appear similar to ESCs, morphology alone cannot be used to identify pluripotent stem cells. It is important to fully characterise cells in order to identify true, fully reprogrammed iPSCs with which to conduct further experiments.

3.2 Methods

3.2.1 Generation of induced pluripotent stem cells

iPSCs were generated using retroviral infection of HFF1 fibroblasts, which were obtained from the UK Stem Cell Bank, as described in section 2.1. Briefly, fibroblasts were infected first with a lentiviral vector carrying the mSlc7a1 retroviral receptor, and subsequently with retroviral vectors carrying the four reprogramming factors (OCT4, SOX2, KLF4 and C-MYC) and GFP. Infected cells were treated with valproic acid and seeded onto iSNL feeder cells. Pluripotent stem cell colonies typically appeared 2-3 weeks following retroviral infection. ESC lines were also obtained from the UK Stem Cell Bank. Full details of these ESC lines can be found in the NIH registry (www.grants.nih.gov/stem_cells/registry/current.htm). A full summary of the lines used in this project and the characterisation carried out on them is shown in table 3.1.

3.2.2 Maintenance of pluripotent stem cells in culture

Stem cells were grown either on mitotically-inactivated feeder cells (iSNLs), which were inactivated as described in section 2.2.1, or on the soluble basement membrane extract Matrigel™, which was prepared as described in section 2.2.3. Stem cells grown on feeders were maintained in knockout serum replacement medium, while stem cells grown on Matrigel™ were maintained in mTeSR1™ defined medium. Stem cells were kept at 37°C,

5% CO2. Undifferentiated stem cells were passaged using manual dissection as described in section 2.2.2, and were cryopreserved as described in section 2.2.5.

3.2.2 Characterisation of pluripotent stem cells

In order to confirm the pluripotency of the stem cell lines used in this study, qRT-PCR was used to analyse expression levels of genes associated with pluripotency and differentiation into the three embryonic germ layers. Briefly, total RNA was isolated using the RNeasy Mini Kit. Reverse transcription of total RNA into cDNA was carried out as described in section 2.4.1. qRT-PCR was carried out using the cycling conditions described in section 2.4.3.

Expression of pluripotency genes was also confirmed using immunocytochemistry. The antibodies used are listed in table 2.16. Cells were either fixed with 4% paraformaldehyde before staining or were stained using a live cell stain, as described in section 2.4.4. Alkaline phosphatase expression levels were visualised using the alkaline phosphatase detection kit. Colonies were visualised using a Nikon Eclipse TS100 inverted microscope.

Cell Line	Type	Official NIBSC Reference	Cell of Origin	Date Generated	Culture Method	Characterisation	Comments
H9	ESC	n/a	n/a	n/a	Matrigel mTeSR-1™	TRA-1-60 live cell stain qRT-PCR for pluripotency genes Immunocytochemistry for pluripotency genes qRT-PCR for exogenous transgenes qRT-PCR for spontaneous differentiation. Alkaline phosphatase staining	Used as positive control in characterisation experiments (chapter 3). Used in differentiation experiments (chapter 4) Used in miRNA microarray and validation experiments (chapter 5) Used in miRNA functional experiments (chapter 6)
H7	ESC	n/a	n/a	n/a	Matrigel mTeSR-1™	TRA-1-60 live cell stain qRT-PCR for pluripotency genes Immunocytochemistry for pluripotency genes qRT-PCR for exogenous transgenes qRT-PCR for spontaneous differentiation	Used in differentiation experiments (chapter 4) Used in miRNA microarray and validation experiments (chapter 5)
H1	ESC	n/a	n/a	n/a	iSNL feeders KOSR medium	TRA-1-60 live cell stain qRT-PCR for pluripotency genes Immunocytochemistry for pluripotency genes qRT-PCR for exogenous transgenes qRT-PCR for spontaneous differentiation Alkaline phosphatase staining	Used in differentiation experiments (chapter 4) Used in miRNA microarray and validation experiments (chapter 5)
MRC5I	iPSC	NIBSC_001	MRC5 fibroblasts	January 2010	iSNL feeders KOSR medium	TRA-1-60 live cell stain qRT-PCR for pluripotency genes Immunocytochemistry for pluripotency genes qRT-PCR for exogenous transgenes qRT-PCR for spontaneous differentiation Alkaline phosphatase staining	Used in differentiation experiments (chapter 4) Used in miRNA microarray and validation experiments (chapter 5)
MRC9G	iPSC	NIBSC_003	MRC9 fibroblasts	February 2010	iSNL feeders KOSR medium	qRT-PCR for pluripotency genes Immunocytochemistry for pluripotency genes qRT-PCR for exogenous transgenes qRT-PCR for spontaneous differentiation Alkaline phosphatase staining	Used in differentiation experiments (chapter 4) Used in miRNA microarray and validation experiments (chapter 5)
HFFA		n/a	HFF1 fibroblasts	March 2011	iSNL feeders KOSR medium	TRA-1-60 live cell stain	Incompletely reprogrammed Passage discontinued after p6

HFFB		n/a	HFF1 fibroblasts	March 2011	iSNL feeders KOSR medium	TRA-1-60 live cell stain	Incompletely reprogrammed Passage discontinued after p5
Wild7	iPSC	NIBSC_020	HFF1 fibroblasts	February 2013	Matrigel mTeSR-1™	TRA-1-60 live cell stain qRT-PCR for pluripotency genes Immunocytochemistry for pluripotency genes qRT-PCR for exogenous transgenes qRT-PCR for spontaneous differentiation Alkaline phosphatase staining	Used in characterisation experiments (chapter 3) Used in miRNA functional experiments (chapter 6)
Wild11	iPSC	NIBSC_021	HFF1 fibroblasts	February 2013	Matrigel mTeSR-1™	TRA-1-60 live cell stain qRT-PCR for pluripotency genes Immunocytochemistry for pluripotency genes qRT-PCR for exogenous transgenes qRT-PCR for spontaneous differentiation Alkaline phosphatase staining	Used in characterisation experiments (chapter 3)
Wild12	iPSC	NIBSC_022	HFF1 fibroblasts	February 2013	Matrigel mTeSR-1™	TRA-1-60 live cell stain qRT-PCR for pluripotency genes Immunocytochemistry for pluripotency genes qRT-PCR for exogenous transgenes qRT-PCR for spontaneous differentiation Alkaline phosphatase staining	Used in characterisation experiments (chapter 3)
Wild17	iPSC	NIBSC_023	HFF1 fibroblasts	February 2013	Matrigel mTeSR-1™	TRA-1-60 live cell stain qRT-PCR for pluripotency genes Immunocytochemistry for pluripotency genes qRT-PCR for exogenous transgenes qRT-PCR for spontaneous differentiation Alkaline phosphatase staining	Used in characterisation experiments (chapter 3)

Table 3.1. Summary of pluripotent stem cell lines used and characterisation carried out.

3.3 Results

3.3.1 Generation of New iPS Cell Lines

iPSCs were generated from HFF1 fibroblasts by retroviral infection of four reprogramming factors: OCT4, SOX2, KLF4, and C-MYC. During reprogramming, a sub-population of cells was also infected with retroviral vector encoding GFP, which allows monitoring of infection efficiency in both PlatE cells and fibroblasts. Infection efficiency must be high in order for fibroblasts to receive adequate levels of all four of the reprogramming factors. Figure 3.1 shows that both PlatE packaging cells and infected fibroblasts express high levels of GFP using these infection conditions. Expression of GFP is maintained after the infected fibroblasts are passaged onto feeder cells.

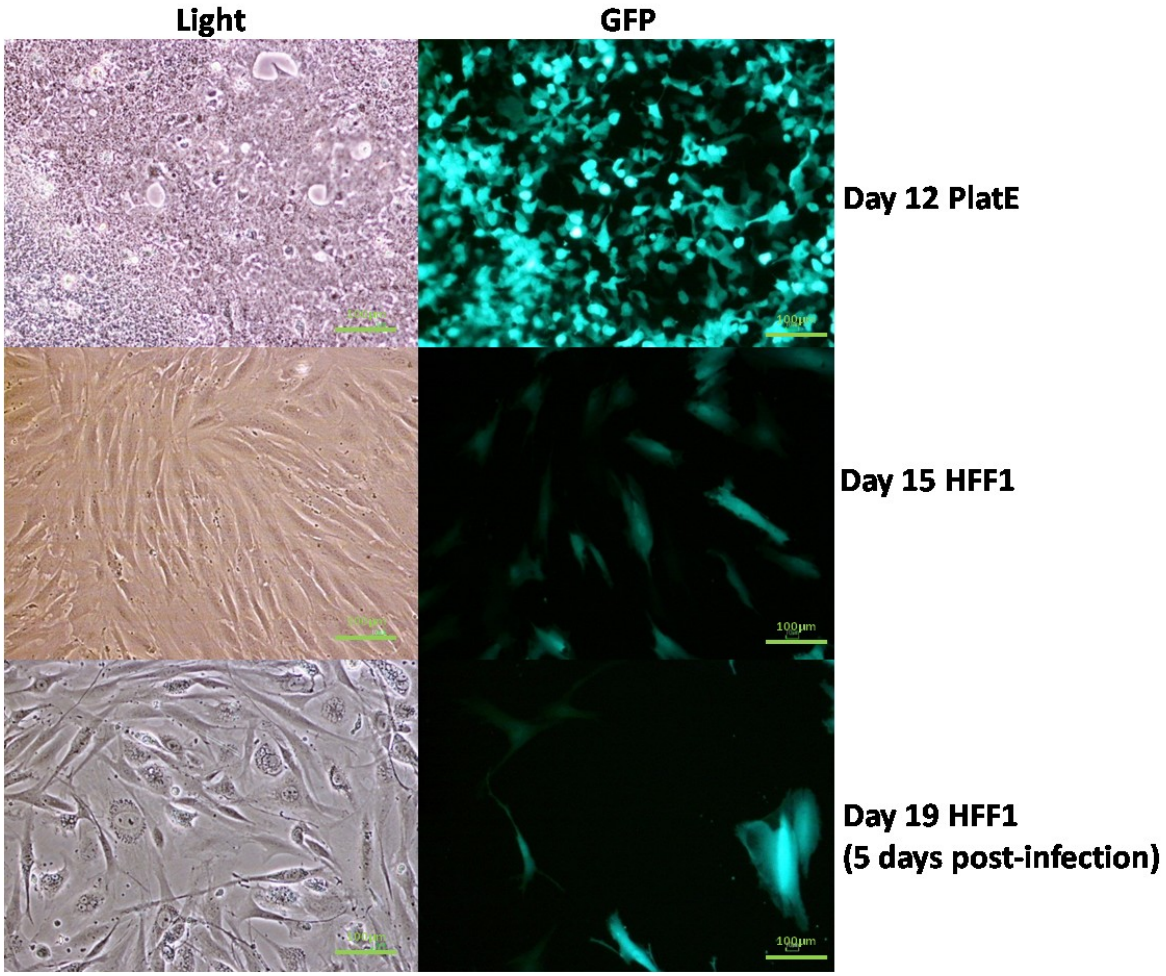

Figure 3.1 GFP expression during reprogramming *shows that both PlatE cells (day 12) and fibroblasts (day 15) have been successfully infected with the GFP plasmid. Expression of GFP is maintained after the infected fibroblasts are passaged onto feeder cells (day 19). 10x magnification, scale bar 100µm.*

During reprogramming, fibroblasts were either infected with both the mouse Slc7a1 retrovirus receptor and the reprogramming factors (mSlc OKSM), reprogramming factors only (neg OKSM) or not infected at all (neg neg). Twelve days post-infection, cells expressing the mSlc receptor which had been infected with the reprogramming factors had changed their morphology and were starting to form colonies that closely resembled pluripotent stem cells, while the cells which did not express the Slc receptor remained fibroblastic in appearance (figure 3.2, panel A). This demonstrates the importance of infecting the fibroblasts with the mouse retrovirus receptor during the reprogramming process, as it allows subsequent infection with the retroviral vectors containing the reprogramming factors. The cells that formed stem cell-like colonies could be cultured and passaged onto new feeders (figure 3.2, panel B) and subsequently maintained in culture as described in section 2.2.

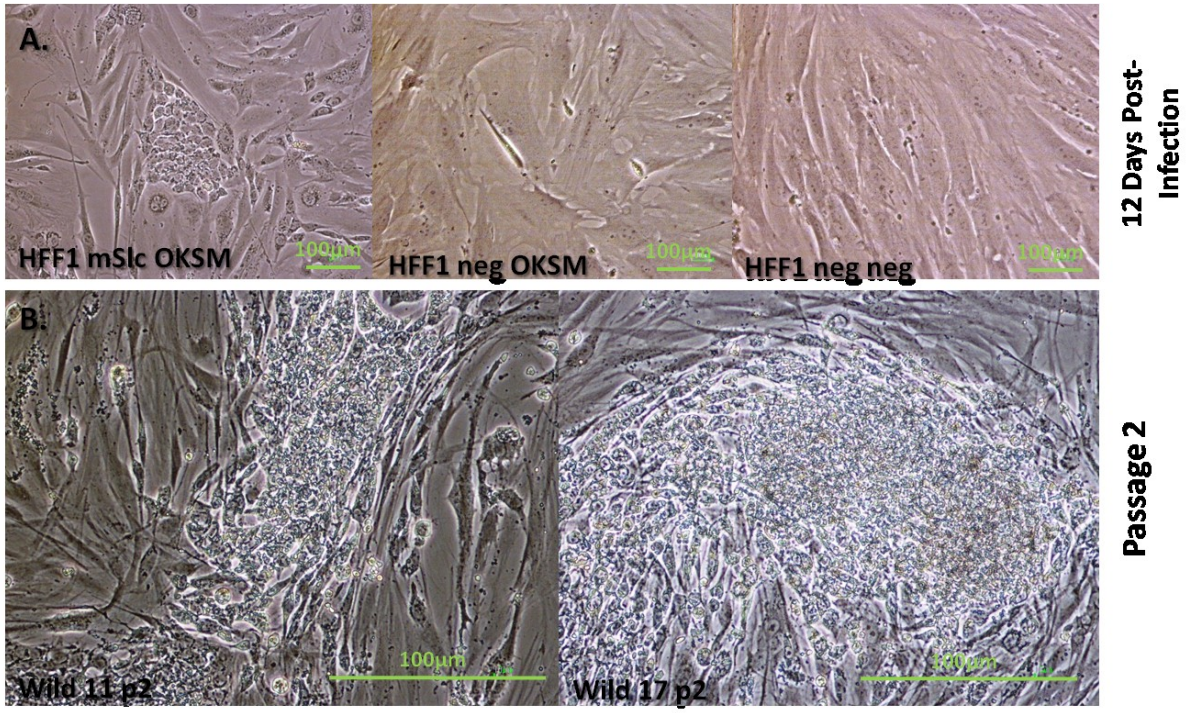

Figure 3.2. Formation of stem cell-like colonies after transfection of HFF1 fibroblasts with reprogramming factors. *A) Fibroblasts expressing the mSlc receptor and infected with the reprogramming factors (mSlc OKSM) formed stem cell-like colonies, while those infected with only the reprogramming factors (neg OKSM) or not infected at all (neg neg) did not. B) Stem cell-like colonies could be passaged. 10x magnification, scale bar 100µm.*

3.3.2 Characterisation of New iPSC Lines

A summary of the pluripotent stem cell lines used in subsequent experiments and the characterisation carried out for each line is shown in table 3.1.

Characterisation of Newly Reprogrammed Lines Identifies Incomplete Reprogramming

iPSCs cannot be identified by their morphology alone, since reprogramming of somatic cells results in a wide range of cells that are at different stages of the reprogramming process. Some are fully reprogrammed true iPSCs, but others are only partially reprogrammed and do not fulfil all of the criteria for pluripotency. This was demonstrated by an early reprogramming experiment, in which fibroblasts were reprogrammed and formed colonies that strongly resembled iPSC colonies (HFFA and HFFB lines), but which on further investigation were found to be incompletely reprogrammed.

Immunocytochemistry

The first check for pluripotency is live cell staining to investigate expression of TRA-1-60, a surface marker characteristic of pluripotent stem cells. Since this marker is expressed on the cell surface, the cells do not need to be fixed and permeabalised prior to staining, allowing continued culture and passage of the cells after staining. This is particularly useful at early stages of characterisation, when numbers of cells available for analysis are limited. Cells were cultured as described in section 2.2.2 for at least 5 days before staining with TRA-1-60 StainAlive dye as described in section 2.4.4. Figure 3.3 shows high levels of TRA-1-60 expression in an ESC line (H1), but no expression in the newly reprogrammed lines HFFA and HFFB, indicating that these cells are not true pluripotent iPSCs.

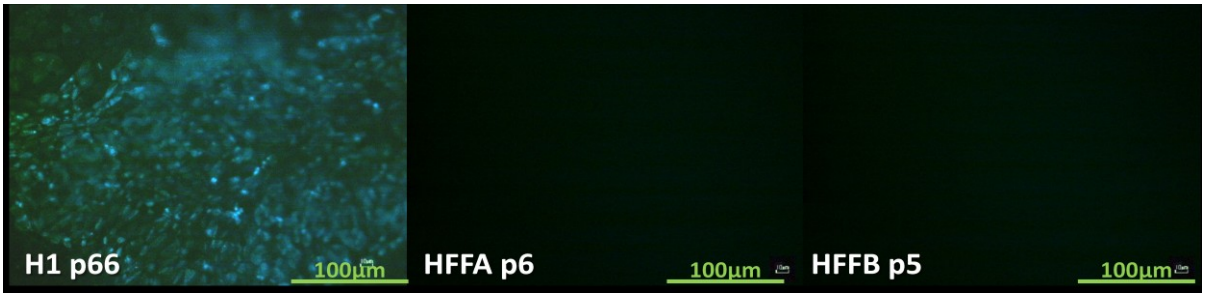

Figure 3.3. Immunocytochemistry for TRA-1-60 expression using live cell stain. *10x magnification, scale bar 10µm*

qRT-PCR

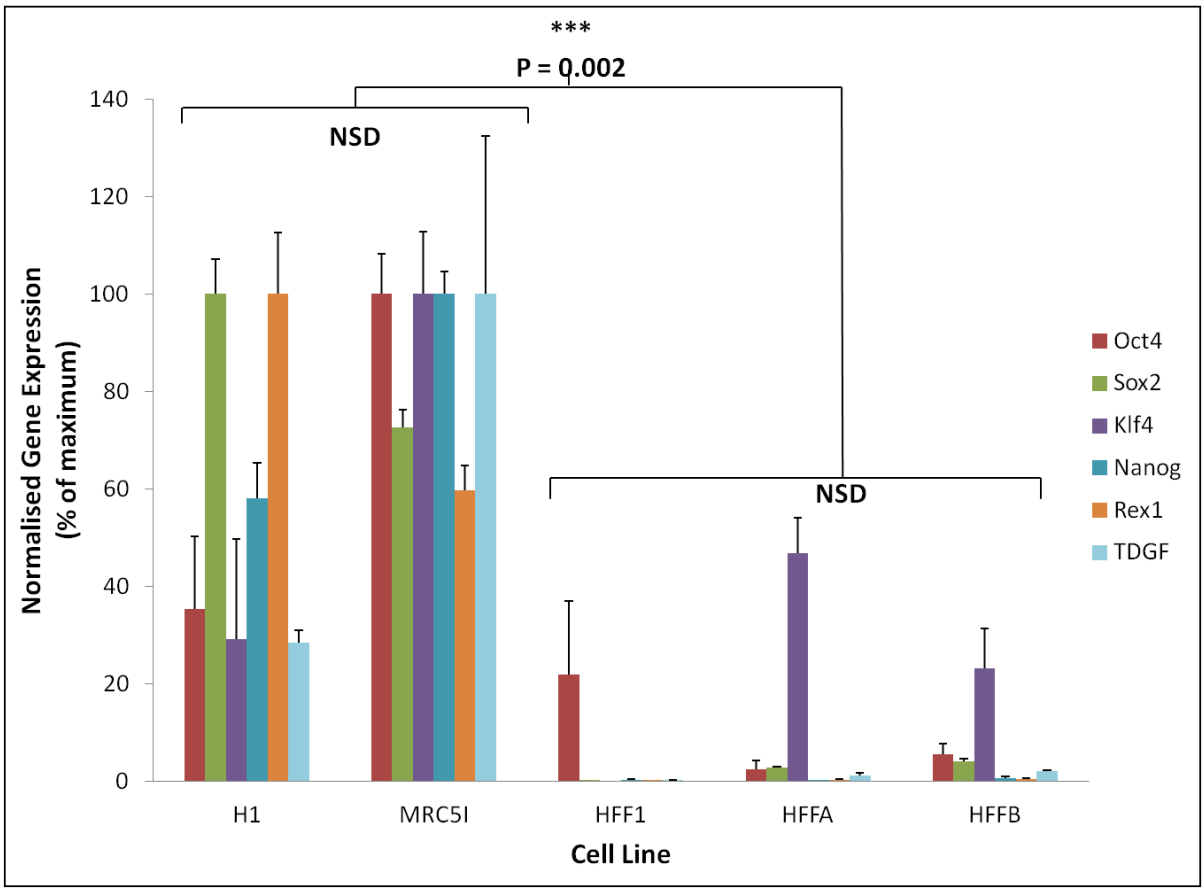

Figure 3.4. qRT-PCR showing expression levels of key pluripotency genes *in an ESC line (H1), an established iPSC line (MRC5I), two newly reprogrammed cell lines (HFFA and HFFB), and the starting fibroblast cell line (HFF1).*

Cells were further characterised using qRT-PCR as described in section 2.4 to investigate expression levels of key pluripotency genes, including the four exogenous reprogramming factors *Oct4*, *Sox2*, *Klf4* and *C-Myc*, as well as the additional pluripotency-associated genes *Rex1*, *Nanog*, and *Tdgf*. Analysis of variance with Fisher's *a priori* test was performed to determine significant differences between groups.

Figure 3.4 shows that the partially reprogrammed cells (HFFA and HFFB) both have significantly ($p = 0.002$, $n = 3$) lower levels of expression of pluripotency genes than either an ESC line (H1) or established iPSC line (MRC5I), which both have similar high expression levels of pluripotency genes. The expression levels of pluripotency genes in the newly

reprogrammed cells are not significantly different from those in the starting fibroblast population (HFF1), showing that they have failed to activate endogenous pluripotency genes.

Characterisation of Completely Reprogrammed iPSCs

Cells generated in a subsequent reprogramming experiment were fully characterised and found to fulfil pluripotency criteria, and these were used in subsequent experiments. All iPS cell lines used in these experiments had a stable, ESC-like morphology and growth pattern (figure 3.5). They could be expanded in culture for more than 10 passages, and made viable frozen stocks.

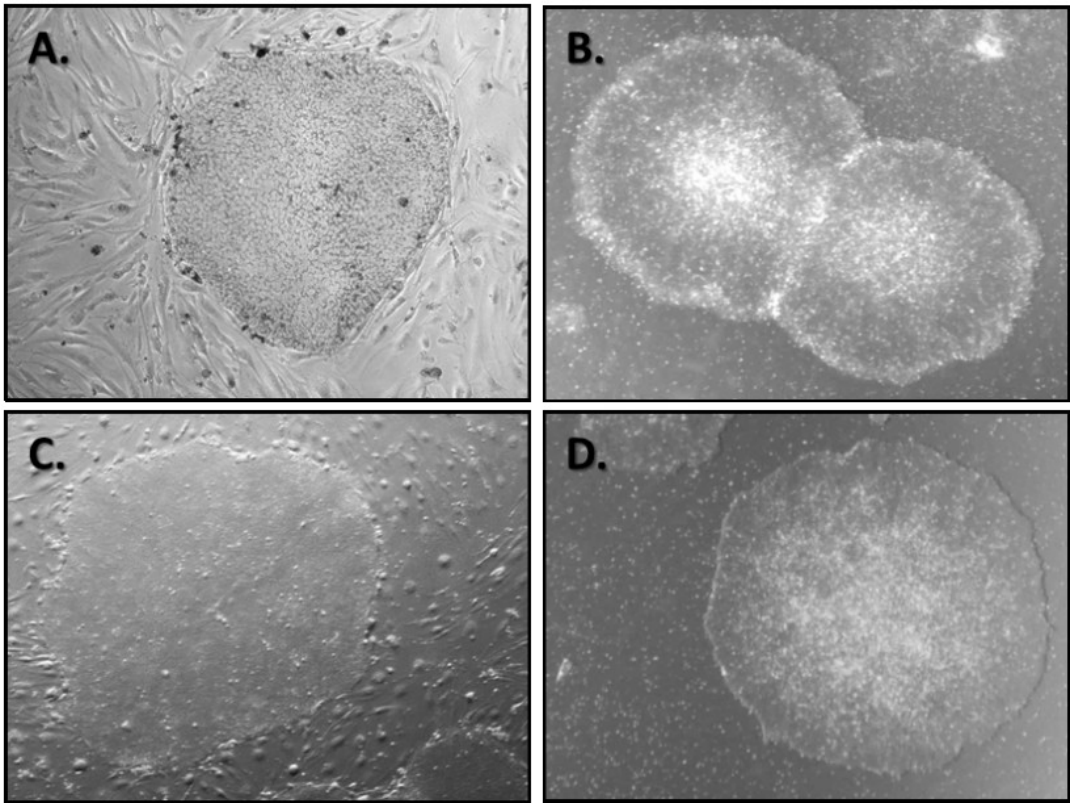

Figure 3.5 Morphology of stem cell colonies. *A) iPSC colony maintained on iSNL feeders, B) iPSC colony maintained on Matrigel™, C) ESC colony maintained on iSNL feeders, D) ESC colony maintained on Matrigel™.*

Immunocytochemistry

Cells were characterised at an early stage (p2) by TRA-1-60 live cell staining to determine whether the colonies generated were true pluripotent stem cell colonies. Figure 3.6 shows high levels of TRA-1-60 expression in an ESC line (H1), two established iPSC lines (MRC5I and MRC9G) and three newly reprogrammed cell lines (Wild7, Wild12 and Wild17). This

indicates that these colonies are likely to be fully reprogrammed. TRA-1-60^{+} colonies were selected for further passaging, while TRA-160^{-} colonies were discarded.

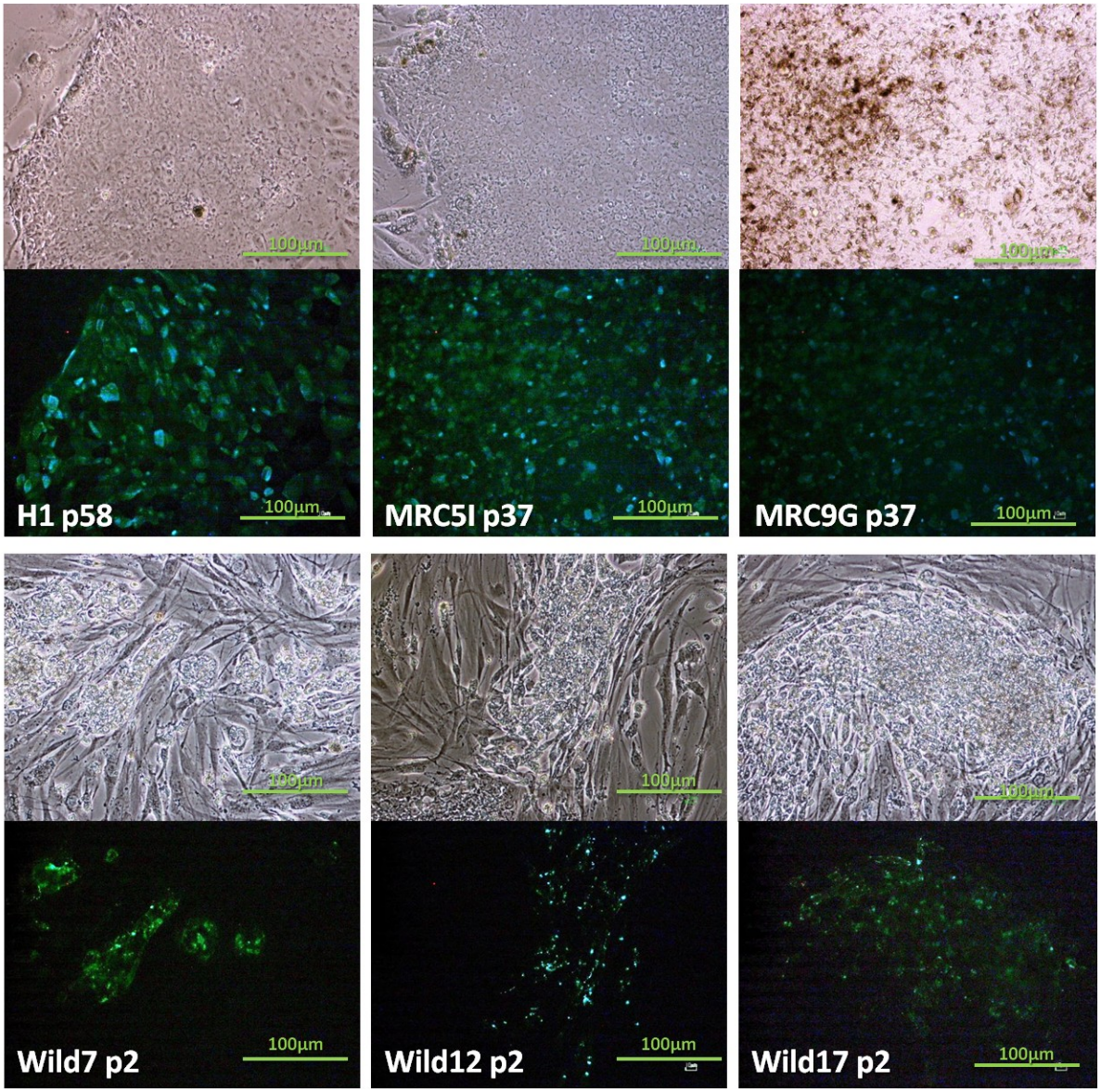

Figure 3.6. Immunocytochemistry for TRA-1-60 expression *using live cell stain in an ESC line (H1), two established iPSC lines (MRC5I and MRC9G), and three newly reprogramed iPSC lines (Wild7, Wild12 and Wild17).10x magnification, scale bar 100µm.*

Undifferentiated pluripotent stem cells are also characterised by high levels of alkaline phosphatase, an enzyme that is ubiquitously expressed in undifferentiated human pluripotent stem cells, embryonic carcinoma cells and embryonic germ cells (Thomson et al, 1998). It can be stained using a Napthol/Fast Red Violet staining solution. Figure 3.7 shows that both an ESC line (H1) and two iPSC lines (MRC9G and MRC5I) all have high levels of alkaline phosphatase, while a differentiated stem cell colony does not.

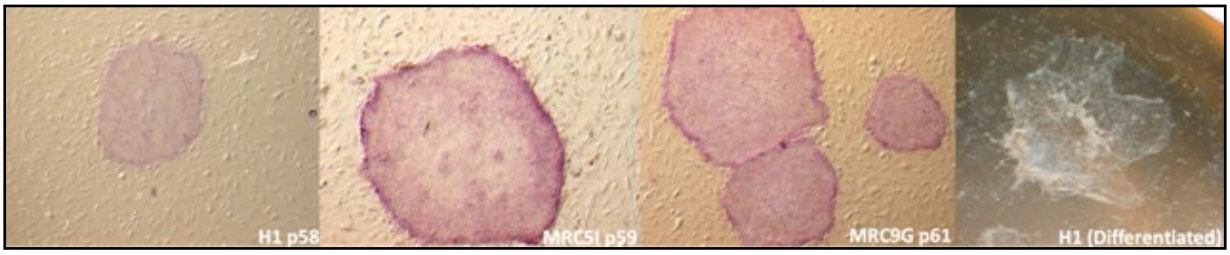

Figure 3.7. Alkaline phosphatase staining *in an ES cell line (H1) and two iPS cell lines (MRC5I and MRC9G) grown on iSNL feeders. 2x magnification.*

Once adequate cell numbers had been obtained, immunocytochemistry for several other pluripotency markers could be carried out on cells that had been fixed and permeabalised, as described in section 2.4.4. Control dishes were prepared at the same time, using both a secondary antibody only and isotype control (data not shown). Cells were counter-stained with Hoescht DNA stain. The cells were visualised using a Nikon Eclipse TS100 inverted microscope. Figures 3.8, 3.9, 3.10 and 3.11 show immunocytochemistry for several important pluripotency markers in four newly reprogrammed iPSC lines, including the transcription factors OCT4 and NANOG, as well as two stem cell-specific surface markers, SSEA-3 and TRA-1-81. All four cell lines show high expression of all of these markers at the protein level, confirming their pluripotency.

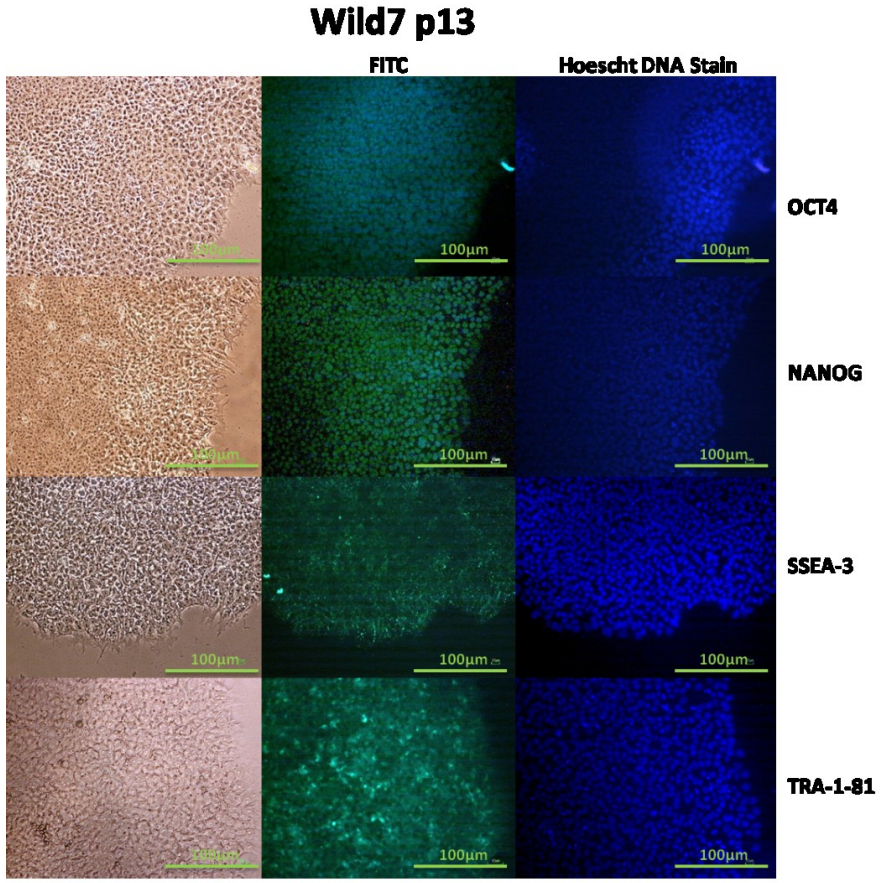

Figure 3.8. Expression of the pluripotency markers OCT4, NANOG, SSEA-3 and TRA-1-81 in a newly reprogrammed iPSC line*. 10x magnification, scale bar 100µm.*

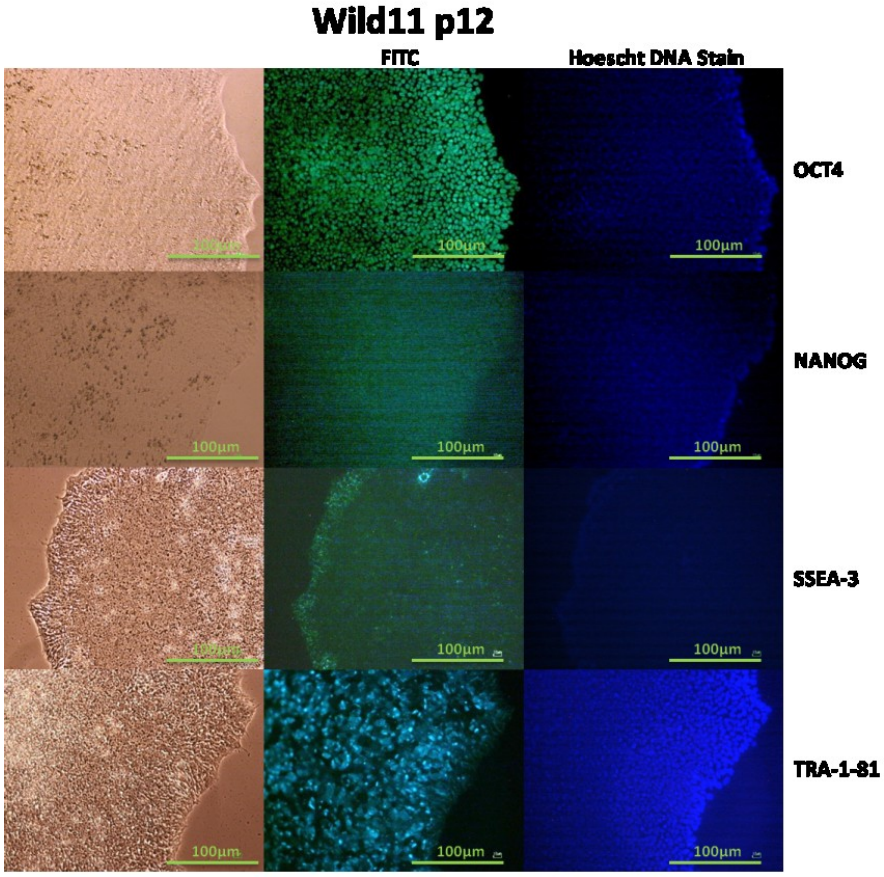

Figure 3.9. Expression of the pluripotency markers OCT4, NANOG, SSEA-3 and TRA-1-81 in a newly reprogrammed iPSC line*. 10x magnification, scale bar 100µm.*

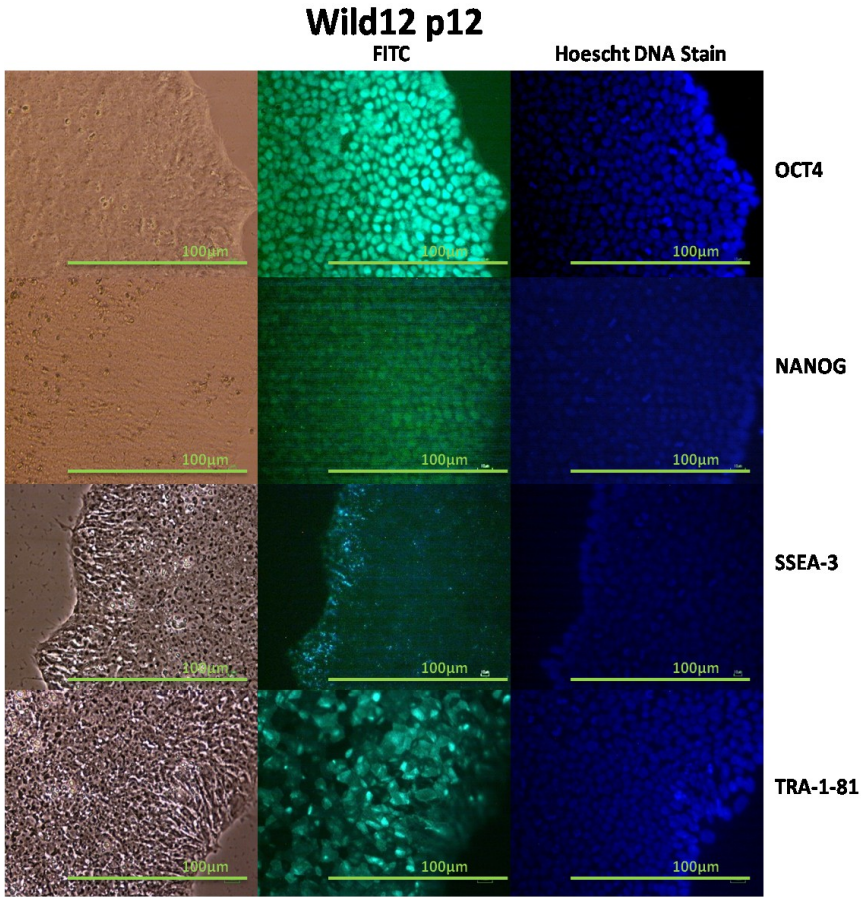

Figure 3.10. Expression of the pluripotency markers OCT4, NANOG, SSEA-3 and TRA-1-81 in a newly reprogrammed iPSC line*. 10x magnification, scale bar 100µm.*

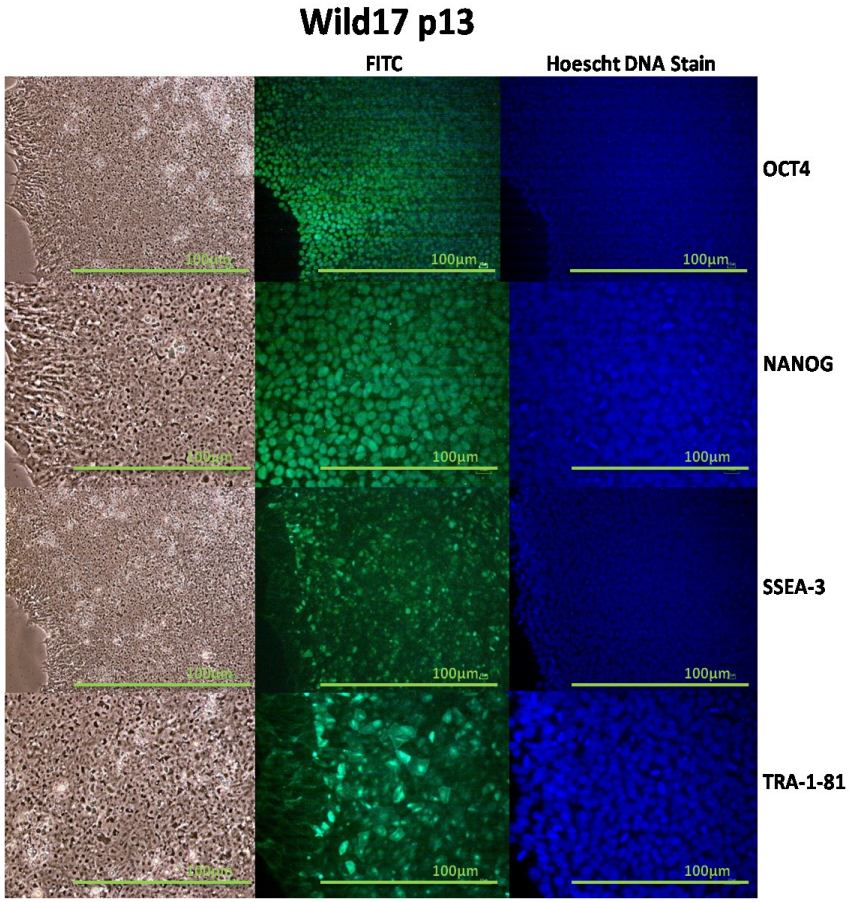

Figure 3.11 Expression of the pluripotency markers OCT4, NANOG, SSEA-3 and TRA-1-81 in a newly reprogrammed iPSC line*. 10x magnification, scale bar 100µm.*

qRT-PCR

Expression levels of key pluripotency genes were investigated in undifferentiated ESCs and iPSCs using qRT-PCR as described in section 2.4.3. Analysis of variance with Fisher's *a priori* test was performed to determine significant differences between groups.

Figure 3.12 shows that all four newly reprogrammed cell lines (Wild7, Wild11, Wild12 and Wild17) have significantly ($p < 0.005$, $n = 3$) higher expression of several pluripotency-associated genes, including *Oct4, Sox2, Klf4, C-Myc, Nanog, Rex1,* and *Tdgf,* than the starting fibroblast line (HFF1). Expression levels in new iPSC lines are not significantly different from an ESC line (H9), confirming the pluripotency of these cells.

Another requirement for a cell line to be considered pluripotent is silencing of the exogenous transgenes used in reprogramming (Okita et al, 2007; Wernig et al, 2007), demonstrating that the cell line is capable of up-regulating endogenous pluripotency gene expression and therefore maintaining pluripotency independently of the exogenous reprogramming factors. Expression levels of the factors used in reprogramming were investigated in HFF1 fibroblasts (HFF1 uninfected), HFF1 cells that had received both the mouse retrovirus receptor and the reprogramming factors (HFF1 mSlc OKSM), HFF1 cells that had received the mouse retrovirus receptor only (HFF1 mSlc neg), four new iPSC lines (Wild7, Wild11, Wild12 and Wild17) and an ESC line (H9). Figure 3.13 shows that there is little or no expression of exogenous OCT4, SOX2, KLF4 and C-MYC in either the untransfected cells or the cells that received the receptor alone, but levels in HFF1 cells, 12 days after infection with the reprogramming factors, are significantly higher ($p < 0.001$, $n = 3$). The transgenes are then silenced in the newly reprogrammed iPSC lines at an early passage (cells tested at p12). There is also no expression of the transgenes in an ESC line (H9).

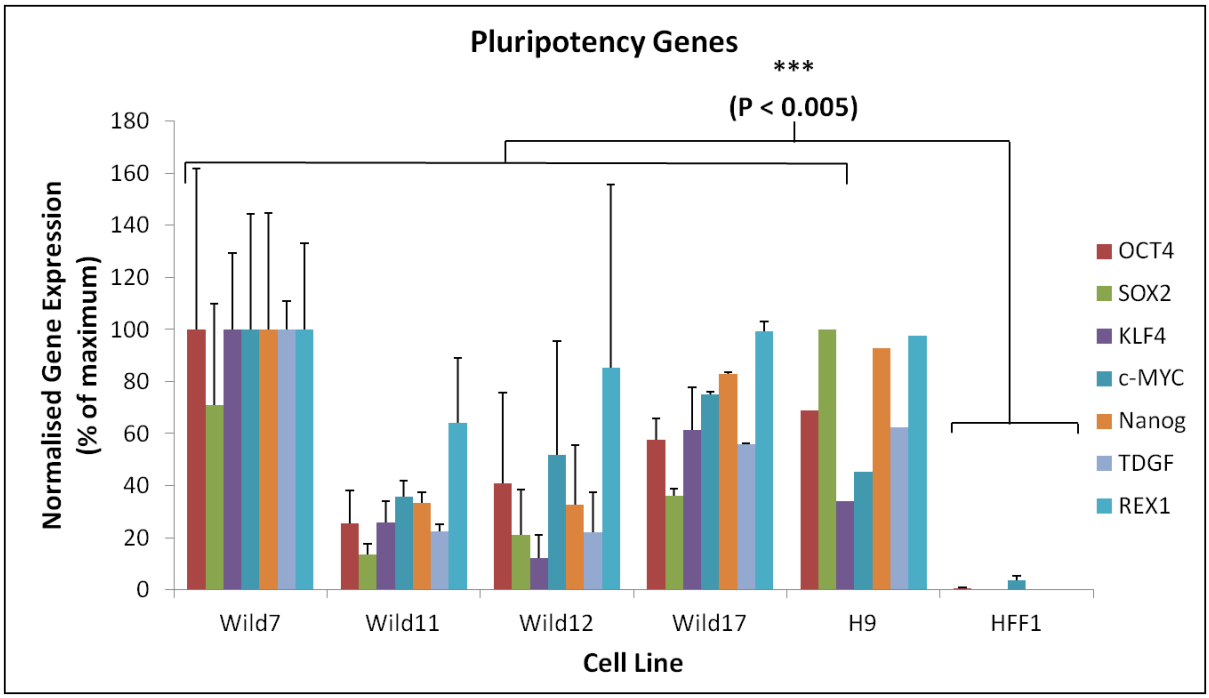

Figure 3.12. qRT-PCR showing expression levels of key pluripotency genes *in an ESC line (H9), four newly reprogrammed iPSC lines (Wild7, Wild11, Wild12 and Wild17, and the starting fibroblast cell line HFF1).*

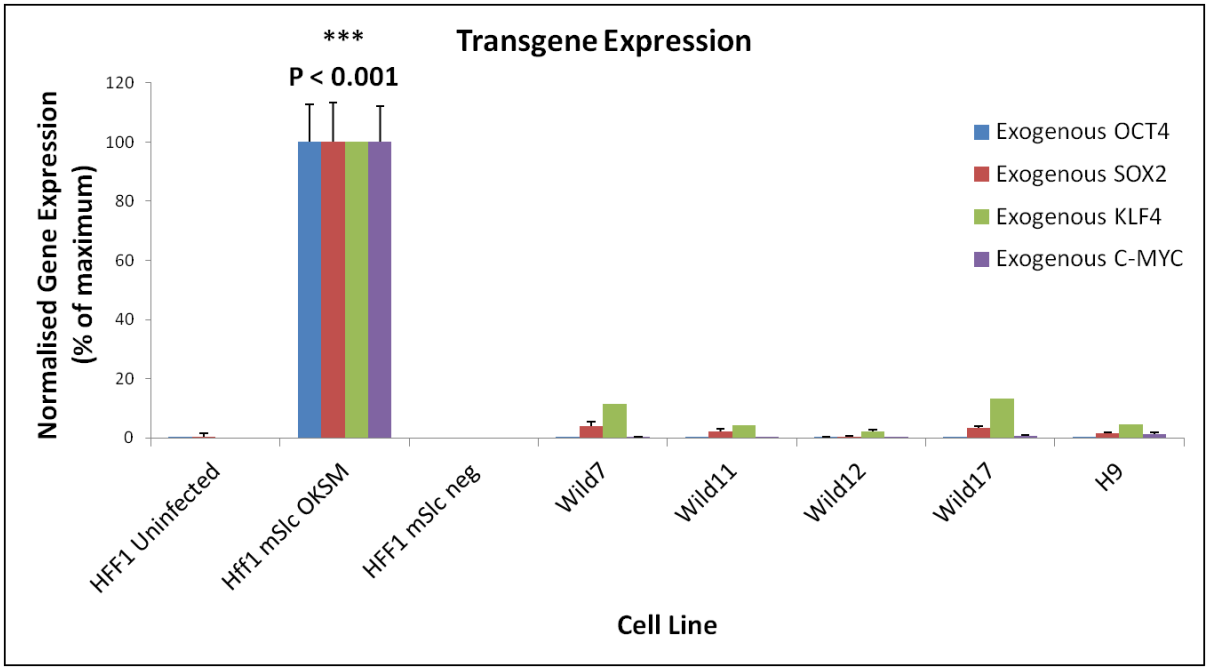

Figure 3.13. qRT-PCR showing expression levels of exogenous genes *used in reprogramming in uninfected fibroblasts (HFF1 uninfected), fibroblasts infected with both the mSlc receptor and the OKSM reprogramming factors (HFF1 mSlc OKSM), fibroblasts infected with both the mSlc receptor only (HFF1 mSlc neg), four newly reprogrammed iPSC lines (Wild7, Wild11, Wild12 and Wild17), and an ESC line (H9).*

Cell lines were also characterised by their ability to differentiate into cells derived from each of the three embryonic germ layers, a defining characteristic of pluripotent stem cells. Cells were allowed to spontaneously differentiate over a period of 7 days before RNA isolation.

Figures 3.14, 3.15 and 3.16 show that four newly generated iPSC lines (Wild7, Wild11, Wild12 and Wild17) as well as an ESC line (H9) were able to give rise to all three germ layers, as shown by upregulation of genes characteristic of endoderm (*Sox17*, *Cxcr4*, *Foxa2* and *Cer1*), mesoderm (*Brachyury*) and ectoderm (*Nestin*, *Pax6*, and *Tubulin*) in spontaneously differentiated cells compared to undifferentiated cells, confirming their pluripotency and ability to differentiate *in vitro*. Expression of these markers was significantly increased in spontaneously differentiated cells compared to undifferentiated cells, which showed very low levels of expression of these genes which is likely due to the presence of some differentiated cells in the colonies.

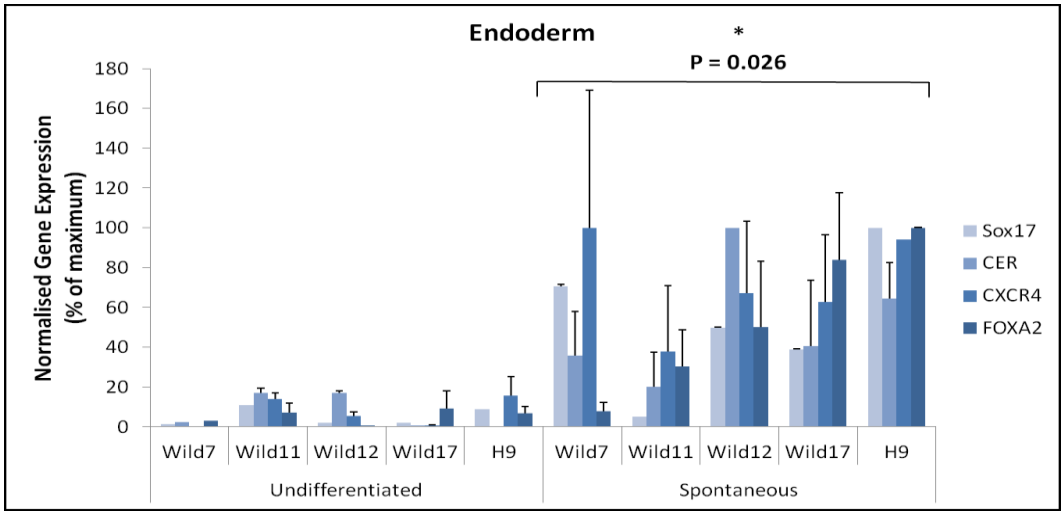

Figure 3.14.qRT-PCR data showing expression of marker genes characteristic of endoderm

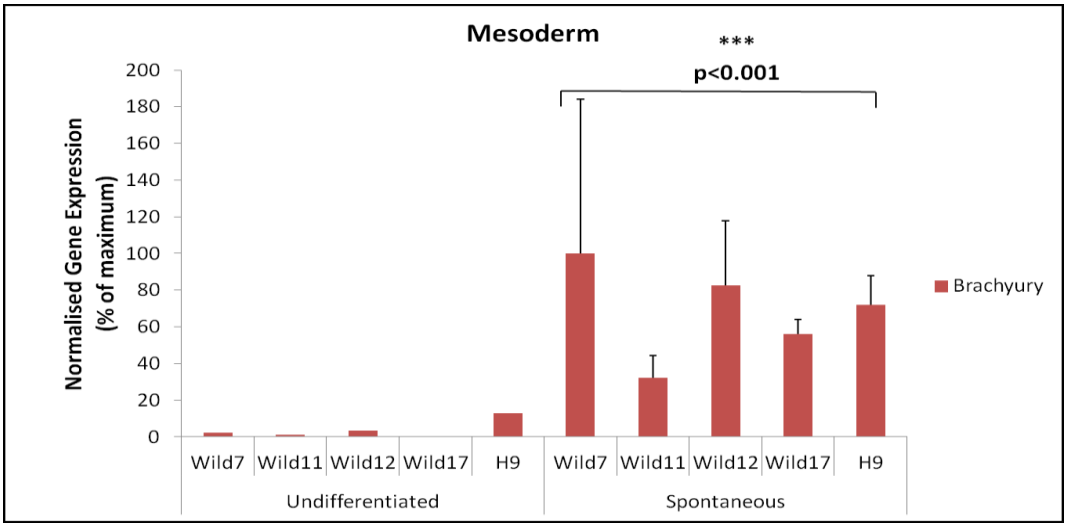

Figure 3.15. qRT-PCR data showing expression of a marker gene characteristic of mesoderm.

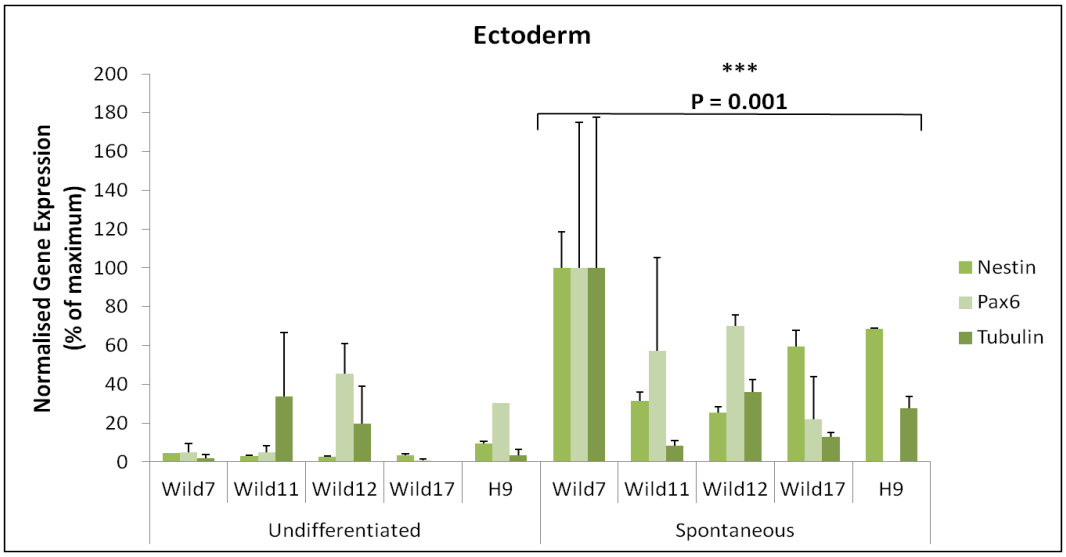

Figure 3.16. RT-PCR data showing expression of marker genes characteristic of ectoderm.

3.4 Discussion

Initial reprogramming experiments resulted in the production of cells which closely resembled iPSCs in terms of their morphology and growth characteristics, and which could be maintained in culture over a number of passages, but which, when further characterised, showed inadequate expression of pluripotency markers. These cells are likely to represent a population of incompletely reprogrammed cells which have been previously described (Meissner *et al*, 2007; Chan *et al*, 2009). These cells can be maintained in a relatively stable state, and can express some stem cell markers such as SSEA-1, and can even occasionally go on to give rise to fully reprogrammed cells, either spontaneously or through treatment with DNA demethylating agents (Meissner *et al*, 2008). However in many cases these cells would never be able to generate fully pluripotent cells. This shows the importance of fully characterising new iPSC lines based on molecular markers rather than cell morphology, as these cells are likely to behave very differently to true iPSCs.

One of the reasons for the appearance of these incompletely reprogrammed populations may be due to the reprogramming method used, which relies on delivery of the reprogramming factors in four separate viral vectors. This makes it difficult to control the dose of each reprogramming factor that each individual cell will receive, and while some will receive an adequate dose of all four reprogramming factors and go on to be fully reprogrammed into iPSCs, others may receive an inadequate dose of the reprogramming factors to allow them to completely alter their cell fate to that of a pluripotent cell, but which may be sufficient to allow partial reprogramming. Further complicating matters is the fact that excess expression levels of OCT4 (Niwa *et al*, 2000) or SOX2 (Kopp *et al*, 2008) are actually detrimental to the maintenance of pluripotency. High levels of C-MYC and KLF4 together with low levels of SOX2 and OCT4 may lead to the generation of rapidly growing non-iPSCs which have been transformed rather than reprogrammed (Yamanaka, 2007). Although a separate vector encoding GFP was used in these experiments to monitor efficiency of delivery of the reprogramming factors, this is a crude method that simply ascertains whether transfection is working, and not the dose of each reprogramming factor that the cells receive. However, this method is still one of the most efficient methods of reprogramming, which is why it was chosen for this study.

Early reprogramming studies relied on a selection system where iPSCs were isolated by antibiotic selection using a neomycin resistance gene inserted into a pluripotency gene locus.

When cells were treated with neomycin, only cells that had reactivated ESC-specific genes could survive. Takahashi & Yamanaka (2006) used selection for *Fbx15*, a gene that is specifically expressed in pluripotent stem cells, but that is not essential for maintenance of pluripotency. However, although these cells expressed molecular markers of pluripotency, and were able to give rise to all three embryonic germ layers as assessed by teratoma formation, these cells were not able to give rise to live chimeras, indicating restricted developmental potential compared to ESCs. Subsequent studies used selection for genes essential to pluripotency, such as *Nanog* and *Oct4* (Maherali et al, 2007; Wernig et al, 2007), to select reprogrammed cells, and this was more successful in producing cells that could be transmitted through the germline, although all-iPSC mice were not generated through tetraploid complementation until later (Okita et al, 2007). While it was initially thought that iPSCs could not be isolated without some sort of selection, a study later showed that unmodified fibroblasts could be reprogrammed, although the presence of partially reprogrammed cells was noticed (Meissner et al, 2007).

Another technique to eliminate partially reprogrammed colonies is to use doxycycline-inducible reprogramming factors that can be turned on and off by the addition and removal of doxycycline from the medium (Wernig *et al*, 2008). When doxycycline is removed from the medium, fully reprogrammed cells that have started to express pluripotency factors endogenously, and therefore do not require expression of the reprogramming factors, can continue to grow, but partially reprogrammed cells do not. “Secondary iPSCs” can then be generated from the differentiated progeny of these iPSCs simply through the addition of doxycycline, allowing high-efficiency, high-throughput reprogramming (Maherali *et al*, 2008) However, it is worth noting that cells that have been manipulated in this way, while useful for improving reprogramming techniques, are unlikely to ever be considered clinically relevant. The same group later generated “reprogrammable mice”, in which genes for pluripotency are placed under the control of doxycycline and inserted into a genomic locus which is constitutionally active in all tissues of the mouse. These mice allow iPSCs to be generated from any tissue by simply isolating the cells and adding doxycycline to the growth medium (Hanna *et al*, 2008). A subsequent study described transgenic mice which carried all possible combinations of the four reprogramming factors, allowing further elucidation of the role of each factor in reprogramming, as well as allowing screening for chemicals that can replace the reprogramming factors (Markoulaki *et al*, 2009).

Reprogramming methods which use a single polycistronic vector to deliver all four reprogramming factors (Sommer et al, 2009; Carey et al, 2009) may reduce the number of incompletely reprogrammed cells, as all the cells which are successfully transfected will receive all the factors, making it more likely that the cells will be fully reprogrammed. This method also reduces the number of integration sites of the viral vector into the host genome, as this can also affect reprogramming efficiency. The method used in our study results in random integration of viral vectors into the host cell genome, usually at several different loci. Genomic integration has been shown to alter gene function (Kustikova *et al*., 2005), and random integration may influence the molecular signature of iPSCs by disrupting regulatory regions in the genome. Transcriptional analysis has showed that transgene expression from incompletely silenced viral vectors could perturb global gene expression in iPSCs (Soldner *et al*, 2009). In addition, vectors which integrate into the genome are retained when the cells divide. For this reason, non-integrating methods of reprogramming (summarised in table 1.1) may provide an advantage over early retroviral and lentiviral reprogramming techniques in terms of the generation of fully reprogrammed cells, in addition to their advantages in terms of safety, which is an important consideration for generating clinically-relevant iPSC lines.

The presence of these incompletely reprogrammed cells show the importance of fully characterising new iPSC lines before they are used in subsequent experiments to ensure the quality of the lines used. Guidelines for the characterisation of iPSCs have now been well-established (Maherali & Hochedlinger, 2008), and are largely based on guidelines first used for the characterisation of ESCs (Brivanlou *et al*, 2003). The first step in identifying cells that have been successfully reprogrammed is observation of morphological changes, although as these results show, this is an inadequate way of assessing whether cells have been fully reprogrammed, and further characterisation of molecular and cellular markers is needed. These markers include the transcription factors OCT4 and NANOG, which are essential for the maintenance of pluripotency. The results described here showed that these genes are expressed in newly reprogrammed iPSCs both at the mRNA and the protein level. Expression of several other pluripotency-associated genes, including *Sox2*, *Klf4*, *c-Myc*, *Tdgf* and *Rex1* were also assessed by qRT-PCR, and were found to be expressed at high levels in new iPSC lines as well as in an ESC line, confirming the pluripotency of these cell lines. These markers were chosen as they are known to be expressed in pluripotent stem cells, but ideally a high-throughput method would be used to screen for expression of a larger panel of pluripotency-associated genes. The presence of the stem cell surface antigens TRA-1-60, TRA-1-81 and

SSEA-3 was also assessed by immunocytochemistry, and all markers were expressed in both iPSCs and ESCs, confirming their pluripotency. Expression levels of pluripotency genes were analysed throughout the culture of these cells, and figure 3.17 shows that two iPSC lines (MRC5I and MRC9G) maintained expression of these genes at later passages, with no significant difference in the expression of OCT4 and NANOG in either cell line ($p = 0.378$ for MRC5I, $p = 0.931$ for MRC9G) at any passage number.

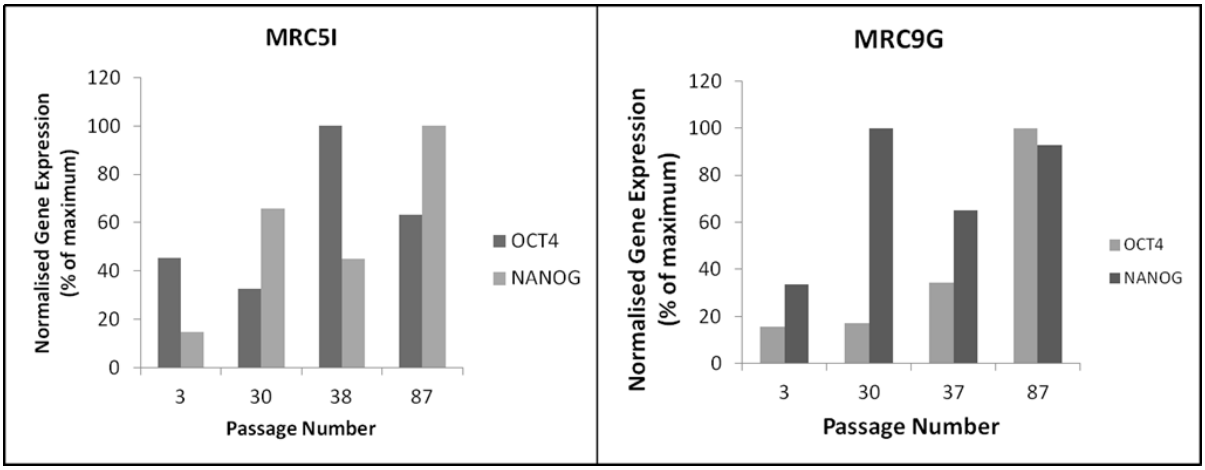

Figure 3.17 qRT-PCR data showing the expression of the pluripotency genes OCT4 and NANOG in two iPSC lines (MRC5I and MRC9G) at several different passages.

The newly reprogrammed iPSCs were assessed for silencing of the transgenes used in reprogramming, an essential criteria of truly reprogrammed cells. Transgene expression was high in fibroblasts at day 17 of the reprogramming process, but was subsequently silenced in iPSCs at early passages (~p10), showing that these iPSCs are capable of self-renewal in culture independent of transgene expression. Expression of endogenous *Oct4*, *Sox2*, *Klf4* and c-*Myc* was found to be high in iPSC lines, suggesting that they have successfully reactivated the network of gene expression necessary for the maintenance of pluripotency.

Finally, an *in vitro* assessment of the differentiation potential of iPSCs was carried out. Cells were allowed to spontaneously differentiate for 7 days and the expression of germ layer-specific genes was assessed. All iPSC lines and an ESC line showed expression of genes representative of each embryonic germ layer, confirming their pluripotency and ability to differentiate.

The most stringent test of pluripotency available for human pluripotent stem cells is the teratoma assay. However, this is an expensive and time-consuming assay, and there has been debate over how useful it actually is, with many pointing out the qualitative nature of this assay, and the fact that it does not accurately assess whether the iPSCs are able to give rise to every cell in the body, nor does it assess the contribution of the iPSCs to the germline: many mouse iPSC lines are able to give rise to teratomas containing cells derived from all three germ layers, but are not able to contribute to all-iPSC mice. In terms of this study, it was decided that an *in vitro* differentiation assay was sufficient to determine the pluripotency of these cell lines, as the cells were subsequently to be used for investigation of *in vitro* differentiation. There has been a shift in recent years towards an increased focus on *in vitro* testing of new iPSC lines: the TaqMan® hPSC Scorecard developed by Life Technologies aims to allow researchers to fully characterise their pluripotent stem cell lines in a single PCR-based assay, which allows measurement of gene expression in order to determine whether cell lines are able to give rise to all three embryonic germ layers. In addition, PluriTest (www.pluritest.org), an open access resource, allows researchers to compare their gene expression data with a large panel of established pluripotent stem cell lines (Muller et al, 2008). These methods are particularly relevant for lines which are generated primarily for research use, where the requirement for pluripotency is less stringent than in cells which could potentially be used in a therapeutic setting.

There are several other considerations to be taken into account before iPSCs can be considered for clinical applications. These include the starting cell type, the reprogramming factors used, and the method of delivering those reprogramming factors. In addition to these considerations, there are still numerous technical difficulties standing in the way of generating clinical-grade iPSCs, such as the requirement for GMP conditions and reagents, and the need for large-scale production. These factors are discussed further in chapter 1.

In conclusion, this chapter demonstrates the stochastic nature of the reprogramming process, and the requirement for robust characterisation of new iPSC lines. In addition, this chapter demonstrates the successful reprogramming of HFF1 fibroblasts into iPSCs for use in future experiments.

Chapter 4

Differentiation of Pluripotent Stem Cells into Definitive Endoderm

Contents

List of Tables

List of Figures

Chapter 4: Differentiation of Pluripotent Stem Cells into Definitive Endoderm

4.1 Introduction

Transplantation of both whole pancreas and islets has shown proof of principle for cell replacement therapy to treat T1DM. There are numerous different sources of tissue that could potentially be used, but there are certain requirements in order for them to be suitable for transplantation: first, they must be able to synthesise and store insulin, and release sufficient amounts in response to various secretagogues, including circulating glucose. Second, any source of replacement β-cells will also need to be able to generate large numbers of cells *in vitro* to meet the requirement of transplantation therapy, but this proliferative capacity must be tightly regulated to avoid post-transplantation expansion. Thirdly, the replacement β-cells must show persistent engraftment and survival. Lastly, they must satisfy safety concerns. Much of the focus has been on pluripotent stem cells, as they have the capacity to self-renew indefinitely in culture and can potentially give rise to any cell in the body, but directing stem cells to differentiate *in vitro* into a pancreatic lineage has proved challenging.

To date, the most successful protocols for generating insulin-expressing cells *in vitro* have been those which recapitulate the signalling pathways which are important during *in vivo* pancreatic development, including TGF-β, Wnt, Hedgehog and Notch signalling (see section 1.3). The first step in the differentiation of pluripotent stem cells into β-cells is the production of definitive endoderm (DE), and this has been shown to be a critical stage in the differentiation process, demonstrating the need for protocols which generate DE at a high efficiency (Courtney et al., 2010). One of the most widely adopted protocols for the differentiation of ESCs into DE was published in 2005, and involved treatment with high concentrations of Activin A and low serum levels, resulting in production of a population containing approximately 80% SOX17-expressing cells. Other markers of DE, including FOXA2, CXCR4, and CER1 were also upregulated, while SOX7, a marker of visceral endoderm, was not (D'Amour et al, 2005). Other studies (particularly Nostro et al., 2011) have gone on to further elucidate the signalling molecules and exposure times required for the most efficient differentiation into DE in vitro.

While most differentiation protocols have relied on recombinant growth factors to provide the necessary signals to the differentiating cells, the use of small molecules which may be able to replace some or all of these factors has been of great interest, as they can be chemically synthesised and purified. This could potentially result in less variability in biological activity than is seen with growth factors. In addition, they are cheaper to produce, which is an important consideration given that these cells will need to be produced on a large scale if they are to be used for clinical applications to treat a large number of patients. A study by Borowiak *et al.* (2009) described the use of the small molecules IDE1 and IDE2, which are thought to be activators of the TGF-β signalling pathway, to improve the efficiency of DE formation in comparison to an Activin A-based protocol (D'Amour et al, 2005).

4.2 Methods

Given the large numbers of protocols published for the differentiation of pluripotent stem cells into DE, the aim of this chapter is to evaluate three widely quoted protocols to determine which results in the highest yield of DE. This is an important consideration for subsequent experiments, since these will be focussed on the DE stage of the differentiation process. If the yield of DE is low then results may be affected by the presence of large numbers of other cell types in the differentiated samples. In addition, high differentiation efficiency at the DE stage is important for the successful differentiation of stem cells into insulin-expressing cells (Courtney et al., 2010). This chapter will also compare the differentiation propensities of several ESC and iPSC lines in order to determine whether there are any differences between these two cell types.

4.2.1 Differentiation of Pluripotent Stem Cells into Definitive Endoderm

The use of small molecules in directed differentiation protocols is of great interest, as they can be chemically synthesised and purified, making them cheaper to produce. A study by Borowiak *et al.* (2009) described the use of the small molecules IDE1 and IDE2 to improve the efficiency of DE formation from several ESC lines in comparison to an Activin A-based protocol. Following on from this, I have investigated the effect of IDE1 on the efficiency of DE formation in iPSC and ESC lines. In 2011, Nostro *et al.* reported an optimised protocol for the differentiation of pluripotent stem cells into DE, and I have also investigated the efficiency of this protocol for DE formation in iPSC and ESC lines. Both of these protocols were compared with the protocol described in the initial report by D'Amour et al (2005).

Stem cells were maintained in culture as previously described in section 2.2. Before the start of the differentiation protocol on day 1, samples of undifferentiated colonies were taken. Colonies undergoing differentiation were washed with PBS before being exposed to either the D'Amour protocol, the Nostro protocol, or IDE1 according to the time scale shown in table 4.1. Cells that were allowed to spontaneously differentiate over the same timescale were also included in the comparison as a control and were treated with RPMI only, to show whether the protocols for directed differentiation result in a higher efficiency of DE formation than cells which were not treated with a differentiation protocol.

Stage 1 Formation of DE							
Day	**1**	**2**	**3**	**4**	**5**	**6**	**7**
D'Amour	RPMI L-glutamine (2mM) Activin A (100ng/ml) Wnt3A (25ng/ml)	RPMI L-glutamine (2mM) Activin A (100ng/ml) Foetal Calf Serum (0.2%)					
Borowiak	RPMI L-glutamine (2mM) FCS (0.2%) IDE1 (400nM)						
Nostro	RPMI Wnt3A (3ng/ml) Activin A (100ng/ml) BMP4 (5ng/ml) VEGF (20ng/ml) bFGF (2.5ng/ml)						

Table 4.1. Reagents and concentrations used to direct differentiation of stem cells to DE.

4.2.2. Characterisation of Differentiated Cells

Quantitative Reverse Transcription-Polymerase Chain Reaction (qRT-PCR)

RNA was isolated from undifferentiated and differentiated cells by mechanical dissection of colonies into 350µl lysis buffer containing Buffer RLT and 1% (v/v) β-mercaptoethanol. Total RNA was isolated using the RNeasy mini kit before being reverse transcribed into cDNA. 9.5µl total RNA, 1µl Oligo$(dT)_{15}$ primers (0.5µg/µl) and 1µl random primers (0.5µg/µl) were mixed and incubated at 70°C for 5min then immediately cooled on ice for 5min. A master mix was prepared according to table 4.2 and 8.5µl was added to each sample. The samples were then incubated for 50min at 42°C, followed by 15min at 70°C.

Reverse Transcription Mix	
Reagent	*Volume Needed For 1 Reaction (µl)*
5x Reaction Buffer	4
10mM dNTPs	1
40U/µl RNAsin	2
400mM DTT	0.5
MMLV-RT	1
Total	**8.5**

Table 4.2. Reagents and volume needed for the reverse transcription reaction.

PCR reactions were set up in 20µl reactions as shown in table 4.3. Serial dilutions of a standard cDNA were prepared as described in section 2.4.2 and run together with the samples of interest to allow quantification of gene expression. A negative control in which the cDNA was replaced with H_2O was also included in each reaction. PCR reactions were carried out on the Qiagen Rotor-Gene™ 6000 using the following PCR cycling conditions: 10min at 95°C; followed by 40 repeated cycles of 5sec at 95°C, 15sec at 58°C and 10sec at 72°C. The fluorescence was acquired at the end of the elongation step. A melt curve analysis was carried out to ensure amplification of a single product, and the products were run on a gel to confirm the amplicon was the correct size (table 2.12).

qRT-PCR Reaction Mix	
Reagent	*Volume Needed For 1 Reaction (µl)*
2x Sensimix Plus *(containing reaction buffer, heat-activated Taq Polymerase, dNTPs, 6mM $MgCl_2$, internal reference and stabilisers, and SYBR green intercalating dye)*	10
10µM Forward Primer	1
10µM Reverse Primer	1
cDNA	2
PCR-grade H_2O	6
Total	**20**

Table 4.3. qRT-PCR reaction mix.

Immunocytochemistry

Medium was aspirated and the cell colonies were washed twice with PBS. The cells were fixed with 500µl of 4% paraformaldehyde and incubated for 20min at room temperature before being washed 3 times with PBST, and then incubated in PBST at room temperature for 15min. 500µl of primary antibody diluted appropriately in PBST (table 2.16), was added to the cells which were then incubated in the dark overnight at 4°C. The cells were washed 3 times with PBST and 500µl secondary antibody at the appropriate dilution (table 2.16) was then added to the cells and left to incubate at room temperature for 1 hour. 200µl of 5µg/ml Hoescht DNA stain solution was added to the cells and incubated for 1min at room temperature. The cells were then washed twice in PBST. Control dishes were also prepared, by omitting the primary antibody to ascertain the specificity of the antibody, and also using an isotype control together with the secondary antibody to ascertain the degree of non-specific binding. The cells were visualised using a Nikon Eclipse TS100 inverted microscope.

4.3 Results

4.3.1 Comparison of Differentiation Protocols

Given that the successful generation of pancreatic progenitors, and subsequently the development of more mature hormone-expressing cells, is critically dependent on the efficiency of DE formation, I have compared three differentiation protocols, together with spontaneously differentiated cells, for their ability to form DE. Borowiak *et al.* (2009) used the small molecule IDE1 to improve the efficiency of DE formation in comparison to an Activin A-based protocol, while in 2011, Nostro *et al.* reported an optimised protocol for the differentiation of pluripotent stem cells into DE. Differentiated cells and undifferentiated cells were characterised using immunocytochemistry and qRT-PCR.

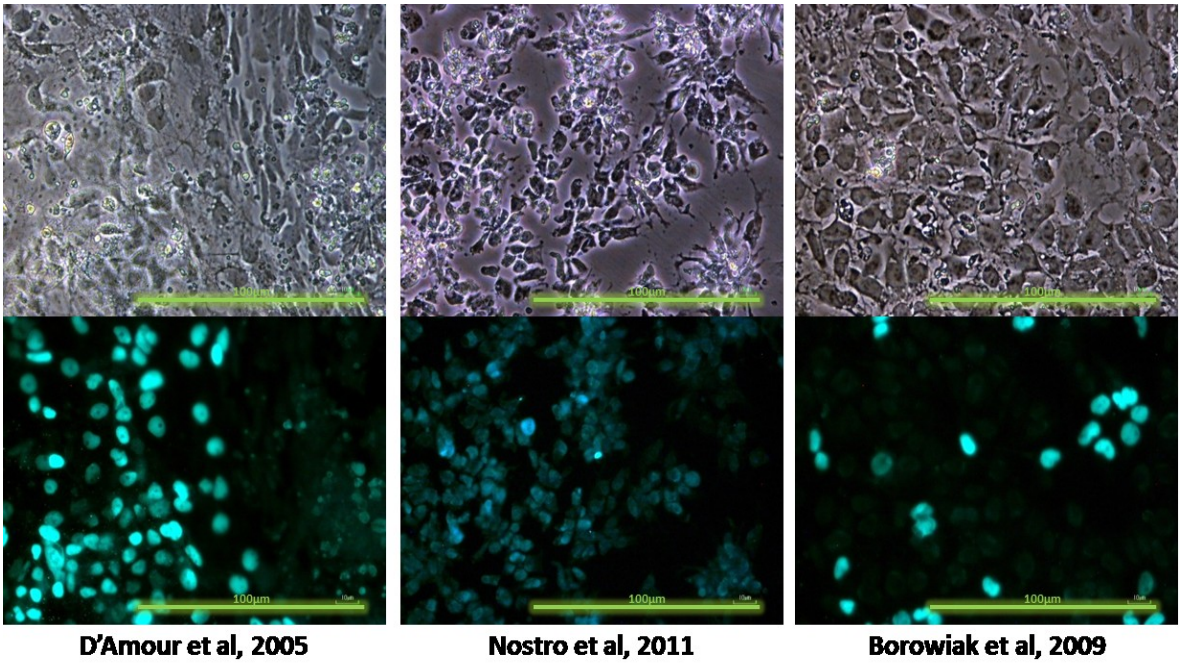

Figure 4.1.Immunocytochemistry for SOX17 *was carried out on an iPSC line (MRC5I) directed to differentiate into DE using one of three differentiation protocols published by D'Amour et al, 2005; Nostro et al, 2011; or Borowiak et al, 2009. 10x magnification, scale bar 100µm.*

Immunocytochemistry was used to determine the levels of SOX17, an early marker of DE, in differentiated cells. A representative example using the MRC5I iPSC line is shown in figure 4.1, which shows that immunofluorescence of the SOX17 protein was highest in cells differentiated using the D'Amour protocol (~70%), less in those differentiated using the Nostro protocol (~59%) and lowest in the cells differentiated using the Borowiak protocol (~18%), indicating that the D'Amour protocol results in the highest efficiency of differentiation into DE.

qRT-PCR was used to investigate gene expression in five cell lines: two iPS (MRC5I and MRC9G) and three ES (H1, H7 and H9). Four genes were used as markers of DE: *Sox17*, *Cxcr4*, *Foxa2* and *Cer1*. Gene expression was compared in cells which had been directed to differentiate using the D'Amour, Nostro or Borowiak protocols, or which had been allowed to spontaneously differentiate. Fold change in gene expression between differentiated and undifferentiated cells was calculated and analysis of variance with Fisher's *a priori* test was used to determine statistical differences between protocols.

Figure 4.2 shows the fold change in *Sox17* expression in differentiated cells compared to undifferentiated cells harvested at the start of the differentiation protocol. Gene expression is significantly ($p < 0.001$, $n = 5$) higher in cells treated with the D'Amour protocol than with either the Nostro or Borowiak protocols, which do not result in significantly higher gene expression than seen in the spontaneously differentiated cells, although some upregulation is noticeable in cells treated with the Nostro protocol. *Sox17* is one of the earliest markers of DE (Lewis & Tam, 2006), and is essential for endoderm formation, as mice lacking SOX17 do not develop gut endoderm from DE (Kanai-Azuma *et al*., 2002).

Figure 4.3 shows the fold change in *Cxcr4* expression in differentiated cells compared to undifferentiated cells. Gene expression is significantly ($p < 0.001$, $n = 5$) higher in both the D'Amour and Nostro protocols than in the Borowiak protocol or spontaneously differentiated cells. There is no significant difference in gene expression between either the D'Amour and Nostro protocols, or the Borowiak protocol and spontaneously differentiated cells. The expression of *Cxcr4* is observed on the cell surface of DE cells (McGrath *et al.*, 1999; Nair & Schilling, 2008), but not visceral endoderm cells (Yasunaga et al, 2005).

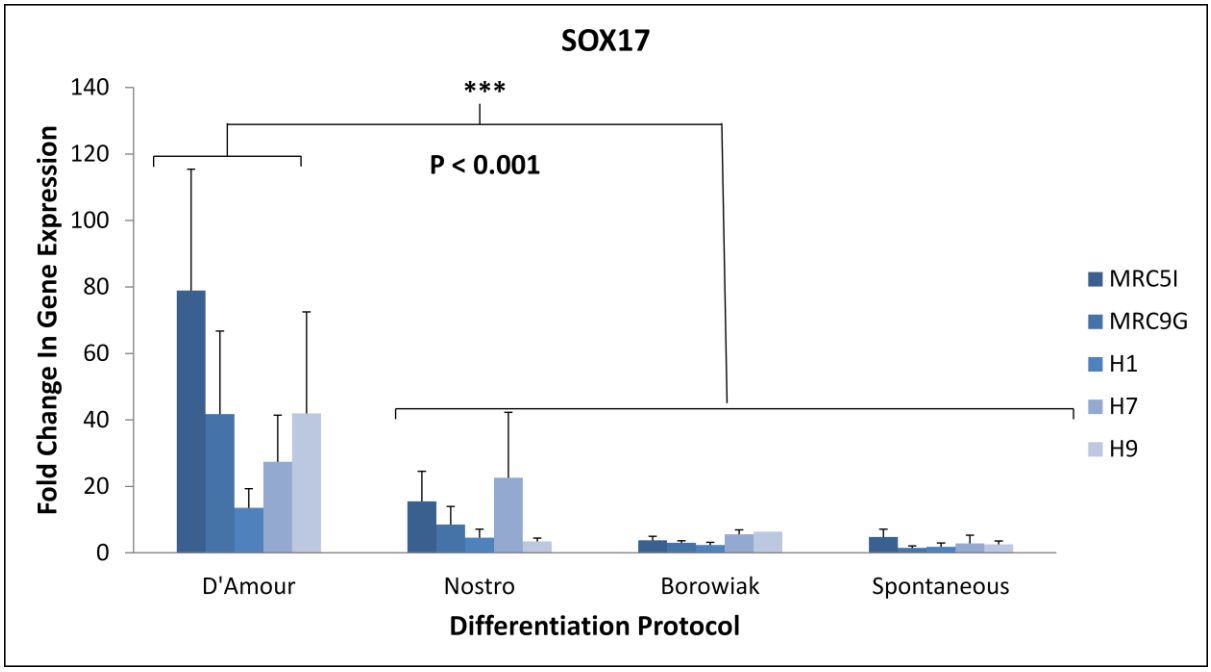

Figure 4.2.Expression of Sox17 *assessed by qRT-PCR in iPS (MRC5I and MRC9G) and ES (H1, H7 and H9) cell lines either directed to differentiate using the D'Amour, Nostro or Borowiak protocols, or allowed to spontaneously differentiate.*

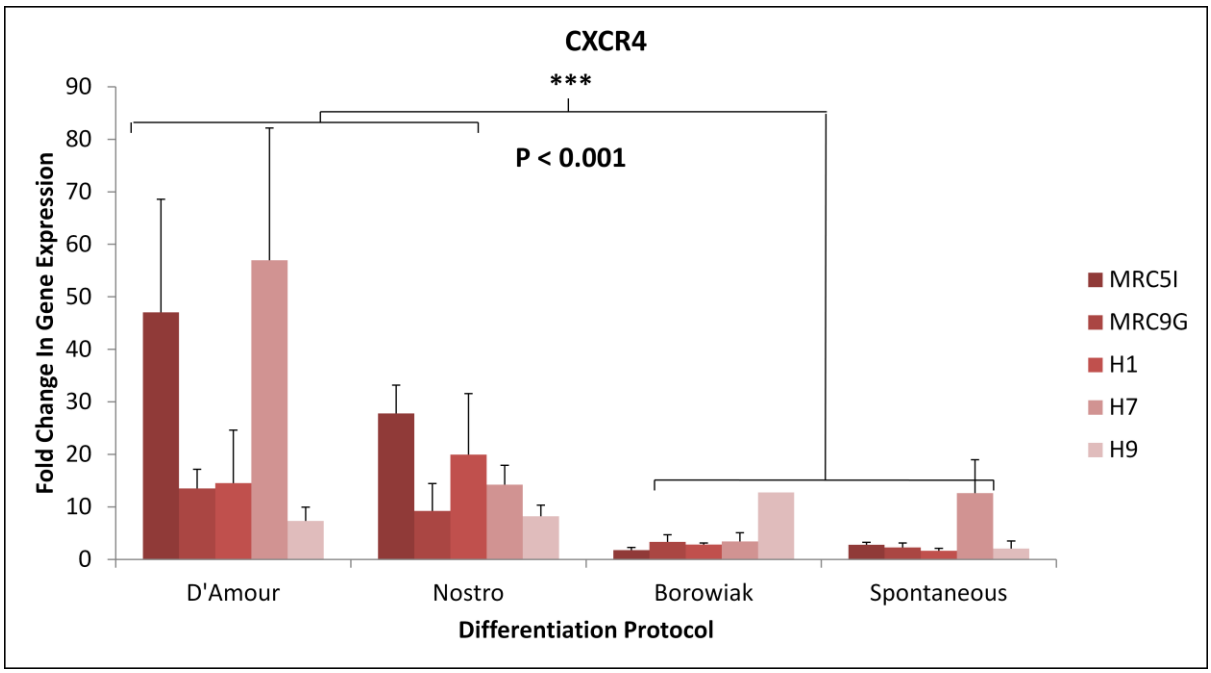

Figure 4.3. Expression of Cxcr4 *assessed by qRT-PCR in iPS (MRC5I and MRC9G) and ES (H1, H7 and H9) cell lines either directed to differentiate using the D'Amour, Nostro or Borowiak protocols, or allowed to spontaneously differentiate.*

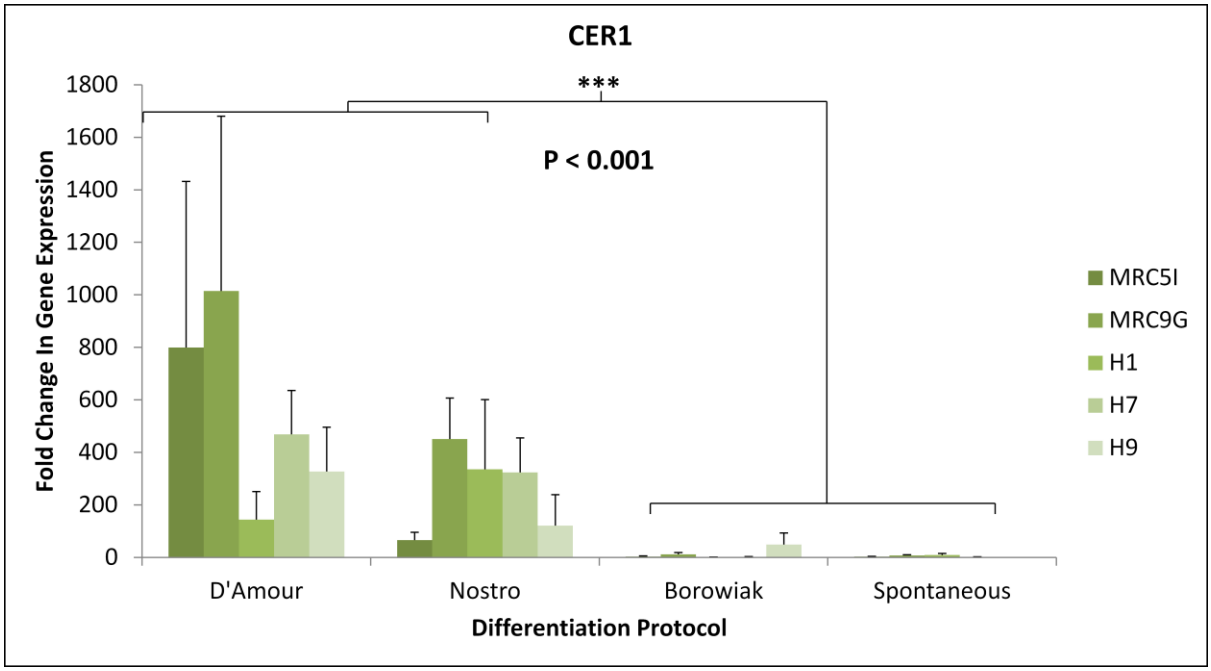

Figure 4.4. Expression of Cer1 *assessed by qRT-PCR in iPS (MRC5I and MRC9G) and ES (H1, H7 and H9) cell lines either directed to differentiate using the D'Amour, Nostro or Borowiak protocols, or allowed to spontaneously differentiate.*

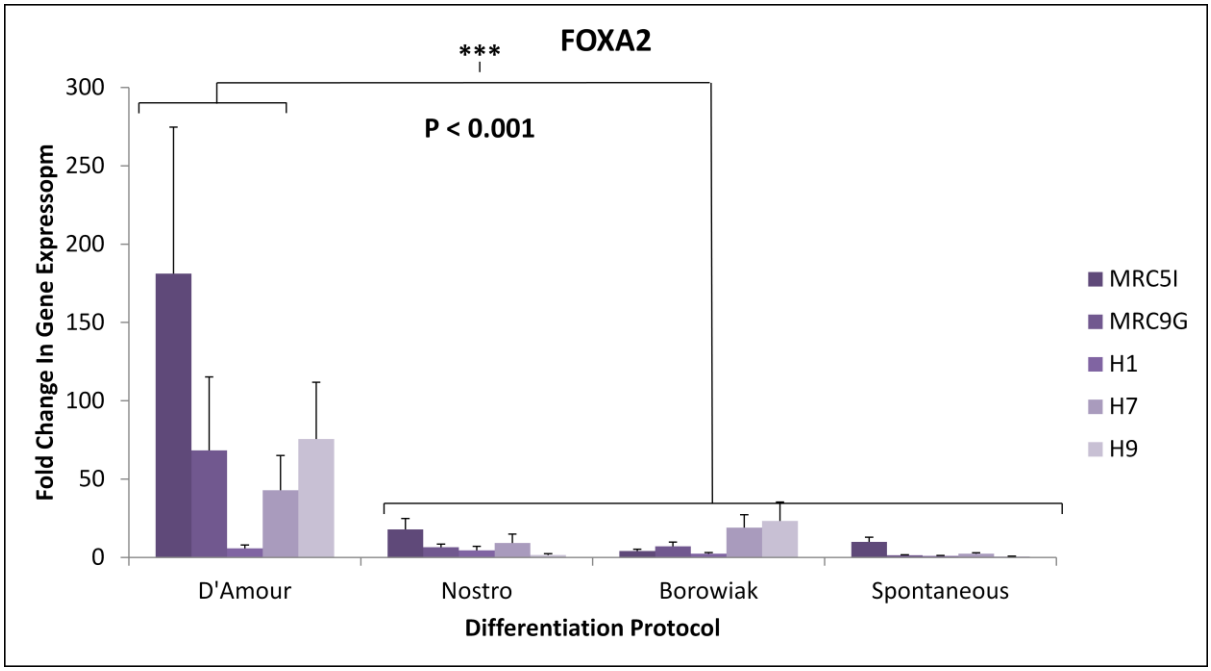

Figure 4.5.Expression of Foxa2 *assessed by qRT-PCR in iPS (MRC5I and MRC9G) and ES (H1, H7 and H9) cell lines either directed to differentiate using the D'Amour, Nostro or Borowiak protocols, or allowed to spontaneously differentiate.*

Figure 4.4 shows the fold change in *Cer1* expression in differentiated cells compared to undifferentiated cells. Gene expression is significantly ($p < 0.001$, $n = 5$) higher in both the D'Amour and Nostro protocols than in the Borowiak protocol or spontaneously differentiated cells. There is no significant difference in gene expression between either the D'Amour and Nostro protocols, or the Borowiak protocol and spontaneously differentiated cells. *Cer1* is another early marker of DE, and its expression is transient and restricted to anterior DE (Belo *et al*., 1997; Biben *et al*., 1998).

Figure 4.5 shows the fold change in *Foxa2* expression in differentiated cells compared to undifferentiated cells. Gene expression is significantly ($p < 0.001$, $n = 5$) higher in cells treated with the D'Amour protocol than with either the Nostro or Borowiak protocols, which do not result in significantly higher gene expression than seen in the spontaneously differentiated cells. FOXA2 is essential for DE formation, as homozygous knockout in mice is lethal due to a lack of endoderm (Sasaki & Hogan, 1993; Monaghan *et al*., 1993).

Together, these results clearly show that the D'Amour protocol consistently results in the most upregulation of DE-associated genes compared to any of the other protocols, and significantly more than in spontaneously differentiated cells. Although the Nostro protocol results in similar upregulation of the *Cxcr4* and *Cer1* genes to the D'Amour protocol, the fact that *Sox17* and *Foxa2* expression are not significantly upregulated in cells treated with this protocol suggests that differentiation to DE may be less complete. These results are in contrast to those published by Nostro et al (2011), when they reported that this protocol produced a population of cells containing approximately 70% SOX17-expressing cells. Cells treated with the Borowiak protocol fail to upregulate any DE-associated genes at a higher efficiency than spontaneously differentiating cells, suggesting that this protocol is having little or no effect in directing the cell lines used in this study towards an endodermal fate.

4.3.2 Comparison of Cell Lines

During experiments carried out to compare the efficiency of different protocols in forming DE, it became evident that different cell lines had different propensities to form DE, which were consistent over several experiments, with some cell lines showing higher upregulation of DE genes than others. To investigate this further, differentiation of individual cell lines using the D'Amour protocol, was compared. Differentiated cells and undifferentiated cells were characterised using immunocytochemistry and qRT-PCR for *Sox17, Cxcr4, Cer1* and *Foxa2*. Fold change in gene expression between differentiated and undifferentiated cells was calculated and analysis of variance with Fisher's *a priori* test was performed to determine significant differences between groups.

Figure 4.6 shows images of two iPSC lines (MRC5I and MRC9G) and three ESC lines (H1, H7 and H9) taken at the beginning (undifferentiated) and end (D'Amour and spontaneous) of differentiation. Undifferentiated colonies have a characteristic morphology, and are typically smooth, round and largely homogenous in appearance. The cells are small and round, with a large nucleus, and are densely packed. In contrast, after differentiation to DE, the morphology changes and colonies appear grainy, and are less homogenous. The cells are sparser, which is probably due to the large amounts of cell death noticed after the start of differentiation. The cells have begun to change shape and increase in size. In addition, spontaneously differentiated colonies are distinguishable from colonies induced to differentiate using the D'Amour protocol. There is no obvious difference in morphology between different cell lines, either ESCs (H1, H7 and H9) or iPSCs (MRC5I and MRC9G), or between colonies grown on feeders (MRC5I, MRC9G and H1) and those grown on Matrigel™ (H7 and H9).

Figure 4.7 shows immunocytochemistry for SOX17 expression in two iPSC lines (MRC5I and MRC9G) and three ESC lines (H1, H7 and H9) induced to differentiate into DE using the D'Amour protocol. While H1 shows noticeably lower levels of SOX17 expression, there is no obvious difference between the other cell lines, which all show high expression of SOX17, indicating a high efficiency of differentiation. The approximate percentage of SOX17-expressing cells in the each cell line is: MRC5I = 75%, MRC9G = 74%, H1 = 42%, H7 = 62%, H9 = 69%.

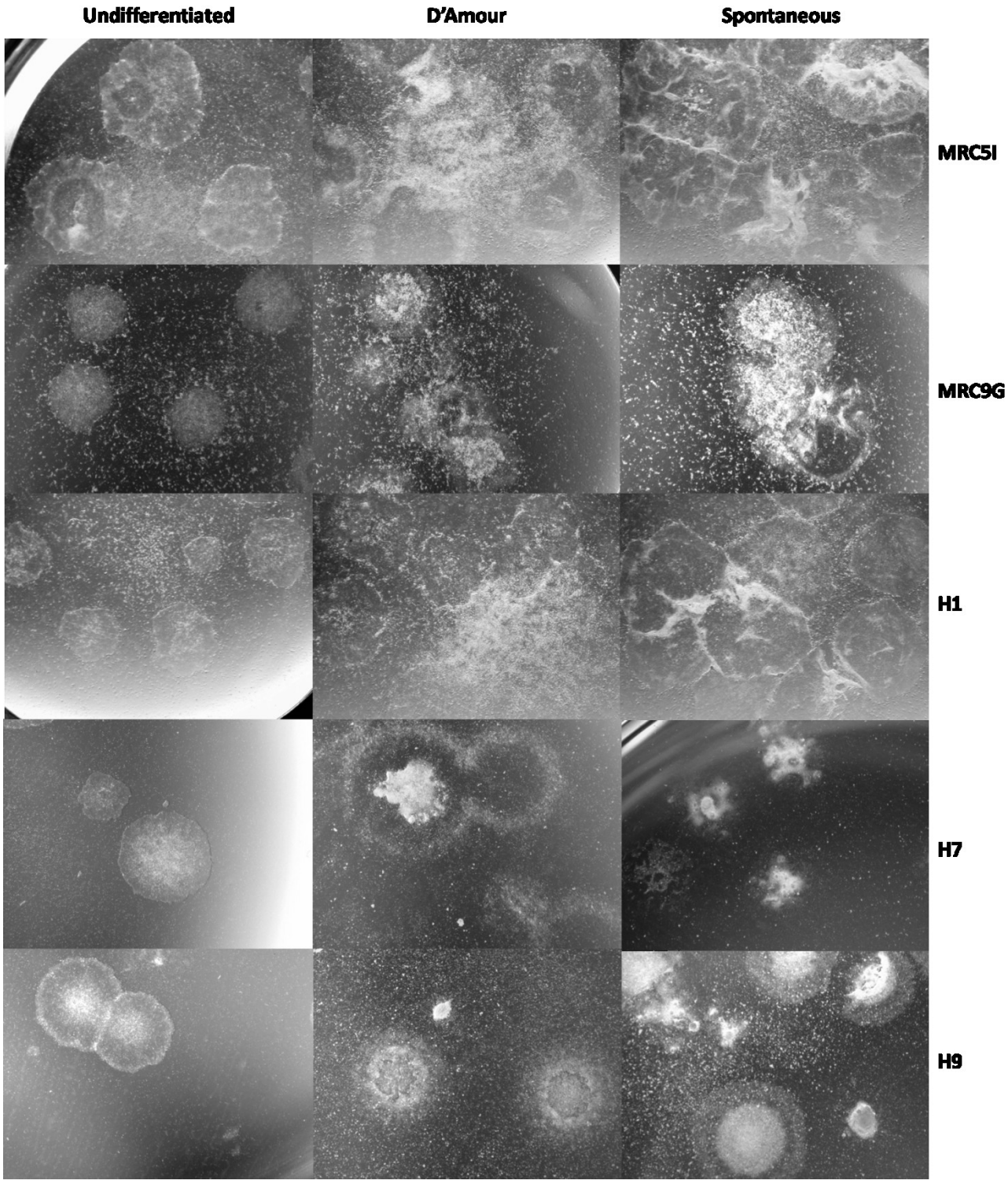

Figure 4.6.Change in colony morphology during differentiation to DE. *Images of ESCs (H1, H7 and H9) and iPSCs (MRC5I and MRC9G) were taken at the start and end of differentiation in order to observe changes in colony morphology. 2x magnification.*

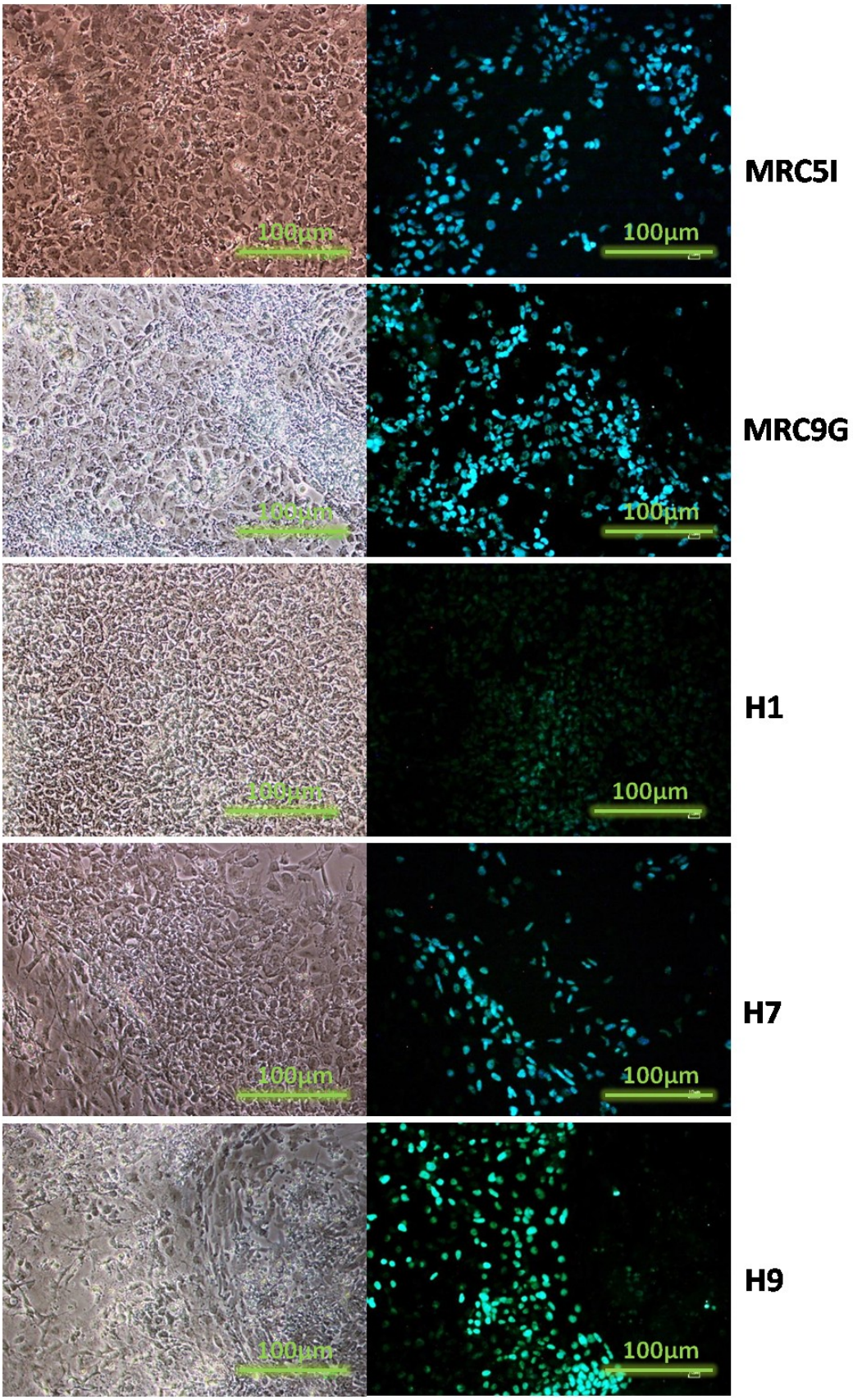

Figure 4.7.Immunocytochemistry for SOX17 expression *was carried out on two iPSC lines (MRC5I and MRC9G) and three ESC lines (H1, H7 and H9) induced to differentiate into DE using the D'Amour protocol. 10x magnification, scale bar 100µm.*

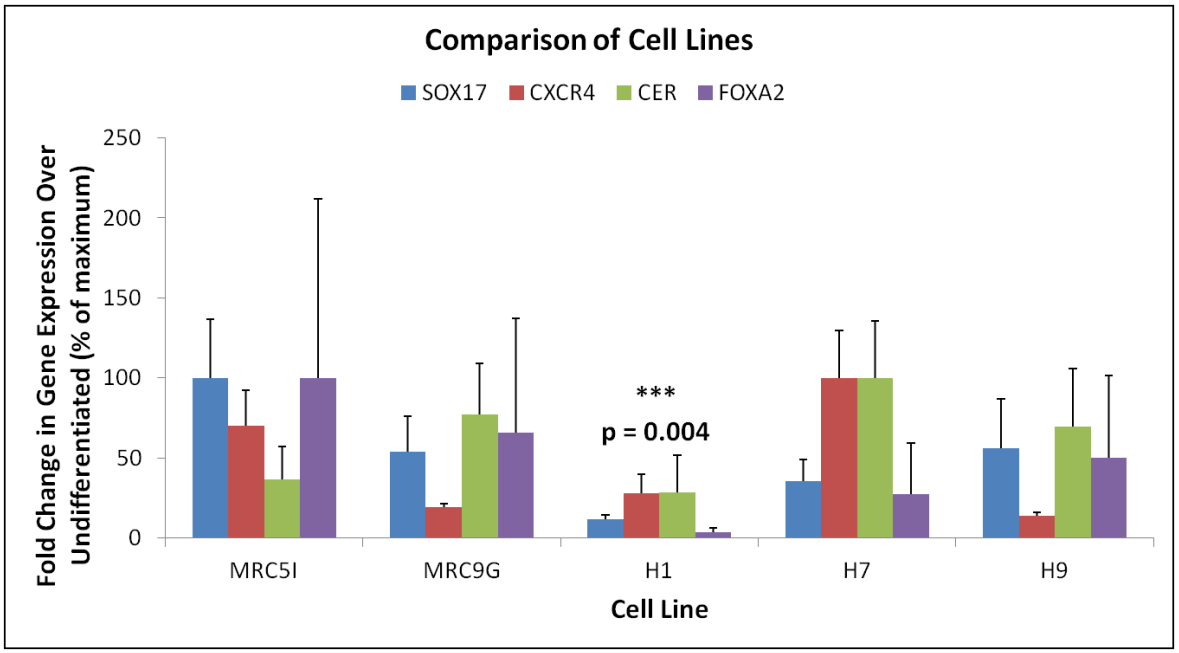

Figure 4.8. Comparison of gene expression in cell lines directed to differentiate into DE.

Figure 4.8 shows the expression of genes characteristic of DE (*Sox17*, *Cxcr4*, *Cer1* and *Foxa2*) in two iPSC lines (MRC5I and MRC9G) and three ESC lines (H1, H7 and H9). There is variability between cell lines in the expression of these genes which is most noticeable in the H1 cell line, which shows significantly lower ($p = 0.004$, $n = 5$) expression of DE genes than any of the other cell lines tested, suggesting that this cell line is more resistant to differentiation into DE.

Initial experiments were carried out using two iPSC lines generated at NIBSC, MRC5I and MRC9G, and a well-characterised ESC line (H9). However, the iPSC lines are grown on iSNL feeders, and the H9 line is grown on Matrigel™. For this reason, the H1 cell line was included in differentiation experiments, as this is an ESC line grown on iSNL feeders, and was therefore expected to represent an ESC line grown in a comparable culture format to the available iPSC lines. However, data from my own results (figure 4.8) and others (Bock et al, 2011; Nostro et al, 2011) indicate that this line has a low propensity for forming DE. The H7 ESC line, also grown on Matrigel™, was also included in differentiation experiments.

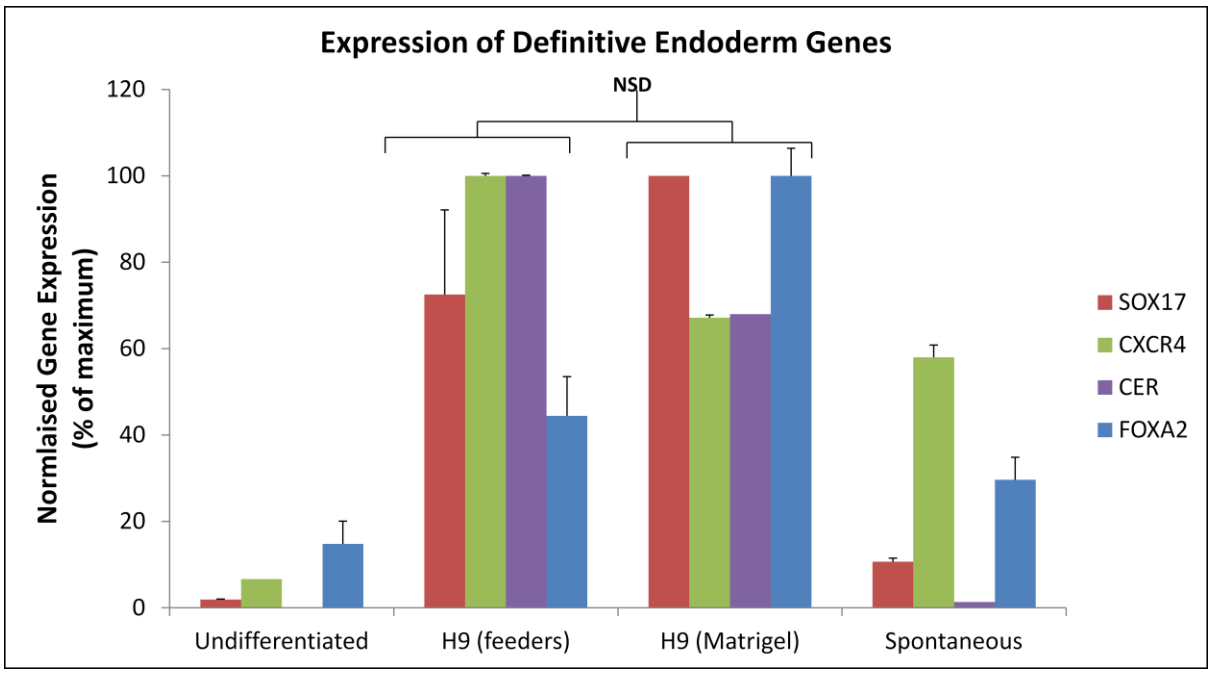

Figure 4.9. Expression of Sox17, CXCR4, CER-1 and FOXA2 in H9 cells cultured on either iSNL feeders or Matrigel™ and directed to differentiate using the D'Amour protocol.

Investigation of whether the low differentiation efficiency observed in the H1 cell line was a result of the culture conditions (i.e. being grown on feeders, compared to the other ESC lines which were all grown on Matrigel™) could not be carried out, as the H1 cell line could not be maintained on Matrigel™. For this reason, the differentiation efficiency of the H9 ESC line grown on both feeders and Matrigel™ was compared to determine whether growing ESC lines on feeders was generally detrimental to their ability to differentiate into DE. Differentiation into DE was carried out using the D'Amour protocol as described above, and the cells were assessed for their expression of *Sox17*, *Cxcr4*, *Foxa2*, and *Cer-1*. Figure 4.9 shows that there is no significant difference ($p = 0.789$, $n = 3$) in the differentiation efficiency between cells cultured on feeders and cells cultured on Matrigel™. Expression of DE genes is significantly upregulated in cells treated with the D'Amour protocol, in both culture formats ($p < 0.001$, $n = 3$), indicating that H9 is capable of differentiating into DE whether cultured on iSNL feeders or Matrigel™. There is no significant difference in expression of these genes in spontaneously differentiated samples compared to undifferentiated samples, although CXCR4 is upregulated in these samples. However, this gene is not specific to DE and is upregulated in many other tissue types.

4.4 Discussion

The aim of the work described in this chapter was to compare three differentiation protocols in order to optimise the yield of DE for further experiments. Numerous protocols have been developed for the differentiation of pluripotent stem cells into DE, and subsequently insulin-expressing cells (see section 1.7.1), but the most successful protocols have been those which have recapitulated the signalling pathways which are important during *in vivo* pancreatic development, including the influential protocol originally published by D'Amour et al (2005) and later optimised by Kroon et al (2008).

However, there are limitations to these protocols. The cells produced are frequently polyhormonal and unable to release insulin in response to glucose (although they can release insulin in response to other secretagogues). These features are indicative of an immature phenotype, a common problem for many different *in vitro* differentiation protocols. While increasing the length of time in culture did allow the cells to mature somewhat, they were still very different from the equivalent adult cells, suggesting that increased time in culture alone is not enough to drive maturation, and that additional signals are needed (Patterson *et al.,* 2012). However, Kroon *et al.* (2008) succeeded in producing a population of pancreatic progenitor cells which, several weeks after transplantation into immunodeficient mice, were shown to release C-peptide and insulin in response to glucose, demonstrating that unknown factors in the *in vivo* environment were capable of inducing these cells to complete differentiation into a more mature β-cell phenotype. This result was confirmed by another group using a different protocol (Shim *et al.*, 2007). Since then, there have been efforts to elucidate the factors needed for β-cell maturation (Blum et al, 2012; Xie et al, 2013). However, the precise factors involved in β-cell maturation remain unknown, and, until they are elucidated, it is unlikely that fully functional mature β-cells will be generated *in vitro*.

Another problem facing *in vitro* differentiation protocols is the variability in the cells produced. One factor affecting variability is the fact that many of the published protocols rely on the use of recombinant growth factors to direct differentiation. These can vary considerably in their biological activity from batch to batch, and in addition are very expensive to produce. Both of these problems render these protocols unsuitable for large-scale production of differentiated cells for cell replacement therapy. For this reason, there has been a large amount of research carried out on the use of low molecular weight, cell

permeable, bioactive small molecules to enhance the activity of, or replace, growth factors in differentiation protocols. These have the advantages of being specific and dose-dependent. They have rapid and reversible effects, and synthetic chemistry allows almost unlimited diversity to control molecular interactions.

There have been several reports of small molecules which can increase the efficiency of growth factor-directed differentiation into a pancreatic lineage, but which are not able to replace growth factors in the protocol. LY294002 and TKT-II inhibit PI3K signalling (which is important for maintaining pluripotency), increasing the efficiency of differentiation to DE when used in combination with Activin A (McLean et al., 2007). However these chemicals were not able to replace Activin A in this protocol but only enhance its effects. This study also described the link between serum levels and Activin signalling, suggesting that components of serum, such as IGF or insulin, block differentiation through increased PI3K signalling. Another study showed that sodium butyrate could increase the efficiency of endoderm formation in cells treated with Activin A, but that it was insufficient to induce differentiation itself (Jiang et al, 2007). This is likely to be due to its effects on the epigenetics of the cells, making them more receptive to differentiation signals. Shim et al (2007) used retinoic acid in combination with Activin A to increase the efficiency of differentiation.

Borowiak et al (2009) carried out a large-scale screen of over 4000 compounds in the absence of Activin A signalling in an attempt to identify chemicals that could induce DE formation, rather than simply enhance it. They identified two compounds, termed inducer of DE (IDE)-1 and IDE-2, which were reported to give a yield of DE of over 80%. These small molecules are thought to act through activation of the TGF-β pathway, as Smad phosphorylation was observed following exposure of cells to both IDE1 and IDE2. Unfortunately, despite these promising results, I did not observe any upregulation of DE markers when using IDE1 beyond what is seen in spontaneously differentiating cells, suggesting that this protocol had little or no effect on differentiation on the cell lines used in this study. This may be due to the different cell lines used, which may have different abilities to respond to these chemicals, as the study by Borowiak et al (2009) used three ESC lines (HUES4, HUES8 and HUES9), while I used three different ESC lines (H1, H7 and H9) in addition to two iPSC lines (MRC5I and MRC9G). It is worth mentioning that I only tested IDE-1 and not IDE-2, although they are very similar compounds which are thought to act in the same way. Culture conditions

were very similar between the two studies, with both using cells grown on mouse feeders in KOSR medium, although I also used cell lines grown on Matrigel™. Another recent study comparing the effects of several small molecules in DE formation also found IDE1 and IDE2 to be ineffective (Tahamtani et al, 2013), and to date there are no other reports that I am aware of using these compounds to successfully induce endoderm differentiation, which suggests that other researchers have not found them to be efficient either. Indeed, a more recent publication by the same group that published the IDE results used Activin A to induce differentiation into DE, not IDE1 or IDE2 (Chetty et al, 2013).

Since then, several other groups have reported small molecules which can enhance or replace growth factors in differentiation. Kunisada et al. (2012) replaced Wnt3A with CHIR00921 (a GSK3β inhibitor) in DE formation, while Sekine et al (2012) reported increased efficiency of endoderm formation using media supplemented with B27. Norrman et al (2012) reported that endodermal cells produced using this protocol were the most functionally mature when comparing gene expression in single cells differentiated using three different Activin A-based protocols. Another study reported enhancement of differentiation efficiency by priming cells with Rapamycin prior to differentiation (Tahamtani et al, 2013). In addition to small molecules which induce endoderm differentiation, some small molecules promote differentiation generally through down-regulation of pluripotency, such as stauprimide, which results in increased DE formation in both mouse and human ESCs (Zhu et al, 2009).

Another problem facing these differentiation protocols is the low yield of cells obtained. In an attempt to increase the efficiency of differentiation, Nostro et al (2011) optimised differentiation conditions. They further elucidated the temporal requirements for TGF-β, Wnt and BMP4 signalling, and showed that BMP4, bFGF and VEGF could replace serum in differentiation protocols. Being able to carry out differentiation in serum-free conditions is an advantage, both in terms of reproducibility and in being more clinically relevant. However, my results show that although treatment with the protocol developed by Nostro et al (2011) results in significant upregulation of *Cer1* and *Cxcr4* to a level comparable to that seen when using the D'Amour protocol, there is no significant upregulation of *Sox17* and *Foxa2* over what is seen in spontaneously differentiating cells, suggesting that these cells may not be fully representative of DE. This in in contrast to the results published by Nostro et al (2011), who reported over 70% of cells expressing *Sox17*. This study used two of the same cell lines used here (H1 and H9), as well as the HES2 and HES3 ESC lines and an iPSC line, and used

very similar culture methods. Interestingly, they also noticed the poor differentiation of H1 into a pancreatic lineage, suggesting that this is a property inherent to this stem cell line.

Culture conditions also play an important role in differentiation to DE, as demonstrated by Kubo et al (2004), who showed that high serum concentrations resulted in a lower efficiency of differentiation into DE, due to the presence of undefined growth factors and hormones in serum. Reduction of serum concentration to 0.2-0.5% was shown to be the most successful for generating DE (D'Amour et al, 2005). Whether the cells are differentiated in monolayer culture or as embryoid bodies is also likely to affect the differentiation process, and alter the necessary dose and timing of signalling factors (Sulzbacher et al, 2009). In an attempt to recreate more physiological conditions, and improve the quality of the cells produced, Van Hoof et al (2011) developed a protocol for the differentiation of ESCs into DE in adherent, size-controlled, clusters, which were able to give rise to homogenous clusters of pancreatic progenitors. However, Nostro et al (2011) found that differentiation kinetics were faster in monolayer cultures vs. embryoid bodies, with endodermal gene expression arising by day 3 in monolayer culture but not until day 5 in embryoid bodies, although embryoid body culture has the advantage of being more easily scaled up. Chetty et al (2013) found that the differentiation efficiency varied depending on the density of the cells at the start of the protocol, with cells plated at higher densities forming more DE than those at low densities, suggesting that contact-mediated growth inhibition might improve the cells' response to differentiation signals. Treating low density cells with DMSO prior to differentiation improved the efficiency of differentiation to a level comparable with high-density cells, through its effects on the cell cycle. This was true for differentiation into several lineages, including DE, and should be taken into consideration when optimising differentiation protocols. Even when efficiency of differentiation into DE is high, cell yield is generally low due to the high levels of cell death during differentiation. One explanation for this is that differentiation occurs in a selective manner, where only a few cells in the starting population are competent to respond to exogenous signals and so differentiate, while other cells are not competent to differentiate and so die. This model is supported by a study carried out by Ungrin *et al* (2012), who monitored cell proliferation during differentiation, and found that differentiation resulted in an initial wave of cell death followed by proliferation and differentiation of a sub-population of cells. They then went on to identify ways in which this initial wave of cell death could be prevented in order to increase the yield of differentiated cells, and found that optimising initial embryoid body formation resulted in a 36-fold increase

in yield during differentiation to DE. This method may also allow scale-up of production to the levels needed to produce adequate numbers of differentiated cells for cell therapy.

In an attempt to make differentiation protocols more reproducible, as well as allowing for scale up, commercial products for differentiation are now becoming available. These have the additional advantages of being both fully defined, as well as serum- and animal product-free, making them more clinically applicable. While these products will not change the fact that different cell lines have different propensities for differentiation, they should cut down on experimental variation arising from the reagents used.

Although protocols have now been developed which allow the production of DE at high efficiencies, it is still difficult to achieve a completely pure population. This is important both for efficiency of the later stages of pancreatic differentiation and for studying DE without contamination by other cell types. Cell sorting based on expression of extracellular markers to achieve a pure population requires the use of a combination of markers in order to select for DE cells alone, as many markers are expressed by other cell types, and results in loss of cells and reduced cell function. Efforts have been made to further characterise DE cells in an attempt to identify better markers that will allow purification of an endoderm population (Wang et al, 2012).

An appealing alternative may be the generation of stable endodermal progenitor cell lines from pluripotent stem cells. An early study by Tada et al (2005) used a mouse ESC line with a GFP reporter in the locus of the *Goosecoid* gene to identify DE cells which could be expanded in culture while still maintaining expression of endodermal markers. However, although some further differentiation was observed, these cells could not give rise to mature pancreatic or hepatic cells, suggesting that they are functionally restricted. A subsequent study, also in mice, used a reporter gene in the *Hex* locus to purify endodermal cells, and demonstrated that these could be expanded in culture and could give rise to AFP-expressing hepatic cells, as well as pancreatic progenitors (Morrison et al, 2008). More recently, self-renewing endodermal progenitor cell lines have been generated from human ESCs and iPSCs, which express key endodermal genes such as *Foxa2* and *Sox17* and are able to differentiate *in vitro* into hepatic, pancreatic and intestinal lineages. Despite their extensive proliferative capacity *in vitro*, when injected into immunodeficient mice they did not form teratomas (Cheng et al, 2012). Another report demonstrated that endoderm cells derived from both mouse and human ESCs could be expanded in culture by co-culturing with mouse

mesenchymal cells. These cells could spontaneously give rise to glucose-responsive insulin-producing cells after transplantation into mice (Sneddon et al, 2012).

Another strategy for differentiation into DE is to overexpress transcription factors known to be important at this stage of development, such as SOX17 (Qu et al, 2008; Seguin et al, 2008). One of these studies used SOX17 overexpression to establish a stable endoderm progenitor line which could be maintained in culture over several passages, and retained the ability to differentiate into hepatic and pancreatic cells (Seguin et al, 2008). While this cell line is likely to be useful to study definitive endoderm, it is unlikely that cells which have been genetically manipulated in this way will ever be considered for clinical use.

During the comparison of differentiation protocols, it was noticed that different cell lines had different propensities to differentiate into DE. To investigate this further, a comparison of both ESCs and iPSCs was carried out and it was found that in all experiments, the H1 ESC line gave rise to significantly lower amounts of DE than any of the other cell lines. There were also statistically non-significant differences between other cell lines tested, with the iPSC line MRC5I and the ESC line H9 both differentiating to DE at high efficiencies. These differences were consistent over several experiments, suggesting that they are due to intrinsic properties of the cell line, and not random variation. This variability between cell lines has been described before (Abeyta *et al*, 2004; Martinez *et al*, 2012; Osafune *et al*, 2008). However, while some studies have shown that iPSCs have lower efficiencies of differentiation into haematopoietic, neuroepithelial, neuronal and cardiac lineages than ESCs (Feng *et al*, 2010; Hu *et al*, 2010; Narsinh *et al*, 2011), this was not found with differentiation into DE in this study. Although there was a significant difference in DE gene expression between ESCs and iPSCs ($p = 0.049$), this is likely to be due to the very low levels of gene expression in the H1 ESC line, which has previously been noticed to be inefficient at endoderm formation (Bock et al, 2011), as when this cell line was excluded from the analysis there was no longer a significant difference between the two cell types ($p = 0.672$). In this study, iPSCs were able to form DE at least as well as ESCs. This suggests that differences between cell lines may not be due to systematic differences between ESCs and iPSCs, but are instead a result of inherent variability between pluripotent stem cell lines. One study has suggested that variability between cell lines in their ability to differentiate into endoderm is due to their differing levels of endogenous Wnt signalling, with cells that have high levels

forming predominantly endodermal and cardiac cells, and cells that have low levels forming predominantly ectodermal cells. It has been reported that manipulating the levels of Wnt signalling removes much of the heterogeneity observed between different cell lines (Blauwkamp et al, 2012). Others have suggested that differences in BMP4 signalling may influence differentiation into DE (Christodoulou et al, 2011; Nostro et al, 2011).

To investigate whether the low differentiation efficiency observed in the H1 cell line was a result of the culture conditions (i.e. being grown on feeders, compared to the other ESC lines which were all grown on Matrigel™), or an inherent property of the cell line itself, I wanted to compare the differentiation efficiency of H1 cultured on both feeders and Matrigel™. However, I was unable to maintain H1 on iSNL feeders, so I tested the differentiation efficiency of the H9 ESC line grown on both feeders and Matrigel™ to determine whether growing ESC lines on feeders was generally detrimental to their ability to differentiate into DE. I found no differences in differentiation efficiency between H9 cells cultured on feeders and H9 cultured on Matrigel™, although there was some variability in gene expression between them. There was significant upregulation of DE genes in the treated samples compared to the undifferentiated samples, demonstrating that the H9 cell line can differentiate into DE at high efficiencies irrespective of the culture conditions used.

The subject of whether there are any systematic differences between ESCs and iPSCs has been hotly debated. While early studies comparing the gene expression signatures of ESCs and iPSCs found that they were very similar, and distinct from other cell types (Muller *et al*, 2008), more recent studies have identified numerous subtle differences between the two cell types, including differences in differentiation propensity, gene expression, miRNA expression, and epigenetics (see section 1.9). In addition, iPSCs are generally less successful in generating high percentage chimaeras and live mice through tetraploid complementation than ESCs (Stadtfeld et al, 2010). These studies have led to questions about whether iPSCs are indeed identical to ESCs, and if they are not, what the functional implications of this might be for the use of iPSCs for *in vitro* applications such as disease modelling and drug screening, as well as their *in vivo* use for cell replacement therapy. One study has been carried out to investigate whether the differences observed between mouse ESCs and iPSCs affect their differentiation into DE. This study compared 4 iPSC lines and 2 ESC lines, and found considerable variability between cell lines in their propensity to form DE, although no overall difference between ESCs and iPSCs (Christodoulou et al, 2011). In addition, they

compared genetically identical mouse ESCs and iPSCs, and tried to link aberrant imprinting of the *Dlk1-Dio3* region of chromosome 12, which has been previously described as a difference commonly seen between genetically identical ESCs and iPSCs (Stadtfeld et al, 2010), with differences in endoderm formation. Although they found significant differential expression of BMP4 between ESCs and iPSCs during differentiation, possibly linked to this aberrant imprinting, this did not appear to be associated with any reduction in differentiation propensity (Christodoulou et al, 2011). This suggests that although transcriptional differences may exist between ESCs and iPSCs, they may not adversely affect the function of the cells.

It is becoming clear that the heterogeneity and behaviour of pluripotent stem cells is more complex than previously thought, and that iPSCs and ESCs are neither identical nor completely distinct populations; instead, they seem to overlap, with greater variability within each population than is observed between the populations. This variability is likely to have functional implications for the cells, so it may be more useful to consider each cell line in terms of its quality and utility, in order to choose the best cell line for the application. With a view to this end, some researchers have generated a bioinformatics assay for pluripotency (Soldner *et al*, 2009), while others have produced a 'scorecard' to evaluate the character of both iPSCs and ESCs, and to predict the quality and utility of any pluripotent cell in a high-throughput manner: a large-scale study by Bock *et al* (2011) used gene expression profiling, DNA methylation mapping, and a high-throughput quantitative differentiation assay to establish genome-wide reference maps for patterns of gene expression and DNA methylation in order to provide a baseline against which comparisons of new ES and iPS cell lines can be made in order to assess their quality and utility.

In conclusion, comparison of several differentiation protocols demonstrated that the protocol published by D'Amour et al (2006) was the most efficient, and this was used in all further experiments. In addition, this chapter demonstrated that both ESCs and iPSCs are able to give rise to DE, although at varying efficiencies. The mechanisms underlying these differences will be investigated further in subsequent chapters.

Chapter 5

Identification of miRNAs that play a Role in the Formation of Definitive Endoderm

Contents

List of Figures

List of Tables

Chapter 5: Identification of miRNAs that play a Role in the Formation of Definitive Endoderm

5.1 Introduction

MicroRNAs (miRNAs) are small (18-25 nucleotides), non-coding RNAs that regulate gene expression post-transcriptionally. Over 21,000 miRNAs have now been identified in numerous different species, according to miRBase version 19 (www.mirbase.org), and they are involved in almost every biological process, including development, metabolism, and ageing, as well as in many diseases (Slack, 2010). Cells express a miRNA signature characteristic of their lineage and developmental state (Strauss et al., 2006). Although the number of miRNAs is small compared to that of mRNAs (the human genome is thought to encode around 1,000 miRNAs, while the number of mRNAs is estimated to be around 30,000), a single miRNA is able to target hundreds of mRNAs, and one mRNA may be targeted by several miRNAs. As a result, the potential for fine tuning of mRNA translation by miRNAs is enormous.

5.1.1 The role of miRNAs in Pluripotency and Differentiation

The importance of miRNAs in the maintenance of stem cell pluripotency and the control of cellular proliferation and differentiation was first identified through experiments disrupting the miRNA processing enzymes DICER and DGCR8, as DICER- and DGCR8-null ESCs show defects in proliferation and differentiation (Kanellopoulou *et al*., 2005; Wang *et al.*, 2007). Sequencing experiments have since shown that a small number of miRNAs are specifically expressed in pluripotent stem cells, and these are quickly down-regulated upon differentiation. These miRNAs are highly conserved between mice and humans, and include the miR-290 family (miR-290, miR-291a, miR-291b, miR-292, miR-293, miR-294, miR-295) in mice (Houbaviy et al. 2003), its human homologues the miR-371 family (miR-371, miR-372, miR-373), and the miR-302 cluster (miR-302a, miR-302b, miR-302c, miR-302d, miR-367) which is present in both mice and humans (Suh et al. 2004). Many stem cell-specific miRNAs are co-transcribed as polycistronic transcripts, suggesting common

upstream regulation and co-ordinated expression patterns and, like the miR-520 and miR-302 clusters (Ren *et al.*, 2009), they often have similar seed sequences, suggesting common mRNA targets (Martinez & Gregory, 2010). Such a high degree of functional redundancy suggests that they play a critical role in maintaining the pluripotent state, and the targets of these miRNAs include many genes with promoters bound by the stem cell-associated transcription factors NANOG, SOX2 and OCT4, suggesting an important role in the regulation of gene expression to maintain pluripotency and self-renewal.

Although capable of continuous self-renewal and maintenance of pluripotency, pluripotent stem cells are permanently poised to differentiate when the appropriate cues arise. This massive transformation of phenotype poses a major challenge to the cell, as the entire network of transcription factors and signalling pathways must be changed within a short developmental window. The ability of miRNAs to regulate hundreds of target genes post-transcriptionally makes them ideally suited to bring about rapid transformations in cell phenotype. The switch from pluripotency to differentiation requires down-regulation of pluripotency-associated genes, a reduction in the signalling for self-renewal and the activation of early lineage-specific gene expression. These changes can be rapidly effected by miRNAs through the silencing of pluripotency genes and the inhibition of genes that negatively regulate early differentiation. Indeed, the first miRNAs to be discovered, let-7 and lin-4 (the mammalian homologue of which is miR-125), have been shown to play a key role in the transition from pluripotency to differentiation. Thus, let-7 promotes the exit from the cell cycle by the repression of a number of cell cycle-regulating genes (Mallanna & Rizzino, 2010) and directly targets the 3'UTR of pluripotency genes such as *c-Myc* and *Lin28* in order to regulate their expression negatively (Melton *et al*., 2010). Subsequent studies have identified an increasing number of miRNAs that repress components of the stem cell/pluripotency control network. These include miR-34, miR-134, miR-145, miR-296 and miR-470 which, through a range of interactions, target the pluripotency genes *Oct4, Sox2*, *Klf4* and *Nanog* (Tay *et al*., 2008; Xu *et al*., 2009). Cells express a miRNA signature characteristic of their lineage and developmental state, with an overall trend of increasing complexity in miRNA expression during differentiation, with pluripotent stem cells expressing relatively few miRNAs, and more differentiated cells expressing an increased number of miRNAs. This is in contrast to the pattern of mRNA expression, which has been shown to be restricted with increasing development, with fewer genes expressed in more mature cells (Strauss *et al*., 2006).

A number of studies have now demonstrated the importance of miRNAs in commitment to a specific embryonic germ layer during development. In mouse ESCs, elevated levels of miR-296 and miR-134 enhanced the expression of genes associated with differentiation toward ectoderm (Tay *et al.*, 2008). In human ESCs, members of the miR-30 family have been shown to regulate the embryonic ectoderm development protein which is important in the development of the neural tube (Song *et al*., 2011), and miR-125 isoforms have been shown to potentiate early neural specification (Boissart *et al.,* 2012). The miR-302 family is expressed in undifferentiated ESCs and has been shown to be critical in the maintenance of pluripotency but a study by Rosa *et al.* (2011) demonstrated that this family also plays a role in the earliest stages of differentiation, as in humans, repression of the *Lefty1* and *Lefty2* genes by miR-302 results in increased TGF-β signalling, leading to increased endoderm and mesoderm formation, and decreased neuroectoderm formation (Rosa & Brivanlou, 2011). The miR-17-92 cluster plays an important role in mesoderm differentiation and knockout in mouse embryos results in severe developmental defects affecting the heart and lungs (Ventura *et al.* 2008). Foshay *et al.* (2009) demonstrated the importance of one member of this cluster, miR-93, in the regulation of both endoderm and mesoderm differentiation through its targeting of *Stat3*. miR-124 is highly expressed in human ESCs and targets two regulators of cytoskeletal rearrangement important in cell migration during mesodermal differentiation (Berardi *et al.*, 2012). Another cluster of pluripotency-associated miRNAs, the miR-290 cluster, regulates *Pax6* expression in early lineage commitment, and knockout of these miRNAs in mice resulted in an increased propensity to form ectoderm at the expense of mesoderm and endoderm (Kaspi *et al,.* 2013).

One early event that is critical in the commitment of differentiating cells to an endodermal or mesodermal lineage is epithelial-to-mesenchymal transition (EMT). EMT is important in multiple events during early embryogenesis such as the formation of the placenta and the primitive streak. It ultimately leads to the generation of mesendoderm, which then gives rise to both the mesoderm and the endoderm. Members of the miR-200 family of miRNAs have been implicated in EMT and, as an example, have been shown to be important in targeting SIP1 (Chng *et al.,* 2010; Korpal & Kang, 2008), which promotes neuroectoderm formation at the expense of mesendoderm. Therefore, the inhibition of SIP1 by the miR-200 family has a positive effect on the commitment to mesendoderm (Chng *et al.,* 2010). In addition, downregulation of miR-200a is important in definitive endoderm formation, as it results in reduced repression of ZEB2, leading to downregulation of E-cadherin and subsequently

allowing EMT (Liao *et al.*, 2013). Other miR-200 family members also regulate EMT through their targets ZEB1 and ZEB2 (Gregory *et al.*, 2008).

5.1.2 The Role of miRNAs in Definitive Endoderm Formation

Numerous miRNAs have been identified as playing a role in pancreatic development. A limited number of studies have investigated miRNA expression at the definitive endoderm (DE) stage using an Activin A-based protocol to induce pluripotent stem cells to differentiate. These studies have shown that DE is characterised by a unique miRNA expression signature (Tzur *et al.*, 2008; Hinton *et al.,* 2010; Tsai *et al.,* 2010; Kim *et al.,* 2011; Porciuncula *et al.,* 2013). Although there is relatively little overlap in the miRNAs identified in these studies, which is likely to be due to the wide range of cell lines and differentiation protocols used, some miRNAs, including miR-375, miR-708 and miR-744, were identified in several of the studies, providing strong evidence for their involvement in DE formation (Hinton *et al*., 2010). However, few of these studies went on to elucidate a function for these miRNAs in DE formation. early study identified *Timm8a* as a target of miR-375 but didn't elucidate a function for this pathway in the formation of endoderm (Hinton *et al.*, 2010) and more recently, a study identified a direct relationship between miR-200a and *Sox17* (Liao et al., 2013) which is known to be critical in the development of DE.

5.1.3 Differences in miRNA Expression between ESCs and iPSCs

There have been several studies which have investigated differences in miRNA expression between ESCs and iPSCs, but although there has been some overlap in the miRNAs identified, so far there is little evidence for consistent differences between ESCs and iPSCs. Wilson *et al.* (2009) demonstrated that, although ESCs and iPSCs had very similar miRNA expression profiles, there were subtle differences between the two cell types. Specifically, although some stem cell-specific miRNAs were highly expressed in both ESCs and iPSCs, others, including the miR-520 and miR-371 clusters, were more highly expressed in ESCs than in iPSCs. In addition, some miRNAs, including miR-886-5p and let-7a, which are associated with differentiated cells, were more highly expressed in iPSCs than ESCs, suggesting that these miRNAs may be incompletely silenced during reprogramming. This

study observed a downward trend in miRNA expression when transitioning from fibroblast to iPSC, but the expression levels of many miRNAs do not decrease to the levels observed in ESCs. However, this study did not report the passage numbers of the cells used. Other studies have reported that these differences may be alleviated after extended passaging (Chin *et al*., 2009).

In a meta-analysis focussing on differences in gene expression between ESCs and iPSCs, Chin *et al*. (2009) also carried out miRNA profiling, and found little difference in miRNA expression between ESCs and iPSCs; however a few miRNAs were consistently differentially expressed between ESCs and iPSCs. These included some of the miRNAs identified by Wilson *et al*. (2009). Another study, using a larger sample group found that pluripotent stem cells could be divided into two distinct groups based on their miRNA expression levels. However, this segregation was not dependent on cell type as both iPSCs and ESCs were present in both groups: one group consisted of mainly ESCs with some virally-reprogrammed iPSCs, while the other consisted of iPSCs generated in a number of other ways, as well as the ESC line H9. This group did not report differences in expression of miR-302, miR-371 or miR-502 clusters between ESCs and iPSCs, as reported by the other studies mentioned (Chin *et al.*, 2009; Wilson *et al*., 2009), although there was some variability in their expression between cell lines.

Some of the most compelling evidence of differences in miRNA expression between ESCs and iPSCs comes from studies in isogenic murine ESCs and iPSCs investigating the imprinted *Dlk1-Dio3* region of the genome, which has been shown to be the only conserved difference between genetically identical mouse ESCs and iPSCs. This region, containing 26 miRNAs, is imprinted, and is maternally expressed in mammals. However, it was strongly repressed in most iPSCs compared to ESCs, and only iPSC lines that expressed this region were capable of giving rise to all-iPSC mice (Stadtfeld *et al*., 2010; Liu *et al*., 2010). However, it is unclear whether these differences are also seen in human stem cells, as they differ significantly from murine stem cells.

While there have been several studies investigating differences in miRNA expression between undifferentiated ESCs and iPSCs, there have been fewer investigating differences in their differentiated progeny. Wilson *et al*. (2009) reported some differences in miRNA expression in embryoid bodies derived from ESCs and iPSCs. However, the functional significance of these differences is still unclear.

Aims

The aim of this chapter is to investigate miRNAs which a) are up- or down-regulated upon differentiation to DE, and which could therefore be involved in this process; and b) are differentially expressed between ESCs and iPSCs, and which may lead to differences in the propensity of different cell lines to differentiate into DE, with a view to narrowing down the pool of candidate miRNAs to take forward into studies of possible function.

5.2 Methods

5.2.1. Choice of Cell Lines

Initial experiments were carried out using two iPSC lines generated at NIBSC, MRC5I and MRC9G, and a well-characterised ESC line (H9). However, the iPSC lines are grown on iSNL feeders, and the H9 line is grown on Matrigel™. For this reason, the H1 cell line was included in differentiation experiments and microarray analysis, as this is an ESC line grown on iSNL feeders, and was therefore expected to represent an ESC line grown in a comparable culture format to the available iPSC lines.

5.2.2 *In Vitro* Differentiation of Pluripotent Stem Cells into Definitive Endoderm

Stem cells were maintained in culture as previously described in section 2.2. Before the start of the differentiation protocol on day 1, samples of colonies containing undifferentiated cells were taken. Colonies undergoing differentiation were washed with PBS before the medium was changed to medium 1A (table 5.1). On day 2, the medium was changed to medium 1B, and differentiated colonies were harvested at the end of day 3.

Stage 1 **Formation Of Definitive Endoderm**			
Day	**1**	**2**	**3**
Reagents	**Medium 1A** RPMI L-glutamine (2mM) Activin A (100ng/ml) Wnt3A (25ng/ml)	**Medium IB** RPMI L-glutamine (2mM) Activin A (100ng/ml) Foetal Calf Serum (0.2%)	

Table 5.1. Reagents and concentrations used to direct differentiation of stem cells to definitive endoderm, as described in D'Amour et al (2006).

5.2.3 Microarray Analysis

Microarray analysis was used to analyse microRNA expression in samples of undifferentiated ESCs and iPSCs, as well as in ESCs and iPSCs that had undergone differentiation to the DE

stage. Samples were taken in triplicate, and total RNA, including the small RNA fraction, was isolated as described in section 2.5.2. Microarray hybridisation and image scanning were carried out by Exiqon Services, Denmark, who then supplied the raw data to be analysed at NIBSC. Background correction, normalisation and miRNA expression analysis was carried out using Nexus Expression software as described in section 2.5.1.

5.2.4 qRT-PCR For miRNA Expression

Microarray results were confirmed by qRT-PCR, which is highly recommended given the high chance of false positive results occurring during microarray analysis, even when a multiple testing correction is performed. Isolation of total RNA, reverse transcription and qRT-PCR for the analysis of miRNA expression were carried out as described in 2.4.3.

In order to select a suitable reference RNA to normalise the expression of the miRNAs of interest, the quantitation data from PCR analysis of several potential reference RNAs were compared and analysed using GenEx software across a range of samples from both differentiated and undifferentiated ES and iPS cells. SNORD43 was found to have the least variability between samples (figure 5.1, highlighted in red), so was chosen as a reference gene. SNORA66 was analysed in 2 separate assays (16.06.2011 and 17.06.2011), and had similar variation in both. miR-367, a pluripotency-associated miRNA, which is expected to vary between undifferentiated and differentiated samples, has the highest variation. This demonstrates the validity of this approach for choosing a reference gene for miRNA analysis.

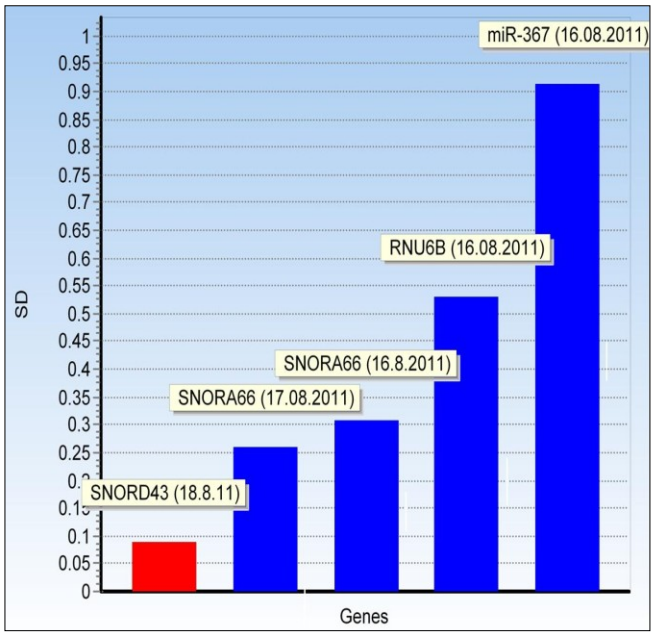

Figure 5.1 A plot of the standard deviation of quantitation data for miRNAs, *analysed using GenEx software (Exiqon).*

5.3 Results

5.3.1 Analysis of Microarray Results

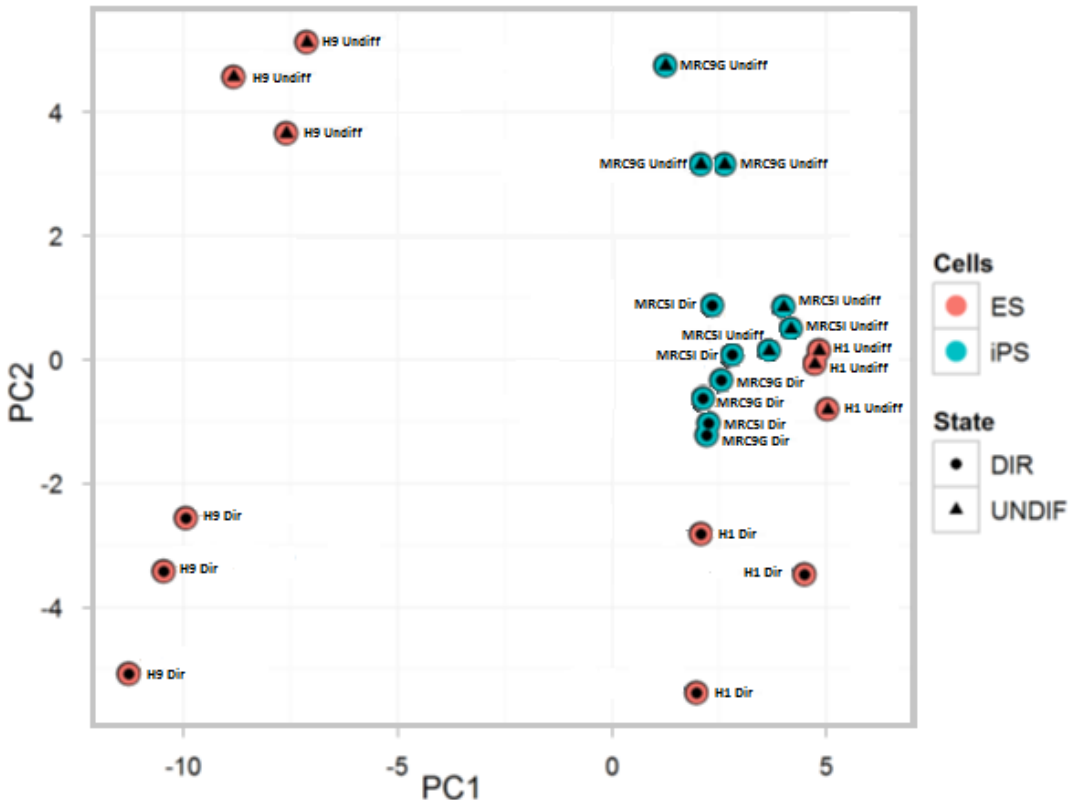

Figure 5.2.Principal component analysis *allows identification of the factors associated with differences in miRNA expression between different samples. Samples which cluster closely together have more similar miRNA expression profiles, while samples which are further apart have more dissimilar miRNA expression profiles.*

Principal component analysis (PCA) is a way of looking at variance between samples. If biological differences between the samples are pronounced, this will be a principal component of the variation and the samples will separate along the axis accordingly. PCA allows visual identification of differences in miRNA expression related to biological or technical factors, providing a useful summary of the microarray data. For example, in figure 5.2, differentiated cells (circles) cluster together, separately from undifferentiated cells (triangles). ESCs (red) and iPSCs (blue) also cluster separately. H9 cells cluster separately from the other cell lines (H1, MRC5I and MRC9G). Although the possibility exists that H9 cells genuinely have a significantly different miRNA expression profile to other pluripotent stem cells, it is important to note that these cells are grown on Matrigel™ rather than iSNL feeders which support the growth of H1, MRC5I and MRC9G cell lines, which suggests that culture conditions may also have an impact on miRNA expression.

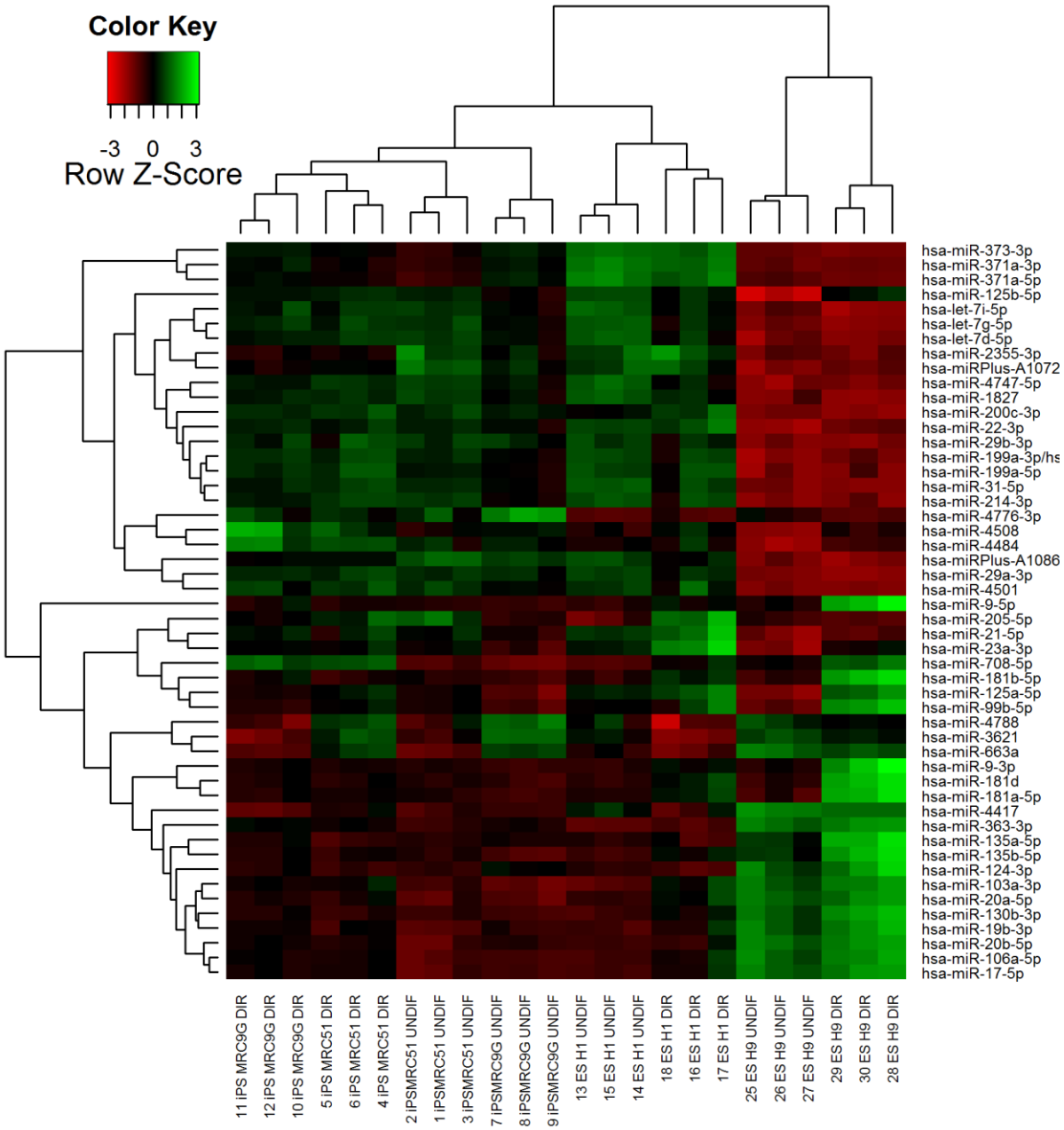

Figure 5.3. Unsupervised hierarchical clustering analysis *based on the expression of the 50 miRNAs with the highest variation between samples. miRNAs which are expressed at lower levels than in the reference sample are shown in red, while miRNAs which are expressed at higher levels are shown in green.*

Unsupervised hierarchical clustering analysis (figure 5.3) looks at the 50 miRNAs with the highest variation between samples to see how either samples (horizontal clusters) or miRNAs (vertical clusters) cluster together based on this variance. Hierarchical clustering treats each data point as a single cluster, and then successively merges clusters based on their similarity until all points have been merged into a single cluster. This is represented by the

dendrograms. This allows quick visual identification of the samples which are most similar in terms of their miRNA expression profile.

Figure 5.3 shows that the two iPSC lines (MRC5I and MRC9G) cluster together, with undifferentiated samples distinct from differentiated samples. This shows that these iPSC lines are highly similar in terms of their miRNA expression, both at the undifferentiated and the DE stage. The H1 ESC line, which is grown on feeders, is the most similar to these iPSC lines, but clusters separately from them, indicating differences in miRNA expression. The H9 cell line, which is grown on Matrigel™, clusters separately both from the iPSC lines and the ESC line on feeders, again suggesting that different culture conditions may result in differences in miRNA expression patterns in this cell line. Figure 5.3 clearly shows that many of the 50 miRNAs shown are differentially expressed in H9 compared to the other cell lines, confirming the likely distinctions identified in the PCA.

Figure 5.3 also shows that miRNAs from the same family cluster together: for example, miR-373-3p, miR-371a-3p, and miR-371a-5p cluster together at the top of the list. These miRNAs are all part of the same family, whose members are highly expressed in pluripotent stem cells, where they play an important role in maintaining pluripotency (Suh et al, 2004). Several members of the let-7 family, which play a role in differentiation (Morin et al., 2008) are also seen to cluster together. Several members (miR-17-5p, miR-106a-5p, miR-20b-5p, miR-19b-3p, miR-20a-5p) of the miR-17~25 cluster, composed of two miRNA families, which are involved in maintaining pluripotency and controlling differentiation (Laurent *et al.*, 2008), also cluster together at the bottom of the list. These results suggest that miRNAs from the same families, which are likely to be present in the same regions of the genome, and to be co-regulated and co-transcribed, have similar expression patterns in these samples. In addition, the 3p and 5p forms of some miRNAs also cluster together e.g. miR-371a-3p and -5p; miR-135a-3p and -5p. These different forms of the miRNA are derived from different ends of the same pre-miRNA molecule, and are separated as the miRNAs are processed (see figure 1.10), which means that their expression patterns are highly similar, although they are likely to have different target genes. This indicates that the results of the microarray follow expected biological patterns and offers confidence in the validity of the results.

miRNAs Differentially Expressed Between Undifferentiated & Differentiated Samples

miRNAs Upregulated In DE Formation												
miRNA	ESC		iPSC		MRC5I		MRC9G		H9		H1	
	p-value	*logFC*	*p-value*	*logFC*	*p-value*	*logFC*	*p-value*	*logFC*	*p-value*	*logFC*	*p-value*	*logFC*
hsa-miR-375	0.019	0.502	0.042	0.876	0.017	0.620	**0.116**	**1.210**	0.092	0.479	0.048	0.404
hsa-miR-708-5p	0.050	1.066	0.000	1.702	0.021	1.082	0.048	1.464	0.020	1.835	0.083	1.095
hsa-miR-744-5p	0.029	0.471	0.038	0.655	0.064	0.548	0.085	0.713	**0.234**	**0.659**	0.084	0.430
hsa-miR-27b-3p	0.036	0.846	0.070	0.307	0.077	0.550	**0.489**	**0.213**	0.085	0.308	0.084	1.400
hsa-miR-26b-5p	0.021	0.574	0.066	0.356	0.065	0.569	**0.747**	**0.108**	0.092	0.499	0.084	0.659
hsa-miR-30b-5p	0.050	0.699	0.068	0.252	0.065	0.719	**0.565**	**0.177**	**0.108**	**0.278**	**0.104**	**0.701**
hsa-miR-4530	0.019	0.687	0.051	0.929	0.025	0.742	**0.118**	**1.078**	**0.155**	**1.002**	**0.109**	**0.648**
hsa-miR-151a-3p	0.094	0.406	0.019	0.188	**0.173**	**0.167**	**0.299**	**0.166**	**0.106**	**0.195**	0.084	0.780
hsa-miR-151a-5p	0.066	0.507	0.020	0.181	**0.173**	**0.167**	**0.201**	**0.189**	**0.165**	**0.174**	0.084	0.824
hsa-miR-191-5p	0.023	0.453	0.081	0.304	**0.234**	**0.423**	**0.331**	**0.289**	**0.273**	**0.260**	0.084	0.511

miRNAs Downregulated In DE Formation												
miRNA	ESC		iPSC		MRC5I		MRC9G		H9		H1	
	p-value	*logFC*	*p-value*	*logFC*	*p-value*	*logFC*	*p-value*	*logFC*	*p-value*	*logFC*	*p-value*	*logFC*
hsa-miR-3941	0.065	-0.480	0.042	-0.376	0.059	-0.646	**0.163**	**-0.244**	**0.236**	**-0.391**	0.077	-0.462
hsa-miR-3148	0.066	-0.302	0.041	-0.216	0.101	-0.323	**0.226**	**-0.167**	**0.250**	**-0.263**	**0.104**	**-0.363**
hsa-miR-124-5p	0.015	-0.207	0.063	-0.148	**0.116**	**-0.267**	**0.464**	**-0.053**	**0.249**	**-0.139**	0.084	-0.438
hsa-miR-4285	0.045	-0.555	0.066	-0.578	**0.118**	**-0.692**	**0.335**	**-0.541**	**0.979**	**-0.007**	0.083	-0.868
hsa-miR-3935	0.050	-0.236	**0.128**	**-0.190**	**0.351**	**-0.177**	**0.287**	**-0.224**	**0.218**	**-0.180**	0.089	-0.336
hsa-miR-378a-3p	0.021	-0.309	**0.313**	**-0.165**	**0.986**	**0.008**	**0.335**	**-0.238**	0.086	-0.356	0.075	-0.184
hsa-miR-4451	0.035	-0.351	**0.163**	**-0.200**	**0.729**	**-0.060**	**0.155**	**-0.339**	0.015	-0.442	**0.148**	**-0.268**
hsa-miR-516b-5p	0.039	-0.232	**0.145**	**-0.125**	**0.116**	**-0.155**	**0.583**	**-0.081**	0.072	-0.276	0.088	-0.190
hsa-miR-4436b-5p	**0.269**	**-0.277**	0.037	-0.371	0.056	-0.439	**0.278**	**-0.330**	**0.310**	**-0.112**	0.083	-0.446
hsa-miR-4732-3p	**0.380**	**-0.389**	0.005	-0.719	**0.204**	**-0.618**	0.079	-0.789	**0.952**	**0.011**	0.083	-0.864

Table 5.2. Comparison of miRNA expression in differentiated vs. undifferentiated samples, allowing identification of the 10 miRNAs most highly up- (upper table, highlighted in green) or down- (lower table, highlighted in red) regulated during differentiation of ESCs and iPSCs to definitive endoderm. *Individual comparisons were carried out for ESCs, iPSCs and each individual cell line included in the microarray analysis. Lighter shades denote miRNAs with a p-value of < 0.1, darker shades denote miRNAs with a p-value of < 0.05*

Comparison of differentiated and undifferentiated samples was carried out between ESCs, iPSCs and for each individual cell line included in the microarray analysis. A total of 82 miRNAs were identified as being differentially expressed between differentiated and undifferentiated ESCs (shown in supplementary table 1) and 86 miRNAs were differentially expressed between differentiated and undifferentiated iPSCs (shown in supplementary table 2). These miRNAs may play a role in differentiation to DE. As it is not practical to further investigate so many results, the miRNAs were ranked by p-value, and the top 10 miRNAs which were most highly up- or down-regulated in ESCs and iPSCs were selected for further

analysis and are shown in table 5.2. The expression of these miRNAs in each individual cell line included in the microarray analysis is also shown in table 5.2, showing some variability in expression of these miRNAs between cell lines, but also that the pattern of up- or down-regulation is similar in many of the cell lines. For example, miR-375 is consistently upregulated in all the cell lines tested, although this upregulation is not statistically significant in MRC9G. The same is true for miR-708, which is strongly upregulated in most cell lines, and moderately upregulated in H1. The patterns are less strong for miRNAs which are down-regulated in DE formation, perhaps because this is such an early stage of differentiation. Several miRNAs on the list have been previously described as being expressed at the DE stage, including miR-375 (Poy *et al*., 2004; Kloosterman *et al*., 2007; Tzur *et al*; 2008; Hinton *et al,* 2009; Kim *et al*, 2011), miR-708-5p (Hinton *et al*, 2009; Kim *et al*, 2011), and miR-744 (Kim *et al,* 2011), suggesting that the DE generated in these studies is similar to that generated by other researchers with respect to its miRNA expression profile.

However, the very different propensities of H1 and H9 to give rise to DE (see chapter 4), and the fact that the different culture formats (feeders vs. Matrigel™) may well result in significant differences in miRNA expression, means that combining the data for these cell lines in order to identify miRNAs involved in DE formation could potentially lead to important candidate miRNAs being overlooked for further study. For this reason, the analysis was repeated, firstly excluding H1, which has a low propensity for DE formation, and then excluding H9, which has a unique miRNA expression pattern. The top 10 miRNAs most highly up- and down-regulated in each analysis are shown in tables 5.3 and 5.4, respectively. miRNAs also identified in table 5.2 are shown in bold.

miRNAs Upregulated In DE Formation				
miRNA	iPSC		H9	
	p-value	*logFC*	*p-value*	*logFC*
hsa-miR-708-5p	**0.000**	**1.702**	**0.021**	**1.082**
hsa-miR-302c-3p	0.020	0.558	0.021	-0.654
hsa-miR-4791	0.020	0.373	0.030	-0.419
hsa-miR-4532	0.037	0.589	0.015	0.670
hsa-miR-375	**0.042**	**0.876**	**0.017**	**0.620**
hsa-miR-4484	0.042	0.803	0.017	1.043
hsa-miR-4289	0.017	0.236	0.053	0.570
hsa-miR-491-3p	0.043	-0.679	0.029	0.731
hsa-miR-4530	**0.051**	**0.929**	**0.025**	**0.742**
hsa-miR-93-5p	0.011	0.631	0.066	0.829

miRNAs Downregulated In DE Formation				
miRNA	iPSC		H9	
	p-value	*logFC*	*p-value*	*logFC*
hsa-miR-302c-3p	0.020	0.558	0.021	-0.654
hsa-miR-4791	0.020	0.373	0.030	-0.419
hsa-miR-491-3p	0.043	-0.679	0.029	0.731
hsa-miR-302f	0.020	0.339	0.057	-0.554
hsa-miR-9-3p	0.011	-0.331	0.066	0.257
hsa-miR-659-3p	0.011	-0.331	0.066	0.257
hsa-miR-302a-5p	0.039	0.350	0.065	-0.606
hsa-miR-675-5p	0.081	-0.287	0.026	0.599
hsa-miR-3651	0.098	0.306	0.019	-0.338
hsa-miR-5701	0.098	0.380	0.030	-0.395

Table 5.3 Comparison of differentiated vs. undifferentiated samples from iPSCs and H9 showing the 10 miRNAs most highly up- (left table, highlighted in green) or down- (right table, highlighted in red) regulated during differentiation of H9 and iPSCs to definitive endoderm. *Lighter shades denote miRNAs with a p-value of < 0.1, darker shades denote miRNAs with a p-value of < 0.05. miRNAs which also appear on the top 10 list when all cell lines are included in the analysis are shown in bold.*

Table 5.3 shows the top 10 miRNAs most highly up- or down-regulated upon DE formation when H1 is excluded from the analysis. There are considerable differences between iPSCs and H9, with several miRNAs that are strongly upregulated in iPSCs being downregulated in H9 (e.g. miR-302c-3p, miR-4791) and vice versa (e.g. miR-491-3p, miR-9-3p), with the result that several of these miRNAs appear on both the lists of miRNAs most strongly upregulated in DE formation and most strongly downregulated. These miRNAs are clearly candidates for further study and again shows that differentiating iPSC lines exhibit a distinct miRNA expression pattern to differentiating H9 cells, even though the differentiation propensity is not manifestly different from analyses of mRNA expression (see chapter 4). miRNAs which also appear in table 5.2 are shown in bold, and include miR-708-5p, miR-375 and miR-4530, which are strongly upregulated in both iPSC lines and H9.

miRNAs Upregulated In DE Formation				
miRNA	iPSC		H1	
	p-value	*logFC*	*p-value*	*logFC*
hsa-miR-708-5p	**0.000**	**1.702**	**0.083**	**1.095**
hsa-miR-107	0.010	0.372	0.075	0.807
hsa-miR-375	**0.042**	**0.876**	**0.048**	**0.404**
hsa-miR-93-5p	0.011	0.631	0.083	0.538
hsa-miR-17-5p	0.011	0.544	0.084	0.807
hsa-miR-106a-5p	0.008	0.569	0.088	0.804
hsa-miR-103a-3p	0.013	0.633	0.084	1.001
hsa-miR-92a-3p	0.011	0.464	0.088	0.624
hsa-miR-4289	0.017	0.236	0.084	0.635
hsa-miR-151a-3p	**0.019**	**0.188**	**0.084**	**0.780**

miRNAs Downregulated In DE Formation				
miRNA	iPSC		H1	
	p-value	*logFC*	*p-value*	*logFC*
hsa-miR-4732-3p	**0.005**	**-0.719**	**0.083**	**-0.864**
hsa-miR-9-3p	0.011	-0.331	0.084	0.933
hsa-miR-32-3p	0.005	-0.719	0.090	-0.361
hsa-miR-3941	**0.042**	**-0.376**	**0.077**	**-0.462**
hsa-miR-4436b-5p	**0.037**	**-0.371**	**0.083**	**-0.446**
hsa-miR-3148	**0.041**	**-0.216**	**0.104**	**-0.363**
hsa-miR-3591-5p	0.035	-0.406	**0.111**	**-0.200**
hsa-miR-124-5p	**0.063**	**-0.148**	**0.084**	**-0.438**
hsa-miR-4285	**0.066**	**-0.578**	**0.083**	**-0.868**
hsa-miR-2116-5p	0.044	-0.242	**0.118**	**-0.199**

Table 5.4. Top 10 miRNAs most highly up- (left table, highlighted in green) or down- (right table, highlighted in red) regulated during differentiation of H1 and iPSCs to definitive endoderm. *Lighter shades denote miRNAs with a p-value of < 0.1, darker shades denote miRNAs with a p-value of < 0.05. miRNAs which also appear on the top 10 list when all cell lines are included are shown in bold.*

Table 5.4 shows the top 10 miRNAs most highly up- or down-regulated upon DE formation when H9 is excluded from the analysis. There is considerably more overlap between this analysis and the analysis including all the cell lines than in the analysis excluding H1 (table 5.3), suggesting that, as expected from the PCA and hierarchical clustering, iPSCs and H1 ESCs exhibit more similarity in their miRNA expression patterns than iPSCs vs. H9 ESCs, possibly as a result of both populations being grown on feeders. This is particularly true for miRNAs that are downregulated in DE formation, including miR-4732-3p, miR-3941, miR-4436b-5p, miR-124-5p, and miR-4284. Once again, the miRNAs that are consistently upregulated include miR-708-5p and miR-375, further suggesting a role for these miRNAs in DE formation.

Comparisons of differentiated and undifferentiated samples were also carried out for each individual cell line. 28 miRNAs were significantly upregulated in MRC5I, while 5 were downregulated (supplementary table 3). 24 miRNAs were significantly upregulated in MRC9G, while 34 were downregulated (supplementary table 4). 106 miRNAs were significantly upregulated in H9, while 48 were downregulated (supplementary table 5). 52 miRNAs were significantly upregulated in H1, while 16 were downregulated (supplementary table 6). The top 10 most highly up- and down-regulated miRNAs for each cell line are shown in tables 5.5, 5.6, 5.7 and 5.8. miRNAs also identified in table 5.2 are shown in bold.

Once again, the miRNAs that appear most frequently on the lists of miRNAs up-regulated in DE formation include miR-708-5p and miR-375, which is a strong indication that these miRNAs play an important role in DE formation. Many of the miRNAs downregulated in DE formation also appear in these individual analyses, although these are less consistent between cell lines. This suggests that the combined analysis is an adequate starting point for the identification of miRNA species for validation by qRT-PCR and ultimately for further studies of function. In addition, miRNA species which are differentially expressed in iPSCs when compared to the ESC cell lines individually (i.e. H1 or H9) will also be investigated by qRT-PCR, e.g. miR-302c-3p, miR-4791, and miR-491-3p.

MRC5I		
miRNA	*p-value*	*logFC*
hsa-miR-3676-3p	0.004	1.357
hsa-miR-4707-5p	0.004	0.776
hsa-miR-1469	0.009	0.761
hsa-miR-585	0.046	0.829
hsa-miR-4488	0.046	1.029
hsa-miR-4463	0.046	0.812
hsa-miR-708-5p	**0.048**	**1.464**
hsa-miR-4792	0.048	0.975
hsa-miR-638	0.048	0.707
hsa-miR-664-3p	0.048	0.193
hsa-miR-4436b-5p	**0.056**	**-0.439**
hsa-miR-3941	**0.059**	**-0.646**
hsa-miR-1321	0.063	-0.169
hsa-miR-9-3p	0.085	-0.494
hsa-miR-659-3p	0.085	-0.494

Table 5.8. Comparison of differentiated vs. undifferentiated samples, showing the top 10 miRNAs up- and down-regulated upon endoderm formation in the MRC5I cell line. *Lighter shades denote miRNAs with a p-value of < 0.1, darker shades denote miRNAs with a p-value of < 0.05. miRNAs which also appear on the top 10 list when all cell lines are included are shown in bold. Only 5 miRNAs were significantly down-regulated in this cell line.*

MRC9G		
miRNA	*p-value*	*logFC*
hsa-miR-708-5p	**0.020**	**1.835**
hsa-miR-4791	0.079	0.512
hsa-miR-361-5p	0.079	0.335
hsa-miR-106a-5p	0.079	0.546
hsa-miR-93-5p	0.079	0.667
hsa-miR-103a-3p	0.079	0.678
hsa-miR-106b-5p	0.079	0.415
hsa-miR-302c-3p	0.079	0.674
hsa-miR-20a-5p	0.079	0.529
hsa-miR-302a-5p	0.079	0.415
hsa-miR-3183	0.079	-0.318
hsa-miR-4732-3p	**0.079**	**-0.789**
hsa-miR-32-3p	0.079	-0.789
hsa-miR-4456	0.079	-0.572
hsa-miR-4417	0.079	-0.404
hsa-miR-4768-5p	0.079	-0.661
hsa-miR-519	0.079	-0.385
hsa-miR-3158-5p	0.085	-1.273
hsa-miR-665	0.091	-0.551
hsa-miR-361-3p	0.092	-0.462

Table 5.7. Comparison of differentiated vs. undifferentiated samples, showing the top 10 miRNAs up- and down-regulated upon endoderm formation in the MRC9G cell line. *Lighter shades denote miRNAs with a p-value of < 0.1, darker shades denote miRNAs with a p-value of < 0.05. miRNAs which also appear on the top 10 list when all cell lines are included are shown in bold.*

H1		
miRNA	*p-value*	*logFC*
hsa-miR-375	**0.048**	**0.404**
hsa-miR-107	0.075	0.807
hsa-miR-125a-5p	0.075	0.706
hsa-miR-181a-5p	0.075	1.216
hsa-miR-193b-3p	0.075	1.116
hsa-miR-708-5p	**0.083**	**1.095**
hsa-miR-93-5p	0.083	0.538
hsa-miR-320a	0.083	0.859
hsa-miR-23a-3p	0.083	1.284
hsa-miR-181d	0.083	1.284
hsa-miR-3158-5p	0.048	-0.849
hsa-miR-4639-3p	0.063	-0.443
hsa-miR-3621	0.075	-0.799
hsa-miR-378a-3p	0.075	-0.184
hsa-miR-3941	**0.077**	**-0.462**
hsa-miR-4732-3p	**0.083**	**-0.864**
hsa-miR-4436b-5p	**0.083**	**-0.446**
hsa-miR-4285	**0.083**	**-0.868**
hsa-miR-302b-3p	0.084	-0.278
hsa-miR-124-5p	**0.084**	**-0.438**

Table 5.6. Comparison of differentiated vs. undifferentiated samples, showing the top 10 miRNAs up- and down-regulated upon endoderm formation in the H1 cell line. *Lighter shades denote miRNAs with a p-value of < 0.1, darker shades denote miRNAs with a p-value of < 0.05. miRNAs which also appear on the top 10 list when all cell lines are included are shown in bold.*

H9		
miRNA	*p-value*	*logFC*
hsa-miR-99b-5p	0.015	2.187
hsa-miR-125a-5p	0.015	2.040
hsa-miR-181d	0.015	2.631
hsa-miR-181b-5p	0.016	2.074
hsa-miR-181a-5p	0.017	3.151
hsa-miR-483-5p	0.017	0.702
hsa-miR-375	**0.017**	**0.620**
hsa-miR-708-5p	**0.021**	**1.082**
hsa-miR-4306	0.026	0.579
hsa-miR-5572	0.029	0.460
hsa-miR-302b-3p	0.015	-0.727
hsa-miR-3651	0.019	-0.338
hsa-miR-5701	0.030	-0.395
hsa-miR-5002-5p	0.032	-0.347
hsa-miR-3621	0.050	-0.392
hsa-miR-516b-5p	**0.072**	**-0.276**
hsa-miR-378a-3p	**0.086**	**-0.356**
hsa-miR-664-3p	0.092	-0.098
hsa-miR-200c-3p	0.092	-0.295

Table 5.5. Comparison of differentiated vs. undifferentiated samples, showing the top 10 miRNAs up- and down-regulated upon endoderm formation in the H9 cell line. *Lighter shades denote miRNAs with a p-value of < 0.1, darker shades denote miRNAs with a p-value of < 0.05. miRNAs which also appear on the top 10 list when all cell lines are included are shown in bold. Only 9 miRNAs were significantly down-regulated in this cell line.*

miRNAs Differentially Expressed Between iPSCs & ESCs: Undifferentiated Samples

When undifferentiated iPSCs and undifferentiated ESCs were compared, there were no miRNAs that were significantly differently expressed, suggesting that starting populations of these cell populations are equivalent in terms of their miRNA expression. This was a rather surprising result and is in contrast to several previous studies (Wilson et al, 2009; Chin et al, 2009; Neveu et al, 2010), which reported miRNAs that were differentially expressed between undifferentiated iPSCs and ESCs. Since this result was unexpected, and since we know that H9 and H1 ESC lines exhibit different miRNA expression profiles, undifferentiated iPSCs were also compared with undifferentiated H1 and H9 ESC lines individually.

When undifferentiated iPSCs were compared to undifferentiated H1, only 5 miRNAs were significantly differentially expressed: 3 were down regulated in iPSCs compared to H1, and 2 were upregulated (supplementary table 7). The fact that so few miRNAs are differentially expressed between H1 and the two iPSC lines suggest that they are very similar in the undifferentiated state. However, when undifferentiated iPSCs were compared to undifferentiated H9, 186 miRNAs were differentially expressed: 107 were downregulated in iPSCs compared to H9 and 79 were upregulated (supplementary table 8). The large number of miRNAs differentially expressed between H9 and the iPS cell lines suggests that even in the undifferentiated state, culture conditions can influence miRNA expression levels. This is despite the extensive characterisation carried out on these cells which showed that these cell lines are equivalent in terms of their expression of pluripotency genes (see chapter 3).

miRNAs Differentially Expressed Between iPSCs & ESCs: Differentiated Samples

91 miRNAs were significantly ($p < 0.05$) differentially expressed between differentiated iPSCs and ESCs (table 5.9). 61 miRNAs were downregulated in iPSCs vs. ESCs (red), while 30 were upregulated (green). Several of these miRNAs were also identified as being up- or down-regulated in differentiated samples compared to undifferentiated samples (table 5.2), and these are shown in bold.

Comparison of miRNA Expression in Differentiated iPSCs vs. Differentiated ESCs		
miRNA	**p-value**	**LogFC**
hsa-miR-151a-5p	**0.0035**	**-0.737**
hsa-miR-151b	0.0035	-1.3088
hsa-miR-99b-5p	0.0035	-1.3789
hsa-miR-125a-5p	0.0036	-0.8069
hsa-miR-130a-3p	0.0043	-0.6481
hsa-miR-181a-5p	0.0043	-2.0135
hsa-miR-191-5p	0.0043	-0.6342
hsa-miR-197-3p	0.0043	-0.3174
hsa-miR-30b-5p	**0.0043**	**-0.9451**
hsa-miR-30c-5p	0.0043	-0.8929
hsa-miR-320a	0.0043	-0.6468
hsa-miR-320c	0.0043	-0.9285
hsa-miR-335-5p	0.0043	-0.9878
hsa-miR-3653	0.0043	-0.5547
hsa-miR-3685	0.0043	0.2931
hsa-miR-423-3p	0.0043	-0.8026
hsa-miR-20a-5p	0.0046	-0.9107
hsa-miR-27b-3p	**0.005**	**-0.9839**
hsa-miR-15b-5p	0.0054	-0.4956
hsa-miR-320b	0.0054	-0.7598
hsa-miR-181d	0.0056	-1.9588
hsa-miR-30d-5p	0.0056	-0.5706
hsa-miR-204-3p	0.0064	0.7034
hsa-miR-4419b	0.0064	0.4019
hsa-miR-4505	0.0064	0.7786
hsa-miR-454-3p	0.0064	-0.8979
hsa-miR-92a-3p	0.0064	-0.6663
hsa-miR-103a-3p	0.0104	-0.7494
hsa-miR-151a-3p	**0.0104**	**-0.4902**
hsa-miR-16-5p	0.0107	-0.5991
hsa-miR-4484	0.0116	1.1915
hsa-miR-17-5p	0.0122	-0.8922
hsa-miR-3646	0.0122	0.3478
hsa-miR-106a-5p	0.0143	-0.8975
hsa-miR-30e-5p	0.0143	-0.6218
hsa-miR-4532	0.0143	0.6558
hsa-miR-93-5p	0.0143	-0.6366
hsa-miR-130b-3p	0.0149	-1.2139
hsa-miR-181b-5p	0.0152	-1.446
hsa-miR-3687	0.0152	-1.0384
hsa-miR-26b-5p	**0.0154**	**-0.4076**
hsa-miR-4299	0.0154	0.4855
hsa-miR-711	0.0165	0.9775
hsa-miR-107	0.018	-0.7897
hsa-miR-19a-3p	0.0183	-0.6169
hsa-miR-106b-5p	0.0191	-0.8247
hsa-miR-1587	0.0191	0.7501
hsa-miR-34a-5p	0.0191	-0.6565
hsa-miR-378a-3p	**0.0191**	**0.23**
hsa-miR-4780	0.0191	-0.3121
hsa-miR-135b-5p	0.0193	-1.6118
hsa-miR-4284	0.0195	-0.2847
hsa-miR-17-3p	0.0197	-0.7416
hsa-miR-513a-5p	0.0199	0.4952
hsa-miR-342-3p	0.0218	-0.9452
hsa-miR-19b-3p	0.0219	-0.8428
hsa-miR-4657	0.0248	-0.6466
hsa-miR-331-3p	0.0255	-0.6998
hsa-miR-4792	0.0255	0.8258
hsa-miR-150-5p	0.0261	0.1916
hsa-miR-4286	0.0263	-0.6974
hsa-miR-4301	0.0263	-0.5148
hsa-miR-9-3p	0.0263	-2.1832
hsa-miR-335-3p	0.0277	-0.2762
hsa-miR-92b-3p	0.0277	-0.9395
hsa-miR-18b-5p	0.0284	-0.8408
hsa-miR-1273g-3p	0.0305	-0.334
hsa-miR-4417	0.0305	-1.487
hsa-miR-4714-5p	0.0305	-0.3333
hsa-miR-4530	**0.0314**	**0.7268**
hsa-miR-4289	0.0321	-0.9101
hsa-miR-4508	0.0338	0.8684
hsa-miR-4726-5p	0.0343	0.4853
hsa-miR-1909-3p	0.035	0.2531
hsa-let-7c	0.036	0.8174
hsa-miR-149-3p	0.0378	0.5623
hsa-miR-25-3p	0.0378	-0.4055
hsa-miR-548an	0.0378	-0.2313
hsa-miR-9-5p	0.0378	-2.257
hsa-miR-654-3p	0.0394	0.7269
hsa-miR-3201	0.0413	0.6605
hsa-miR-15a-5p	0.0427	-0.6342
hsa-miR-936	0.0433	0.9696
hsa-miR-4747-5p	0.0434	0.8154
hsa-miR-4667-5p	0.0447	0.328
hsa-miR-30c-1-3p	0.0465	0.3686
hsa-let-7g-5p	0.0475	1.1879
hsa-miR-4712-3p	0.0479	-0.2099

Table 5.9. miRNAs significantly ($p < 0.05$) differentially expressed between iPSCs and ESCs *that had been directed to differentiate into definitive endoderm, ranked in order of significance. miRNAs that are upregulated in iPSCs vs. ESCs are highlighted in green, while miRNAs that are downregulated in iPSCs vs. ESCs are highlighted in red. The miRNAs that are also included on the lists of miRNAs most highly up- or down-regulated during endoderm formation are in bold.*

However, given the very different propensities of H1 and H9 to give rise to DE, and the fact that the different culture methods appear to result in significant differences in miRNA expression, combining the data for these cell lines in order to identify miRNAs differently expressed between ESCs and iPSCs is problematic. For this reason, the analysis was repeated, firstly excluding H9 to eliminate any differences derived from culture methods (table 5.10), and secondly excluding H1, which does not have a high propensity for DE formation (table 5.11). Finally, a comparison of differentiated H1 and H9 was carried out (table 5.12).

Comparison of miRNA Expression in Differentiated iPSCs vs. Differentiated H1		
miRNA	**LogFC**	**P-Value**
hsa-miR-371a-5p	1.8662	0.018
hsa-miR-4288	0.4336	0.018
hsa-miR-3653	0.6953	0.018
hsa-miR-23b-3p	1.4598	0.018
hsa-miR-24-3p	1.1901	0.018
hsa-miR-23a-3p	1.428	0.018
hsa-miR-181a-5p	1.0852	0.018
hsa-miR-27b-3p	**1.2984**	**0.019**
hsa-miR-34a-5p	1.0237	0.019
hsa-miR-125a-5p	0.9247	0.0213
hsa-miR-222-3p	0.4793	0.0213
hsa-miR-363-3p	-0.8662	0.0213
hsa-miR-373-3p	1.7474	0.0213
hsa-miR-151a-5p	**0.6916**	**0.0213**
hsa-miR-27a-3p	1.2937	0.0213
hsa-miR-2355-3p	0.756	0.0359
hsa-miR-371a-3p	1.6578	0.0403
hsa-miR-708-5p	**-0.6927**	**0.0408**
hsa-miR-30c-5p	0.5951	0.0408
hsa-miR-22-3p	0.7858	0.042
hsa-miR-99b-5p	1.1732	0.042
hsa-miR-30b-5p	0.7343	0.042
hsa-miR-1909-3p	-0.3628	0.042
hsa-miR-518b	-0.3893	0.042
hsa-miR-151a-3p	**1.2359**	**0.042**
hsa-miR-302d-5p	-0.2972	0.042
hsa-miR-181d	1.0152	0.0426
hsa-miR-21-5p	1.117	0.0478

Table 5.10.miRNAs significantly (p < 0.05) differentially expressed between differentiated iPSCs and differentiated H1. *miRNAs that are upregulated in iPSCs vs. H1 are highlighted in green, while miRNAs that are downregulated in iPSCs vs. H1 are highlighted in red. The miRNAs that are also included on the lists of miRNAs most highly up- or down-regulated during endoderm formation are in bold.*

Comparison of miRNA Expression in Differentiated iPSCs vs. Differentiated H9		
miRNA	**LogFC**	**P-Value**
hsa-miR-29a-3p	-3.3296	0
hsa-miR-17-5p	1.2259	0
hsa-miR-200c-3p	-2.6019	0
hsa-miR-181a-5p	2.6324	0
hsa-miR-106a-5p	1.2456	0
hsa-miR-107	1.1469	0
hsa-miR-130b-3p	1.7291	0
hsa-miR-92b-3p	1.4253	0
hsa-miR-31-5p	-1.7119	0
hsa-miR-92a-3p	0.8751	0
hsa-let-7d-5p	-1.5536	0
hsa-miR-4288	-0.6438	0
hsa-miR-3687	1.4838	0
hsa-miR-20b-5p	1.2647	0
hsa-miR-1827	-1.9765	0
hsa-miR-151a-5p	**0.7672**	**0**
hsa-miR-4747-5p	-1.2849	0
hsa-miR-181d	2.5878	0
hsa-miR-4501	-3.2632	0
hsa-miR-4289	1.3984	0
hsa-miR-93-5p	0.8917	0
hsa-miR-936	-1.517	0
hsa-miR-103a-3p	0.9847	0
hsa-miR-363-3p	1.5835	0
hsa-miR-9-3p	3.2816	0
hsa-miR-423-3p	0.993	0
hsa-let-7g-5p	-1.9037	0
hsa-miR-199a-3p	-1.7248	0
hsa-miR-22-3p	-1.5781	0
hsa-miR-135b-5p	2.3319	0
hsa-miR-30c-5p	1.0914	0
hsa-miR-373-3p	-2.6799	0.0001
hsa-miR-106b-5p	1.1741	0.0001
hsa-miR-320c	1.1126	0.0001
hsa-miR-4500	-0.488	0.0001
hsa-miR-20a-5p	1.1417	0.0001
hsa-miR-342-3p	1.373	0.0001
hsa-miR-135a-5p	2.1674	0.0001
hsa-miR-361-5p	0.9664	0.0001
hsa-miR-124-3p	1.533	0.0001
hsa-miR-17-3p	1.0524	0.0001
hsa-miR-513a-5p	-0.6881	0.0001
hsa-miR-30b-5p	**1.0857**	**0.0001**
hsa-let-7i-5p	-1.6503	0.0001
hsa-miR-34c-5p	-1.0879	0.0001
hsa-let-7c	-1.2486	0.0001
hsa-miR-9-5p	3.494	0.0001
hsa-miR-130a-3p	0.7934	0.0001
hsa-miR-18b-5p	1.249	0.0001
hsa-miR-4484	-1.5818	0.0001
hsa-miR-371a-5p	-1.8858	0.0001
hsa-miR-454-3p	1.0835	0.0001
hsa-miR-25-3p	0.6245	0.0001
hsa-miR-141-3p	-1.5852	0.0001
hsa-miR-214-3p	-1.3622	0.0001
hsa-miR-221-3p	-0.8206	0.0002
hsa-miR-4419b	-0.5062	0.0002
hsa-miR-199a-5p	-1.6738	0.0002
hsa-miR-181b-5p	2.0125	0.0002
hsa-miR-19b-3p	1.1852	0.0002
hsa-miR-320a	0.7557	0.0003
hsa-miR-4532	-0.8479	0.0004
hsa-miR-222-3p	-0.5013	0.0004
hsa-miR-99b-5p	1.5161	0.0004
hsa-miR-331-3p	1.0019	0.0005
hsa-miR-29b-3p	-2.0941	0.0005
hsa-miR-30e-5p	0.8051	0.0006
hsa-miR-877-5p	0.4864	0.0006
hsa-miR-27b-3p	**0.7742**	**0.0007**
hsa-miR-320b	0.8703	0.0007
hsa-let-7b-5p	-1.1435	0.0007
hsa-miR-4299	-0.6563	0.0007
hsa-miR-4284	0.3662	0.0007
hsa-miR-151a-5p	**1.3574**	**0.0008**
hsa-miR-371a-3p	-1.8188	0.0008
hsa-miR-335-5p	1.0916	0.0008
hsa-miR-4726-5p	-0.6872	0.0009
hsa-miR-191-5p	**0.7445**	**0.001**
hsa-miR-340-5p	0.9998	0.0011
hsa-miR-18a-3p	0.6234	0.0017
hsa-miR-15b-5p	0.5813	0.0017
hsa-miR-2355-3p	-0.6624	0.0018
hsa-let-7a-5p	-0.778	0.002
hsa-miR-3653	0.4609	0.002
hsa-miR-302a-3p	0.4046	0.0022
hsa-miR-4521	0.8567	0.0022
hsa-miR-15a-5p	0.9395	0.0023
hsa-miR-197-3p	0.353	0.0023
hsa-miR-19a-3p	0.7923	0.0024

hsa-miR-125a-5p	0.7283	0.0026
hsa-miR-16-5p	0.7372	0.0026
hsa-miR-302c-3p	0.415	0.0028
hsa-miR-424-5p	-0.6669	0.0038
hsa-miR-4505	-0.8631	0.0039
hsa-miR-151a-3p	**0.4639**	**0.0041**
hsa-miR-302b-5p	0.4751	0.0042
hsa-miR-4687-3p	0.4767	0.0042
hsa-miR-21-5p	-0.9605	0.0043
hsa-miR-3201	-0.8782	0.0044
hsa-miR-4417	2.0142	0.0046
hsa-miR-711	-1.1984	0.0048
hsa-miR-3685	-0.2898	0.0057
hsa-miR-4800-5p	-0.9898	0.0066
hsa-miR-365a-3p	-0.2632	0.0067
hsa-miR-548an	0.3144	0.008
hsa-miR-205-5p	-0.9881	0.008
hsa-miR-3611	-0.8365	0.0083
hsa-miR-106b-3p	0.5294	0.0085
hsa-miR-30d-5p	0.5454	0.009
hsa-miR-4475	-0.7388	0.0091
hsa-miR-34a-5p	0.4117	0.01
hsa-miR-204-3p	-0.6755	0.01
hsa-miR-4780	0.3391	0.0102
hsa-miR-26b-5p	**0.4187**	**0.0102**
hsa-miR-659-3p	-0.3752	0.0104
hsa-miR-302d-5p	0.2053	0.0108
hsa-miR-302f	0.2863	0.0109
hsa-miR-4508	-1.1427	0.0112
hsa-miR-585	0.5165	0.0126
hsa-miR-302a-5p	0.3851	0.0131
hsa-miR-378a-3p	**-0.2707**	**0.0142**
hsa-miR-1587	-0.8565	0.0154
hsa-miR-148b-3p	0.5833	0.0154
hsa-miR-374b-5p	0.7289	0.0167
hsa-miR-520c-3p	0.4135	0.0167
hsa-miR-3646	-0.3297	0.0193
hsa-miR-155-5p	-0.5876	0.0194
hsa-miR-548ap-5p	0.3937	0.02
hsa-miR-3924	-0.5755	0.0203
hsa-miR-5002-5p	-0.3993	0.0203
hsa-miR-299-5p	0.5036	0.0206
hsa-miR-1275	-0.4743	0.0235
hsa-miR-302e	0.3684	0.0246
hsa-miR-1285-3p	0.3768	0.0262
hsa-miR-4657	0.6929	0.0263
hsa-miR-4478	0.3008	0.0273
hsa-miR-1915-3p	0.4853	0.0284
hsa-miR-143-3p	-0.7689	0.0297
hsa-miR-3183	0.3325	0.0327
hsa-miR-302b-3p	0.3255	0.0333
hsa-miR-4530	**-0.795**	**0.0348**
hsa-miR-182-5p	-0.3582	0.0349
hsa-miR-4792	-0.8801	0.0358
hsa-miR-3656	0.282	0.0359
hsa-miR-634	0.1949	0.0359
hsa-miR-665	0.4129	0.0396
hsa-miR-335-3p	0.2642	0.0421
hsa-miR-4286	0.7085	0.0425
hsa-miR-4714-5p	0.3298	0.0473
hsa-miR-3182	0.7784	0.0486
hsa-miR-3156-3p	0.2378	0.0486
hsa-miR-4301	0.4719	0.0488
hsa-miR-140-3p	0.5838	0.0488
hsa-miR-23a-3p	-0.2527	0.0488

Table 5.11. miRNAs significantly ($p < 0.05$) differentially expressed between differentiated iPSCs and differentiated H9. *miRNAs that are upregulated in iPSCs vs. H9 are highlighted in green, while miRNAs that are downregulated in iPSCs vs. H9 are highlighted in red. The miRNAs that are also included on the lists of miRNAs most highly up- or down-regulated during endoderm formation are in bold.*

Comparison of miRNA Expression in Differentiated H1 vs. Differentiated H9		
miRNA	**Log FC**	**p-value**
hsa-miR-340-3p	-0.4991	0.0017
hsa-miR-31-5p	1.9208	0.0053
hsa-miR-363-3p	-2.4497	0.0053
hsa-miR-371a-3p	3.4766	0.0053
hsa-miR-373-3p	4.4272	0.0053
hsa-miR-584-5p	-1.0799	0.0053
hsa-miR-371a-5p	3.752	0.0053
hsa-miR-4288	1.0774	0.0053
hsa-miR-3941	**0.2222**	**0.0053**
hsa-miR-642b-3p	-1.235	0.0053
hsa-miR-29a-3p	3.4275	0.0053
hsa-miR-4454	0.7002	0.0053
hsa-miR-375	**-0.4237**	**0.0053**
hsa-miR-135a-5p	-2.5126	0.0062
hsa-miR-22-3p	2.3639	0.0063
hsa-miR-221-3p	1.034	0.0063
hsa-miR-3116	-0.5492	0.0063
hsa-miR-664-3p	-0.3895	0.0067
hsa-miR-34c-5p	1.3234	0.0071
hsa-miR-29b-3p	2.0456	0.0075
hsa-miR-361-5p	-0.966	0.0078
hsa-miR-200c-3p	3.0395	0.0091
hsa-miR-1908	-0.4547	0.0098
hsa-miR-21-5p	2.0775	0.0104
hsa-miR-665	-0.6988	0.0105
hsa-miR-518b	-0.2902	0.0106
hsa-miR-1915-3p	-1.1319	0.0118
hsa-let-7g-5p	1.7895	0.0118
hsa-miR-1827	1.9613	0.0118
hsa-miR-182-5p	0.8887	0.0121
hsa-miR-4687-3p	-0.723	0.0121
hsa-miR-3158-5p	-0.6867	0.0123
hsa-miR-574-3p	0.4143	0.0123
hsa-miR-302c-3p	-0.7537	0.0127
hsa-miR-302b-5p	-0.555	0.0127
hsa-miR-4788	-0.931	0.0134
hsa-miR-423-5p	-0.1657	0.0134
hsa-let-7i-5p	1.5926	0.0134
hsa-miR-141-3p	2.0581	0.014
hsa-miR-3195	-1.0326	0.014
hsa-miR-23a-3p	1.6807	0.014
hsa-miR-424-5p	0.935	0.014
hsa-miR-205-5p	2.0712	0.014
hsa-miR-199a-3p	1.9192	0.0151
hsa-miR-222-3p	0.9806	0.0151
hsa-miR-92b-3p	-1.2144	0.0151
hsa-miR-4289	-1.2209	0.0151
hsa-miR-2355-3p	1.4183	0.0151
hsa-miR-4750	-0.6062	0.0151
hsa-miR-4732-5p	-1.3053	0.0154
hsa-miR-214-3p	1.6785	0.0158
hsa-miR-124-3p	-1.8288	0.0158
hsa-miR-3924	1.1683	0.0158
hsa-miR-3656	-0.6829	0.0158
hsa-miR-4674	-0.9545	0.0158
hsa-miR-4501	3.4363	0.0158
hsa-miR-99b-3p	-0.3322	0.0158
hsa-miR-4725-3p	-0.6959	0.0166
hsa-miR-130b-3p	-1.2881	0.0174
hsa-miR-4429	0.51	0.0174
hsa-miR-4707-5p	-0.6952	0.0174
hsa-let-7b-5p	1.0822	0.0179
hsa-miR-181b-5p	-1.4162	0.0183
hsa-miR-3940-5p	-0.7796	0.0183
hsa-miR-199a-5p	1.9867	0.0183
hsa-miR-492	-0.5573	0.0183
hsa-miR-4436b-5p	**0.2129**	**0.0189**
hsa-miR-107	-0.893	0.019
hsa-miR-155-5p	1.145	0.019
hsa-miR-18b-5p	-1.0206	0.019
hsa-miR-3687	-1.1135	0.019
hsa-miR-3621	-0.8171	0.019
hsa-miR-454-3p	-0.7424	0.019
hsa-miR-4475	0.6704	0.019
hsa-miR-93-5p	-0.6378	0.019
hsa-miR-9-5p	-3.0924	0.019
hsa-miR-25-3p	-0.5475	0.019
hsa-miR-181a-5p	-1.5472	0.019
hsa-miR-302f	-0.4464	0.019
hsa-miR-135b-5p	-1.8001	0.0198
hsa-miR-4500	0.5582	0.0198
hsa-miR-663a	-1.2821	0.0198
hsa-let-7d-5p	1.4661	0.0201
hsa-miR-585	-0.7841	0.0203
hsa-miR-302b-3p	-0.4837	0.0206
hsa-miR-4747-5p	1.1736	0.0206
hsa-miR-15a-5p	-0.7632	0.0207
hsa-miR-24-3p	1.1876	0.0211
hsa-miR-943	-0.6964	0.0217

hsa-miR-302a-5p	-0.542	0.0221
hsa-miR-299-5p	-0.5708	0.0225
hsa-miR-181d	-1.5727	0.023
hsa-miR-3162-3p	-0.7361	0.0233
hsa-miR-23b-3p	1.2015	0.0236
hsa-miR-516b-5p	**-0.1073**	**0.0238**
hsa-miR-20b-5p	-1.1864	0.0238
hsa-miR-302e	-0.5172	0.0238
hsa-miR-302d-5p	-0.5025	0.0238
hsa-miR-96-5p	0.4278	0.0244
hsa-miR-3591-5p	0.4059	0.0244
hsa-miR-520c-3p	-0.5002	0.0247
hsa-miR-3183	-0.6093	0.0247
hsa-miR-9-3p	-2.7459	0.0254
hsa-miR-193b-3p	0.6602	0.0265
hsa-miR-4299	0.427	0.0272
hsa-miR-936	1.3686	0.0272
hsa-miR-660-3p	-0.287	0.0273
hsa-miR-4768-5p	-0.7331	0.0273
hsa-miR-5100	0.7137	0.0273
hsa-miR-1323	-0.2217	0.0276
hsa-let-7c	1.078	0.028
hsa-miR-331-3p	-0.7551	0.0289
hsa-miR-877-5p	-0.4579	0.0292
hsa-miR-4451	**-0.1507**	**0.0303**
hsa-miR-25-5p	-0.8015	0.0303
hsa-miR-140-3p	-0.2999	0.0305
hsa-miR-27a-3p	1.3614	0.0305
hsa-miR-342-3p	-1.0695	0.0307
hsa-miR-1299	0.1187	0.0312
hsa-miR-106b-3p	-0.4973	0.0313
hsa-miR-574-5p	0.2894	0.0314
hsa-miR-340-5p	-0.9276	0.0314
hsa-miR-4258	-0.4585	0.0322
hsa-miR-874	-1.0039	0.0322
hsa-miR-302a-3p	-0.6523	0.0323
hsa-miR-1285-3p	-0.3082	0.033
hsa-miR-106b-5p	-0.8736	0.0333
hsa-miR-638	-0.6155	0.0333
hsa-miR-5572	-0.3801	0.0339
hsa-miR-552	-0.5116	0.0345
hsa-miR-17-3p	-0.777	0.0357
hsa-miR-365a-3p	0.6473	0.037
hsa-miR-1290	-0.6166	0.037
hsa-miR-4516	-0.6803	0.037
hsa-miR-483-5p	0.3528	0.0373
hsa-miR-4417	-1.3181	0.0388
hsa-miR-373-5p	0.7357	0.039
hsa-miR-148b-3p	-0.4799	0.0405
hsa-miR-652-5p	-0.5021	0.0408
hsa-miR-548ap-5p	-0.6468	0.0422
hsa-miR-4795-3p	-0.2856	0.0429
hsa-miR-3679-3p	0.328	0.0433
hsa-miR-1184	-0.7149	0.0434
hsa-miR-106a-5p	-0.8702	0.0438
hsa-miR-5187-3p	-0.4782	0.0439
hsa-miR-4508	0.6858	0.0448
hsa-miR-634	-0.3676	0.0448
hsa-miR-4632	-0.5121	0.045
hsa-miR-675-5p	0.2231	0.045
hsa-miR-19b-3p	-0.8562	0.0456
hsa-miR-4653-3p	-0.4738	0.0456
hsa-miR-17-5p	-0.8343	0.0456
hsa-miR-4800-5p	0.8808	0.0456
hsa-miR-4484	0.9759	0.0456
hsa-miR-708-5p	**-0.8074**	**0.0456**
hsa-miR-18a-3p	-0.5474	0.0486

Table 5.12 miRNAs significantly ($p < 0.05$) differentially expressed between differentiated H1 and differentiated H9. *miRNAs that are upregulated in H1 vs. H9 are highlighted in green, while miRNAs that are downregulated in H1 vs. H9 are highlighted in red. The miRNAs that are also included on the lists of miRNAs most highly up- or down-regulated during endoderm formation are in bold.*

Table 5.10 shows the comparison of differentiated iPSCs with differentiated H1 (excluding H9). 28 miRNAs are significantly differentially expressed between iPSCs and H1: 23 are expressed at a lower level in iPSCs than in H1 (red), and 5 are expressed at a higher level (green). Interestingly, miR-708-5p, which was identified as one of the miRNAs strongly upregulated in DE formation, is expressed at higher levels in iPSCs than in H1. This may partially explain why the H1 cell line has a lower propensity to form DE. However, miR-27b, also identified as potentially being involved in DE formation, is expressed at lower levels in iPSCs than in H1.

Table 5.11 shows the comparison of differentiated iPSCs with differentiated H9 (excluding H1). 154 miRNAs were significantly differentially expressed between iPSCs and H9, with 60 expressed at a higher level in iPSCs compared to H9 (green) and 94 expressed at a lower level (red). The fact that a much greater number of miRNAs are differentially expressed between iPSCs vs. H9 compared to iPSCs vs. H1 again suggests that culture conditions have having an effect on miRNA expression.

Several miRNAs, including miR-151a-3p, miR-151a-5p, and miR-27b were identified in both the overall comparison of iPSC vs. ESC, as well as the individual comparison of iPSC vs. H1 and iPSC vs. H9. These miRNAs are more likely to represent conserved differences between iPSCs and ESCs which are not cell line-specific, although this is difficult to determine in an analysis of only four cell lines.

Table 5.12 shows the comparison of differentiated H1 with differentiated H9.156 miRNAs were significantly differentially expressed between the two cell lines, with 98 miRNAs downregulated in H1 vs. H9 (red), and 58 miRNAs upregulated in H1 vs. H9 (green). Two miRNAs identified as being strongly upregulated in DE formation, miR-375 and miR-708-5p, are expressed at lower levels in H1 than in H9, which again may partly explain the reduced propensity for differentiation into DE observed in the H1 cell line. The miRNAs identified as being differentially expressed between iPSCs and ESCs were not differentially expressed between H1 and H9.

5.3.2 Validation of Microarray Results by qRT-PCR

As microarray analysis is prone to give false positive results due to the high numbers of comparisons being performed in parallel, even when a statistical correction is used, it is important to validate the results using qRT-PCR, as this is a more reliable method of analysing miRNA expression levels. The expression levels of miRNAs were analysed in undifferentiated samples, samples directed to differentiate into DE, and samples that were allowed to spontaneously differentiate. Spontaneously differentiated cells were not included in the microarray analysis, as it was thought that they would result in too much background noise and make data analysis too complicated, but were included in qRT-PCR analysis, allowing investigation of whether changes in miRNA expression are specific to DE formation, or are also involved in more general differentiation. Samples were taken in triplicate and at least two separate differentiation experiments were analysed for expression of each miRNA. In addition to the four cell lines investigated using microarray (MRC5I, MRC9G, H1 and H9), an additional ESC line (H7) was included in qRT-PCR analysis. This is another ESC line grown on Matrigel™, and was included to provide additional information as to whether the differences noticed in miRNA expression between cell lines grown on feeders and the H9 cell line grown on Matrigel™ were due to the culture methods. Including another cell line also indicates whether the array results are applicable to other cell lines than those analysed in the array. miRNA expression was analysed using Students' T-test to determine whether the fold change in miRNA expression in differentiated samples from undifferentiated samples was significant.

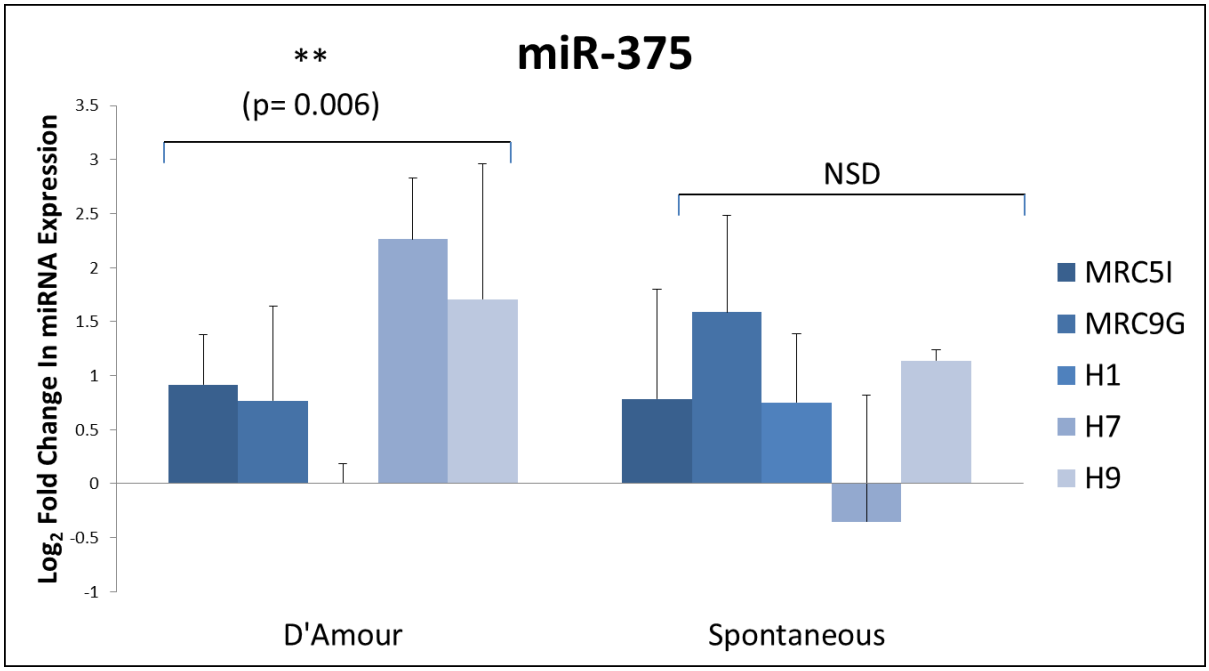

Figure 5.4. qRT-PCR results showing log_2 fold change in expression levels of miR-375 in differentiated samples compared to undifferentiated samples. *miR-375 is significantly (p = 0.006, n = 3) upregulated in cells directed to differentiate using the D'Amour protocol, but not in spontaneously differentiated cells.*

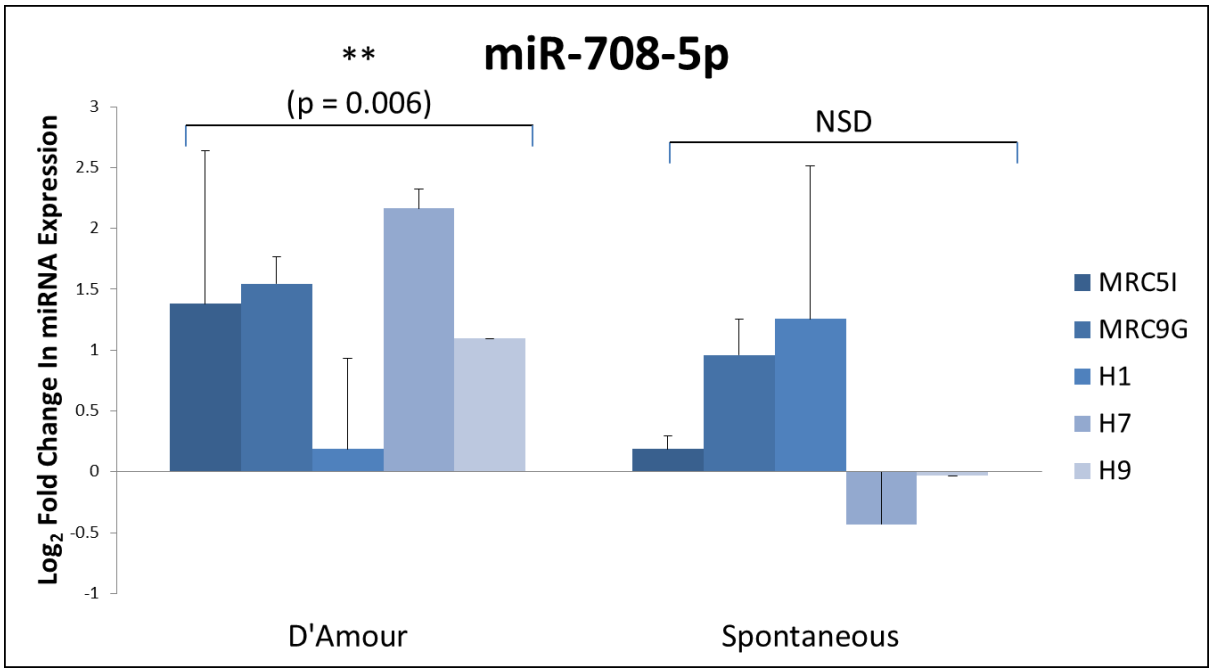

Figure 5.5. qRT-PCR results showing log_2 fold change in expression levels of miR-708-5p in differentiated samples compared to undifferentiated samples. *miR-708-5p is significantly (p = 0.006, n = 2) upregulated in cells directed to differentiate using the D'Amour protocol, but not in spontaneously differentiated cells.*

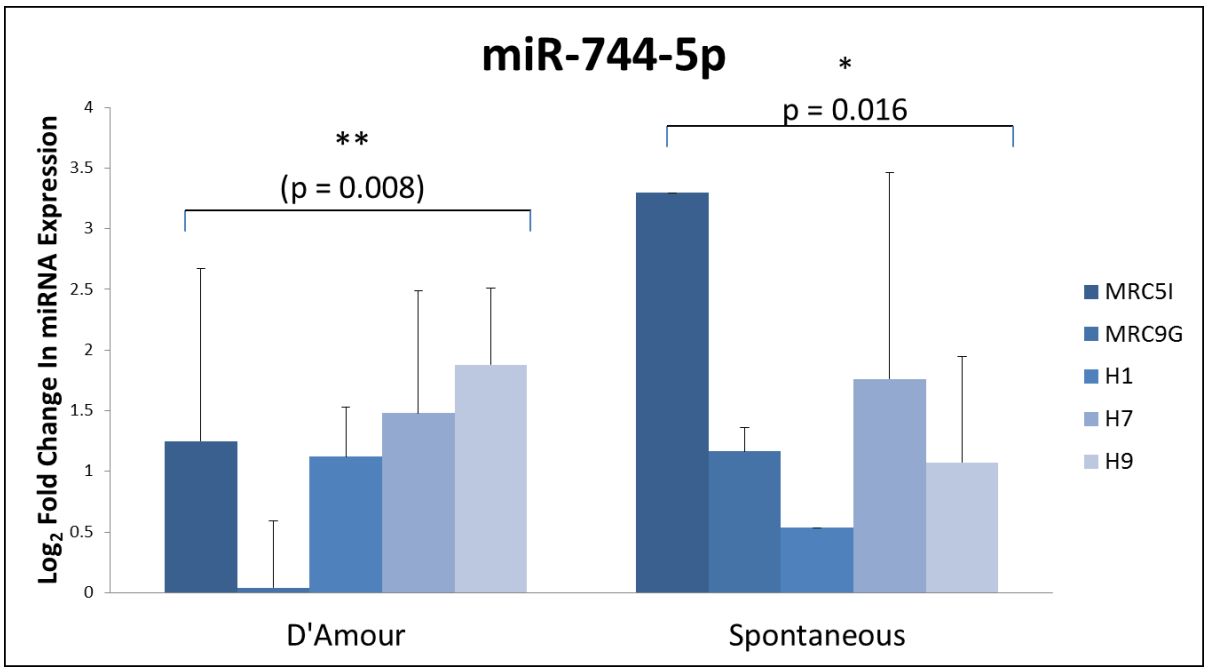

Figure 5.6 qRT-PCR results showing log_2 fold change in expression levels of miR-744-5p in differentiated samples compared to undifferentiated samples. *miR-744-5p is significantly upregulated in both cells directed to differentiate using the D'Amour protocol (p = 0.008, n = 3) and spontaneously differentiated cells (p = 0.016, n = 3).*

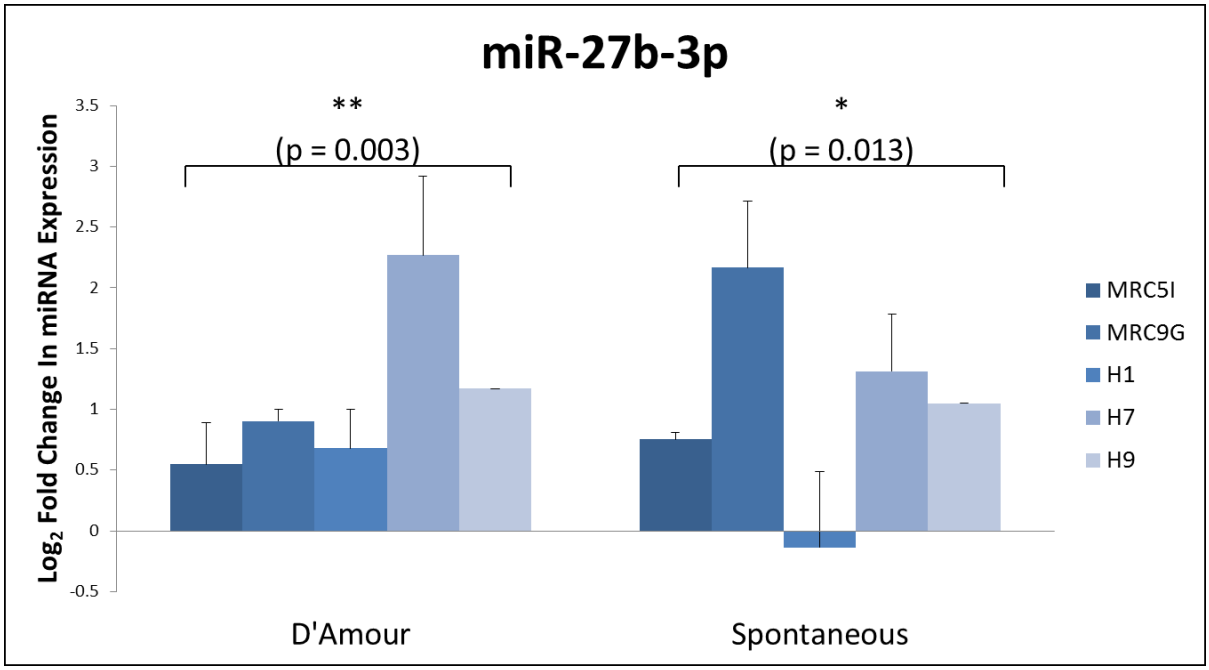

Figure 5.7. qRT-PCR results showing log_2 fold change in expression levels of miR-27b-3p in differentiated samples compared to undifferentiated samples. *miR-27b-3p is significantly upregulated in both samples differentiated using the D'Amour protocol (p = 0.003, n = 2) and spontaneously differentiated cells (p = 0.013, n = 2).*

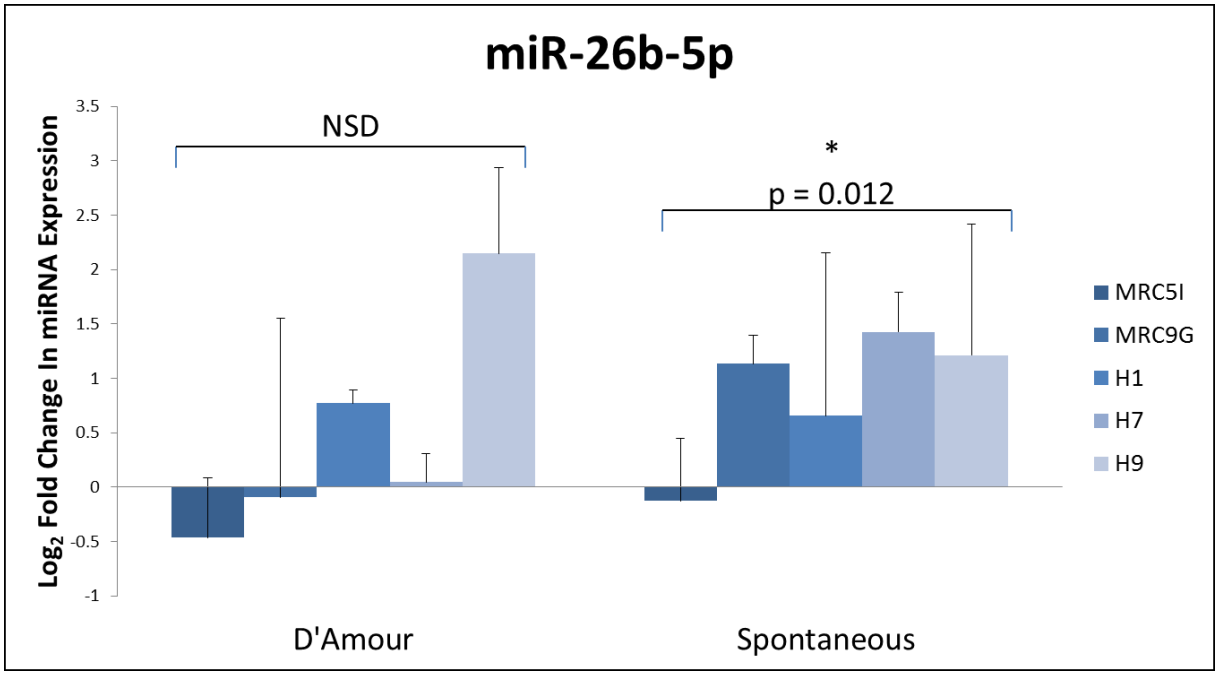

Figure 5.8. qRT-PCR results showing log_2 fold change in expression levels of miR-26b-5p in differentiated samples compared to undifferentiated samples. *miR-26b-5p is significantly upregulated in spontaneously differentiated samples (p = 0.012, n = 3) but not in cells directed to differentiate using the D'Amour protocol.*

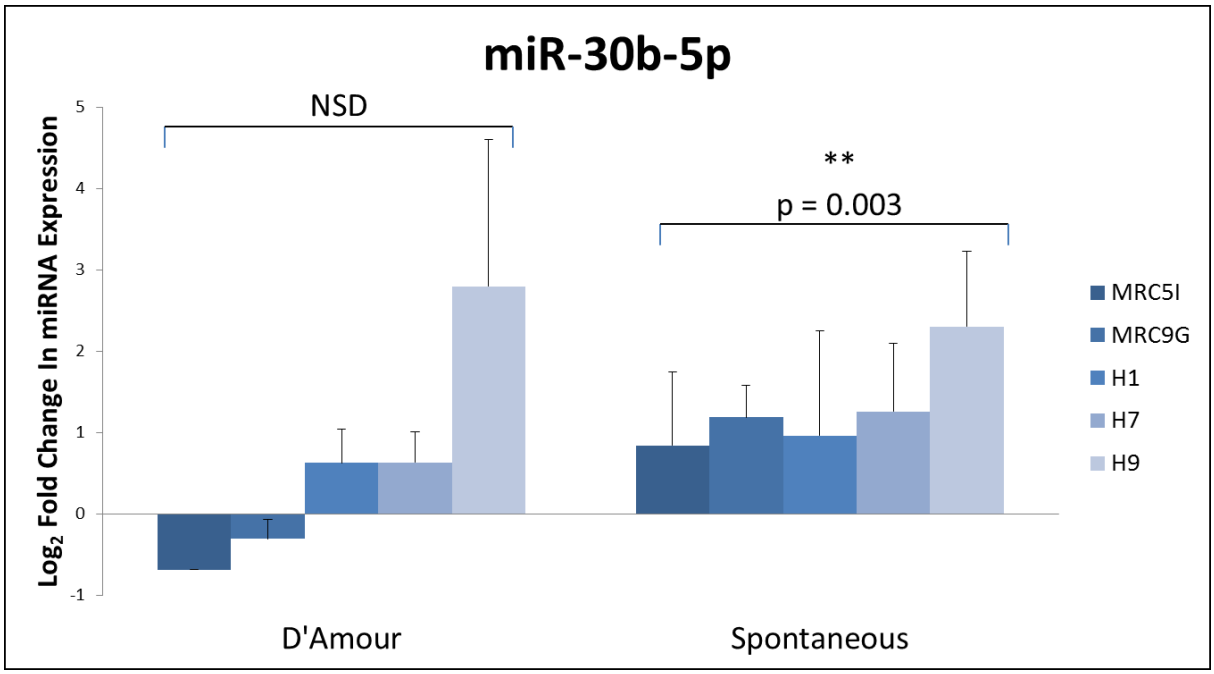

Figure 5.9. qRT-PCR results showing log_2 fold change in expression levels of miR-30b-5p in differentiated samples compared to undifferentiated samples. *miR-30b-5p is significantly upregulated in spontaneously differentiated samples (p = 0.012, n = 4) but not in cells directed to differentiate using the D'Amour protocol.*

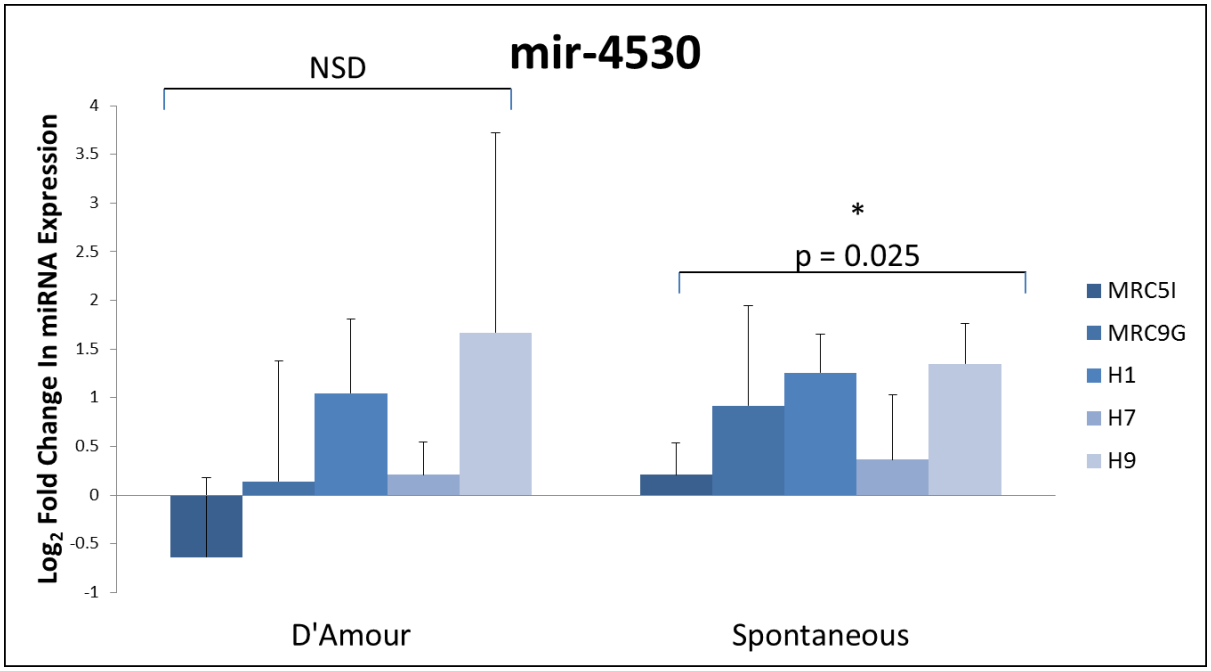

Figure 5.10. qRT-PCR results showing log_2 fold change in expression levels of miR-4530 in differentiated samples compared to undifferentiated samples. *miR-4530 is significantly upregulated in spontaneously differentiated samples (p = 0.025, n = 3) but not in cells directed to differentiate using the D'Amour protocol.*

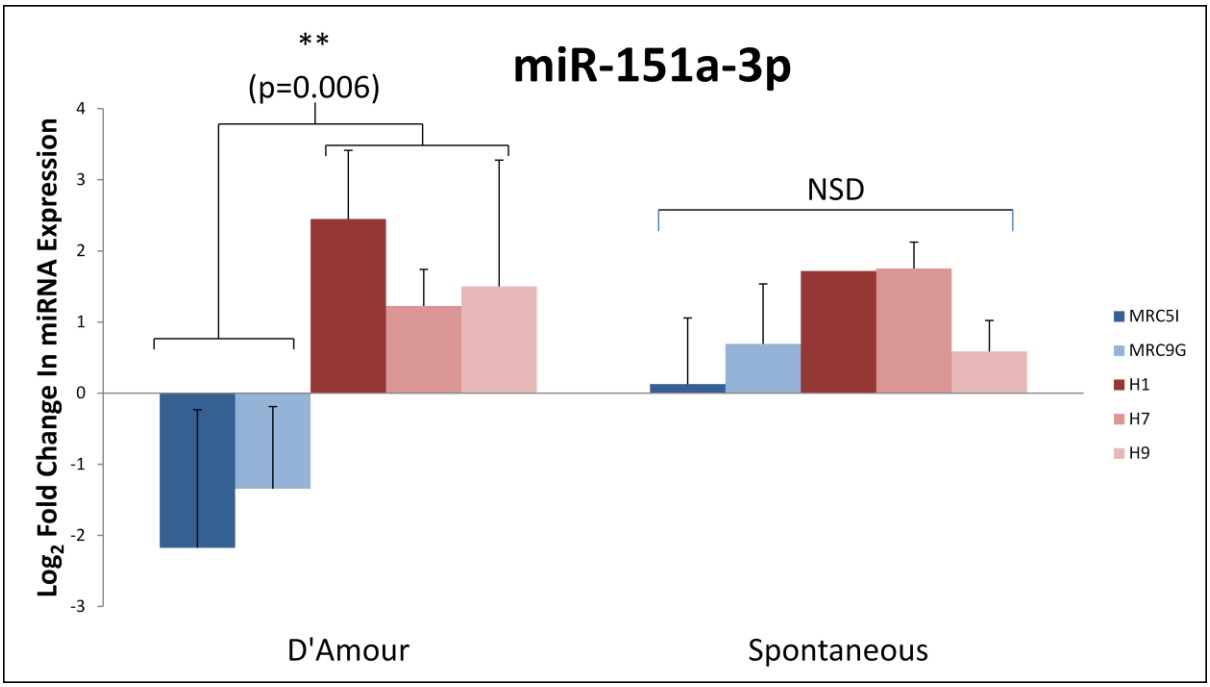

Figure 5.11. qRT-PCR results showing log_2 fold change in expression levels of miR-151a-3p in differentiated samples compared to undifferentiated samples. *miR-151a-3p is significantly (p = 0.006, n = 3) upregulated in three ESC lines (H1, H7 and H9, shown in red) but downregulated in two iPSC lines (MRC5I and MRC9G, shown in blue).*

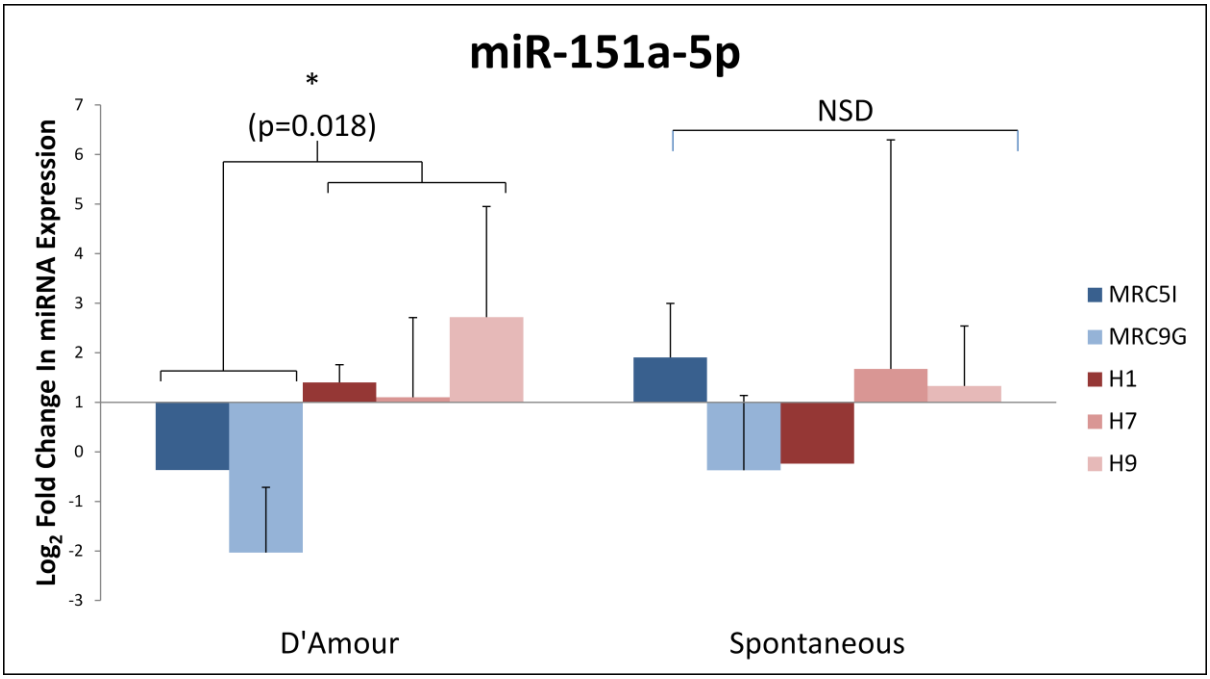

Figure 5.12. qRT-PCR results showing log_2 fold change in expression levels of miR-151a-5p in differentiated samples compared to undifferentiated samples. *miR-151a-5p is significantly (p = 0.018, n = 3) upregulated in three ESC lines (H1, H7 and H9, shown in red) but downregulated in two iPSC lines (MRC5I and MRC9G, shown in blue).*

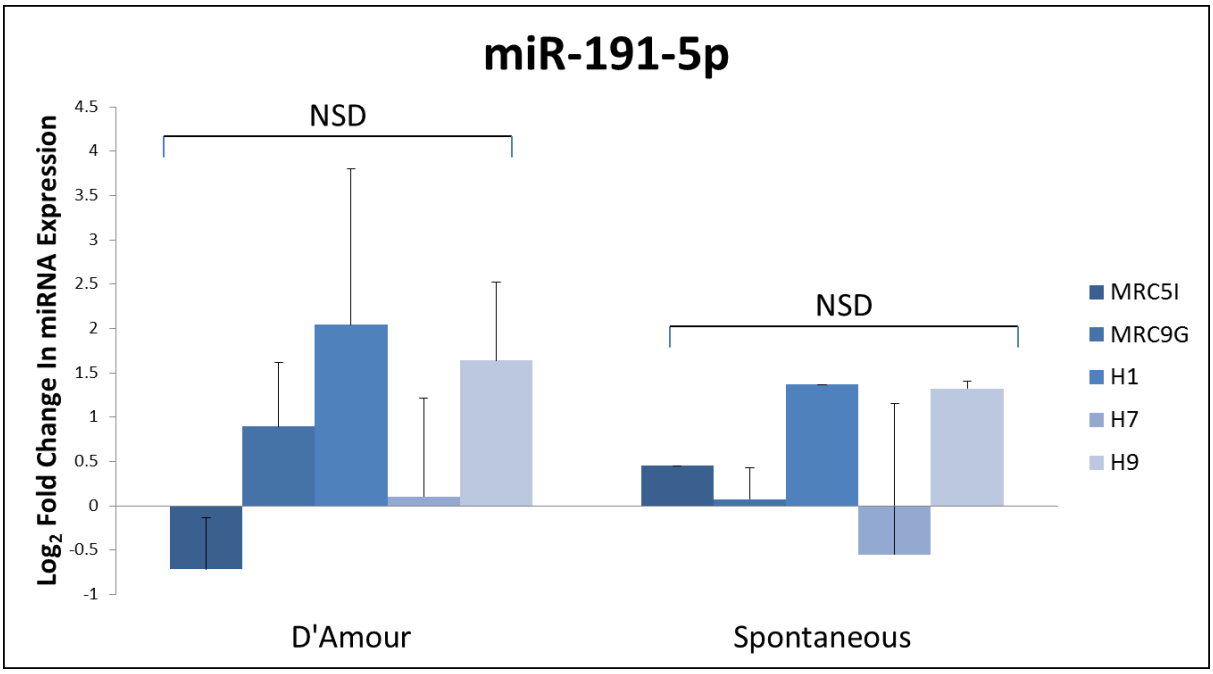

Figure 5.13. qRT-PCR results showing log_2 fold change in expression levels of miR-191-5p in differentiated samples compared to undifferentiated samples. *miR-191-5p is not significantly upregulated in either cells directed to differentiate using the D'Amour protocol or cells allowed to spontaneously differentiate.*

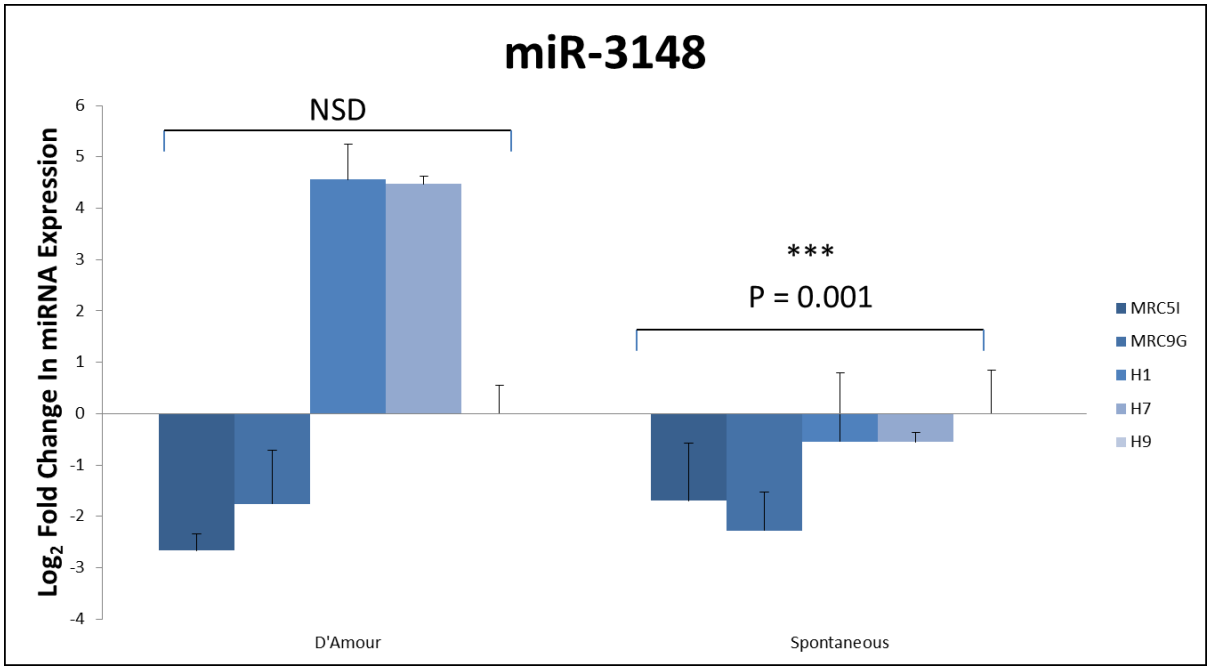

Figure 5.14. qRT-PCR results showing log_2 fold change in expression levels of miR-3148 in differentiated samples compared to undifferentiated samples. *miR-3148 is significantly (p = 0.001, n = 3) upregulated in spontaneously differentiated cells but not in cells directed to differentiate using the D'Amour protocol.*

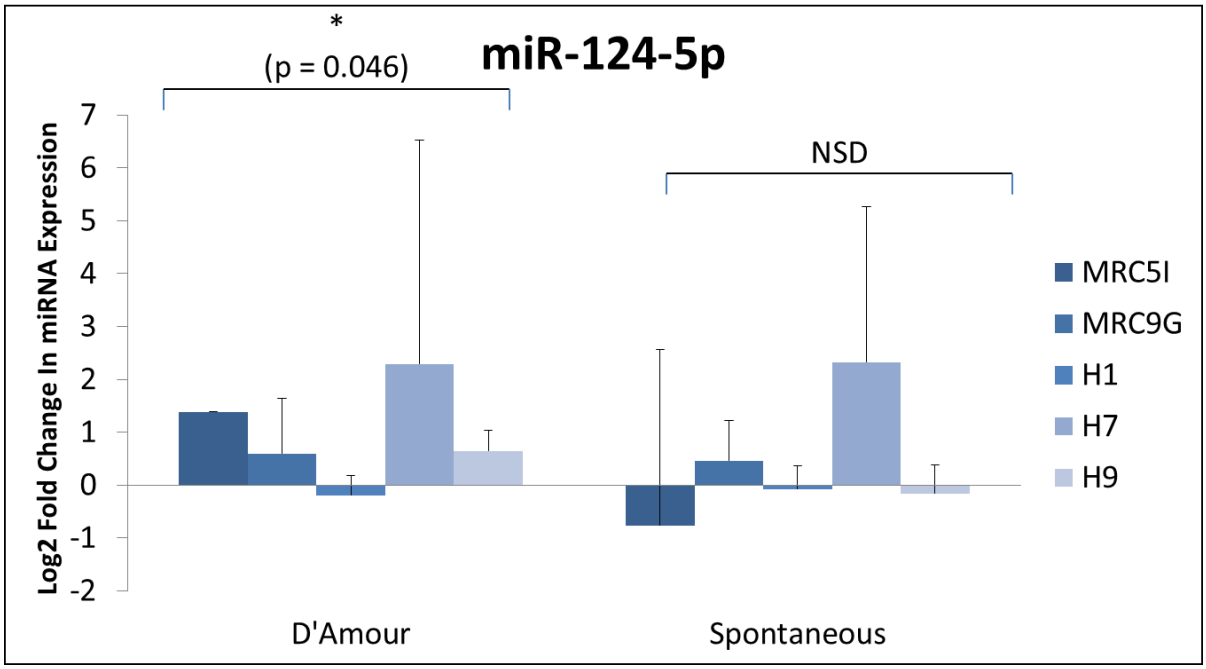

Figure 5.15. qRT-PCR results showing log_2 fold change in expression levels of miR-124-5p in differentiated samples compared to undifferentiated samples. *miR-124-5p is significantly (p = 0.046, n = 3) upregulated in cells directed to differentiate using the D'Amour protocol, but not in spontaneously differentiated cells.*

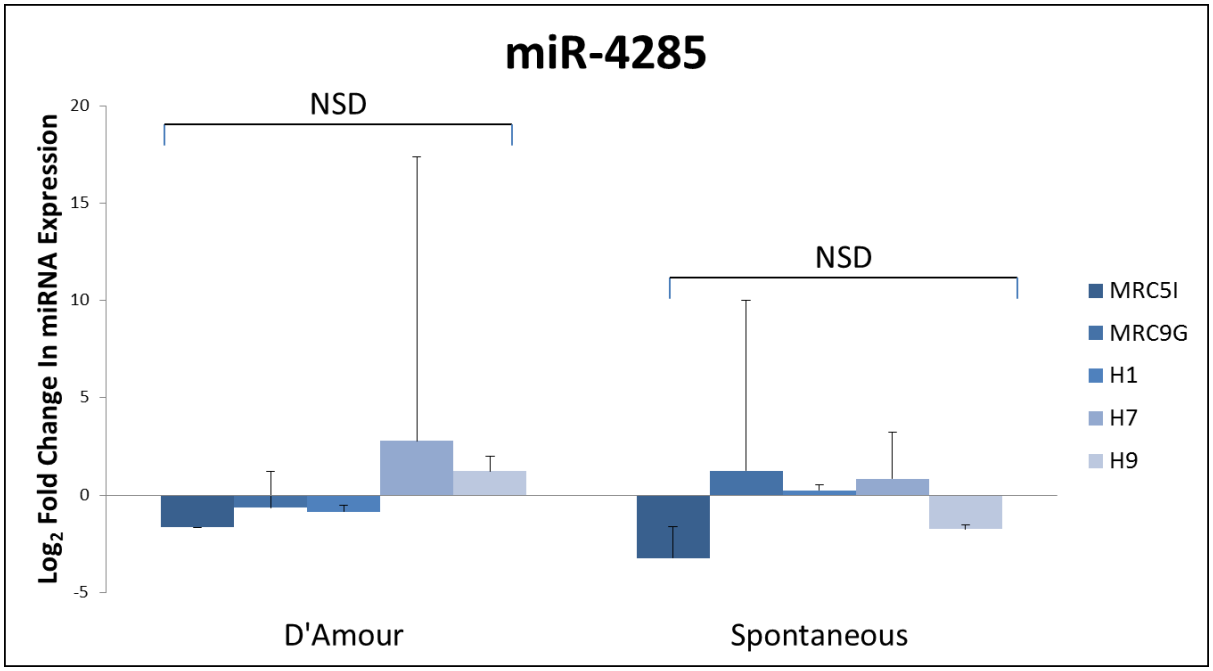

Figure 5.16. qRT-PCR results showing log_2 fold change in expression levels of miR-4285 in differentiated samples compared to undifferentiated samples. *miR-4285 is not significantly upregulated in either cells directed to differentiate using the D'Amour protocol or cells allowed to spontaneously differentiate.*

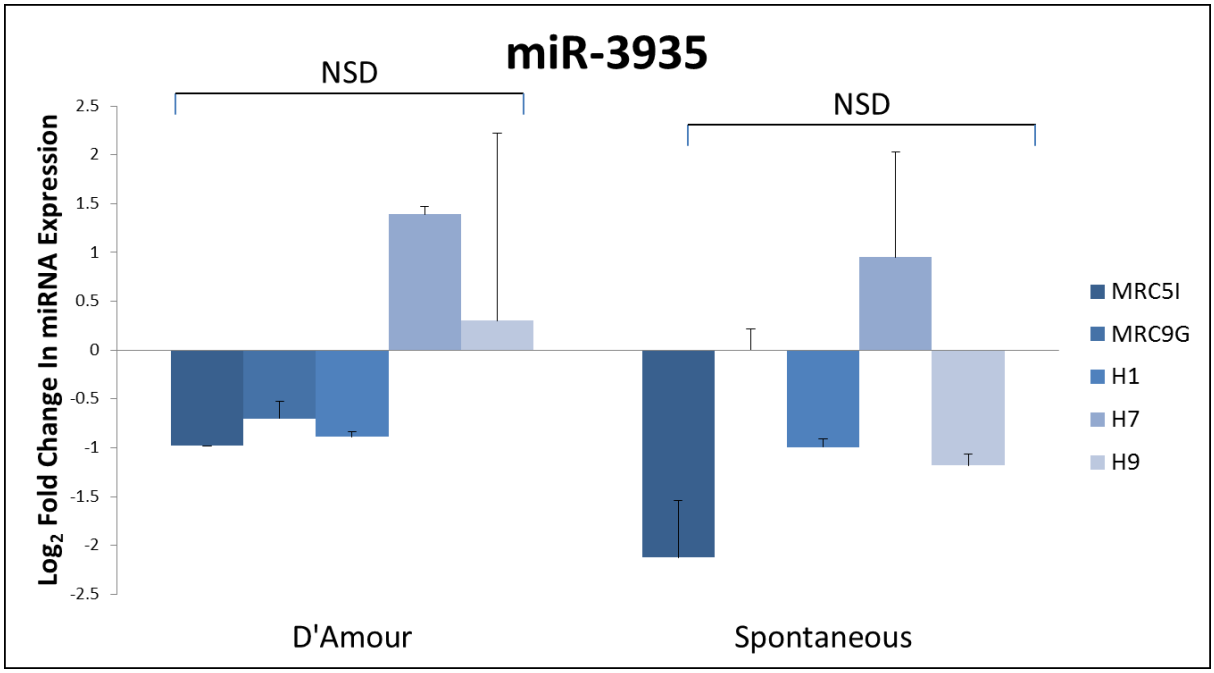

Figure 5.17. qRT-PCR results showing log_2 fold change in expression levels of miR-3935 in differentiated samples compared to undifferentiated samples. *miR-3935 is not significantly upregulated in either cells directed to differentiate using the D'Amour protocol or cells allowed to spontaneously differentiate.*

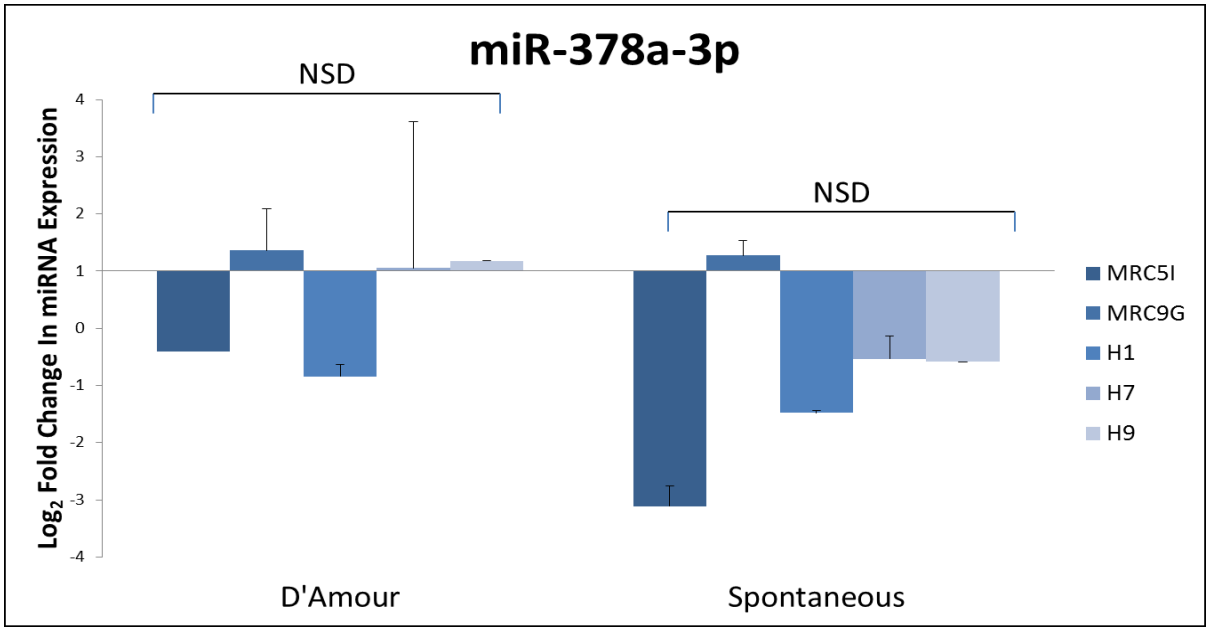

Figure 5.18. qRT-PCR results showing log_2 fold change in expression levels of miR-378a-3p in differentiated samples compared to undifferentiated samples. *miR-378a-3p is not significantly upregulated in either cells directed to differentiate using the D'Amour protocol or cells allowed to spontaneously differentiate.*

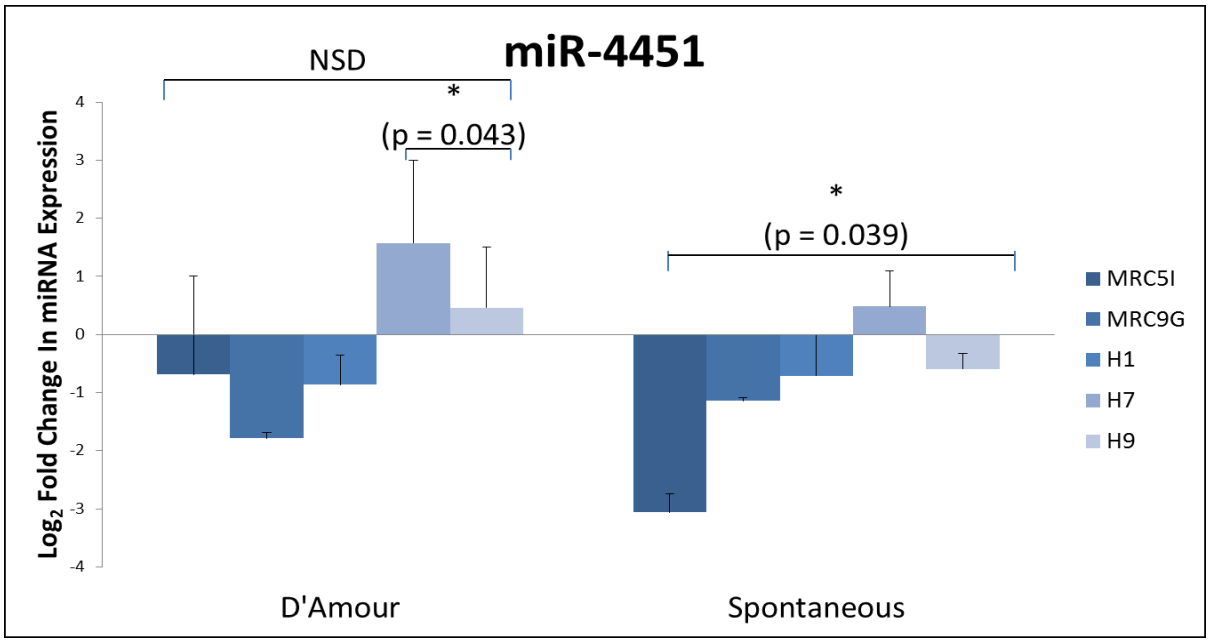

Figure 5.19. qRT-PCR results showing log_2 fold change in expression levels of miR-4451 in differentiated samples compared to undifferentiated samples. *miR-4451 is significantly upregulated only in cells grown on Matrigel™ (H7 and H9), but not those grown on feeders (MRC5I, MRC9G and H1), which are directed to differentiate into definitive endoderm. miR-4451 is also significantly downregulated in spontaneously differentiated samples.*

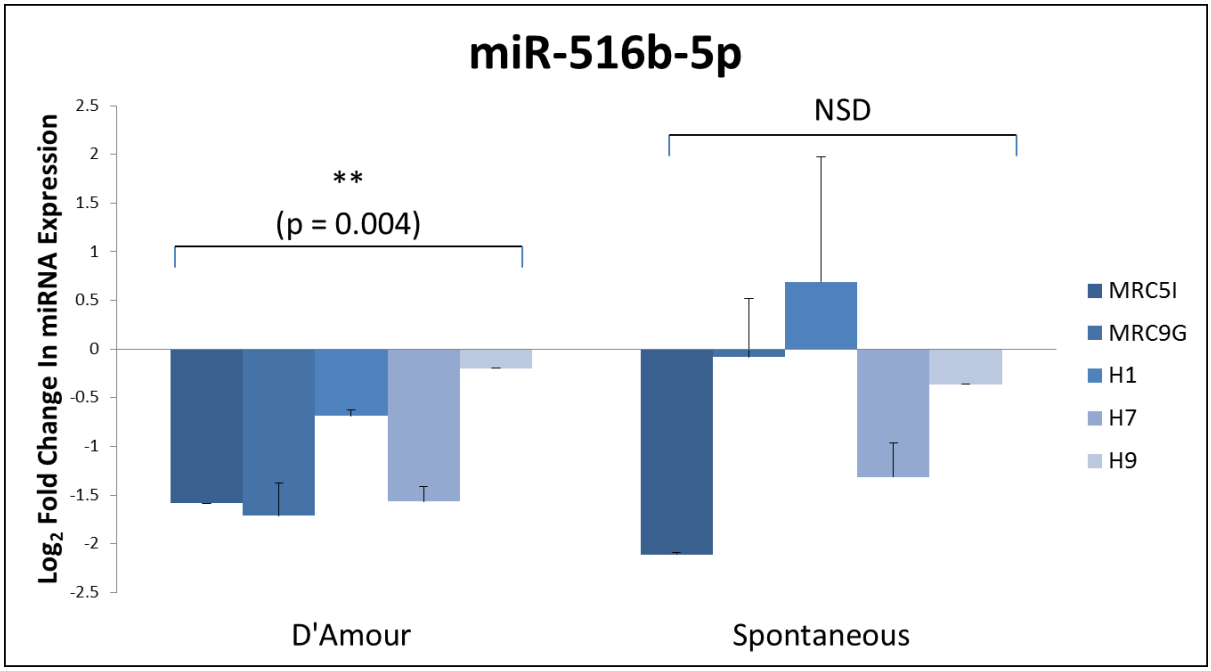

Figure 5.20. qRT-PCR results showing log_2 fold change in expression levels of miR-516b-5p in differentiated samples compared to undifferentiated samples. *miR-516b-5p is significantly ($p = 0.004$, $n = 3$) downregulated in cells directed to differentiate using the D'Amour protocol, but not spontaneously differentiated cells.*

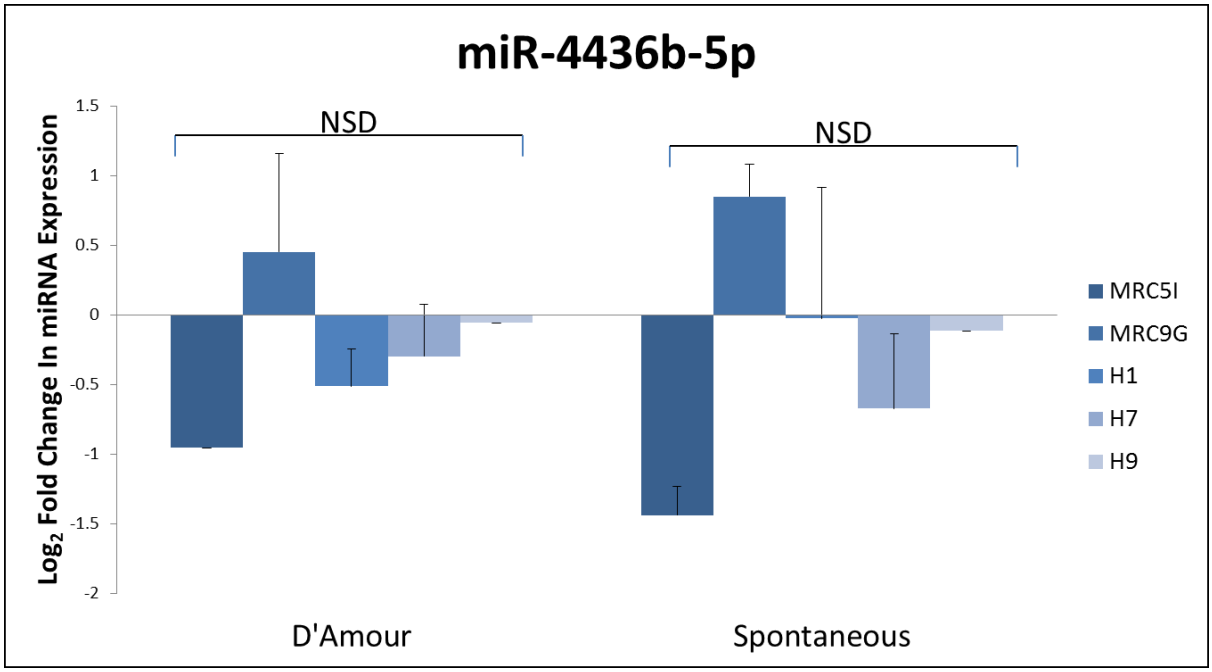

Figure 5.21. qRT-PCR results showing log_2 fold change in expression levels of miR-4436b-5p in differentiated samples compared to undifferentiated samples. *Levels of miR-4436b-5p expression are not significantly changed in either cells directed to differentiate using the D'Amour protocol or cells allowed to spontaneously differentiate.*

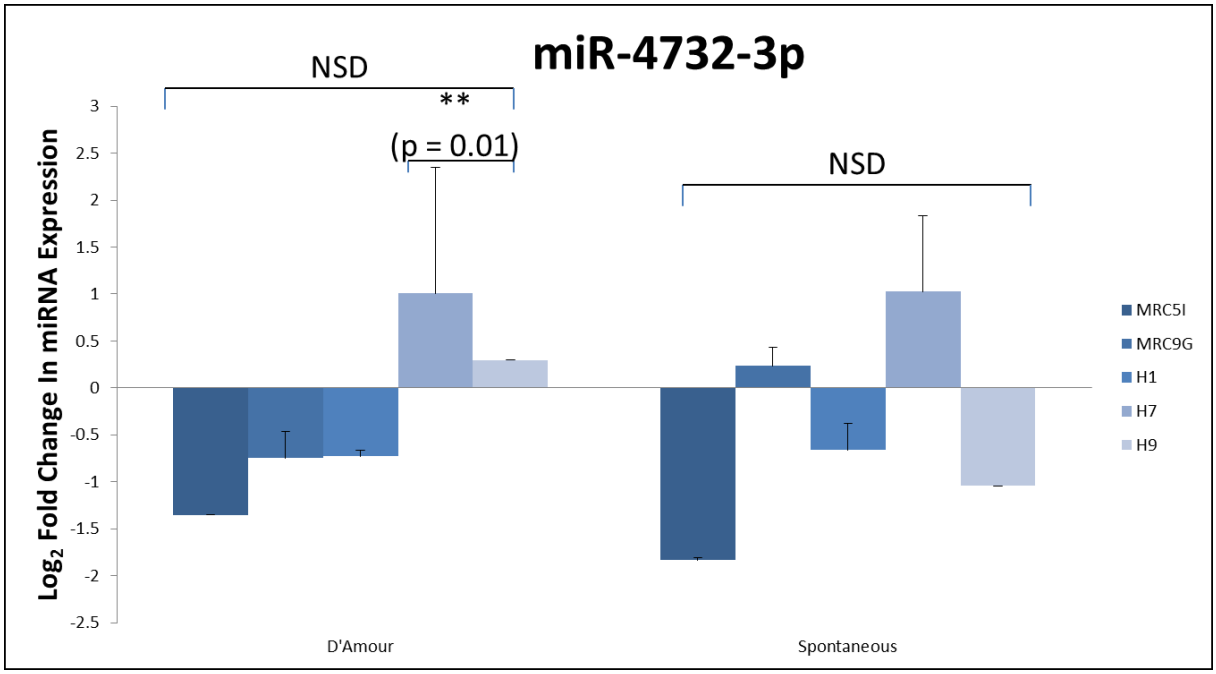

Figure 5.22. qRT-PCR results showing log_2 fold change in expression levels of miR-4732-3p in differentiated samples compared to undifferentiated samples. *miR-4732-3p is significantly upregulated only in cells grown on Matrigel™ (H7 and H9), but not those grown on feeders (MRC5I, MRC9G and H1), which are directed to differentiate into definitive endoderm. There is no significant change in miRNA expression in spontaneously differentiated samples.*

miRNAs Involved in Endoderm Formation

The expression of the top ten miRNAs that were predicted to be the most highly up- or down-regulated during DE formation by microarray analysis was investigated using qRT-PCR. Of the ten miRNAs predicted to be upregulated in DE formation, only two (miR-375 and miR-708-5p) were confirmed to be specifically upregulated in cells undergoing differentiation to DE, and not in spontaneously differentiated cells (figures 5.4 and 5.5, respectively). Although there was some upregulation of these miRNAs in spontaneously differentiated samples, perhaps suggesting a role in more general differentiation, this was not statistically significant. These miRNAs are strongly upregulated in differentiated cells, and their expression mirrors that of DE marker genes such as *Sox17*, with much lower expression in the H1 cell line, which my results have previously shown to have a lower propensity for forming DE (see chapter 4). In addition, both miR-375 (Tzur et al, 2008; Hinton et al, 2010; Kim et al, 2011) and miR-708 (Hinton et al, 2010; Kim et al, 2011) have been previously described to be upregulated in DE formation.

Similarly, of the ten miRNAs predicted to be downregulated in DE formation, only one (miR-516b-5p) was confirmed to be significantly downregulated in cells undergoing differentiation to DE, and not in spontaneously differentiated cells (figure 5.20). This miRNA is strongly downregulated in all five cell lines tested, and its expression inversely mirrors that of DE marker genes such as *Sox17*, with a much reduced downregulation of expression in the H1 cell line which has been previously shown to have a lower propensity for forming DE. miR-3148 was significantly downregulated in spontaneously differentiated cells, but not cells directed to differentiate using the D'Amour protocol (figure 5.14). Expression of miR-3941, predicted to be downregulated in DE formation by the microarray results, could not be detected in any samples analysed by qRT-PCR. miR-124-5p, which was predicted by the array data to be downregulated in DE formation actually appeared to be significantly upregulated when assessed using qRT-PCR (figure 5.15), although the expression levels were highly variable between experiments and the p-value was approaching non-significance ($p = 0.046$).

Two miRNAs, miR-744 and miR-27b, that were predicted by the microarray analysis to be upregulated in DE formation were confirmed by qRT-PCR, but were also found to be significantly upregulated in spontaneously differentiated cells (figure 5.6 and 5.7, respectively). There were no miRNAs that were significantly downregulated in both DE and

spontaneously differentiated cells. Some miRNAs, including miR-26b-5p, miR-30b-5p, and miR-4530, were significantly upregulated only in spontaneously differentiated samples, although some upregulation was seen in samples directed to differentiate into DE (figures 5.8, 5.9 and 5.10, respectively). These results show the importance of including spontaneously differentiated samples in the comparison as a control, as they allow identification of the miRNAs which are specifically upregulated in DE formation. The miRNAs which are also up- or down-regulated in spontaneously differentiated samples may play a more general role in differentiation, rather than in DE formation specifically.

miR-191-5p, which was predicted to be upregulated in DE formation by microarray analysis, showed no significant upregulation in either samples directed to differentiate into DE or spontaneously differentiated samples (figure 5.13). This was also true of some miRNAs predicted to be downregulated in DE formation, including miR-4285, miR-3935, miR-378a-3p, and miR-4436b (figures 5.16, 5.17, 5.18, and 5.21). This shows the importance of validating microarray results with a more robust method such as qRT-PCR, but may also reflect the variability of differentiation efficiency between experiments.

Differential expression of miRNAs between cells grown under different conditions were also noticed during validation of microarray results, with miR-4451and miR-4732-3p upregulated in cells grown on Matrigel (H7 and H9) and downregulated in cells grown on iSNL feeders (MRC5I, MRC9G and H1) (figure 5.19 and 5.22). This again suggests that the culture conditions have some influence on the expression of miRNAs, in support of the microarray results. These differences were not apparent in undifferentiated cells.

miRNAs Differentially Expressed Between ESCs & iPSCs

Differential expression of miRNAs between ESCs and iPSCs was also investigated using qRT-PCR to confirm the microarray results. Of the 91 miRNAs reported to be differentially expressed between DE derived from ESCs and iPSCs, 7 were also on the top ten lists of miRNAs up- or down-regulated during DE formation, including miR-151a-5p, miR-30b-5p, miR-27b-3p, miR-151a-3p, miR-26b-5p, miR-378a-3p, and miR-4530. Of these 7 miRNAs, only two (miR-151a-3p and miR-151a-5p) were found to be significantly differently expressed in either ESCs or iPSCs directed to differentiate into DE (figure 5.11 and 5.12). These miRNAs are upregulated upon differentiation to DE in ESCs and downregulated in iPSCs. There are no differences in the expression of these miRNAs in undifferentiated ESCs and iPSCs.

Because of the apparent differences between the H1 and H9 ESC lines, the expression levels of several other miRNAs which were identified as being significantly differentially expressed in iPSCs when compared to the ESC cell lines individually (i.e. H1 or H9) were also investigated by qRT-PCR. Several miRNAs that are strongly upregulated in iPSCs were predicted to be downregulated in H9 (e.g. miR-302c-3p, miR-4791) and vice versa (e.g. miR-491-3p). However, as no significant differences could be seen in the expression of any of these miRNAs between the different cell lines, these miRNAs were not investigated further.

Something that became quickly apparent from the microarray results was that the culture methods have an impact on miRNA expression pattern, with the principal component analysis (figure 5.2) and hierarchical clustering (figure 5.3) showing that the H9 cell line clusters separately from the other three cell lines included in the analysis. Although the possibility exists that H9 cells genuinely have a significantly different miRNA expression profile to other ESCs, these differences in miRNA expression are more likely to be a result of the different culture conditions. However, no functional effects of these differences were noticed in the analysis of either pluripotency genes (chapter 3) or definitive endoderm genes (chapter 4), indicating that these differences in miRNA expression do not have a significant functional impact on the pluripotency or differentiation of H9 cells. In addition, during PCR validation of microarray results, two miRNAs (miR-4451 and miR-4732-3p; figures 5.19 and 5.22, respectively) were identified as being upregulated in DE formation only in the cell lines cultured on Matrigel (H7 and H9), and not those on iSNL feeders, although these differences were not apparent in undifferentiated cells.

To investigate these culture-specific differences further, H9 cells were grown on both Matrigel™ and iSNL feeders, and subjected to the differentiation protocol described in section 5.2.1. Comparison of these two methods in the same cell line allows us to elucidate differences in miRNA expression that are specific to the culture conditions. The expression of miRNAs identified by qRT-PCR as being a) upregulated in DE formation; b) downregulated in DE formation or c) differentially expressed between cells grown on iSNL feeders and cells grown on Matrigel™ were investigated in H9 cells cultured using both methods. Analysis of variance was used to determine any statistically significant differences between culture methods.

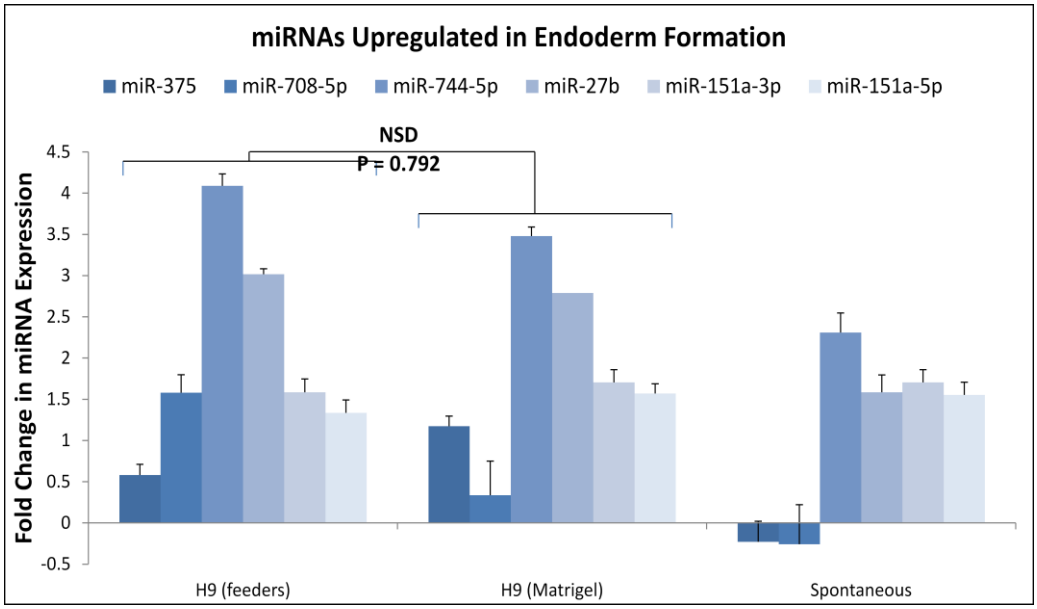

Figure 5.23. Comparison of expression of miRNAs upregulated in DE formation in H9 cells cultured on iSNL feeders or Matrigel™

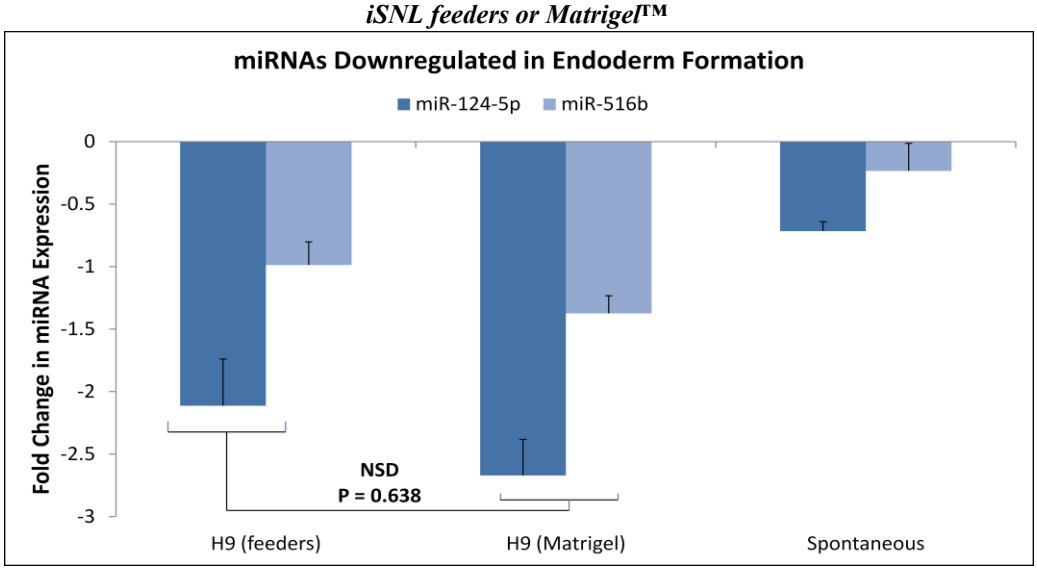

Figure 5.24 Comparison of expression of miRNAs downregulated in DE formation in H9 cells cultured on iSNL feeders or Matrigel™

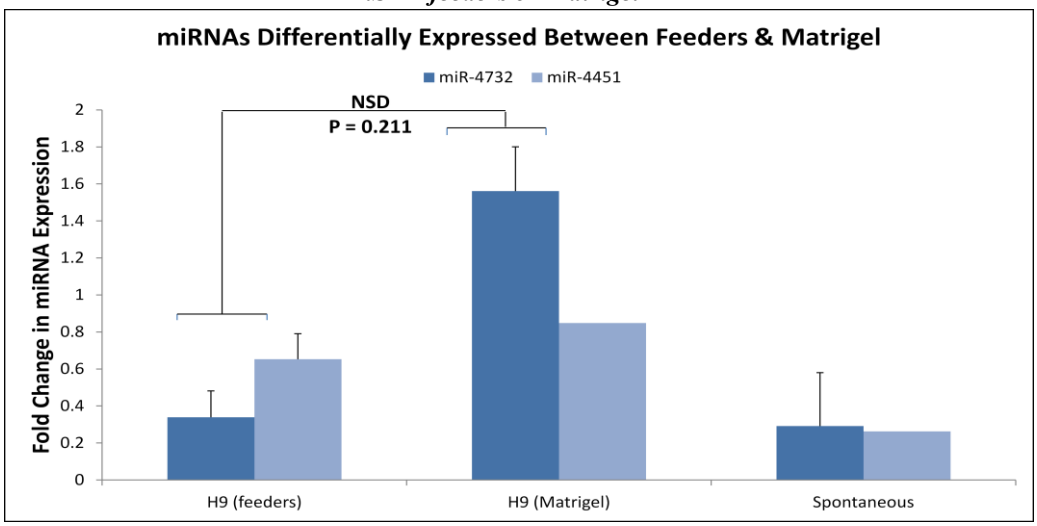

Figure 5.25. Comparison of expression of miRNAs differentially expressed between cells cultured on iSNL feeders or Matrigel™

Figure 5.23 shows the expression of miR-375, miR-708-5p, miR-744-5p, miR-27b, miR-151a-3p and miR-151a-5p in H9 cells cultured on iSNL feeders or Matrigel. These miRNAs have all been previously demonstrated to be consistently upregulated in DE formation in several ESC and iPSC lines, and figure 5.23 shows that they are upregulated in both culture conditions, with no significant difference between them ($p = 0.792$, $n = 3$). In addition, the patterns of expression are similar to those previously observed, with miR-375 and miR-708-5p upregulated only in cells directed to differentiate using the D'Amour protocol, and miR-744-5p, and miR-27b upregulated in both directed and spontaneous differentiation.

Figure 5.24 shows the expression of miR-124-5p and miR-516b. miR-516b was previously demonstrated to be downregulated in DE formation, and this is true for H9 grown on both feeders and Matrigel™. There are no significant differences ($p = 0.638$, $n = 3$) in the expression of these two miRNAs between feeders and Matrigel. Interestingly, miR-124-5p expression was predicted by the microarray to be downregulated in DE formation, and this has been the finding of other studies as well (Tzur *et al.*, 2008; Hinton *et al.*, 2010). However, previous qRT-PCR validation of the microarray results indicated that miR-124-5p expression was in fact upregulated in those samples (figure 5.15), although there was high variability between experiments, and the upregulation was close to being non-significant. In contrast to these findings, the comparison of differentiation on different culture methods again demonstrated downregulation of miR-124-5p in the H9 cell line, despite the use of the same differentiation protocol.

Figure 5.25 shows the expression of miR-4732 and miR-4451. These miRNAs were shown in qRT-PCR validation of microarray results to be differentially expressed between cell lines grown on feeders and Matrigel. Both were upregulated in DE formation only in the cell lines grown on Matrigel, and not those on feeders. This is reflected in the comparison of H9 on feeders and Matrigel, as both miRNAs are expressed at lower levels in H9 grown on feeders, but this is not statistically significant ($p = 0.211$, $n = 3$).

5.4 Discussion

In this chapter, I used microarray analysis to identify miRNAs differentially expressed between undifferentiated and differentiated cells, and which may therefore play a role in DE formation. A comparison of undifferentiated and differentiated cells was carried out for ESCs (H1 and H9 combined), iPSCs (MRC5I and MRC9G combined), as well as for individual cell lines. Although there are some differences in the miRNAs which are up- or down-regulated in these comparisons, several miRNAs show consistency between several cell lines, most notably miR-375 and miR-708-5p, which are consistently and significantly upregulated in all cell lines. These microarray results were validated using qRT-PCR, which confirmed the upregulation of miR-375 and miR-708-5p, suggesting a role for these miRNAs in DE formation. Both of these miRNAs have been previously reported to be upregulated in DE formation (Tzur *et al.*, 2008; Hinton *et al.*, 2010; Kim *et al.*, 2011), although their exact mechanism of action has not been fully elucidated. There was less consistency in the miRNAs predicted to be downregulated in DE formation by the microarray, and this was reflected in the PCR validation, as only one of the ten miRNAs investigated, miR-516b, was found to be significantly downregulated.

For the purpose of narrowing down the candidate miRNAs to focus on a more manageable number to investigate further, only miRNAs which were significantly up- or down-regulated in all cell lines were considered. Other miRNAs are strongly up- or down-regulated in some cell lines but not all e.g. miR-26b-5p, which is strongly upregulated in H9; miR-4436b-5p, which is strongly downregulated in MRC5I. These miRNAs may still be involved in endoderm formation, and may provide information about the different propensities of individual cell lines to form DE, but were disregarded in this study in favour of candidates likely to have a more general role in differentiation to DE.

Interestingly, miR-124-5p, which was predicted by the array data to be downregulated in DE formation, and which has been previously described as being downregulated in DE formation (Tzur *et al*., 2008; Hinton *et al*., 2010) actually appeared to be significantly upregulated when assessed using qRT-PCR for the validation of the microarray results. Upregulation was only seen in cells directed to differentiate, and not in spontaneously differentiating cells, but there was high variation between experiments and the level of upregulation was only just statistically significant. This miRNA is strongly associated with the neuronal lineage, and its

expression in these samples might reflect contamination of endodermal tissue with ectoderm due to inefficient differentiation, but the pattern of expression, with the highest expression in the cell lines with the highest efficiency of making DE, and the lowest in H1, which is poor at making DE, suggests that this may not be the case. miR-124-5p is also known to play a role in later pancreas development through its target genes *Pdx1* and *Neurod1* (Baroukh *et al*., 2007), which may explain its upregulation in these samples. Interestingly, the comparison of miRNA expression in cells cultured on iSNL feeders and Matrigel™ again demonstrated downregulation of miR-124-5p in the H9 cell line, despite the use of the same differentiation protocol. This is not due to a failure of differentiation in any of the samples, as qRT-PCR data shows that DE genes were highly upregulated in all of these samples. However, variability between experiments using the same protocol is a well-known issue, and this demonstrates the importance of fully elucidating the mechanisms involved in differentiation to enable a higher degree of standardisation of differentiation protocols.

The later stages of pancreatic development have been reasonably well-characterised in terms of miRNA expression and the functions of some of the most important of these miRNAs have also been elucidated. However, there have been relatively few studies focussing on the DE stage. I have identified several miRNAs which are either upregulated (miR-375, miR-708-5p) or downregulated (miR-516b) during DE formation and which may therefore play a role in this process. While miR-375 (Tzur *et al*., 2008; Hinton *et al*., 2010; Kim *et al*., 2011) and miR-708 (Hinton *et al*., 2010; Kim *et al*., 2011) have been previously described as being up- or down-regulated during DE formation, miR-516b has not.

A better understanding of the role of miRNAs in DE formation may prove useful both in terms of characterisation and in manipulation of cell fate. The expression of miR-375 and miR-708 in the cell lines tested in this study mirrored the expression of genes used as markers for this germ layer, such as *Sox17*, with the highest expression of these miRNAs in the cell lines previously noted to give rise to DE at higher efficiency, and with the lowest expression in the H1 cell line, which has a poor efficiency of differentiation into DE. In addition to analysis of gene expression, analysis of miRNA expression may also be helpful when characterising the differentiated progeny of stem cells.

Another way in which a better understanding of the miRNA expression patterns during differentiation may be useful is in their potential to influence cell fate, either through replacing transcription factors and signalling molecules, or enhancing differentiation

efficiency. The use of miRNAs for this purpose is appealing, as they have the potential to directly and immediately alter gene transcription, which may reduce the time needed for *in vitro* differentiation and increase the efficiency of differentiation and the yield of the desired cell type. The ability of miRNAs to post-transcriptionally regulate hundreds of target genes makes them ideally suited to bring about rapid transformations in cell phenotype. They can be transiently expressed in a cell, and therefore do not permanently alter gene expression in the target cells, which has been a concern with other methods of altering cell fate, such as the use of lentiviral vectors to overexpress transcription factors. In addition, miRNAs are inherently less variable than the recombinant growth factors commonly used in differentiation protocols.

qRT-PCR was also used to validate differences in miRNA expression between ESCs and iPSCs, with a view to identifying any differences in miRNA expression which may have a functional impact on their ability to differentiate into a pancreatic lineage. Initial comparison of undifferentiated ESCs and iPSCs did not identify any miRNAs differentially expressed between the two cell types. When undifferentiated H1 samples were compared with undifferentiated iPSC samples, only 5 miRNAs were significantly differentially expressed. This suggests that these samples are highly similar in terms of their miRNA expression. However, when undifferentiated H9 samples were compared with undifferentiated iPSC samples, 186 miRNAs were differentially expressed, suggesting either that this cell line genuinely exhibited a significantly different miRNA expression profile, or that the culture conditions have a significant effect on miRNA expression levels. A similar phenomenon was noticed when comparing differentiated samples. When differentiated H9 were compared with differentiated iPSCs, 154 miRNAs were differentially expressed, but when differentiated H1 were compared with differentiated iPSCs, only 28 miRNAs were significantly differentially expressed. The fact that a greater number of miRNAs are differentially expressed between H9 and iPSCs compared to H1 and iPSCs again suggests that culture conditions have having an effect on miRNA expression.

The differences in miRNA expression which were observed in the H9 cell line were not anticipated from either the pluripotency characterisation data (chapter 3) or the differentiation data (chapter 4). For this reason, another cell line grown on Matrigel™ (H7) was included in the PCR validation studies. These experiments identified two miRNAs that appeared to show

differences in their expression possibly as a result of the culture method: miR-4451and miR-4732 were upregulated in cells grown on Matrigel (H7 and H9) and downregulated in cells grown on iSNL feeders (MRC5I, MRC9G and H1). To investigate these differences in miRNA expression, a further differentiation experiment was carried out using H9 cells grown on both iSNL feeders and Matrigel, and the expression of miRNAs previously demonstrated by qRT-PCR validation of the microarray results to be up- or down-regulated in DE formation was investigated in these samples. There were no significant differences in expression of any of these miRNAs between H9 cells grown on iSNL feeders and H9 cells grown on Matrigel. There were some differences in the expression of miR-4732 and miR-4451, which had been previously indicated by the PCR validation result, but this was not statistically significant. These miRNAs may represent conserved differences between cell lines cultured in different ways, but a larger experiment including more cell lines would be needed to investigate this further.

Despite observed dissimilarities between the cell lines and between culture methods, some miRNAs, including miR-151a-3p, miR-151a-5p and miR-27b, previously identified as potentially playing a role in endoderm formation, were also identified as being differentially expressed in both the overall comparison of iPSC vs. ESC, as well as the individual comparison of H1 vs. iPSC and H9 vs. iPSC. These miRNAs are more likely to represent conserved differences between ESCs and iPSCs which are not specific to certain cell lines, although this is difficult to determine in an analysis of only five cell lines. The strongest contrast was seen in expression of miR-151a-3p and miR-151a-5p, which are upregulated upon differentiation to DE in ESCs, and downregulated in iPSCs. This pattern of expression is not seen in spontaneously differentiated samples. These two miRNAs are derived from the same pre-miRNA stem loop, and are therefore likely to be co-regulated and co-transcribed, although as their sequence is different they are likely to have different target genes and therefore different mechanisms of action. The function of these miRNAs will be further investigated in chapter 6.

Chapter 6

Investigation of the Function of miRNAs in Differentiation to Definitive Endoderm

Contents

List of Figures

Chapter 6: Investigation of the Function of miRNAs in Differentiation to Definitive Endoderm

6.1 Introduction

The roles of miRNAs in the later stages of pancreatic development, as well as in mature pancreatic function, have been reasonably well described (section 1.10.4). However, relatively few studies (Tzur *et al.*, 2008; Hinton *et al.*, 2010; Tsai *et al.*, 2010; Kim *et al.*, 2011) have investigated miRNA expression at the definitive endoderm (DE) stage. Although there is relatively little overlap in the miRNAs identified in these studies, which is likely to be due to the different cell lines and differentiation protocols used, some miRNAs, including miR-375, miR-708 and miR-744, were identified in several of the studies, providing strong evidence for their involvement in DE formation. However, few of these studies went on to elucidate a function for these miRNAs in DE formation. An early study in ESCs investigated the effect on differentiation of overexpressing miR-122, but found that, rather than increasing differentiation efficiency into an endodermal lineage, this actually resulted in an overall delay in differentiation (Tzur et al, 2008). A subsequent study identified *TIMM8A* as a target of miR-375 but didn't elucidate a function for this pathway in DE formation (Hinton et al, 2010). Fu *et al.* (2011) showed that overexpression of a panel of miRNAs in differentiating mouse ESCs resulted in upregulation of the definitive endoderm genes *Sox17* and *Foxa2*, suggesting an increase in differentiation efficiency.

miR-375 was also identified as upregulated in DE formation by my own microarray analysis (chapter 5). A role for this miRNA in DE formation has been suggested in ESCs (Hinton et al, 2010), but it is unknown whether this mechanism also applies to iPSCs. In addition, the study carried out by Hinton et al (2010) only investigated the ectopic expression of miR-375, and not inhibition. This can result in false results due to supraphysiological levels of miRNAs, which have the potential to saturate RISC complexes and displace other endogenous miRNAs, consequently causing low affinity target sites to appear functionally important. These problems may be overcome by the use of miRNA silencing with an inhibitor to influence mRNA expression within physiological levels, although this approach

has its own limitations, as miRNAs often exist in families of highly related sequences and this high degree of redundancy poses challenges for loss-of-function analyses, as other family members may compensate for the loss of the miRNA of interest. Ideally, both overexpression and inhibition approaches should be used when investigating the function of miRNAs. In addition, this study only investigated the levels of TIMM8A mRNA following manipulation of miR-375 levels, and did not show any effect on protein levels. This is important as miRNAs do not always mediate translational repression through degradation of target mRNA, they may simply prevent its transcription. For these reasons, the functional relationship between miR-375 and TIMM8A was investigated in both ESCs and iPSCs. As this relationship has already been established, it was thought that this would also be a useful way of validating the new techniques used in this chapter to study miRNA function.

In addition to identifying miRNAs important in DE formation, microarray analysis and subsequent qRT-PCR was also used to identify miRNAs that were also differentially expressed between ESCs and iPSCs, and which may therefore have an impact on their ability to differentiate into DE. Two miRNAs fulfilled this criteria, miR-151a-3p and miR-151a-5p. These miRNAs were shown to be upregulated in DE formation in ESCs, but downregulated in iPSCs. These miRNAs have not been previously implicated in DE formation, although their expression has been noted in foetal liver (Fu et al, 2005) and their function remains unknown, although a few gene targets have been identified (mirtarbase.mbc.nctu.edu.tw/index.php, accessed 03/12/2013). For this reason, these miRNAs were identified for further study through identification of gene targets, computationally and experimentally through the use of a luciferase assay, and through manipulation of miRNA expression levels in ESCs and iPSCs.

6.2 Methods

6.2.1 Identification of Gene Targets of miRNAs

None of the currently available existing target prediction software tools has been able to include all the currently known features of miRNA-target interactions. For this reason, three algorithms were used to generate lists of predicted miRNA targets: TargetScan (www.targetscan.org), miRDB (www.mirdb.org) and Pictar (www.pictar.mdc-berlin.de/cgi-bin/PicTar_vertebrate.cgi). The results are shown in appendix 2.

6.2.2 Luciferase Assay in 293FT Cells

In order to determine whether miRNAs were acting directly by binding to genes which had been identified as putative targets, a luciferase assay was carried out in which a miRNA mimic or inhibitor were co-transduced using lentiviral vectors into 293FT cells together with a reporter plasmid carrying a *Gaussia luciferase* gene and the 3'UTR of the gene of interest, which is where the majority of miRNA-target interactions occur.

Plate Map					
Vector Only	Vector Only	Reagents only	Reagents only	Control Vector Only	Control Vector Only
Vector + Mimic	Vector + Mimic	Vector + Scrambled Mimic	Vector + Scrambled Mimic	Control Vector + Mimic	Control Vector + Mimic
Vector + Inhibitor	Vector + Inhibitor	Vector + Scrambled Inhibitor	Vector + Scrambled Inhibitor	Control Vector + Inhibitor	Control Vector + Inhibitor
Vector + Mimic + Inhibitor	Vector + Mimic + Inhibitor	Untreated	Untreated	Untreated	Untreated

Figure 6.1. Plate map used in luciferase assays.

293FT cells were maintained in culture as described in section 2.1.1. 24hrs prior to infection, 293FT cells were trypsinised and counted before being plated at a density of $3x10^4$ cells per well of a 24-well plate in 293FT medium, giving a confluency of 60-70%. After 24hrs, the

medium was removed from the cells, and they were washed once with Opti-MEM reduced serum medium. 0.5ml of Opti-MEM supplemented with 4µg/ml Polybrene was added to each well, and 10µl of lentivirus or 1µl of control lentivirus (giving a MOI of approximately 10) was added to each well according to the plate map shown in table 6.1. After 5-6 hours, the medium on the cells was replaced with 293FT medium.

72hrs after infection with lentivirus, cells were transfected with the 3'UTR-containing plasmid. Cells were washed once with Opti-MEM and 0.5ml of Opti-MEM was added to the cells. 2.5µl of Lipofectamine 2000 transfection reagent was diluted in 25µl of Opti-MEM per well. 500ng of plasmid DNA was diluted in 50µl of Opti-MEM per well, and the DNA and Lipofectamine were combined and incubated at room temperature for 5mins. The DNA-Lipofectamine mixture was then added dropwise to each well according to the plate map shown in table 6.1. The cells were incubated for 5-6hrs at 37C, 5% CO2 before the medium on the cells was replaced with 293FT medium. 72hrs after plasmid transfection, the culture medium was collected and analysed for luminescence using the SecretePair™ Dual Luminescence Assay Kit as described in section 2.5.4.

6.2.3 Manipulation of miRNA Expression

In order to determine what effects the miRNAs identified in this study might have on differentiation to definitive endoderm, their levels of expression were manipulated using lentiviral vectors carrying either miRNA mimics to overexpress the miRNA of interest, or miRNA inhibitors, to knock down the miRNA of interest. The different protocols used are described in more detail in section 6.3.4.

6.3 Results

6.3.1 Identification of Gene Targets of miRNAs

The results of the target prediction algorithms for miR-151a-3p and miR-151a-5p are shown in appendix 2. There was considerably more overlap in the target predictions from TargetScan and miRDB than from PicTar, which did not overlap with the results from the other two algorithms at all for either of these miRNAs. This is likely to be due to the different criteria measured by each algorithm, as TargetScan and miRDB are more similar to each other than PicTar (see section 2.5.3).

Interestingly, *Sox17* appeared on the list of predicted targets for miR-151a-5p in the results from both TargetScan and miRDB. This gene is one of the earliest markers of DE (Lewis & Tam, 2006), and is essential for endoderm formation, as mice lacking SOX17 do not develop gut endoderm from DE (Kanai-Azuma *et al*., 2002). It is possible that miR-151a-5p has an effect on endoderm formation through regulation of *Sox17* expression, and this was selected for further investigation in this chapter.

6.3.2 Investigation of the Function of miR-375 in Definitive Endoderm Formation

There is strong evidence for the involvement of miR-375 in pancreatic development, both at the DE stage and at later stages. A role for this miRNA in DE formation has been suggested in ESCs (Hinton et al, 2010), but it is unknown whether this mechanism also applies to iPSCs. In addition, this study didn't investigate whether changing levels of miR-375 had an effect on the efficiency of DE formation. For this reason, this miRNA was chosen for further study. As a miRNA-target relationship has already been established, it was thought that this would also be a useful way of validating the new techniques used in this chapter to study miRNA function.

Luciferase Assay in HFF1 Cells

The fibroblast line HFF1 was initially chosen to carry out luciferase assays to determine the miRNA-target relationships, as it is known from earlier reprogramming experiments that this cell line is amenable to transduction using viral vectors. Lentiviral transduction in this cell line was successful (figure 6.2), with the highest expression of GFP/mCherry in the cells infected with the highest amounts of lentivirus and expression reaching maximum levels at 72 hours post-infection. However, attempts to transfect these cells with the plasmid containing the luciferase reporter gene were unsuccessful. For this reason, 293FT cells were selected as an alternative cell line for these experiments, as they were found to be amenable to both lentiviral transduction and plasmid transfection.

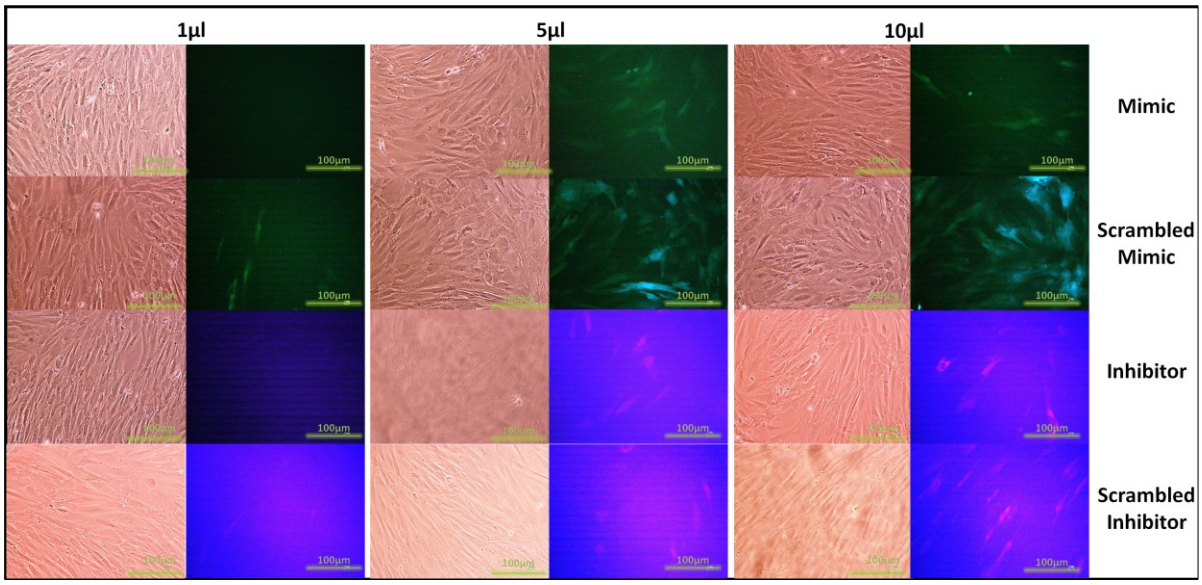

Figure 6.2. Transduction of HFF1 cells with the lentiviral vectors encoding the miR-375 miRNA mimic, the scrambled miRNA mimic, the miR-375 inhibitor or the scrambled miRNA inhibitor. Pictures taken 72hrs post-infection, 10x magnification, scale bar 100µm.

Luciferase Assay in 293FT Cells

Figure 6.3 shows that 293FT cells could be successfully transduced with the lentiviral vectors encoding the miR-375 mimic, miR-375 inhibitor, scrambled control mimic or scrambled control inhibitor, as shown by the expression of the fluorescent reporter genes GFP (mimic and scrambled mimic) and mCherry (inhibitor and scrambled inhibitor). An MOI of 10,

achieved using 10µl of lentiviral particles, was determined to give the highest infection efficiency and was therefore used in subsequent experiments.

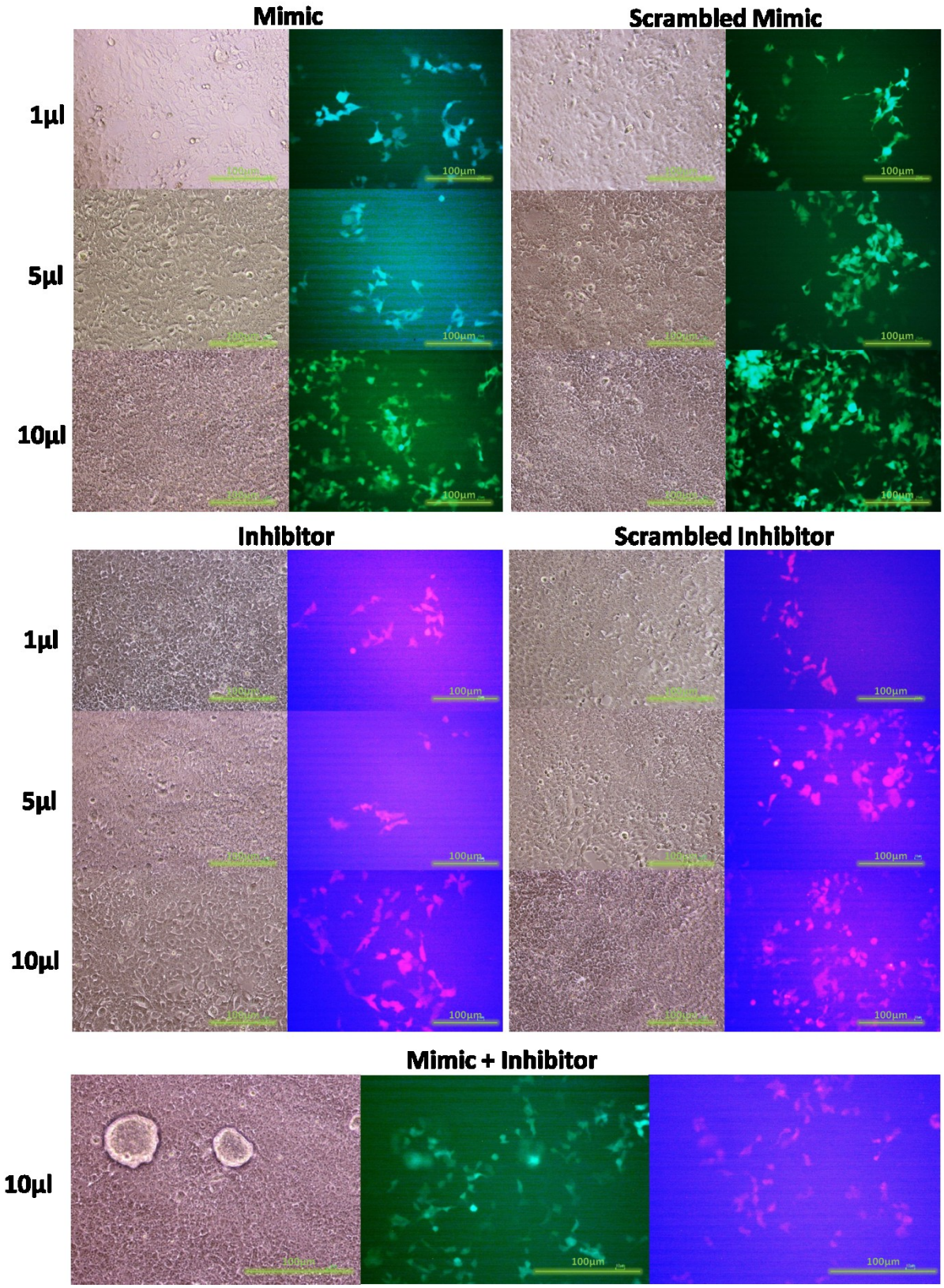

Figure 6.3.Expression of the GFP and mCherry reporter genes in 293FT cells which have been transduced with lentiviral vectors carrying a miR-375 mimic, a miR-375 inhibitor, a scrambled control mimic or a scrambled control inhibitor. *Pictures taken 72hrs post-infection, 10x magnification, scale bar 100µm.*

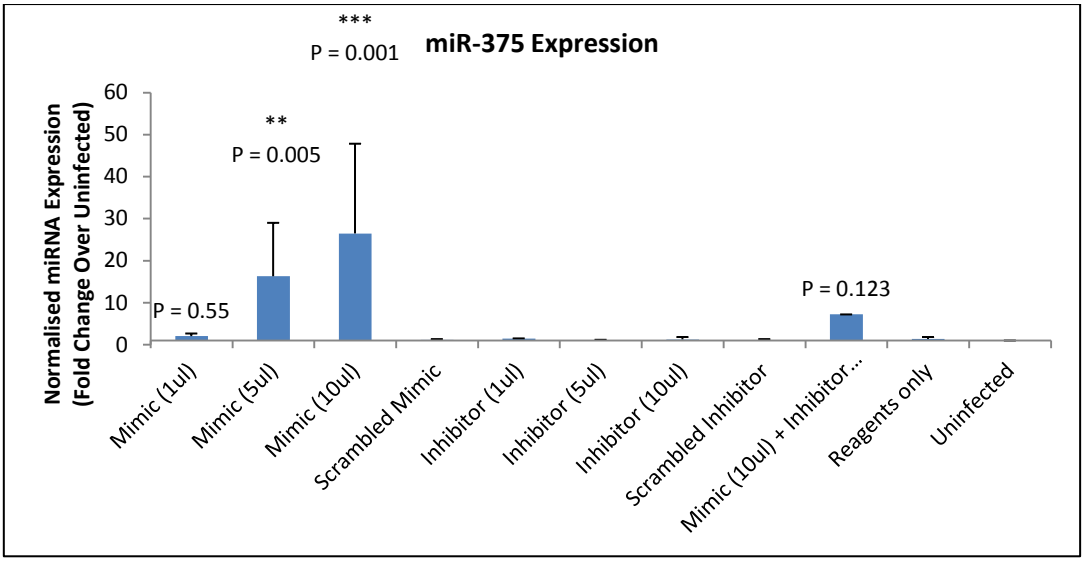

Figure 6.4 qRT-PCR data showing miR-375 expression in transduced 293FT cells

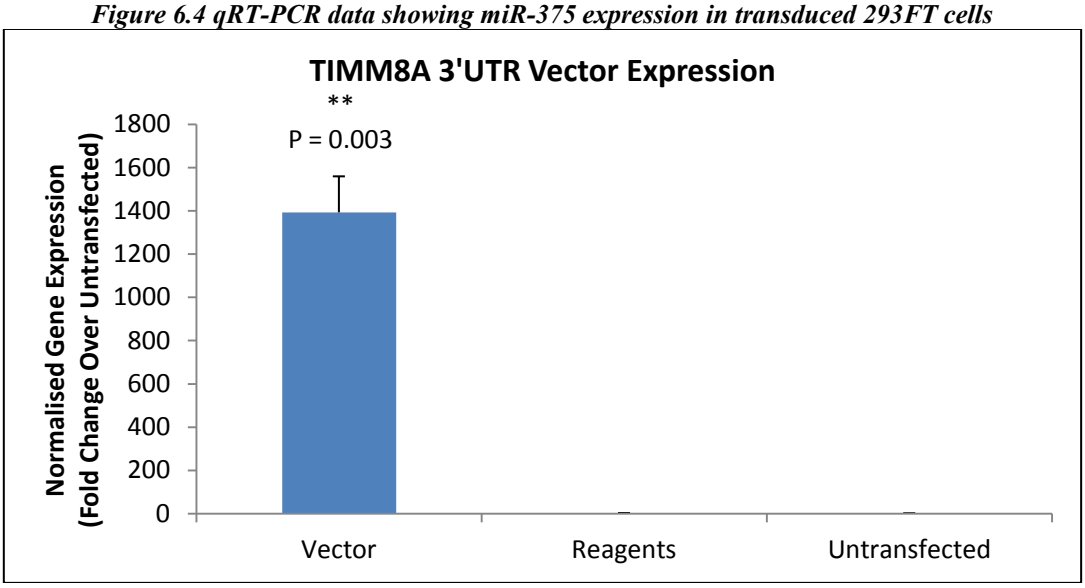

Figure 6.5. qRT-PCR data showing expression of the 3'UTR reporter plasmid in transfected 293FT cells

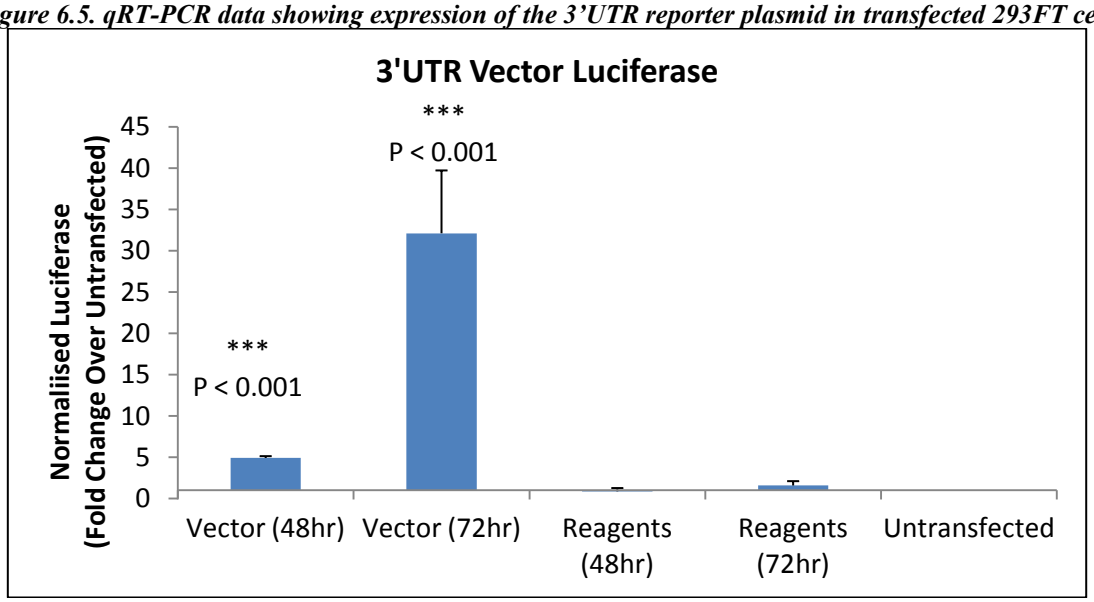

Figure 6.6. Luciferase expression in transduced 293FT cells

Figure 6.4 shows that transduction of 293FT cells with the lentiviral vectors results in altered expression levels of miR-375 in 293FT cells. There is a significant increase over the untreated cells in miR-375 expression in cells which have been transduced with either 5ul ($p = 0.005$, $n = 2$) or 10µl ($p = 0.001$, $n = 2$) of the miR-375 mimic, with the greatest increase in the cells receiving the highest amount of lentivirus. The cells transduced with 1µl of lentivirus showed a slight increase in miR-375 expression, although this was not statistically significant. There is no miR-375 expression in cells transduced with the scrambled control mimic. As there is no endogenous miR-375 expression in 293FT cells, no reduction in miR-375 levels could be seen in cells transduced with the miR-375 inhibitor alone. However, cells transduced with both the miR-375 mimic and inhibitor show lower levels of miR-375 expression than cells transduced with the miR-375 mimic alone, although this difference is not statistically significant ($p = 0.738$, $n = 2$), confirming the activity of the inhibitor lentiviral vector. Figure 6.5 shows qRT-PCR data showing expression of the luciferase plasmid in 293FT cells. There is significantly higher ($p = 0.003$, $n = 2$) expression of the plasmid in transfected cells than in the untreated cells or the cells treated with the transfection reagents only. Figure 6.6 shows luminescence in transfected 293FT cells. There is a significant increase in luciferase levels in cells transfected with the plasmid at both 48hrs and 72hrs ($p < 0.001$, $n = 4$) compared with untreated cells, although luciferase is highest at 72hrs. These results confirm successful transfection of the luciferase plasmid into the 293FT cells.

Once experimental conditions had been optimised, luciferase assays were carried out in 293FT cells to determine whether we could also detect a direct relationship between miR-375 and its target gene *Timm8a*, as has been previously described (Hinton et al, 2010). 293FT cells were transduced with lentiviral vectors containing the miR-375 mimic, inhibitor or scrambled controls. After 72 hours, the cells were transfected with a vector containing the 3'UTR region of the *Timm8a* gene, together with a luciferase reporter gene. As a control, some cells were transfected with a plasmid containing a scrambled 3'UTR which does not correspond to any known gene sequence. Cells were analysed using a luciferase assay and qRT-PCR. Analysis of variance with Fisher's *a priori* test was used to determine differences between sample groups.

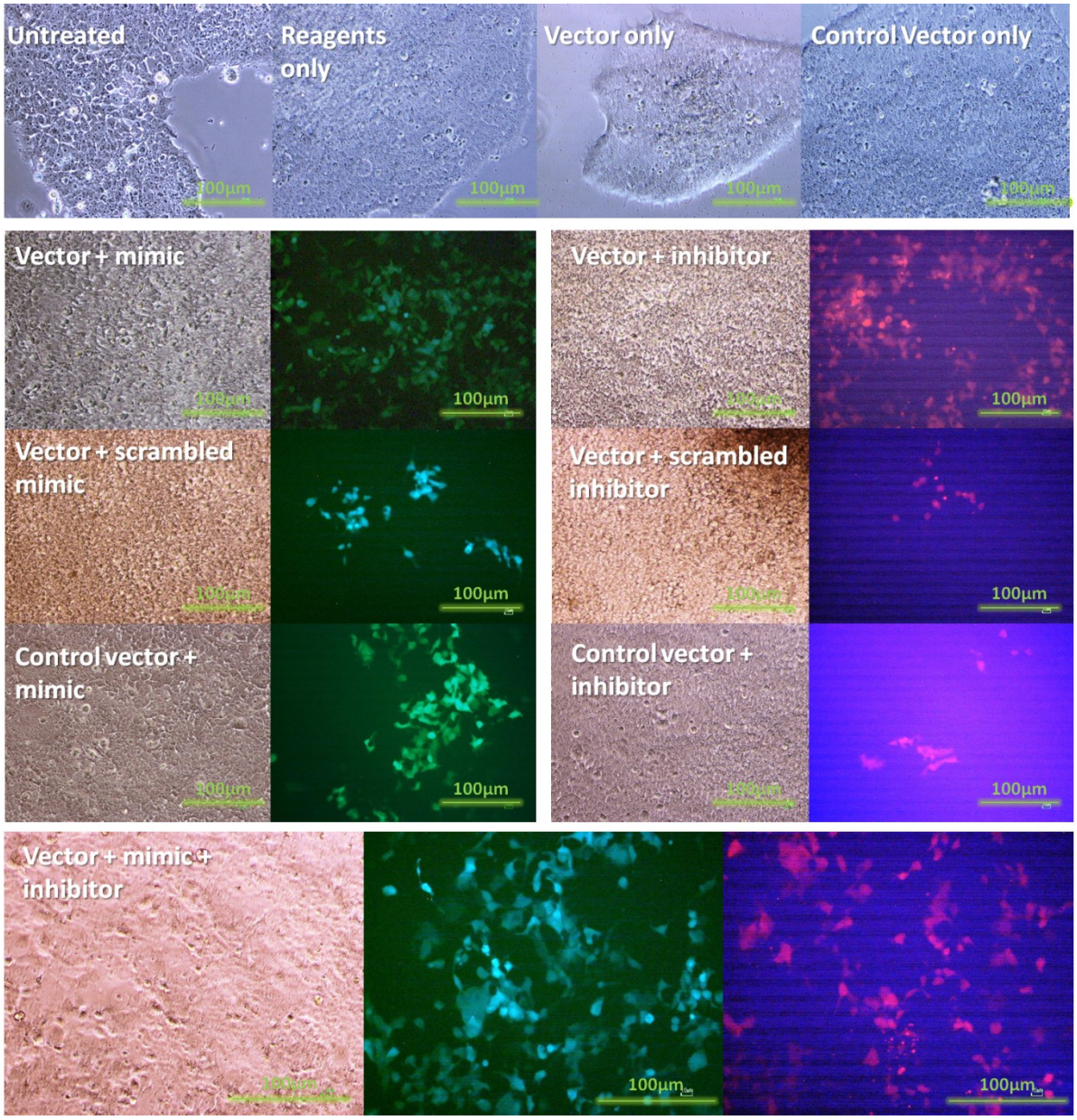

Figure 6.7. Expression of GFP and mCherry reporter genes in 293FT cells transduced with lentiviral vectors carrying a miR-375 mimic, miR-375 inhibitor, scrambled control mimic or scrambled control inhibitor. Cells were also transfected with a luciferase reporter plasmid containing either the* Timm8a *3'UTR or a scrambled control sequence. *10x magnification, scale bar 100µm.*

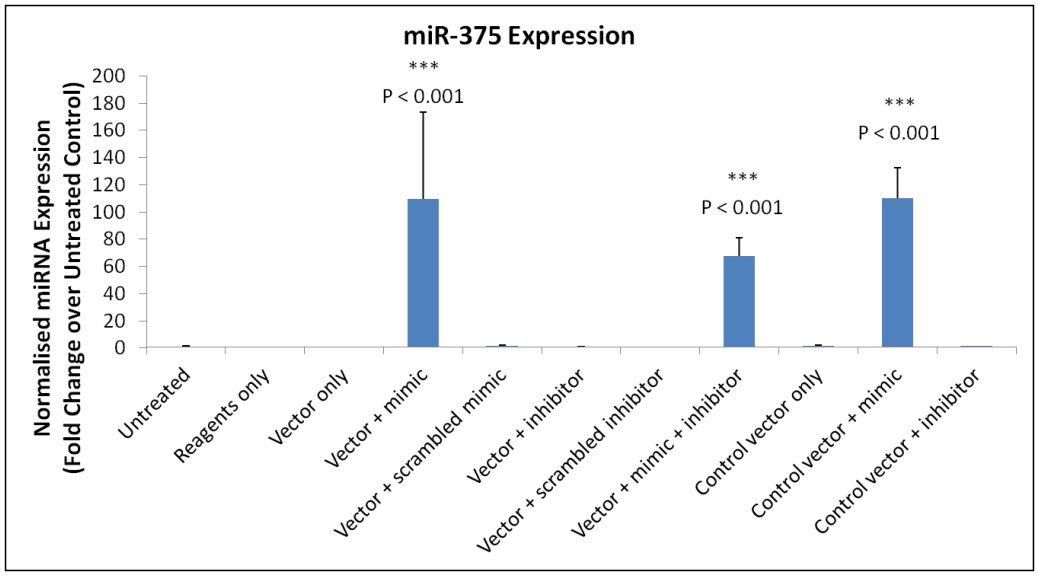

Figure 6.8 qRT-PCR data showing expression of miR-375 in transduced 293FT cells.

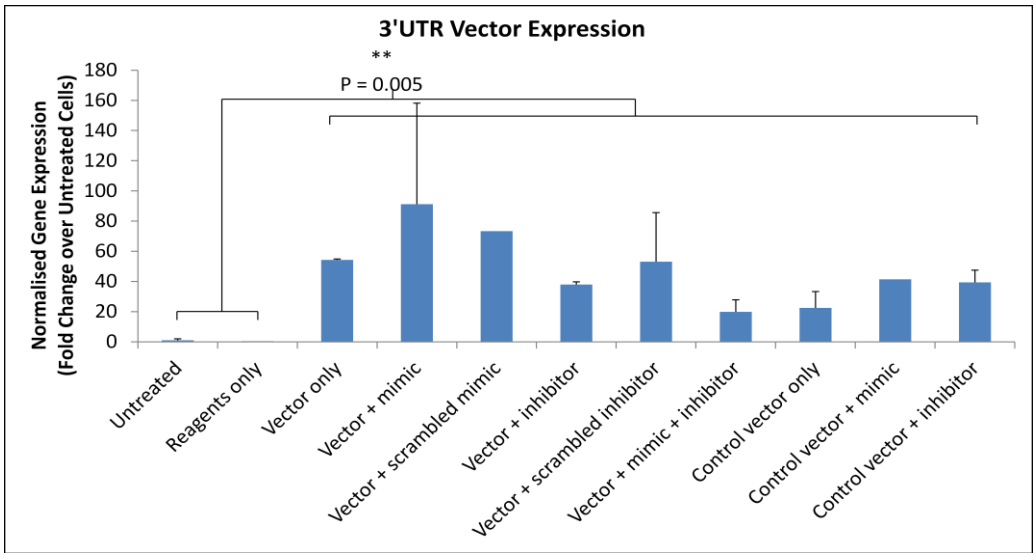

Figure 6.9 qRT-PCR data showing expression of the 3'UTR vector in transfected 293FT cells.

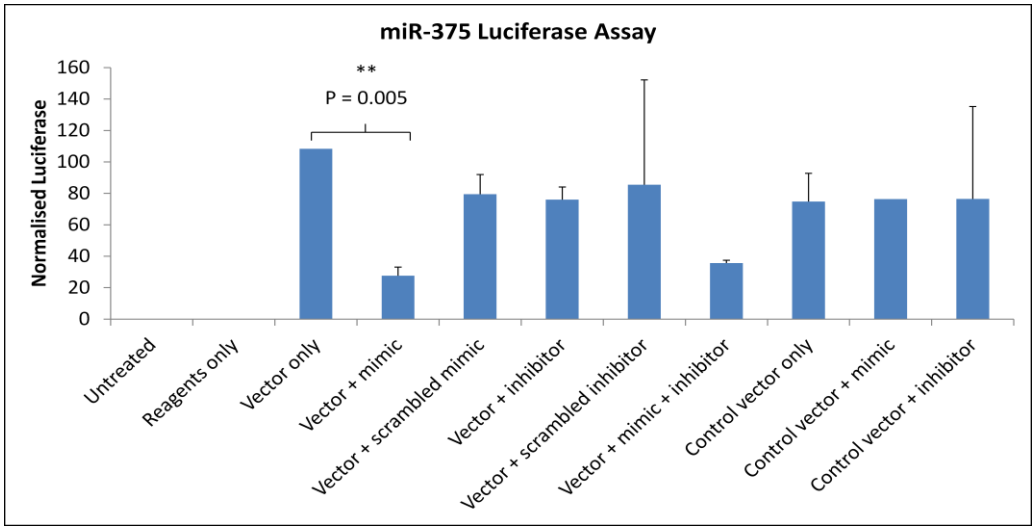

Figure 6.10.Luciferase expression in transduced 293FT cells.

Figure 6.7 shows successful transduction of 293FT cells with lentiviral vectors carrying a miR-375 mimic, miR-375 inhibitor, scrambled control mimic or scrambled control inhibitor, as demonstrated by the expression of the GFP or mCherry reporter genes. Untreated cells, as well as cells treated with reagents only or vector only, were included as negative controls. Figure 6.8 shows expression of miR-375 in transduced 293FT cells. There is a significant ($p=0.008$, $n=2$) increase in miR-375 expression in cells infected with the lentiviral vector encoding the miR-375 mimic compared with untreated cells. The levels of miR-375 expression are reduced in 293FT cells also infected with the lentiviral vector encoding the miR-375 inhibitor, confirming its activity, although this reduction is not statistically significant ($p = 0.711$, $n = 2$). There is no increase in miR-375 expression in cells treated with the scrambled control lentivirus. Figure 6.9 shows expression of the 3'UTR vector in transfected 293FT cells. There is a significant increase ($p = 0.005$, $n = 2$) in 3'UTR vector expression in wells transfected with either the TIMM8A 3'UTR plasmid or the scrambled control plasmid compared to the untransfected cells. Although there is some variability in the expression of the 3'UTR vector between wells, this is not statistically significant, and the use of a dual reporter luciferase assay system means that the luminescence reading can be normalised between samples. Figure 6.10 shows luminescence readings from transduced 293FT cells. There is high luciferase in cells transfected with either the TIMM8A 3'UTR vector or the scrambled control vector. The luciferase levels are significantly ($p = 0.005$, $n = 2$) reduced in cells infected with the miR-375 lentivirus compared to cells treated with the 3'UTR vector only. This demonstrates a direct binding relationship between miR-375 and the 3'UTR region of the TIMM8A gene, confirming the reliability of this assay. There is no reduction in luciferase in cells transduced with the scrambled control vector when the miR-375 lentiviral vector is added, showing that this is not due to non-specific effects, and there is also no reduction when scrambled control lentiviral vectors are used. There is some reduction in luciferase in cells transduced with both the miR-375 mimic and inhibitor vectors, showing that the inhibitor is not present in sufficient amounts to entirely prevent miR-375 binding to the TIMM8A 3'UTR, but this is less than in the cells transduced with the miR-375 mimic alone, and the reduction in luciferase is not statistically significant ($p = 0.239$, $n = 2$).

6.3.3 Investigation of the Function of miR-151a-5p in Definitive Endoderm Formation

Luciferase assay in 293FT Cells

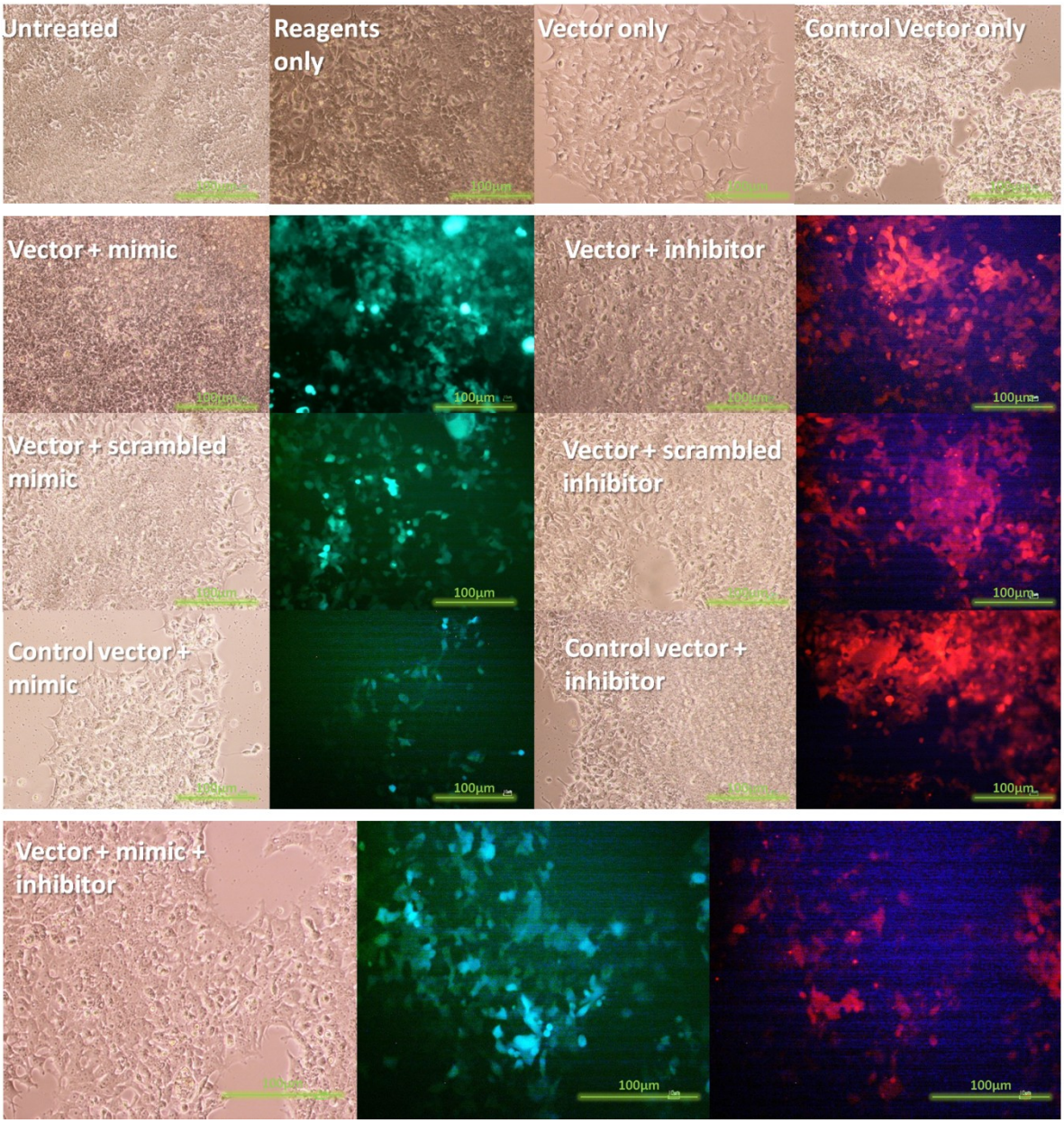

Figure 6.11. Expression of the GFP and mCherry reporter genes in 293FT cells which have been transduced with lentiviral vectors carrying a miR-151a-5p mimic, a miR-151a-5p inhibitor, a scrambled control mimic or a scrambled control inhibitor. Cells were also transfected with a luciferase reporter plasmid containing either the* Sox17 *3'UTR or a scrambled control sequence. *Pictures taken 72hrs post-infection, 10x magnification, scale bar 100µm.*

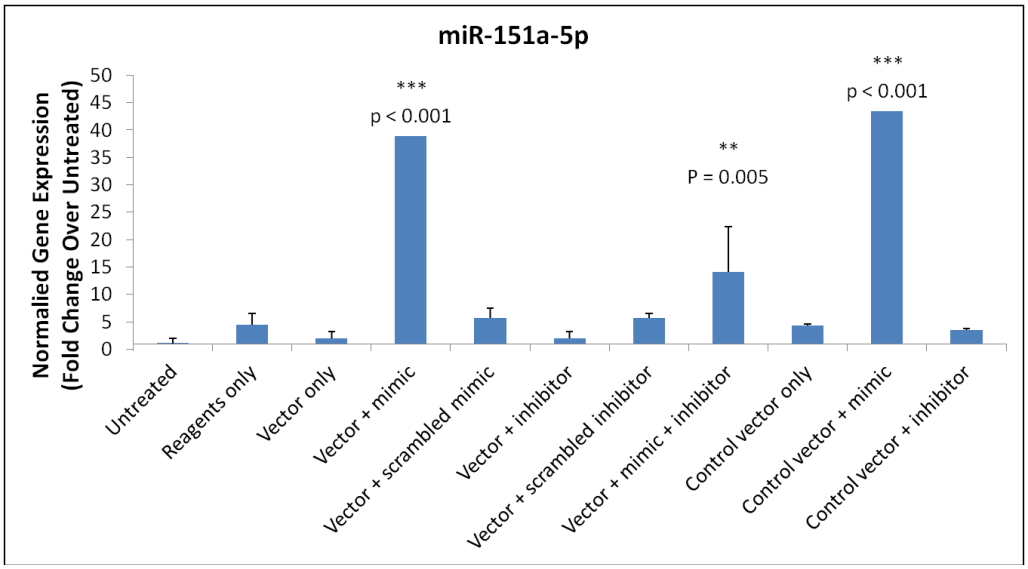

Figure 6.12. qRT-PCR data showing expression of miR-151a-5p in transduced 293FT cells.

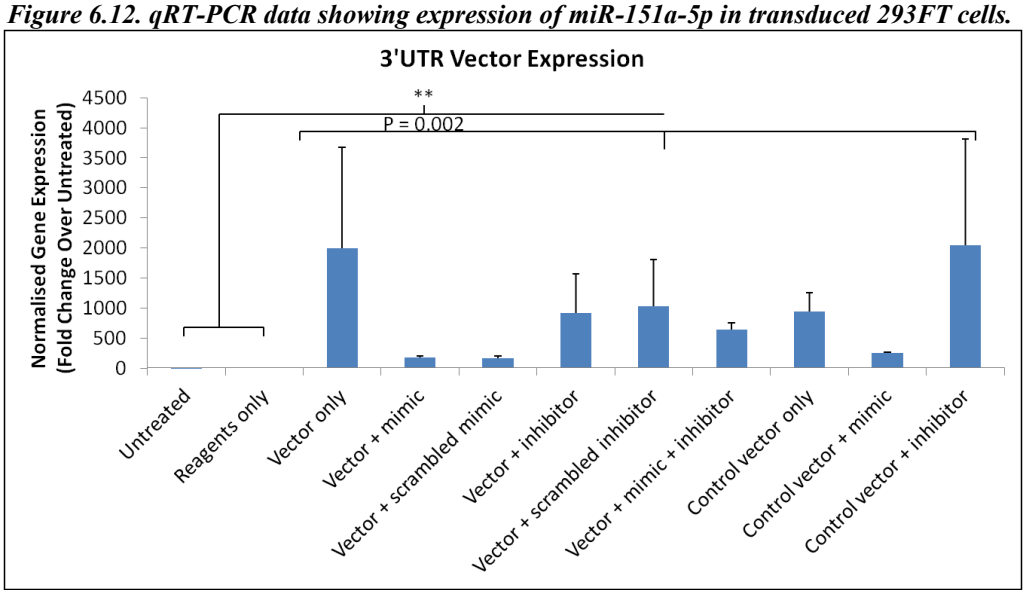

Figure 6.13. qRT-PCR data showing expression of the 3'UTR vector in transduced 293FT cells.

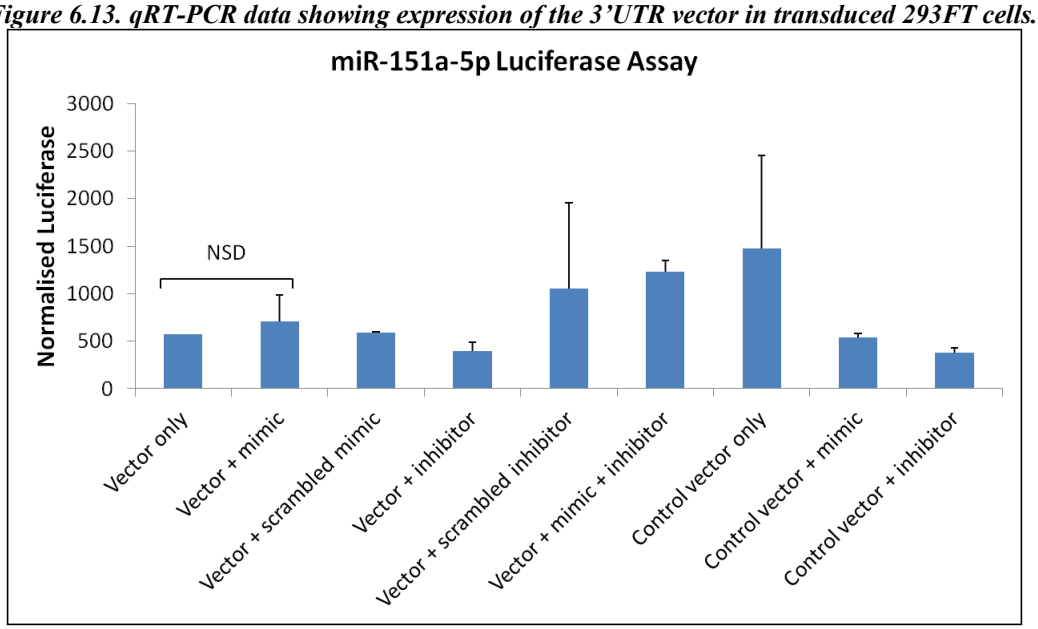

Figure 6.14. Luciferase assay in transduced 293FT cells.

The luciferase assay was next used to determine whether there was a direct relationship between miR-151a-5p and its predicted target gene *Sox17*. Following transduction with the lentiviral vectors and 3'UTR plasmid as described previously, cells were analysed using a luciferase assay and qRT-PCR. Analysis of variance with Fisher's *a priori* test was used to determine differences between sample groups.

Figure 6.11 shows successful transduction of 293FT cells with lentiviral vectors carrying a miR-151a-5p mimic, miR-151a-5p inhibitor, scrambled control mimic or scrambled control inhibitor, as demonstrated by their expression of the GFP or mCherry reporter genes. Untreated cells, as well as cells treated with reagents only or vector only, were included as negative controls. Figure 6.12 shows expression levels of miR-151a-5p in transduced 293FT cells. There is a significant ($p < 0.001$, $n = 2$) increase in miR-151a-5p expression in cells treated with the miR-151a-5p mimic compared to untreated cells. The levels of miR-151a-5p expression are reduced in 293FT cells also infected with the lentiviral vector encoding the miR-151a-5p inhibitor, confirming its activity, although this difference is not statistically significant ($p = 0.094$, $n = 2$). There is no significant increase in miR-151a-5p expression in cells treated with the scrambled control mimic. Figure 6.13 shows expression levels of the 3'UTR in transfected 293FT cells. There is a significant increase ($p = 0.002$, $n = 2$) in 3'UTR vector expression in wells transfected with either the SOX17 3'UTR plasmid or the scrambled control plasmid compared to the untransfected cells. Although there is some variability in the expression of the 3'UTR vector between wells, this is not statistically significant, and the use of a dual reporter luciferase assay system means that the luminescence reading can be normalised between samples. Figure 6.14 shows luminescence readings from transduced 293FT cells. There is high luciferase in cells transfected with either the SOX17 3'UTR vector or the scrambled control vector. However, there is no reduction in luciferase in cells treated with the miR-151a-5p mimic, indicating that there is no direct binding relationship between miR-151a-5p and its predicted target *Sox17*.

6.3.4 Manipulation of miRNA Expression in Pluripotent Stem Cells

Several different protocols were tested to determine which would give the most efficient transduction of pluripotent stem cells with the lentiviral vectors carrying either the miRNA mimic or inhibitor. In the first protocol, cells were passaged on Matrigel™ as previously described. 24hrs prior to lentiviral infection, H9 ESCs were treated with Accutase™ to obtain a single cell suspension. This was added to a Matrigel™-coated 24-well plate at a concentration of $1x10^5$ cells per well in mTesr-1 medium supplemented with Y-27632 ROCK inhibitor at a concentration of 10μM, as this has been previously shown to prevent dissociation-induced apoptosis of stem cells (Watanabe et al, 2007). After 24hrs, the medium was changed to 50% Opti-MEM low serum medium and 50% mTeSR-1, supplemented with 4μg/ml Polybrene. 10μl of lentivirus was added to each well, and the cells were incubated at 37°C for 6hrs. As a control, untreated cells and cells treated with the transfection reagents only were included in the experiment. The medium was then changed to mTeSR-1. However, 24hrs post-infection, there was a large amount of cell death (figure 6.15), despite the presence of the Rock inhibitor, which was not seen in the untreated cells, and there was no expression of the fluorescent reporter gene, suggesting that this protocol is not optimised for pluripotent stem cells.

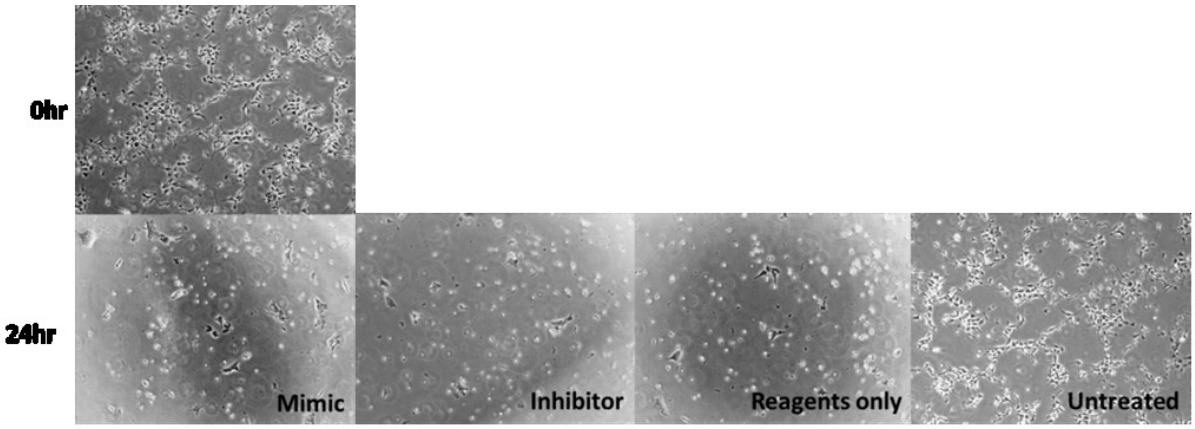

Figure 6.15 Images before (0hr) and after (24hr) lentiviral transfection show poor cell survival. *4x magnification.*

In an attempt to try and optimise this protocol for use with pluripotent stem cells, an experiment was carried out in which the multiplicity of infection (MOI) used and the time for which the cells were treated with the lentivirus were varied, as shown in figure 6.16.

Plate Map				
MOI				**Treatment Time**
5	10	20	50	6 hrs
5	10	20	50	12hrs
5	10	20	50	24hrs

Figure 6.16. Plate map used in lentiviral transduction of stem cells.

As ESCs and iPSCs are known to be vulnerable to increased apoptosis following dissociation into a single-cell suspension (Watanabe et al, 2007), whole colonies were treated instead. H9 ESCs were passaged as normal into a 24-well Matrigel-coated plated approximately 5 days prior to transduction. On the day of infection, cells were washed, and the medium was replaced with 50% Opti-MEM low serum medium and 50% mTeSR-1, supplemented with 4µg /ml Polybrene. The amount of lentivirus indicated in figure 6.16 was added to each well, and the cells were incubated at 37°C for the time shown. As a control, untreated cells and cells treated with the reagents only were included in the experiment. Cell pictures were taken prior to infection, and then at 24, 48 and 72hrs post-infection. While cell viability was improved using this method, there was still a significant amount of cell death at 72hrs post-infection (see figure 6.17) which was not seen in either the untreated cells or the cells treated with the reagents alone, and there was no expression of the fluorescent reporter gene at any time point.

In case the lack of transduction observed was due to the cell line used, this experiment was repeated with the Wild7 iPSC line. However, the results were very similar to those seen when using H9 cells (figure 6.18), with increased cell death in the wells treated with the lentivirus, and no expression of the fluorescent reporter gene at any time point. This suggests that the inefficiency is not due to the cell lines used, but instead due to inadequate protocol optimisation.

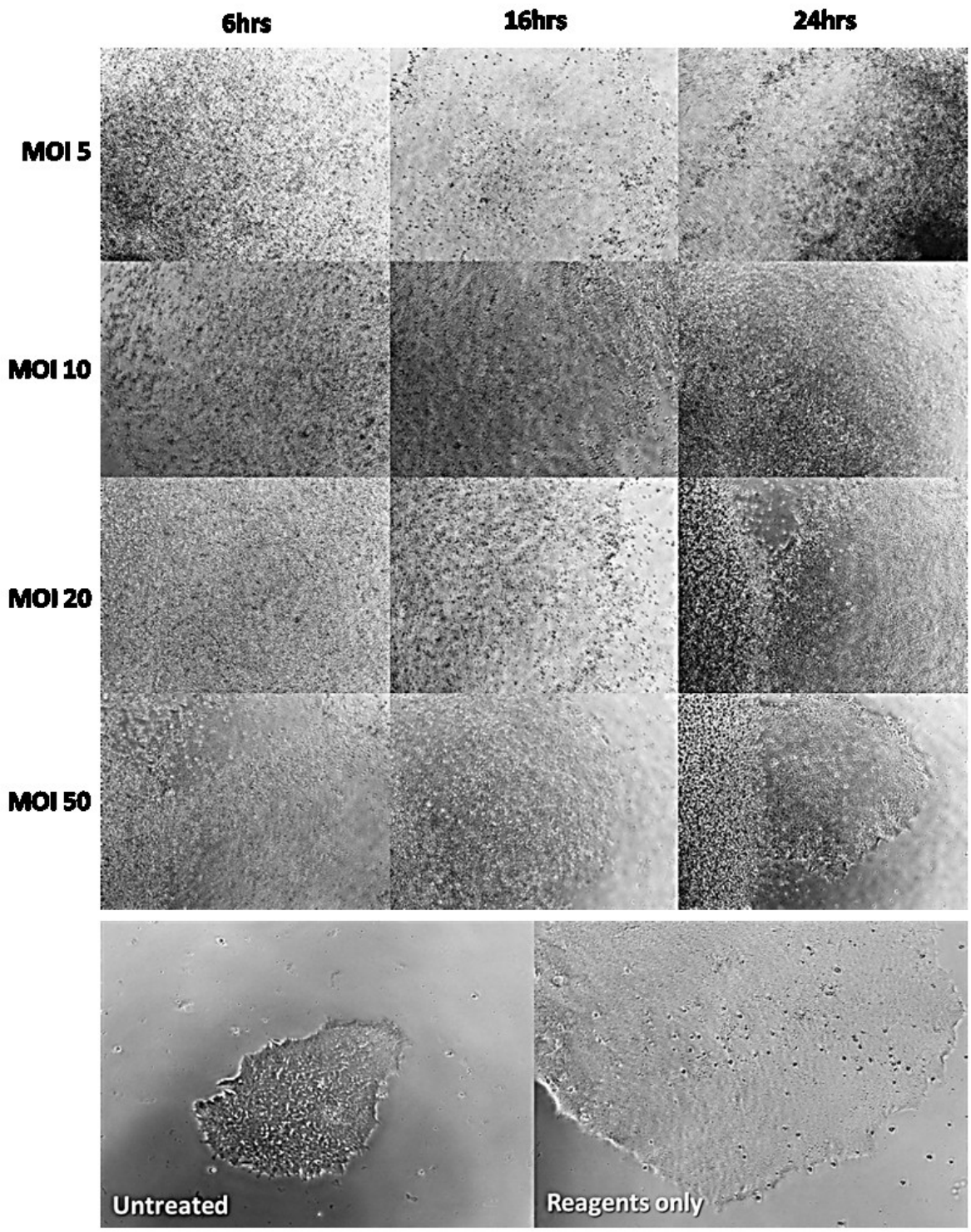

Figure 6.17 Images of H9 ES cells treated with different MOI of lentivirus for different time periods taken 72hrs post-infection. *4x magnification.*

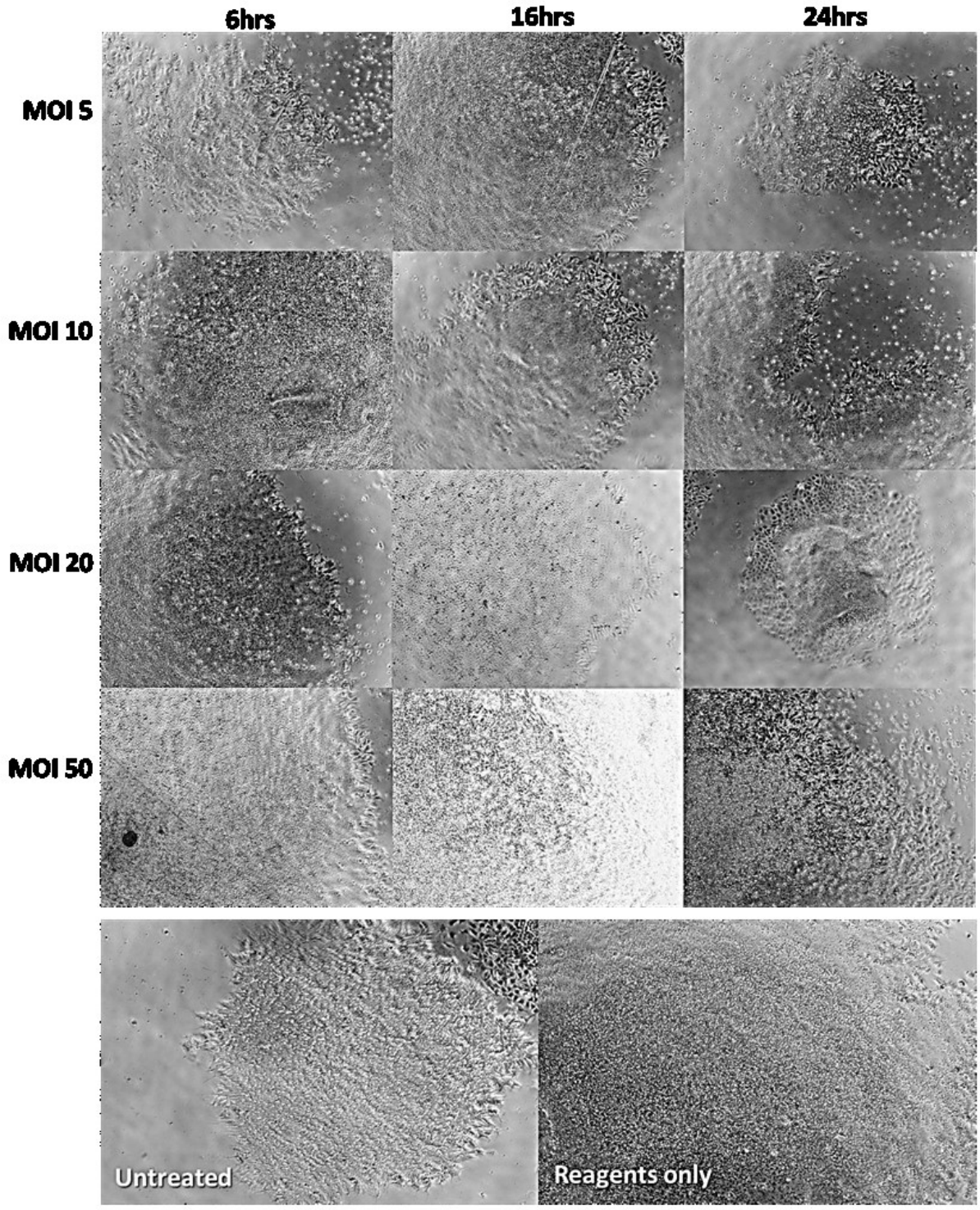

Figure 6.18. Images of Wild7 iPS cells treated with different MOI of lentivirus for different time periods taken 72hrs post-infection. *4x magnification.*

In a further attempt to transduce the stem cells with the lentiviral vectors, a different protocol was used (published by Sigma-Aldrich, available online at http://www.sigmaaldrich.com/life-science/functional-genomics-and-rnai/shrna/learning-center/hesc-transduction.html, accessed 23/02/2014). The cells were passaged as normal in a Matrigel-coated 6-well plate. Prior to infection, the medium on the cells was replaced with mTeSR-1 supplemented with 10µM Y-27632 ROCK inhibitor. The cells were incubated at 37°C for 1hr. Next, the cells were washed twice with DMEM and 1ml of either Accutase™ (to give a single cell suspension) or Dispase (to passage the cells in clumps) was added to the cells, as per the manufacturer's instructions. Once the cells had detached, they were transferred to 15ml tubes and centrifuged at 480 x g for 3min, and the supernatant was removed. The cells were resuspended in 1ml mTeSR-1 supplemented with 4µg/ml Polybrene, and 10µl of lentivirus was added to each tube. The cells were incubated at 37°C for 1 hr, before being washed with mTeSR-1, centrifuged at 480 x g for 3min, resuspended in mTeSR-1 supplemented with 10µM Y-27632 ROCK inhibitor and transferred to a Matrigel-coated 6-well plate. Cell pictures were taken prior to infection, and then at 24hrs, 48hrs and 72hrs post-infection. The amount of cell death was significantly reduced using this protocol (figure 6.19); however, no expression of the fluorescent reporter gene was seen at any time point, demonstrating poor transduction efficiency.

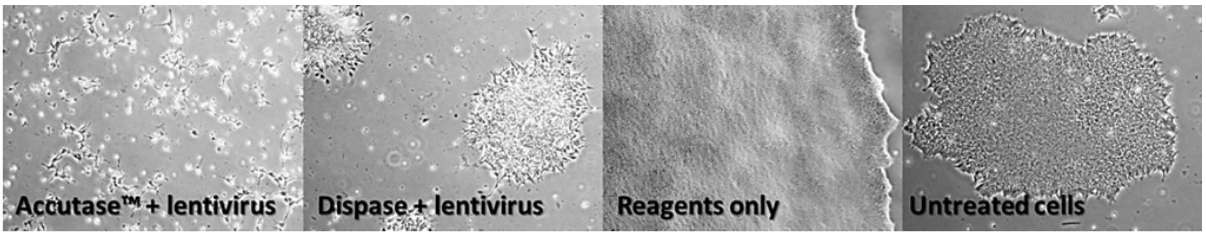

Figure 6.19. Images of cells 72hrs following lentiviral transduction. *4x magnification.*

Unfortunately, due to time constraints, further optimisation of these protocols was not possible, and the functional effects of manipulation of miRNA expression in pluripotent stem cells remain as work to be continued after this PhD.

6.4 Discussion

miR-151a-5p was previously identified using microarray analysis as being upregulated during DE formation, as well as being differentially expressed between ESCs and iPSCs (see chapter 5). The aim of this chapter was to elucidate a role for this miRNA in DE formation through identification of its target genes, and to discover whether manipulation of its expression levels in both ESCs and iPSCs had an effect on differentiation to DE. Although several miRNAs have been previously identified as playing a role in DE formation (Tzur *et al.*, 2008; Hinton *et al.*, 2010; Tsai *et al.*, 2010; Kim *et al.*, 2011), few of these studies went on to elucidate a function for these miRNAs in endoderm formation, and the target genes of many of these miRNAs are still largely unknown.

Identifying miRNA targets is challenging, largely due the limited complementarity between miRNAs and their target mRNAs, and the fact that the interactions between miRNAs and their target genes are still not entirely understood. The lack of high throughput experimental methods for miRNA target identification means there is a reliance on target prediction algorithms. Several independent groups have established computational algorithms designed to predict target genes of miRNAs (reviewed in Rajewsky, 2006). Target prediction software was used to identify potential target genes of both miR-151a-3p and miR-151a-5p, the results of which can be seen in appendix 2. TargetScan, miRDB and Pictar were chosen, as in a comparison of target prediction programs, these three used either alone or in combination had the best trade-off between sensitivity and specificity (Witkos *et al.*, 2011). There was considerably more overlap in the target predictions from TargetScan and miRDB than from PicTar, which did not overlap with the other two algorithms at all for either of these miRNAs. This is likely to be due to the different criteria measured by each algorithm, as TargetScan and miRDB have more similar requirements to each other than PicTar. Interestingly, *Sox17* appeared on the list of predicted targets for miR-151a-5p in the results from both TargetScan and miRDB. This gene is one of the earliest markers of DE (Lewis & Tam, 2006), and is essential for endoderm formation, as mice lacking *Sox17* do not develop gut endoderm from DE (Kanai-Azuma *et al.*, 2002). It is possible that miR-151a-5p has an effect on endoderm formation through regulation of *Sox17* expression, and that the differences in expression of miR-151a-5p between ESCs and iPSCs may be important in DE formation through a pathway involving the regulation of *Sox17* expression.

Although few miRNA-target relationships have been identified in the context of DE formation, Hinton et al (2010) previously identified *Timm8a* as a target of miR-375. This known relationship was used to optimise the experimental conditions of the luciferase assay. Initial experiments demonstrated that although HFF1 cells could be transduced using the lentiviral vectors encoding the miRNA mimic or inhibitor, as demonstrated by their expression of the GFP and mCherry reporter genes, transfection of the 3'UTR plasmid was unsuccessful. For this reason, 293FT cells were used for further experiments, as they were amenable to transduction with the lentiviral vectors encoding the miRNA mimic or inhibitor, as demonstrated by their expression of the GFP and mCherry reporter genes, as well as by qRT-PCR demonstrating upregulation of miRNA expression. In addition, transfection of the 3'UTR plasmid could be carried out at high efficiency, as demonstrated by qRT-PCR and luciferase measurements. Co-transfection of the miR-375 mimic and a luciferase reporter plasmid containing the *Timm8a* 3'UTR (where the majority of miRNA binding occurs) demonstrated that this assay could be used to identify miRNA-gene target binding, as a reduction in luciferase was evident in the presence of the miR-375 mimic.

Subsequently, the luciferase assay was used to determine whether there was a direct interaction between miR-151a-5p and the *Sox17* 3'UTR. However, no such interaction could be identified, as no reduction in luciferase was seen in the presence of the miR-151a-5p mimic. This demonstrates the importance of using experimental approaches to validate predicted target genes of miRNAs, as algorithms used for predicting targets typically predict large numbers of target genes for each miRNA, most of which are not true targets (Thomas *et al.*, 2010). The false positive rate of prediction algorithms has been variously reported as between 24-70% (Thomson *et al.*, 2011).

It is possible that there is an indirect relationship between miR-151a-5p and *Sox17* through a larger regulatory network of miRNAs and genes. This could be investigated further through manipulation of miR-151a-5p expression levels, as this would demonstrate whether the expression levels of *Sox17* are affected by changes in miR-151a-5p expression levels. This experiment would be best carried out in ESCs or iPSCs, as they are known to express both miR-151a-5p and *Sox17*. Alternatively, there may be no relationship between miR-151a-5p and *Sox17*, and miR-151a-5p may mediate its effects on DE formation through another mechanism. One way to investigate this further would be to use microarray analysis to identify changes in gene expression that inversely correlate with experimentally induced

changes in miR-151a-5p expression levels, as this would indicate gene silencing by miRNA binding to its target genes. This would also be useful for identifying target genes not predicted by target prediction algorithms, as many genes identified by overexpression or silencing of miRNAs in combination with microarray analysis were not predicted by target prediction software (Krutzfeldt *et al.*, 2006).

The second aim of this chapter was to investigate the effects of manipulating miR-151a-5p expression in pluripotent stem cells using lentiviral vectors encoding a miRNA mimic and a miRNA inhibitor, to determine whether changes in the expression of miR-151a-5p result in changes in endodermal gene expression, particularly *Sox17*, or altered differentiation efficiency. As miR-151a-5p is differentially expressed between ESCs and iPSCs, it would be interesting to see how manipulating expression of miR-151a-5p affects each cell type. Unfortunately, attempts to transduce human pluripotent stem cells with the lentiviral vectors encoding the miRNA mimic and inhibitor were unsuccessful, demonstrating a need for further optimisation of the experimental conditions, which, due to time constraints, could not be investigated as part of this study.

Several previous studies have attempted to manipulate miRNA expression in order to identify their roles in DE formation. The overexpression of a panel of miRNAs in mouse ESCs resulted in the up-regulation of the DE markers *Sox17* and *Foxa2,* suggesting an increase in differentiation efficiency (Fu et al., 2011). Another study, using human ESCs, attempted to influence endodermal, and particularly hepatic, differentiation using miR-122, which is known to be important in liver development and function. While its overexpression did result in the up-regulation of some hepatic genes, the overall effect was to delay differentiation (Tzur et al., 2008). Hinton *et al.* (2010), again using human ESCs, showed that manipulation of miR-375 resulted in altered expression of *TIMM8A*, although these authors did not observe any change in differentiation efficiency. Similarly, Wei *et al.* (2013) elucidated the temporal expression of several miRNAs known to be involved in pancreatic development, and correlated these with genes known to be expressed at specific stages of differentiation. The latter study showed that manipulation of miR-375 expression in human ESCs resulted in altered expression of *Hnf1β* and *Sox9*, transcription factors important in pancreatic development, but the authors did not demonstrate whether this was a direct interaction between miR-375 and these genes. More recently, miR-375 overexpression has been shown to promote pancreatic endocrine differentiation in human ESCs and iPSCs in the absence of

any extrinsic factors (Lahmy et al., 2013; Lahmy et al, 2014). These results demonstrate the importance of miRNAs in pancreatic differentiation, and demonstrate their potential for improving differentiation protocols.

Chapter 7

Discussion

Contents

Chapter 7: General Discussion

The aim of this study was to identify whether iPSCs are a viable alternative to ESCs for generating β-cells *in vitro* for cell replacement therapy to treat type 1 diabetes. In order to achieve this, new iPSC lines were first generated and characterised. Secondly, the differentiation potential of iPSCs and ESCs to give rise to definitive endoderm (DE), a key early step in pancreatic differentiation, was compared. Finally, the role of miRNAs in differentiation to DE, and differences in miRNA expression between ESCs and iPSCs which may have an effect on the differentiation process, were investigated.

7.1 Generation and Characterisation of iPSCs

The first aim of this study was to generate new iPSC lines, and characterise them to assess their pluripotency. Although several iPSC lines were already available in the lab for use in differentiation experiments, generating and characterising my own iPSC lines allowed me to gain technical expertise and a greater understanding of the reprogramming process. A table showing all the iPSC and ESC lines used in these experiments is included in chapter 3.

Initial reprogramming experiments resulted in the production of cells which closely resembled pluripotent stem cells in terms of their morphology and growth characteristics, and which could be maintained in culture over a number of passages, but which, when further characterised, showed inadequate expression of pluripotency markers. These cells are likely to represent a population of incompletely reprogrammed cells which have been previously described (Meissner *et al*, 2007; Chan *et al*, 2009). One of the reasons for the appearance of these incompletely reprogrammed populations may be due to the reprogramming method used, which relies on delivery of the reprogramming factors in four separate viral vectors. This makes it difficult to control the dose of each reprogramming factor that each individual cell will receive, and while some will receive an adequate dose of all four reprogramming factors and go on to be fully reprogrammed into iPSCs, others may receive an inadequate dose of the reprogramming factors to allow them to completely alter their cell fate to that of a pluripotent cell, but which may be sufficient to allow partial reprogramming. Further complicating matters is the fact that excess levels of OCT4 (Niwa *et al*, 2000) or SOX2 (Kopp *et al*,

2008) are actually detrimental to the maintenance of pluripotency. High levels of C-MYC and KLF4 together with low levels of SOX2 and OCT4 may lead to the generation of rapidly growing "non-iPSCs" which have been transformed rather than reprogrammed (Yamanaka, 2007). Although a separate vector encoding GFP was used in these experiments to monitor efficiency of delivery of the reprogramming factors, this is a crude method that simply ascertains whether transfection is working, and not the dose of each reprogramming factor that the cells receive. However, this method of reprogramming is still one of the most efficient, which is why it was chosen for this study. The presence of incompletely reprogrammed cells show the importance of fully characterising new iPSC lines before they are used in subsequent experiments to ensure the quality of the lines used. Guidelines for the characterisation of iPSCs have now been well-established (Maherali & Hochedlinger, 2008), and are largely based on guidelines first used for the characterisation of ESCs (Brivanlou *et al*, 2003).

Subsequent reprogramming experiments resulted in iPSCs that closely resembled ESCs in terms of their expression of pluripotency genes as assessed by qRT-PCR and immunocytochemistry, their expression of alkaline phosphatase, and their ability to spontaneously differentiate into cells derived from all three embryonic germ layers. iPSCs were also shown to silence the exogenous transgenes used in reprogramming, demonstrating that they have successfully reactivated the network of gene expression necessary for the maintenance of pluripotency. Although several pluripotency genes were assessed using qRT-PCR and immunocytochemistry, a high-throughput method that could be used to screen for expression of a larger number of pluripotency-associated genes would give a more robust characterisation of new iPSC lines and their ability to spontaneously differentiate into lineages from all three germ layers. In addition, the most stringent test of pluripotency available for human pluripotent stem cells is the teratoma assay. However, this is an expensive and time-consuming assay, and there are ethical concerns over the use of animals for a routine test which would need to be carried out for every iPSC line generated. There has also been debate over how useful it actually is, with many pointing out the qualitative nature of this assay, and the fact that it does not accurately assess whether the iPSCs are able to give rise to every cell in the body, nor does it assess the contribution of the iPSCs to the germline. Many mouse iPSC lines are able to give rise to teratomas containing cells derived from all three germ layers, but are not able to contribute to all-iPSC mice. There has been an increased focus in recent years on *in vitro* testing of new iPSC lines: for example, the TaqMan® hPSC Scorecard developed by Life Technologies aims to allow researchers to fully characterise their pluripotent stem cell lines in a single PCR-based assay. In addition, PluriTest (www.pluritest.org), an open access resource, allows researchers to compare their gene expression

data with a large panel of established pluripotent stem cell lines (Muller et al, 2008). These methods are particularly relevant for lines which are generated primarily for research use, where the requirement for pluripotency is less stringent than in cells which could potentially be used in a therapeutic setting. In terms of this study, it was decided that an *in vitro* differentiation assay was sufficient to determine the pluripotency of these cell lines, as the cells were subsequently to be used for investigation of *in vitro* differentiation.

7.2 Differentiation of Pluripotent Stem Cells into Definitive Endoderm

The next stage of the study was to investigate the differentiation efficiency of the ESC and iPSC lines. The first step in the differentiation of pluripotent stem cells into β-cells is the production of definitive endoderm (DE), and this has been shown to be a critical stage in the differentiation process, demonstrating the need for protocols which generate DE at a high efficiency (Courtney et al., 2010). Differences in differentiation efficiency between ESCs and iPSCs at this stage are likely to impact the differentiation efficiency of later stages of the differentiation process. In addition, the full differentiation protocol for the generation of insulin-expressing cells in vitro is a lengthy (17 days) and inefficient process, in contrast to the generation of DE, which takes 3 days and routinely results in a yield of around 70% *Sox17*-expressing cells (D'Amour et al, 2006). For these reasons, comparison of the differentiation propensities of ESCs and iPSCs was compared at the DE stage rather than the insulin-expressing stage of differentiation.

Given the large numbers of protocols published for the differentiation of pluripotent stem cells into DE, the second aim of this part of the study was to evaluate three widely applied protocols to determine which results in the highest yield of DE in our cells. Obtaining a high yield of DE was important for subsequent experiments, as contamination of DE samples with high levels of other cell types could affect the results of further experiments. These experiments demonstrated that the protocol published by D'Amour et al (2005) resulted in the highest yield of DE compared to any of the other protocols tested, or to spontaneously differentiated cells. Although the protocol published by Nostro et al (2011) results in similar upregulation of the *Cxcr4* and *Cer1* genes to the D'Amour protocol, the fact that *Sox17* and *Foxa2* expression are not significantly upregulated suggests that differentiation to DE may be less complete in the cell lines used in these experiments. These results are in contrast to those published by Nostro et al (2011), where they reported that this protocol produced a population of cells containing approximately 70% *Sox17*-expressing cells. Cells treated

with the protocol published by Borowiak et al (2009) fail to upregulate any DE-associated genes at a higher efficiency than spontaneously differentiating cells, suggesting that this protocol is having little or no effect in directing the cell lines used in this study towards an endodermal fate. This may be due to the use of different cell lines in this study than in the Borowiak or Nostro protocols, and the D'Amour protocol may result in a higher yield simply because it is more applicable across stem cell lines. However, other researchers (Tahamtani et al, 2013; Massumi et al, 2014) have also found the Borowiak protocol to be ineffective. As the D'Amour protocol gave the highest efficiency of differentiation for the cell lines used in these experiments, this was used in all subsequent differentiation experiments.

During the course of these investigations, it was noticed that different cell lines had different propensities to form DE, with one ESC line (H1) forming significantly less DE than any other cell line. These differences between cell lines were consistent over several experiments, suggesting that they are due to intrinsic properties of the cell lines, and not random variation in differentiation efficiency due to other factors. However, a limitation of this study is that the cell lines used are not all cultured in the same way. Initial experiments were carried out using two iPSC lines generated at NIBSC, MRC5I and MRC9G, and a well-characterised ESC line (H9). However, the iPSC lines are grown on iSNL feeders, and the H9 line is grown on Matrigel™. For this reason, the H1 cell line was included in differentiation experiments and microarray analysis, as this is an ESC line grown on iSNL feeders, and was therefore expected to represent an ESC line grown in a comparable culture format to the available iPSC lines. However, data from my own results and others (Bock et al, 2011; Nostro et al, 2011) subsequently indicated that this line has a low propensity for forming DE. The H7 ESC line, also grown on Matrigel™, was also included in differentiation comparisons to provide additional supportive data for ESC lines grown on Matrigel™. To investigate whether the low differentiation efficiency observed in the H1 cell line was a result of being grown on iSNL feeders, compared to the other ESC lines which were all grown on Matrigel™, I tested the differentiation efficiency of the H9 ESC line grown on both feeders and Matrigel™. I found no significant differences in differentiation efficiency between H9 cells cultured on feeders vs. Matrigel™, although there was some slight variability in gene expression between them. There was significant upregulation of DE genes in the treated samples compared to the undifferentiated samples, demonstrating that the H9 cell line can differentiate into DE at high efficiencies irrespective of the culture conditions used.

The subject of whether there are any systematic differences between ESCs and iPSCs has been hotly debated. Disparities were first observed in the differentiation abilities of ESCs and iPSCs, in both teratoma formation and *in vitro* differentiation assays. Just like ESC lines, iPSC lines have considerable heterogeneity and show different propensities for differentiation (Abeyta *et al*, 2004; Martinez *et al*, 2012; Osafune *et al*, 2008), with some studies showing that iPSCs have lower efficiencies of differentiation into haematopoietic, neuroepithelial, neuronal and cardiac lineages compared to ESCs (Feng *et al*, 2010; Hu *et al*, 2010; Narsinh *et al*, 2011). While some researchers argued that these results show that iPSCs have an intrinsically lower capacity to differentiate into certain lineages *in vitro* than ESCs, others have argued that the cell of origin might have an effect on the differentiation propensity of the iPSCs. A study by Boulting *et al* (2011) found that, although there was variation between cell lines in their ability to differentiate, iPSCs were as good as ESCs at generating motor neurones, and that the differentiated cells produced were functionally equivalent. The differences in differentiation propensities could not be correlated with pluripotency markers, karyotype, reprogramming method, or transgene expression, and although some donor-specific differences in gene expression were noticed, these were not statistically significant. A more recent large-scale study compared the propensity of 49 iPSC lines and 10 ESC lines to differentiate into a neural lineage, and found that 7 iPSC lines showed significantly reduced differentiation capacity, which corresponded to differences in the expression levels of several genes. However, the majority of iPSC lines could not be distinguished from ESC lines (Koyanagi-Aoi *et al*, 2013). In contrast, Kwon *et al* (2014) found that iPSCs differentiated into dopaminergic neurons at a higher efficiency than ESCs. These studies have led to questions about whether iPSCs are indeed identical to ESCs, and if they are not, what the functional implications of this might be for the use of iPSCs for *in vitro* applications such as disease modelling and drug screening, and *in vivo* use for cell replacement therapy.

My results showed that the two iPSC lines used in these experiments did not show a lower propensity for differentiating into DE than the 3 ESC lines. In fact, the lowest differentiation efficiency was seen in the H1 cell line. However, there was variability in differentiation efficiency between the cell lines used. These results are similar to a study carried out in mice, which compared 4 iPSC lines and 2 ESC lines, and found considerable variability between cell lines in their propensity to form DE, although no overall differences between ESCs and iPSCs. Although this study found significant differential expression of BMP4 during differentiation, this did not appear to be associated with any reduction in differentiation propensity (Christodoulou et al, 2011). This

suggests that although differences may exist between ESCs and iPSCs, they may not have a functional impact on the cells.

7.3 Investigation of miRNA Expression in DE Formation

MicroRNAs are now known to be important in controlling pluripotency and differentiation of ESCs and iPSCs, and there have been suggestions that there are differences in miRNA expression between these cell types which may have an impact on their ability to differentiate into DE. To investigate this further, miRNA expression patterns in ESCs and iPSCs were investigated at the undifferentiated and DE stage.

There are three main techniques used for profiling miRNA expression: microarray, high-throughput RNA sequencing, and qRT-PCR, each with their own advantages and disadvantages. Microarray analysis has the advantage of being a high-throughput method for profiling miRNA expression, allowing investigation of thousands of miRNAs in parallel. It is also relatively inexpensive. However, limitations of this approach include the fact that the labelling reaction can result in the addition of a large and variable number of nucleotides, and other RNA species may also be labelled, contributing to background noise, and may cross-hybridise with miRNA probes. The variation in the CG content of miRNAs can lead to different annealing temperatures for different miRNAs. To avoid this problem, LNA (locked nucleic acids) can be incorporated into the arrays, as they have an increased affinity for complementary miRNA sequences. In addition, a limited range of quantification, imperfect specificity for some closely related miRNAs, and the inability to absolutely quantify miRNA means that this technique is most suited to carrying out comparisons between two states e.g. differentiated and undifferentiated stem cells. Due to the high chance of false positives occurring during microarray analysis, the results need to be validated by qRT-PCR, as this the most reliable method of analysing miRNA expression. In addition, qRT-PCR is also the cheapest, although the large numbers of miRNAs now known to exist means that this technique is more suited to validation than discovery.

For absolute quantification of miRNAs in a sample, next generation RNA sequencing can be used. This technique has the clear advantage of being the only method allowing discovery of new miRNAs and post-transcriptional modifications of miRNAs. In addition, this method is not adversely affected by variability in melting temperatures, co-expression of nearly identical miRNA family members, or

post-transcriptional modifications. The disadvantages of this method include the high cost; the requirement for high performance bioinformatic infrastructure for data analysis; and the fact that the RNA ligation and PCR amplification steps of sample preparation can lead to bias. Studies comparing microarray and RNA sequencing for analysis of miRNA expression have found that for quantification of miRNAs in a sample, microarray analysis surpasses next generation sequencing in terms of specificity and sensitivity (Willenbrock et al, 2009). For this reason, as well as the cost of the experiments, microarray analysis was chosen to analyse the expression of miRNAs during differentiation to DE. However, it would be interesting to analyse the same samples using RNA sequencing to see if any novel miRNAs could be identified in these samples, and to compare the results of both microarray and RNA sequencing.

Although the later stages of pancreatic development have been reasonably well-characterised in terms of miRNA expression, there have been relatively few studies focussing on the DE stage. Microarray analysis was carried out to identify miRNAs differentially expressed between undifferentiated and differentiated cells, and which may therefore play a role in DE formation. This comparison was carried out for ESCs and iPSCs, as well as for individual cell lines. Although there are some differences in the miRNAs which are up- or down-regulated in these comparisons, several miRNAs show consistency between cell lines, most notably miR-375 and miR-708-5p, which are consistently and strongly upregulated in differentiating cells. These microarray results were validated using qRT-PCR, which again confirmed the strong upregulation of miR-375 and miR-708-5p, suggesting a role for these miRNAs in DE formation. Both of these miRNAs have been previously reported to be upregulated in DE formation (Tzur et al, 2008; Hinton et al, 2010; Kim et al, 2011), although their exact mechanism of action has not been fully elucidated. One study identified TIMM8A as a target of miR-375 but did not elucidate the role of this pathway in DE formation (Hinton et al, 2010). There was less consistency in the miRNAs predicted to be downregulated in DE formation by the microarray, and this was reflected in the PCR validation, as only one of the ten miRNAs investigated (miR-516b) was found to be significantly downregulated.

A better understanding of the role of miRNAs in DE formation may prove useful both in terms of characterisation and in manipulation of cell fate. In addition to analysis of gene expression, analysis of miRNA expression may also be helpful when characterising the differentiated progeny of stem cells. The expression of miR-375 and miR-708 in the cell lines tested in this study mirrored the expression of genes routinely used as markers for this germ layer, such as *Sox17*, with the highest expression of these miRNAs in the cell lines previously noted to give rise to DE at higher efficiency,

and with the lowest expression in the H1 cell line, which has a poor efficiency of differentiation into DE.

7.4 Investigation of Differential miRNA Expression between ESCs and iPSCs

There have been several studies which have investigated differences in miRNA expression between ESCs and iPSCs (Chin et al, 2009; Wilson et al, 2009; Neveu et al, 2010; Razak et al., 2013), with some studies finding differences between the two cell types, and others finding no differences (Koyanagi-Aoi *et al*, 2013). Although there has been some overlap in the miRNAs identified as differentially expressed between ESCs or iPSCs, so far there is little evidence for consistent differences between ESCs and iPSCs. In addition, most of these studies focussed on miRNA expression only in undifferentiated cells. For these reasons, the functional significance of differences between ESCs and iPSCs is still unclear, although as miRNAs are now known to play a significant role in controlling differentiation, differences in miRNA expression between ESCs and iPSCs may affect the cells' ability to differentiate into a pancreatic lineage, rendering them less suitable for use in cell replacement therapy to treat T1DM. With this in mind, microarray analysis was used to investigate differences between ESCs and iPSCs with the aim of identifying any differences in miRNA expression which may have a functional impact on the cells' ability to differentiate into a pancreatic lineage at the DE stage.

Initial comparison of undifferentiated ESCs and iPSCs did not identify any miRNAs differentially expressed between the two cell types. However, when miRNA expression was compared in DE derived from ESCs and iPSCs, several miRNAs were identified as being differentially expressed between the two cell types. In addition, several of these differentially expressed miRNAs, including miR-151a-3p, miR-151a-5p, miR-26b-5p, miR-27b-3p, miR-30b-5p, miR-378a-3p, and miR-4530, were also identified as potentially playing a role in DE formation, as they were also significantly up- or down-regulated in differentiated cells compared to undifferentiated cells. Validation of these results by qRT-PCR identified miR-151a-3p and miR-151a-3p as being significantly differently expressed between ESCs and iPSCs. These two miRNAs are derived from the same pre-miRNA stem loop, and are therefore likely to be co-regulated and co-transcribed, although as their sequence is different they are likely to have different target genes and therefore different mechanisms of action.

Another factor identified by the microarray as having an effect on miRNA expression is the culture conditions that the cells are grown in. Initial experiments were carried out using two iPSC lines generated at NIBSC (MRC5I and MRC9G) and a well-characterised ESC line (H9). However, the iPSC lines are grown on iSNL feeders, and the H9 line is grown on Matrigel™. For this reason, the H1 cell line was also included in differentiation experiments and microarray analysis, as this is an ESC line grown on iSNL feeders, and was therefore expected to represent an ESC line grown in a comparable culture format to the available iPSC lines. However, subsequent experiments demonstrated that this cell line was poor at differentiating into DE. At the time, these were the only cell lines available with which to carry out the analysis, and an obvious improvement to this part of the study would be to include an ESC line cultured on iSNL feeders which was not similarly limited in the efficiency with which it was able to differentiate into DE. During the microarray analysis, it quickly became apparent that the difference in culture conditions may have an effect on miRNA expression, as the H9 cell line was clearly segregated from the other cell lines in the principal component analysis and hierarchical clustering. This was also evident from the comparisons of individual cell lines. When undifferentiated H1 samples were compared with undifferentiated iPSC samples, only 5 miRNAs were significantly differentially expressed. This suggests that these samples are highly similar in terms of their miRNA expression. However, when undifferentiated H9 samples were compared with undifferentiated iPSC samples, 186 miRNAs were differentially expressed. A similar phenomenon was noticed when comparing differentiated samples. When differentiated H9 were compared with differentiated iPSCs, 154 miRNAs were differentially expressed. When differentiated H1 were compared with differentiated iPSCs, only 28 miRNAs were significantly differentially expressed, despite the fact that the H1 cell line has a lower propensity to form DE than the two iPSC lines. In order to ensure that we were not missing any candidate miRNAs which may be differentially expressed between ESCs and iPSCs, the expression of several other miRNAs which were identified as being significantly differentially expressed in iPSCs when compared to the ESC cell lines individually (i.e. H1 or H9) were further investigated further. However, when these miRNAs were assessed using qRT-PCR in five pluripotent stem cell lines, no significant differences in miRNA expression were seen between any of the cell lines tested. While it remains possible that there are differences in the expression of these additional miRNAs due to the different culture conditions, as well as to cell line-specific differences, they were not identified in this study, and these miRNAs were not investigated further.

These differences in miRNA expression in the H9 cell line were not predicted either by the pluripotency characterisation data (chapter 3) or by the differentiation data (chapter 4), as the H9 cell

line showed very similar expression of pluripotency and DE genes to the other cell lines analysed. For this reason, another ESC line (H7) grown on Matrigel™ was included in the PCR validation of the microarray data. These analyses identified two miRNAs that appeared to show differences in their expression depending on the culture method: miR-4451and miR-4732 were upregulated in cells grown on Matrigel (H7 and H9) and downregulated in cells grown on iSNL feeders (MRC5I, MRC9G and H1). To further investigate the possibility that differences in culture conditions had an effect on miRNA expression, a further differentiation experiment was carried out using H9 cells grown on both iSNL feeders and Matrigel™. The expression of miRNAs previously demonstrated by qRT-PCR validation of the microarray results to be a) upregulated in DE formation; b) downregulated in DE formation; or c) differentially expressed between cells grown on iSNL feeders and cells grown on Matrigel™ were investigated. However, there were no significant differences in expression of any of these miRNAs when H9 cells were grown using both culture methods.

The differences observed in miRNA expression in these analyses may be due to a number of factors, including cell type, individual cell lines, culture conditions, and differentiation status. The purpose of this experiment was not to draw any definitive conclusions about differences in miRNA expression between individual cell lines or between different culture conditions, but instead to try to identify miRNAs which may be involved in DE formation, and which are differentially expressed between ESCs and iPSCs. A more complete analysis of global patterns of miRNA expression would require access to a larger number of different ESC and iPSC lines. Furthermore, to elucidate the effects of different culture conditions on miRNA expression, it would be necessary to have access to these cell lines grown on both feeders and Matrigel™, which in many cases, because of specific cell lines requiring certain culture conditions, would not always be possible. For example, despite repeated attempts, we were not able to grow the H1 cell line on Matrigel™, as transfer from iSNL feeders resulted in a failure to grow and loss of pluripotency gene expression.

However, despite the dissimilarities between the cell lines, some miRNAs, including miR-151a-3p, miR-151a-5p and miR-27b, previously identified as potentially playing a role in DE formation, were also identified as being differentially expressed in both the overall comparison of iPSC vs. ESC, as well as the individual comparison of H1 vs. iPSC and H9 vs. iPSC. These miRNAs are more likely to represent conserved differences between ESCs and iPSCs which are not specific to certain cell lines, although this is difficult to determine in an analysis of only four cell lines. The strongest contrast was seen in expression of miR-151a-3p and miR-151a-5p, which are upregulated upon differentiation in ESCs, and downregulated in iPSCs. miR-27b did show differential expression

between the two cell types, being more strongly upregulated in ESCs than in iPSCs, but this difference was not statistically significant in the qRT-PCR validation data.

7.5 Investigation of the Function of miRNAs in DE Formation

miR-375 is known to be essential both for embryonic pancreatic development and mature β-cell function (reviewed in Francis et al., 2014). It has also been reported to be upregulated in DE formation (Tzur et al., 2008; Hinton et al., 2010; Tsai et al., 2010; Kim et al., 2011; Porciuncula et al., 2013) and was also found to be strongly upregulated by my own microarray analysis, but its exact function remains unclear. Hinton et al (2010) demonstrated a direct relationship between miR-375 and its target *Timm8a*, although they did not elucidate a role for this pathway in DE formation. As this relationship has already been established, it was thought that this would be a useful way of validating the methods used to study miRNA function, and indeed, the relationship between miR-375 and its target gene *Timm8a* was confirmed using a luciferase assay.

In addition to identifying miRNAs important in DE formation, microarray analysis and subsequent qRT-PCR was also used to identify miRNAs that were differentially expressed between ESCs and iPSCs, and which may therefore have an impact on the cells' ability to differentiate into DE. Two miRNAs fulfilled this criteria, miR-151a-3p and miR-151a-5p. These miRNAs were shown to be upregulated in DE formation in ESCs, but downregulated in iPSCs. These miRNAs have not been previously implicated in DE formation, although their expression has been noted in foetal liver (Fu et al, 2005) and their function remains unknown, although a few gene targets have been confirmed (mirtarbase.mbc.nctu.edu.tw/index.php, accessed 03/12/2013), although not in the context of pancreatic differentiation. For this reason, these miRNAs were identified for further study.

First, target prediction algorithms were used to identify potential gene targets of miR-151a-3p and miR-151a-5p. The results are included in appendix 2. TargetScan, miRanda and Pictar were chosen, as in a comparison of target prediction programs, these three used either alone or in combination had the best trade-off between sensitivity and specificity (Witkos et al., 2011). Interestingly, *Sox17* appeared on the list of predicted targets for miR-151a-5p in the results from both TargetScan and miRanda. This gene is one of the earliest markers of DE (Lewis & Tam, 2006), and is essential for endoderm formation, as mice lacking *Sox17* do not develop gut endoderm from DE (Kanai-Azuma et al., 2002). A luciferase assay was used to determine whether there is a direct relationship between

miR-151a-5p and *Sox17*, but no binding of this miRNA to the 3'UTR of the *Sox17* mRNA could be detected. Identifying miRNA targets is challenging, largely due the limited complementarity between miRNAs and their target mRNAs, and the fact that the interactions between miRNAs and their target genes are still not entirely understood. This demonstrates the importance of using experimental approaches to validate predicted target genes of miRNAs, as algorithms used for predicting targets typically predict hundreds if not thousands of target genes for each miRNA, most of which are not true targets (Thomas et al., 2010). The false positive rate of prediction algorithms has been variously reported as between 24-70% (Thomson et al., 2011). It is possible that there is an indirect relationship between miR-151a-5p and *Sox17* through a larger regulatory network of miRNAs and genes. This could be investigated further through manipulation of miR-151a-5p expression levels, as this would demonstrate whether the expression levels of *Sox17* are affected by changes in miR-151a-5p expression levels. This experiment would be best carried out in ESCs or iPSCs, as they are known to express both miR-151a-5p and *Sox17*. Alternatively, there may be no relationship between miR-151a-5p and *Sox17*, and miR-151a-5p may mediate its effects on DE formation through another mechanism. One way to investigate this further would be to use microarray analysis to identify changes in gene expression that inversely correlate with changes in miR-151a-5p expression levels, as this would demonstrate gene silencing by miRNA binding to its target genes. This would also be useful for identifying target genes not predicted by target prediction algorithms, as many genes identified by overexpression or silencing of miRNAs in combination with microarray analysis were not predicted by target prediction software, which can largely be explained by the application of stringent criteria regarding the evolutionary conservation of putative binding sites (Krutzfeldt et al., 2006).

The second aim of this chapter was to investigate the effects of manipulating miR-151a-5p expression in pluripotent stem cells using lentiviral vectors encoding a miRNA mimic and a miRNA inhibitor, to determine whether changes in the expression of miR-151a-5p result in changes in endodermal gene expression, particularly *Sox17*, or altered differentiation efficiency. As miR-151a-5p is differentially expressed between ESCs and iPSCs, it would be interesting to see how manipulating expression of miR-151a-5p affects each cell type. However, despite repeated attempts to optimise experimental conditions, successful lentiviral transduction of the cells could not be achieved within the allocated time frame for inclusion in this thesis. Further optimisation of these experiments would include testing lentiviral vectors with different promoters driving gene expression, as a study has demonstrated that different promoters result in differential transgene expression in human ESCs, with some expression from some promoters being suppressed more than

others (Xia et al, 2007). A promising alternative may be the use of lentiviral vectors whose expression is driven by the EF1α promoter, as this has been shown to drive transgene expression more efficiently than the CMV promoter that I used in these experiments (Kim et al, 2007). It is also likely that the experimental conditions, including time of infection and multiplicity of infection, will need to be further optimised.

7.6 Future Perspectives

In recent years, our understanding of the role that miRNAs play in pluripotency and differentiation has greatly increased, and this is a field of research that is likely to continue to expand rapidly in coming years. Increased knowledge of the role of miRNAs in differentiation to different lineages may improve differentiation protocols, as each stage of differentiation is characterised by a unique miRNA signature which can be characterised in the same way that gene expression is currently used for characterisation. In addition, recent studies have demonstrated the ability of miRNAs to influence cell fate.

One of the clearest indications of the power of miRNAs to influence cell fate is through the demonstration that pluripotency-associated miRNAs are able to reprogram somatic cells to iPSCs without the need for additional reprogramming factors (Anokye-Danso et al., 2011). However, miRNAs can also be used to induce differentiation. One of the earliest studies to investigate these effects demonstrated that the overexpression of miRNAs associated with a specific lineage, e.g. miR-1 (muscle) or miR-124 (neurons) in HeLa cells resulted in a shift in gene expression towards those specific lineages (Lim et al., 2005). Since then, there have been a number of similar studies demonstrating, for example, miR-124-mediated conversion of fibroblasts into functional neurons (Ambasudhan et al., 2011; Yoo et al., 2011; Xue et al., 2013) and miRNA-mediated conversion of cardiac fibroblasts into cardiomyocytes, both *in vivo* and *in vitro*, using a combination of the miRNAs miR-1, miR-133, miR-208, and miR-499 (Jayawardena et al., 2012).

To date, there have been few examples of the use of miRNAs to drive or influence pancreatic development, which perhaps reflects the current incomplete understanding of the role of miRNAs in this process. However, miRNAs have been used to drive differentiation into a pancreatic lineage in the absence of any other factors in both ESCs (Lahmy et al, 2013) and iPSCs (Lahmy et al, 2014),

highlighting the potentially important role that miRNAs may play in reaching the ultimate goal of a β-cell replacement therapy for type 1 diabetes.

The use of miRNAs to influence cell fate is appealing for a number of reasons. Firstly, they have the potential to directly and immediately alter gene expression, which may reduce the time required for *in vitro* differentiation of cells and thereby lead to increased efficiency and reduced costs for the manufacture of cell therapy products. Secondly, they can be transiently expressed in a cell and so do not permanently integrate into the cellular genome, which is appealing from a safety perspective. In addition, miRNAs are inherently less variable than the recombinant growth factors commonly used in differentiation protocols. As a result, their use has the potential to reduce the variability currently observed in the differentiation of pluripotent stem cells for cell therapy applications, including the generation of pancreatic β-cells for a therapy for T1DM. However, if the full potential of miRNAs to modulate cellular differentiation is to be realised, a better understanding of the specific role of each miRNA is required. With this in mind, it is likely that the continued development of powerful, high-throughput technologies, such as next generation sequencing and related technologies for identifying miRNAs bound to their targets, such as HTS-CLIP (high-throughput sequencing of RNA isolated by cross-linking immunoprecipitation), are likely to have a significant impact over the next few years.

The potential of miRNAs to alter entire gene networks which are dysregulated in certain disease states also makes them appealing targets for the development of new therapeutics. Based on lessons learned from siRNA technologies, very potent chemically-modified oligonucleotides to repress miRNA expression, as well as synthetic miRNA mimics, are being developed. Despite the short time since the discovery of miRNAs in humans (Pasquinelli et al., 2000), several miRNA therapeutic products have already entered clinical trials (reviewed in van Rooij et al, 2012). Many miRNA therapeutics are currently focussed on the treatment of cancer. However, numerous miRNAs have been identified as playing a role in diabetes and its related complications (reviewed in Kumar et al., 2012), and these may provide targets for the future development of therapeutic miRNAs. At present, however, strategies to achieve cell type-specific delivery of miRNAs into the β-cell are lacking, though the identification of β-cell surface-enriched markers (Stutzer et al. 2012) may provide the basis of new selective delivery approaches in the future.

Finally, miRNAs may also have therapeutic potential when combined with gene therapy. Proof-of-principle has recently been demonstrated in a recent study by Jayawardena et al. (2012), in which delivery of cardiomyocyte-associated miRNAs into ischemic mouse myocardium resulted in the direct conversion of cardiac fibroblasts to cardiomyocytes *in situ*. While this study did not address

any improvement in cardiac function, it does demonstrate the potential of miRNAs to alter cell fate *in vivo*, which may prove a useful therapeutic tool in the future.

7.7 Concluding Remarks

This study demonstrates that iPSC lines are able to form definitive endoderm as well as ESC lines, but that there is variability in efficiency between cell lines. I have investigated the role of miRNAs in differentiation to the definitive endoderm stage, and identified a number of miRNAs as being up- or down-regulated during this process. In addition, I have shown that some of these miRNAs are also differentially expressed between ESCs and iPSCs, and gone on to further investigate the role of one of these, miR-151a-5p. I have confirmed that there is no direct relationship between this miRNA and one of its putative target genes, *Sox17*, demonstrating that this is not the mechanism through which this miRNA is involved in DE formation. The true targets of this miRNA and their role in DE formation have yet to be elucidated. In addition, the functional impact of these differences in miRNA expression on the differentiation of iPSCs and ESCs into a pancreatic lineage requires further investigation. A previous study which identified differences in BMP4 expression levels between ESCs and iPSCs during pancreatic differentiation found that these differences did not appear to be associated with any reduction in differentiation propensity (Christodoulou et al, 2011). This suggests that although differences may exist between ESCs and iPSCs, they may not adversely affect the function of the cells, suggesting that iPSCs may be considered an alternative to ESCs for cell replacement therapy.

It is becoming clear that the heterogeneity and behaviour of stem cells is more complex than had previously been thought, and that iPSCs and ESCs are neither identical nor distinct populations; instead, they seem to overlap, with greater variability within each population than is observed between the populations. This variability is likely to have functional implications for the cells. It may be more useful to consider each cell line in terms of its quality and utility, in order to choose the best cell line for the desired application. With a view to this end, some researchers have produced a 'scorecard' to evaluate the character of both iPSCs and ESCs, and to predict the quality and utility of any pluripotent cell in a high-throughput manner: a large-scale study by Bock *et al* (2011) used gene expression profiling, DNA methylation mapping, and a high-throughput quantitative differentiation assay to establish genome-wide reference maps for patterns of gene expression and DNA

methylation in order to provide a baseline against which comparisons of new ES and iPS cell lines can be made in order to assess their quality and utility.

However, even with a better understanding of the differences between ESCs and iPSCs, and the functional consequences of these differences, there are still numerous hurdles to be overcome before pluripotent stem cells can offer a cure for T1DM. These include the source of the cells used, whether ESC or iPSC. Recently, human ESCs have been generated through somatic cell nuclear transfer (SCNT), providing an alternative source of pluripotent stem cells that does not require the use of fertilised human embryos, although it does still require the use of human oocytes (Tachibana et al, 2014). It is likely that these cells will provide many of the advantages currently associated with iPSCs, including the ability to derive them from donors with known phenotypes, such as specific diseases, as well as providing a source of patient-matched pluripotent stem cells. In addition, these cells could be useful for treating mitochondrial diseases through cell replacement therapy, as the mitochondria in SCNT-ESCs are derived from the donor oocyte. However, careful evaluation of these cells in comparison with traditional ESCs and iPSCs is needed, as iPSCs have demonstrated how different two types of pluripotent stem cell can be. The use of iPSCs for cell replacement therapy requires careful consideration of the cell source and method used for reprogramming (see section 1.8), and the cell lines will need to be derived in a GMP-compliant manner.

There are also issues facing the differentiation protocols used to direct the pluripotent stem cells towards a desired cell fate *in vitro*. While some cell types have proved relatively easy to generate *in vitro*, others, such as pancreatic β-cells, have proved more difficult, with problems of inefficiency of differentiation and limited function in comparison to their *in vivo* counterparts. However, this is the topic of much current research, with steps being made towards a full understanding of the differentiation signals. In addition, the use of small molecules in differentiation has the potential to increase efficiency and reproducibility, and steps have been made towards GMP-compliant culture, expansion and differentiation of cells *in vitro*. Furthermore, it has been shown that pancreatic progenitors derived from pluripotent stem cells undergo further maturation following transplantation, and these cells will form the basis of clinical trials being carried out by Viacyte for the treatment of T1DM (viacyte.com/press-releases/jdrf-to-provide-additional-support-for-upcoming-clinical-trial-of-viacytes-encapsulated-cell-therapy-for-type-1-diabetes, accessed 16/04/2014).

Finally, there are concerns over the safety of transplanted cells derived from pluripotent stem cells. Undifferentiated cells remaining in the graft may result in teratomas, so efforts have been made to exclude these cells (see section 1.8), although a recent study found that even autologous

undifferentiated ESCs were efficiently rejected following transplantation through natural killer cell-mediated killing, due to their low expression of MHC class I molecules (Perez-Cunningham et al, 2014). However, a recent study in rhesus macaques demonstrated that transplantation of undifferentiated iPSCs did result in teratoma formation, although only when administered in high doses, and there was some evidence of an inflammatory immune response (Hong et al, 2014). Another safety concern is the risk of immunogenicity, as recent research has shown that even autologous iPSCs may be subject to immune rejection (Zhao et al, 2011), although other studies have since cast doubt on these findings (Araki et al, 2013; Guha et al, 2013), including one very recent study in a non-human primate model, the first of this type of study not carried out in mice (Hong et al, 2014). However, due to the time it takes to derive and differentiate a patient's own cells for cell replacement therapy, and the expense of making these cells, it may be more appropriate to turn to banking of iPSCs. With the diversity of starting material available from blood donors, it is considerably easier to generate cell lines of a known genetic background that would provide a beneficial immunological match for most of the population, although it is unlikely that even the best possible match from an iPSC bank will be able to entirely escape immune recognition following transplant, making it likely that some form of immunosuppressive treatment may be needed. In addition, the autoimmune nature of T1DM means that if the transplanted cells are not protected in some way from immune attack, perhaps through immune modulation or encapsulation (see section 1.8), the graft will fail.

Bibliography

Aasen, T., Raya, A., Barrero, M.J., Garreta, E., Consiglio, A., Gonzalez, F., Vassena, R., Bilic, J., Pekarik, V., Tiscornia, G., Edel, M., Boue, S., & Izpisua Belmonte, J.C. 2008. Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nat.Biotechnol.*, 26, (11) 1276-1284 available from: PM:18931654

Abeyta, M.J., Clark, A.T., Rodriguez, R.T., Bodnar, M.S., Pera, R.A., & Firpo, M.T. 2004. Unique gene expression signatures of independently-derived human embryonic stem cell lines. *Hum.Mol.Genet.*, 13, (6) 601-608 available from: PM:14749348

Ahlgren, U., Jonsson, J., & Edlund, H. 1996. The morphogenesis of the pancreatic mesenchyme is uncoupled from that of the pancreatic epithelium in IPF1/PDX1-deficient mice. *Development*, 122, (5) 1409-1416 available from: PM:8625829

Ahlgren, U., Pfaff, S.L., Jessell, T.M., Edlund, T., & Edlund, H. 1997. Independent requirement for ISL1 in formation of pancreatic mesenchyme and islet cells. *Nature*, 385, (6613) 257-260 available from: PM:9000074

Akinci, E., Banga, A., Greder, L.V., Dutton, J.R., & Slack, J.M. 2012. Reprogramming of pancreatic exocrine cells towards a beta (beta) cell character using Pdx1, Ngn3 and MafA. *Biochem.J*, 442, (3) 539-550 available from: PM:22150363

Alipio, Z., Liao, W., Roemer, E.J., Waner, M., Fink, L.M., Ward, D.C., & Ma, Y. 2010. Reversal of hyperglycemia in diabetic mouse models using induced-pluripotent stem (iPS)-derived pancreatic beta-like cells. *Proc.Natl.Acad.Sci.U.S.A*, 107, (30) 13426-13431 available from: PM:20616080

Aly, H., Rohatgi, N., Marshall, C.A., Grossenheider, T.C., Miyoshi, H., Stappenbeck, T.S., Matkovich, S.J., & McDaniel, M.L. 2013. A novel strategy to increase the proliferative potential of adult human beta-cells while maintaining their differentiated phenotype. *PLoS.One.*, 8, (6) e66131 available from: PM:23776620

Ambasudhan, R., Talantova, M., Coleman, R., Yuan, X., Zhu, S., Lipton, S.A., & Ding, S. 2011. Direct reprogramming of adult human fibroblasts to functional neurons under defined conditions. *Cell Stem Cell*, 9, (2) 113-118 available from: PM:21802386

Amit, M., Laevsky, I., Miropolsky, Y., Shariki, K., Peri, M., & Itskovitz-Eldor, J. 2011. Dynamic suspension culture for scalable expansion of undifferentiated human pluripotent stem cells. *Nat.Protoc.* , 6, (5) 572-579 available from: PM:21527915

Andang, M., Moliner, A., Doege, C.A., Ibanez, C.F., & Ernfors, P. 2008. Optimized mouse ES cell culture system by suspension growth in a fully defined medium. *Nat.Protoc.*, 3, (6) 1013-1017 available from: PM:18536648

Anderson, M.S. & Bluestone, J.A. 2005. The NOD mouse: a model of immune dysregulation. *Annu.Rev.Immunol.*, 23, 447-485 available from: PM:15771578

Ang, S.L. & Rossant, J. 1994. HNF-3 beta is essential for node and notochord formation in mouse development. *Cell*, 78, (4) 561-574 available from: PM:8069909

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

Aoi, T., Yae, K., Nakagawa, M., Ichisaka, T., Okita, K., Takahashi, K., Chiba, T., & Yamanaka, S. 2008. Generation of pluripotent stem cells from adult mouse liver and stomach cells. *Science*, 321, (5889) 699-702 available from: PM:18276851

Apelqvist, A., Li, H., Sommer, L., Beatus, P., Anderson, D.J., Honjo, T., Hrabe de, A.M., Lendahl, U., & Edlund, H. 1999. Notch signalling controls pancreatic cell differentiation. *Nature*, 400, (6747) 877-881 available from: PM:10476967

Asfari, M., Janjic, D., Meda, P., Li, G., Halban, P.A., & Wollheim, C.B. 1992. Establishment of 2-mercaptoethanol-dependent differentiated insulin-secreting cell lines. *Endocrinology*, 130, (1) 167-178 available from: PM:1370150

Assady, S., Maor, G., Amit, M., Itskovitz-Eldor, J., Skorecki, K.L., & Tzukerman, M. 2001. Insulin production by human embryonic stem cells. *Diabetes*, 50, (8) 1691-1697 available from: PM:11473026

Atkinson, M.A., Maclaren, N.K., Riley, W.J., Winter, W.E., Fisk, D.D., & Spillar, R.P. 1986. Are insulin autoantibodies markers for insulin-dependent diabetes mellitus? *Diabetes*, 35, (8) 894-898 available from: PM:3525287

Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., & Lovell-Badge, R. 2003. Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev.*, 17, (1) 126-140 available from: PM:12514105

Baeyens, L., Lemper, M., Leuckx, G., De, G.S., Bonfanti, P., Stange, G., Shemer, R., Nord, C., Scheel, D.W., Pan, F.C., Ahlgren, U., Gu, G., Stoffers, D.A., Dor, Y., Ferrer, J., Gradwohl, G., Wright, C.V., Van de Casteele, M., German, M.S., Bouwens, L., & Heimberg, H. 2014. Transient cytokine treatment induces acinar cell reprogramming and regenerates functional beta cell mass in diabetic mice. *Nat.Biotechnol.*, 32, (1) 76-83 available from: PM:24240391

Bagge, A., Clausen, T.R., Larsen, S., Ladefoged, M., Rosenstierne, M.W., Larsen, L., Vang, O., Nielsen, J.H., & Dalgaard, L.T. 2012. MicroRNA-29a is up-regulated in beta-cells by glucose and decreases glucose-stimulated insulin secretion. *Biochem.Biophys.Res.Commun.*, 426, (2) 266-272 available from: PM:22940552

Bagge, A., Dahmcke, C.M., & Dalgaard, L.T. 2013. Syntaxin-1a is a direct target of miR-29a in insulin-producing beta-cells. *Horm.Metab Res.*, 45, (6) 463-466 available from: PM:23315993

Baker, D.E., Harrison, N.J., Maltby, E., Smith, K., Moore, H.D., Shaw, P.J., Heath, P.R., Holden, H., & Andrews, P.W. 2007. Adaptation to culture of human embryonic stem cells and oncogenesis in vivo. *Nat.Biotechnol.*, 25, (2) 207-215 available from: PM:17287758

Ballinger, W.F. & Lacy, P.E. 1972. Transplantation of intact pancreatic islets in rats. *Surgery*, 72, (2) 175-186 available from: http://www.ncbi.nlm.nih.gov/pubmed/4262169

Banting, F.G., Best, C.H., Collip, J.B., Campbell, W.R., & Fletcher, A.A. 1922. Pancreatic extracts in the treatment of diabetes mellitus. *Can Med Assoc J*, 12, (3) 141-146 available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1524425/pdf/canmedaj00414-0043.pdf

Bar-Nur, O., Russ, H.A., Efrat, S., & Benvenisty, N. 2011. Epigenetic memory and preferential lineage-specific differentiation in induced pluripotent stem cells derived from human pancreatic islet beta cells. *Cell Stem Cell*, 9, (1) 17-23 available from: PM:21726830

Bar, M., Wyman, S.K., Fritz, B.R., Qi, J., Garg, K.S., Parkin, R.K., Kroh, E.M., Bendoraite, A., Mitchell, P.S., Nelson, A.M., Ruzzo, W.L., Ware, C., Radich, J.P., Gentleman, R., Ruohola-Baker, H., & Tewari, M. 2008. MicroRNA discovery and profiling in human embryonic stem cells by deep sequencing of small RNA libraries. *Stem Cells*, 26, (10) 2496-2505 available from: PM:18583537

Bar, Y., Russ, H.A., Sintov, E., Anker-Kitai, L., Knoller, S., & Efrat, S. 2012. Redifferentiation of expanded human pancreatic beta-cell-derived cells by inhibition of the NOTCH pathway. *J Biol.Chem.*, 287, (21) 17269-17280 available from: PM:22457355

Baroukh, N., Ravier, M.A., Loder, M.K., Hill, E.V., Bounacer, A., Scharfmann, R., Rutter, G.A., & Van, O.E. 2007. MicroRNA-124a regulates Foxa2 expression and intracellular signaling in pancreatic beta-cell lines. *J Biol.Chem.*, 282, (27) 19575-19588 available from: PM:17462994

Bartel, D.P. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, 116, (2) 281-297 available from: PM:14744438

Bartel, D.P. 2009. MicroRNAs: target recognition and regulatory functions. *Cell*, 136, (2) 215-233 available from: PM:19167326

Barton, F.B., Rickels, M.R., Alejandro, R., Hering, B.J., Wease, S., Naziruddin, B., Oberholzer, J., Odorico, J.S., Garfinkel, M.R., Levy, M., Pattou, F., Berney, T., Secchi, A., Messinger, S., Senior, P.A., Maffi, P., Posselt, A., Stock, P.G., Kaufman, D.B., Luo, X., Kandeel, F., Cagliero, E., Turgeon, N.A., Witkowski, P., Naji, A., O'Connell, P.J., Greenbaum, C., Kudva, Y.C., Brayman, K.L., Aull, M.J., Larsen, C., Kay, T.W., Fernandez, L.A., Vantyghem, M.C., Bellin, M., & Shapiro, A.M. 2012. Improvement in outcomes of clinical islet transplantation: 1999-2010. *Diabetes Care*, 35, (7) 1436-1445 available from: PM:22723582

Beattie, G.M., Montgomery, A.M., Lopez, A.D., Hao, E., Perez, B., Just, M.L., Lakey, J.R., Hart, M.E., & Hayek, A. 2002. A novel approach to increase human islet cell mass while preserving beta-cell function. *Diabetes*, 51, (12) 3435-3439 available from: PM:12453897

Belo, J.A., Bouwmeester, T., Leyns, L., Kertesz, N., Gallo, M., Follettie, M., & De Robertis, E.M. 1997. Cerberus-like is a secreted factor with neutralizing activity expressed in the anterior primitive endoderm of the mouse gastrula. *Mech.Dev.*, 68, (1-2) 45-57 available from: PM:9431803

Ben-David, U. & Benvenisty, N. 2011. The tumorigenicity of human embryonic and induced pluripotent stem cells. *Nat.Rev.Cancer*, 11, (4) 268-277 available from: PM:21390058

Ben-David, U., Nudel, N., & Benvenisty, N. 2013. Immunologic and chemical targeting of the tight-junction protein Claudin-6 eliminates tumorigenic human pluripotent stem cells. *Nat.Commun.*, 4, 1992 available from: PM:23778593

Benetti, R., Gonzalo, S., Jaco, I., Munoz, P., Gonzalez, S., Schoeftner, S., Murchison, E., Andl, T., Chen, T., Klatt, P., Li, E., Serrano, M., Millar, S., Hannon, G., & Blasco, M.A. 2008. A mammalian microRNA cluster controls DNA methylation and telomere recombination via Rbl2-dependent regulation of DNA methyltransferases. *Nat.Struct.Mol.Biol.*, 15, (9) 998 available from: PM:18769471

Ber, I., Shternhall, K., Perl, S., Ohanuna, Z., Goldberg, I., Barshack, I., Benvenisti-Zarum, L., Meivar-Levy, I., & Ferber, S. 2003. Functional, persistent, and extended liver to pancreas transdifferentiation. *J Biol.Chem.*, 278, (34) 31950-31957 available from: PM:12775714

Berardi, E., Pues, M., Thorrez, L., & Sampaolesi, M. 2012. miRNAs in ESC differentiation. *Am.J Physiol Heart Circ.Physiol* , 303, (8) H931-H939 available from: PM:22886416

Bergstrom, R., Strom, S., Holm, F., Feki, A., & Hovatta, O. 2011. Xeno-free culture of human pluripotent stem cells. *Methods Mol.Biol.*, 767, 125-136 available from: PM:21822871

Bhushan, A., Itoh, N., Kato, S., Thiery, J.P., Czernichow, P., Bellusci, S., & Scharfmann, R. 2001. Fgf10 is essential for maintaining the proliferative capacity of epithelial progenitor cells during early pancreatic organogenesis. *Development*, 128, (24) 5109-5117 available from: PM:11748146

Biben, C., Stanley, E., Fabri, L., Kotecha, S., Rhinn, M., Drinkwater, C., Lah, M., Wang, C.C., Nash, A., Hilton, D., Ang, S.L., Mohun, T., & Harvey, R.P. 1998. Murine cerberus homologue mCer-1: a candidate anterior patterning molecule. *Dev.Biol.*, 194, (2) 135-151 available from: PM:9501024

Binot, A.C., Manfroid, I., Flasse, L., Winandy, M., Motte, P., Martial, J.A., Peers, B., & Voz, M.L. 2010. Nkx6.1 and nkx6.2 regulate alpha- and beta-cell formation in zebrafish by acting on pancreatic endocrine progenitor cells. *Dev.Biol.*, 340, (2) 397-407 available from: PM:20122912

Blauwkamp, T.A., Nigam, S., Ardehali, R., Weissman, I.L., & Nusse, R. 2012. Endogenous Wnt signalling in human embryonic stem cells generates an equilibrium of distinct lineage-specified progenitors. *Nat.Commun.*, 3, 1070 available from: PM:22990866

Blelloch, R., Venere, M., Yen, J., & Ramalho-Santos, M. 2007. Generation of induced pluripotent stem cells in the absence of drug selection. *Cell Stem Cell*, 1, (3) 245-247 available from: PM:18371358

Bluestone, J.A., Herold, K., & Eisenbarth, G. 2010. Genetics, pathogenesis and clinical interventions in type 1 diabetes. *Nature*, 464, (7293) 1293-1300 available from: PM:20432533

Blum, B., Hrvatin, S.S., Schuetz, C., Bonal, C., Rezania, A., & Melton, D.A. 2012. Functional beta-cell maturation is marked by an increased glucose threshold and by expression of urocortin 3. *Nat.Biotechnol.*, 30, (3) 261-264 available from: PM:22371083

Bock, C., Kiskinis, E., Verstappen, G., Gu, H., Boulting, G., Smith, Z.D., Ziller, M., Croft, G.F., Amoroso, M.W., Oakley, D.H., Gnirke, A., Eggan, K., & Meissner, A. 2011. Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell*, 144, (3) 439-452 available from: PM:21295703

Boissart, C., Nissan, X., Giraud-Triboult, K., Peschanski, M., & Benchoua, A. 2012. miR-125 potentiates early neural specification of human embryonic stem cells. *Development*, 139, (7) 1247-1257 available from: PM:22357933

Bonner-Weir, S. 2000. Life and death of the pancreatic beta cells. *Trends Endocrinol.Metab*, 11, (9) 375-378 available from: PM:11042468

Bonner-Weir, S., Taneja, M., Weir, G.C., Tatarkiewicz, K., Song, K.H., Sharma, A., & O'Neil, J.J. 2000. In vitro cultivation of human islets from expanded ductal tissue. *Proc.Natl.Acad.Sci.U.S.A*, 97, (14) 7999-8004 available from: PM:10884429

Borowiak, M., Maehr, R., Chen, S., Chen, A.E., Tang, W., Fox, J.L., Schreiber, S.L., & Melton, D.A. 2009. Small molecules efficiently direct endodermal differentiation of mouse and human embryonic stem cells. *Cell Stem Cell*, 4, (4) 348-358 available from: PM:19341624

Boyd, A.S. & Fairchild, P.J. 2010. Approaches for immunological tolerance induction to stem cell-derived cell replacement therapies. *Expert.Rev.Clin.Immunol.*, 6, (3) 435-448 available from: PM:20441429

Brambrink, T., Foreman, R., Welstead, G.G., Lengner, C.J., Wernig, M., Suh, H., & Jaenisch, R. 2008. Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell*, 2, (2) 151-159 available from: PM:18371436

Brereton, H.C., Carvell, M.J., Asare-Anane, H., Roberts, G., Christie, M.R., Persaud, S.J., & Jones, P.M. 2006. Homotypic cell contact enhances insulin but not glucagon secretion. *Biochem.Biophys.Res.Commun.*, 344, (3) 995-1000 available from: PM:16643853

Brissova, M., Fowler, M.J., Nicholson, W.E., Chu, A., Hirshberg, B., Harlan, D.M., & Powers, A.C. 2005. Assessment of human pancreatic islet architecture and composition by laser scanning confocal microscopy. *J.Histochem.Cytochem.*, 53, (9) 1087-1097 available from: PM:15923354

Brivanlou, A.H., Gage, F.H., Jaenisch, R., Jessell, T., Melton, D., & Rossant, J. 2003. Stem cells. Setting standards for human embryonic stem cells. *Science*, 300, (5621) 913-916 available from: PM:12738841

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

Brown, M.E., Rondon, E., Rajesh, D., Mack, A., Lewis, R., Feng, X., Zitur, L.J., Learish, R.D., & Nuwaysir, E.F. 2010. Derivation of induced pluripotent stem cells from human peripheral blood T lymphocytes. *PLoS.One.*, 5, (6) e11373 available from: PM:20617191

Bruin, J.E., Rezania, A., Xu, J., Narayan, K., Fox, J.K., O'Neil, J.J., & Kieffer, T.J. 2013. Maturation and function of human embryonic stem cell-derived pancreatic progenitors in macroencapsulation devices following transplant into mice. *Diabetologia*, 56, (9) 1987-1998 available from: PM:23771205

Burns, C.J., Minger, S.L., Hall, S., Milne, H., Ramracheya, R.D., Evans, N.D., Persaud, S.J., & Jones, P.M. 2005. The in vitro differentiation of rat neural stem cells into an insulin-expressing phenotype. *Biochem.Biophys.Res.Commun.*, 326, (3) 570-577 available from: PM:15596137

Butler, A.E., Janson, J., Bonner-Weir, S., Ritzel, R., Rizza, R.A., & Butler, P.C. 2003. Beta-cell deficit and increased beta-cell apoptosis in humans with type 2 diabetes. *Diabetes*, 52, (1) 102-110 available from: PM:12502499

Cabrera, O., Berman, D.M., Kenyon, N.S., Ricordi, C., Berggren, P.O., & Caicedo, A. 2006. The unique cytoarchitecture of human pancreatic islets has implications for islet cell function. *Proc.Natl.Acad.Sci.U.S.A*, 103, (7) 2334-2339 available from: PM:16461897

Cai, J., Yu, C., Liu, Y., Chen, S., Guo, Y., Yong, J., Lu, W., Ding, M., & Deng, H. 2010. Generation of homogeneous PDX1(+) pancreatic progenitors from human ES cell-derived endoderm cells. *J Mol.Cell Biol.*, 2, (1) 50-60 available from: PM:19910415

Cao, F., Drukker, M., Lin, S., Sheikh, A.Y., Xie, X., Li, Z., Connolly, A.J., Weissman, I.L., & Wu, J.C. 2007. Molecular imaging of embryonic stem cell misbehavior and suicide gene ablation. *Cloning Stem Cells*, 9, (1) 107-117 available from: PM:17386018

Cao, L.Z., Tang, D.Q., Horb, M.E., Li, S.W., & Yang, L.J. 2004. High glucose is necessary for complete maturation of Pdx1-VP16-expressing hepatic cells into functional insulin-producing cells. *Diabetes* , 53, (12) 3168-3178 available from: PM:15561947

Card, D.A., Hebbar, P.B., Li, L., Trotter, K.W., Komatsu, Y., Mishina, Y., & Archer, T.K. 2008. Oct4/Sox2-regulated miR-302 targets cyclin D1 in human embryonic stem cells. *Mol.Cell Biol.*, 28, (20) 6426-6438 available from: PM:18710938

Carey, B.W., Markoulaki, S., Hanna, J., Saha, K., Gao, Q., Mitalipova, M., & Jaenisch, R. 2009. Reprogramming of murine and human somatic cells using a single polycistronic vector. *Proc.Natl.Acad.Sci.U.S.A*, 106, (1) 157-162 available from: PM:19109433

Cavelti-Weder, C., Shtessel, M., Reuss, J.E., Jermendy, A., Yamada, T., Caballero, F., Bonner-Weir, S., & Weir, G.C. 2013. Pancreatic Duct Ligation After Almost Complete beta-Cell Loss: Exocrine Regeneration but No Evidence of beta-Cell Regeneration. *Endocrinology*, 154, (12) 4493-4502 available from: PM:24029238

Chakrabarti, S.K., James, J.C., & Mirmira, R.G. 2002. Quantitative assessment of gene targeting in vitro and in vivo by the pancreatic transcription factor, Pdx1. Importance of chromatin structure in directing promoter binding. *J.Biol.Chem.*, 277, (15) 13286-13293 available from: PM:11825903

Chan, E.M., Ratanasirintrawoot, S., Park, I.H., Manos, P.D., Loh, Y.H., Huo, H., Miller, J.D., Hartung, O., Rho, J., Ince, T.A., Daley, G.Q., & Schlaeger, T.M. 2009. Live cell imaging distinguishes bona fide human

iPS cells from partially reprogrammed cells. *Nat.Biotechnol.*, 27, (11) 1033-1037 available from: PM:19826408

Chen, C., Ridzon, D., Lee, C.T., Blake, J., Sun, Y., & Strauss, W.M. 2007. Defining embryonic stem cell identity using differentiation-related microRNAs and their potential targets. *Mamm.Genome*, 18, (5) 316-327 available from: PM:17610011

Chen, C., Chai, J., Singh, L., Kuo, C.Y., Jin, L., Feng, T., Marzano, S., Galeni, S., Zhang, N., Iacovino, M., Qin, L., Hara, M., Stein, R., Bromberg, J.S., Kyba, M., & Ku, H.T. 2011. Characterization of an in vitro differentiation assay for pancreatic-like cell development from murine embryonic stem cells: detailed gene expression analysis. *Assay.Drug Dev.Technol.*, 9, (4) 403-419 available from: PM:21395400

Chen, S., Borowiak, M., Fox, J.L., Maehr, R., Osafune, K., Davidow, L., Lam, K., Peng, L.F., Schreiber, S.L., Rubin, L.L., & Melton, D. 2009. A small molecule that directs differentiation of human ESCs into the pancreatic lineage. *Nat.Chem.Biol.*, 5, (4) 258-265 available from: PM:19287398

Chen, Y.J., Finkbeiner, S.R., Weinblatt, D., Emmett, M.J., Tameire, F., Yousefi, M., Yang, C., Maehr, R., Zhou, Q., Shemer, R., Dor, Y., Li, C., Spence, J.R., & Stanger, B.Z. 2014. De Novo Formation of Insulin-Producing "Neo-beta Cell Islets" from Intestinal Crypts. *Cell Rep.*, 6, (6) 1046-1058 available from: PM:24613355

Cheng, X., Ying, L., Lu, L., Galvao, A.M., Mills, J.A., Lin, H.C., Kotton, D.N., Shen, S.S., Nostro, M.C., Choi, J.K., Weiss, M.J., French, D.L., & Gadue, P. 2012. Self-renewing endodermal progenitor lines generated from human pluripotent stem cells. *Cell Stem Cell*, 10, (4) 371-384 available from: PM:22482503

Chetty, S., Pagliuca, F.W., Honore, C., Kweudjeu, A., Rezania, A., & Melton, D.A. 2013. A simple tool to improve pluripotent stem cell differentiation. *Nat.Methods*, 10, (6) 553-556 available from: PM:23584186

Chin, M.H., Mason, M.J., Xie, W., Volinia, S., Singer, M., Peterson, C., Ambartsumyan, G., Aimiuwu, O., Richter, L., Zhang, J., Khvorostov, I., Ott, V., Grunstein, M., Lavon, N., Benvenisty, N., Croce, C.M., Clark, A.T., Baxter, T., Pyle, A.D., Teitell, M.A., Pelegrini, M., Plath, K., & Lowry, W.E. 2009. Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell*, 5, (1) 111-123 available from: PM:19570518

Chintinne, M., Stange, G., Denys, B., Ling, Z., In, '., V, & Pipeleers, D. 2012. Beta cell count instead of beta cell mass to assess and localize growth in beta cell population following pancreatic duct ligation in mice. *PLoS.One.*, 7, (8) e43959 available from: PM:22952825

Chng, Z., Teo, A., Pedersen, R.A., & Vallier, L. 2010. SIP1 mediates cell-fate decisions between neuroectoderm and mesendoderm in human pluripotent stem cells. *Cell Stem Cell*, 6, (1) 59-70 available from: PM:20074535

Cho, C.H., Hannan, N.R., Docherty, F.M., Docherty, H.M., Joao, L.M., Trotter, M.W., Docherty, K., & Vallier, L. 2012. Inhibition of activin/nodal signalling is necessary for pancreatic differentiation of human pluripotent stem cells. *Diabetologia*, 55, (12) 3284-3295 available from: PM:23011350

Cho, Y.M., Lim, J.M., Yoo, D.H., Kim, J.H., Chung, S.S., Park, S.G., Kim, T.H., Oh, S.K., Choi, Y.M., Moon, S.Y., Park, K.S., & Lee, H.K. 2008. Betacellulin and nicotinamide sustain PDX1 expression and induce pancreatic beta-cell differentiation in human embryonic stem cells. *Biochem.Biophys.Res.Commun.*, 366, (1) 129-134 available from: PM:18060872

Choi, J.B., Uchino, H., Azuma, K., Iwashita, N., Tanaka, Y., Mochizuki, H., Migita, M., Shimada, T., Kawamori, R., & Watada, H. 2003. Little evidence of transdifferentiation of bone marrow-derived cells into pancreatic beta cells. *Diabetologia*, 46, (10) 1366-1374 available from: PM:12898006

Choo, A.B., Tan, H.L., Ang, S.N., Fong, W.J., Chin, A., Lo, J., Zheng, L., Hentze, H., Philp, R.J., Oh, S.K., & Yap, M. 2008. Selection against undifferentiated human embryonic stem cells by a cytotoxic antibody recognizing podocalyxin-like protein-1. *Stem Cells*, 26, (6) 1454-1463 available from: PM:18356574

Chou, B.K., Mali, P., Huang, X., Ye, Z., Dowey, S.N., Resar, L.M., Zou, C., Zhang, Y.A., Tong, J., & Cheng, L. 2011. Efficient human iPS cell derivation by a non-integrating plasmid from blood cells with unique epigenetic and gene expression signatures. *Cell Res.*, 21, (3) 518-529 available from: PM:21243013

Christodoulou, C., Longmire, T.A., Shen, S.S., Bourdon, A., Sommer, C.A., Gadue, P., Spira, A., Gouon-Evans, V., Murphy, G.J., Mostoslavsky, G., & Kotton, D.N. 2011. Mouse ES and iPS cells can form similar definitive endoderm despite differences in imprinted genes. *J Clin.Invest*, 121, (6) 2313-2325 available from: PM:21537085

Clarke, D.L., Johansson, C.B., Wilbertz, J., Veress, B., Nilsson, E., Karlstrom, H., Lendahl, U., & Frisen, J. 2000. Generalized potential of adult neural stem cells. *Science*, 288, (5471) 1660-1663 available from: PM:10834848

Coffinier, C., Barra, J., Babinet, C., & Yaniv, M. 1999. Expression of the vHNF1/HNF1beta homeoprotein gene during mouse organogenesis. *Mech.Dev.*, 89, (1-2) 211-213 available from: PM:10559500

Collombat, P., Mansouri, A., Hecksher-Sorensen, J., Serup, P., Krull, J., Gradwohl, G., & Gruss, P. 2003. Opposing actions of Arx and Pax4 in endocrine pancreas development. *Genes Dev.*, 17, (20) 2591-2603 available from: PM:14561778

Conesa, C., Doss, M.X., Antzelevitch, C., Sachinidis, A., Sancho, J., & Carrodeguas, J.A. 2012. Identification of specific pluripotent stem cell death--inducing small molecules by chemical screening. *Stem Cell Rev.*, 8, (1) 116-127 available from: PM:21617963

Correa-Medina, M., Bravo-Egana, V., Rosero, S., Ricordi, C., Edlund, H., Diez, J., & Pastori, R.L. 2009. MicroRNA miR-7 is preferentially expressed in endocrine cells of the developing and adult human pancreas. *Gene Expr.Patterns.*, 9, (4) 193-199 available from: PM:19135553

Couri, C.E., Oliveira, M.C., Stracieri, A.B., Moraes, D.A., Pieroni, F., Barros, G.M., Madeira, M.I., Malmegrim, K.C., Foss-Freitas, M.C., Simoes, B.P., Martinez, E.Z., Foss, M.C., Burt, R.K., & Voltarelli, J.C. 2009. C-peptide levels and insulin independence following autologous nonmyeloablative hematopoietic stem cell transplantation in newly diagnosed type 1 diabetes mellitus. *JAMA*, 301, (15) 1573-1579 available from: PM:19366777

Courtney, M., Gjernes, E., Druelle, N., Ravaud, C., Vieira, A., Ben-Othman, N., Pfeifer, A., Avolio, F., Leuckx, G., Lacas-Gervais, S., Burel-Vandenbos, F., Ambrosetti, D., Hecksher-Sorensen, J., Ravassard, P., Heimberg, H., Mansouri, A., & Collombat, P. 2013. The Inactivation of Arx in Pancreatic alpha-Cells Triggers Their Neogenesis and Conversion into Functional beta-Like Cells. *PLoS.Genet.*, 9, (10) e1003934 available from: PM:24204325

Courtney, M.L., Jones, P.M., & Burns, C.J. 2010. Importance of quantitative analysis in the generation of insulin-expressing cells from human embryonic stem cells. *Pancreas*, 39, (1) 105-107 available from: PM:20019566

Crasto, W., Jarvis, J., Khunti, K., & Davies, M.J. 2009. New insulins and new insulin regimens: a review of their role in improving glycaemic control in patients with diabetes. *Postgrad.Med J*, 85, (1003) 257-267 available from: PM:19520878

Crook, J.M., Peura, T.T., Kravets, L., Bosman, A.G., Buzzard, J.J., Horne, R., Hentze, H., Dunn, N.R., Zweigerdt, R., Chua, F., Upshall, A., & Colman, A. 2007. The generation of six clinical-grade human embryonic stem cell lines. *Cell Stem Cell*, 1, (5) 490-494 available from: PM:18938745

D'Amour, K.A., Agulnick, A.D., Eliazer, S., Kelly, O.G., Kroon, E., & Baetge, E.E. 2005. Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nat.Biotechnol.*, 23, (12) 1534-1541 available from: PM:16258519

D'Amour, K.A., Bang, A.G., Eliazer, S., Kelly, O.G., Agulnick, A.D., Smart, N.G., Moorman, M.A., Kroon, E., Carpenter, M.K., & Baetge, E.E. 2006. Production of pancreatic hormone-expressing endocrine cells from human embryonic stem cells. *Nat.Biotechnol.*, 24, (11) 1392-1401 available from: PM:17053790

Daneman, D. 2006. Type 1 diabetes. *Lancet* , 367, (9513) 847-858 available from: PM:16530579

de Kort, H., de Koning, E.J., Rabelink, T.J., Bruijn, J.A., & Bajema, I.M. 2011. Islet transplantation in type 1 diabetes. *BMJ*, 342, d217 available from: PM:21257658

de la Tour, D., Halvorsen, T., Demeterco, C., Tyrberg, B., Itkin-Ansari, P., Loy, M., Yoo, S.J., Hao, E., Bossie, S., & Levine, F. 2001. Beta-cell differentiation from a human pancreatic cell line in vitro and in vivo. *Mol.Endocrinol.*, 15, (3) 476-483 available from: PM:11222748

de Vos, P., Hamel, A.F., & Tatarkiewicz, K. 2002. Considerations for successful transplantation of encapsulated pancreatic islets. *Diabetologia*, 45, (2) 159-173 available from: PM:11935147

Demeterco, C., Itkin-Ansari, P., Tyrberg, B., Ford, L.P., Jarvis, R.A., & Levine, F. 2002. c-Myc controls proliferation versus differentiation in human pancreatic endocrine cells. *J Clin.Endocrinol.Metab*, 87, (7) 3475-3485 available from: PM:12107268

Denner, J., Specke, V., Karlas, A., Chodnevskaja, I., Meyer, T., Moskalenko, V., Kurth, R., & Ulrichs, K. 2008. No transmission of porcine endogenous retroviruses (PERVs) in a long-term pig to rat xenotransplantation model and no infection of immunosuppressed rats. *Ann.Transplant.* , 13, (1) 20-31 available from: PM:18344940

Derosa, G. & Sibilla, S. 2007. Optimizing combination treatment in the management of type 2 diabetes. *Vasc.Health Risk Manag.*, 3, (5) 665-671 available from: PM:18078018

Dessimoz, J., Opoka, R., Kordich, J.J., Grapin-Botton, A., & Wells, J.M. 2006. FGF signaling is necessary for establishing gut tube domains along the anterior-posterior axis in vivo. *Mech.Dev.*, 123, (1) 42-55 available from: PM:16326079

Deutsch, G., Jung, J., Zheng, M., Lora, J., & Zaret, K.S. 2001. A bipotential precursor population for pancreas and liver within the embryonic endoderm. *Development*, 128, (6) 871-881 available from: PM:11222142

DiMeglio, L.A., Pottorff, T.M., Boyd, S.R., France, L., Fineberg, N., & Eugster, E.A. 2004. A randomized, controlled study of insulin pump therapy in diabetic preschoolers. *J Pediatr.*, 145, (3) 380-384 available from: PM:15343195

Doi, A., Park, I.H., Wen, B., Murakami, P., Aryee, M.J., Irizarry, R., Herb, B., Ladd-Acosta, C., Rho, J., Loewer, S., Miller, J., Schlaeger, T., Daley, G.Q., & Feinberg, A.P. 2009. Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat.Genet.*, 41, (12) 1350-1353 available from: PM:19881528

Dor, Y., Brown, J., Martinez, O.I., & Melton, D.A. 2004. Adult pancreatic beta-cells are formed by self-duplication rather than stem-cell differentiation. *Nature*, 429, (6987) 41-46 available from: PM:15129273

Draper, J.S., Moore, H.D., Ruban, L.N., Gokhale, P.J., & Andrews, P.W. 2004. Culture and characterization of human embryonic stem cells. *Stem Cells Dev.*, 13, (4) 325-336 available from: PM:15345125

Draper, J.S., Smith, K., Gokhale, P., Moore, H.D., Maltby, E., Johnson, J., Meisner, L., Zwaka, T.P., Thomson, J.A., & Andrews, P.W. 2004. Recurrent gain of chromosomes 17q and 12 in cultured human embryonic stem cells. *Nat.Biotechnol.*, 22, (1) 53-54 available from: PM:14661028

Duncan, S.A., Manova, K., Chen, W.S., Hoodless, P., Weinstein, D.C., Bachvarova, R.F., & Darnell, J.E., Jr. 1994. Expression of transcription factor HNF-4 in the extraembryonic endoderm, gut, and nephrogenic tissue of the developing mouse embryo: HNF-4 is a marker for primary endoderm in the implanting blastocyst. *Proc.Natl.Acad.Sci.U.S.A*, 91, (16) 7598-7602 available from: PM:8052626

Duvivier-Kali, V.F., Omer, A., Parent, R.J., O'Neil, J.J., & Weir, G.C. 2001. Complete protection of islets against allorejection and autoimmunity by a simple barium-alginate membrane. *Diabetes*, 50, (8) 1698-1705 available from: PM:11473027

Efrat, S., Linde, S., Kofod, H., Spector, D., Delannoy, M., Grant, S., Hanahan, D., & Baekkeskov, S. 1988. Beta-cell lines derived from transgenic mice expressing a hybrid insulin gene-oncogene. *Proc.Natl.Acad.Sci.U.S.A*, 85, (23) 9037-9041 available from: PM:2848253

El, Ouamaari A., Baroukh, N., Martens, G.A., Lebrun, P., Pipeleers, D., & Van, O.E. 2008. miR-375 targets 3'-phosphoinositide-dependent protein kinase-1 and regulates glucose-induced biological responses in pancreatic beta-cells. *Diabetes*, 57, (10) 2708-2717 available from: PM:18591395

Elliott, R.B., Escobar, L., Tan, P.L., Muzina, M., Zwain, S., & Buchanan, C. 2007. Live encapsulated porcine islets from a type 1 diabetic patient 9.5 yr after xenotransplantation. *Xenotransplantation.* , 14, (2) 157-161 available from: PM:17381690

Elsner, M., Jorns, A., & Lenzen, S. 2008. Diabetes therapy by lentiviral hepatic insulin gene expression without transformation of liver. *Diabetologia*, 51, (4) 694-695 available from: PM:18239907

Eminli, S., Utikal, J., Arnold, K., Jaenisch, R., & Hochedlinger, K. 2008. Reprogramming of neural progenitor cells into induced pluripotent stem cells in the absence of exogenous Sox2 expression. *Stem Cells*, 26, (10) 2467-2474 available from: PM:18635867

Encyclopaedia Britannica. Structures of the pancreas. 2014. 5-2-2014.
Ref Type: Online Source

Ende, N., Chen, R., & Reddi, A.S. 2004. Transplantation of human umbilical cord blood cells improves glycemia and glomerular hypertrophy in type 2 diabetic mice. *Biochem.Biophys.Res.Commun.*, 321, (1) 168-171 available from: PM:15358230

Eshpeter, A., Jiang, J., Au, M., Rajotte, R.V., Lu, K., Lebkowski, J.S., Majumdar, A.S., & Korbutt, G.S. 2008. In vivo characterization of transplanted human embryonic stem cell-derived pancreatic endocrine islet cells. *Cell Prolif.*, 41, (6) 843-858 available from: PM:19040565

Evans, M.J. & Kaufman, M.H. 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature*, 292, (5819) 154-156 available from: PM:7242681

Expert Committee on the Diagnosis and Classification of Diabetes Mellitus 2003. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care*, 26, (suppl 1) s5-s20 available from: http://care.diabetesjournals.org/content/26/suppl_1/s5.short

Fairchild, P.J. 2010. The challenge of immunogenicity in the quest for induced pluripotency. *Nat.Rev.Immunol.*, 10, (12) 868-875 available from: PM:21107347

Feng, B., Ng, J.H., Heng, J.C., & Ng, H.H. 2009. Molecules that promote or enhance reprogramming of somatic cells to induced pluripotent stem cells. *Cell Stem Cell*, 4, (4) 301-312 available from: PM:19341620

Feng, Q., Lu, S.J., Klimanskaya, I., Gomes, I., Kim, D., Chung, Y., Honig, G.R., Kim, K.S., & Lanza, R. 2010. Hemangioblastic derivatives from human induced pluripotent stem cells exhibit limited expansion and early senescence. *Stem Cells*, 28, (4) 704-712 available from: PM:20155819

Ferber, S., Halkin, A., Cohen, H., Ber, I., Einav, Y., Goldberg, I., Barshack, I., Seijffers, R., Kopolovic, J., Kaiser, N., & Karasik, A. 2000. Pancreatic and duodenal homeobox gene 1 induces expression of insulin genes in liver and ameliorates streptozotocin-induced hyperglycemia. *Nat.Med*, 6, (5) 568-572 available from: PM:10802714

Finch, B.W. & Ephrussi, B. 1967. Retention of multiple developmental potentialities by cells of a mouse testicular teratocarcinoma during prolonged culture in vitro and their extinction upon hybridization with cells of permanent lines. *Proc.Natl.Acad.Sci.U.S.A*, 57, (3) 615-621 available from: PM:16591508

Finnegan, E.F. & Pasquinelli, A.E. 2013. MicroRNA biogenesis: regulating the regulators. *Crit Rev.Biochem.Mol.Biol.*, 48, (1) 51-68 available from: PM:23163351

Fishman, B., Segev, H., Kopper, O., Nissenbaum, J., Schulman, M., Benvenisty, N., Itskovitz-Eldor, J., & Kitsberg, D. 2012. Targeting pancreatic progenitor cells in human embryonic stem cell differentiation for the identification of novel cell surface markers. *Stem Cell Rev.*, 8, (3) 792-802 available from: PM:22592939

Fluri, D.A., Tonge, P.D., Song, H., Baptista, R.P., Shakiba, N., Shukla, S., Clarke, G., Nagy, A., & Zandstra, P.W. 2012. Derivation, expansion and differentiation of induced pluripotent stem cells in continuous suspension cultures. *Nat.Methods*, 9, (5) 509-516 available from: PM:22447133

Fodor, A., Harel, C., Fodor, L., Armoni, M., Salmon, P., Trono, D., & Karnieli, E. 2007. Adult rat liver cells transdifferentiated with lentiviral IPF1 vectors reverse diabetes in mice: an ex vivo gene therapy approach. *Diabetologia*, 50, (1) 121-130 available from: PM:17131142

Foshay, K.M. & Gallicano, G.I. 2009. miR-17 family miRNAs are expressed during early mammalian development and regulate stem cell differentiation. *Dev.Biol.*, 326, (2) 431-443 available from: PM:19073166

Fox, L.A., Buckloh, L.M., Smith, S.D., Wysocki, T., & Mauras, N. 2005. A randomized controlled trial of insulin pump therapy in young children with type 1 diabetes. *Diabetes Care*, 28, (6) 1277-1281 available from: PM:15920039

Francis, N.A., Moore, M., Rutter, G.A., & Burns, C.J. 2014. The Role of MicroRNAs in the Pancreatic Differentiation of Pluripotent Stem Cells. *MicroRNA*, 3(1):54-63.

Fred, R.G., Bang-Berthelsen, C.H., Mandrup-Poulsen, T., Grunnet, L.G., & Welsh, N. 2010. High glucose suppresses human islet insulin biosynthesis by inducing miR-133a leading to decreased polypyrimidine tract binding protein-expression. *PLoS.One.*, 5, (5) e10843 available from: PM:20520763

Frost, R.J. & Olson, E.N. 2011. Control of glucose homeostasis and insulin sensitivity by the Let-7 family of microRNAs. *Proc.Natl.Acad.Sci.U.S.A*, 108, (52) 21075-21080 available from: PM:22160727

Fryer, B.H., Rezania, A., & Zimmerman, M.C. 2013. Generating beta-cells in vitro: progress towards a Holy Grail. *Curr.Opin.Endocrinol.Diabetes Obes.*, 20, (2) 112-117 available from: PM:23426161

Fu, H., Tie, Y., Xu, C., Zhang, Z., Zhu, J., Shi, Y., Jiang, H., Sun, Z., & Zheng, X. 2005. Identification of human fetal liver miRNAs by a novel method. *FEBS Lett.*, 579, (17) 3849-3854 available from: PM:15978578

Fu, S., Fei, Q., Jiang, H., Chuai, S., Shi, S., Xiong, W., Jiang, L., Lu, C., Atadja, P., Li, E., & Shou, J. 2011. Involvement of histone acetylation of Sox17 and Foxa2 promoters during mouse definitive endoderm differentiation revealed by microRNA profiling. *PLoS.One.*, 6, (11) e27965 available from: PM:22132182

Fu, X., Jin, L., Wang, X., Luo, A., Hu, J., Zheng, X., Tsark, W.M., Riggs, A.D., Ku, H.T., & Huang, W. 2013. MicroRNA-26a targets ten eleven translocation enzymes and is regulated during pancreatic cell differentiation. *Proc.Natl.Acad.Sci.U.S.A*, 110, (44) 17892-17897 available from: PM:24114270

Fukushima, A., Okuda, A., Nishimoto, M., Seki, N., Hori, T.A., & Muramatsu, M. 1998. Characterization of functional domains of an embryonic stem cell coactivator UTF1 which are conserved and essential for potentiation of ATF-2 activity. *J Biol.Chem.*, 273, (40) 25840-25849 available from: PM:9748258

Furuyama, K., Kawaguchi, Y., Akiyama, H., Horiguchi, M., Kodama, S., Kuhara, T., Hosokawa, S., Elbahrawy, A., Soeda, T., Koizumi, M., Masui, T., Kawaguchi, M., Takaori, K., Doi, R., Nishi, E., Kakinoki, R., Deng, J.M., Behringer, R.R., Nakamura, T., & Uemoto, S. 2011. Continuous cell supply from a Sox9-expressing progenitor zone in adult liver, exocrine pancreas and intestine. *Nat.Genet.*, 43, (1) 34-41 available from: PM:21113154

Gadue, P., Huber, T.L., Paddison, P.J., & Keller, G.M. 2006. Wnt and TGF-beta signaling are required for the induction of an in vitro model of primitive streak formation using embryonic stem cells. *Proc.Natl.Acad.Sci.U.S.A*, 103, (45) 16806-16811 available from: PM:17077151

Gaglia, J.L., Shapiro, A.M., & Weir, G.C. 2005. Islet transplantation: progress and challenge. *Arch.Med Res.*, 36, (3) 273-280 available from: PM:15925017

Galli, R., Gritti, A., Bonfanti, L., & Vescovi, A.L. 2003. Neural stem cells: an overview. *Circ.Res.*, 92, (6) 598-608 available from: PM:12676811

Gangaraju, V.K. & Lin, H. 2009. MicroRNAs: key regulators of stem cells. *Nat.Rev.Mol.Cell Biol.*, 10, (2) 116-125 available from: PM:19165214

Gaudet, F., Hodgson, J.G., Eden, A., Jackson-Grusby, L., Dausman, J., Gray, J.W., Leonhardt, H., & Jaenisch, R. 2003. Induction of tumors in mice by genomic hypomethylation. *Science*, 300, (5618) 489-492 available from: PM:12702876

Gazdar, A.F., Chick, W.L., Oie, H.K., Sims, H.L., King, D.L., Weir, G.C., & Lauris, V. 1980. Continuous, clonal, insulin- and somatostatin-secreting cell lines established from a transplantable rat islet cell tumor. *Proc.Natl.Acad.Sci.U.S.A*, 77, (6) 3519-3523 available from: PM:6106192

Gerrish, K., Gannon, M., Shih, D., Henderson, E., Stoffel, M., Wright, C.V., & Stein, R. 2000. Pancreatic beta cell-specific transcription of the pdx-1 gene. The role of conserved upstream control regions and their hepatic nuclear factor 3beta sites. *J.Biol.Chem.*, 275, (5) 3485-3492 available from: PM:10652343

Gershengorn, M.C., Hardikar, A.A., Wei, C., Geras-Raaka, E., Marcus-Samuels, B., & Raaka, B.M. 2004. Epithelial-to-mesenchymal transition generates proliferative human islet precursor cells. *Science*, 306, (5705) 2261-2264 available from: PM:15564314

Ghosh, Z., Wilson, K.D., Wu, Y., Hu, S., Quertermous, T., & Wu, J.C. 2010. Persistent donor cell gene expression among human induced pluripotent stem cells contributes to differences with human embryonic stem cells. *PLoS.One.*, 5, (2) e8975 available from: PM:20126639

Giorgetti, A., Montserrat, N., Aasen, T., Gonzalez, F., Rodriguez-Piza, I., Vassena, R., Raya, A., Boue, S., Barrero, M.J., Corbella, B.A., Torrabadella, M., Veiga, A., & Izpisua Belmonte, J.C. 2009. Generation of induced pluripotent stem cells from human cord blood using OCT4 and SOX2. *Cell Stem Cell*, 5, (4) 353-357 available from: PM:19796614

Git, A., Dvinge, H., Salmon-Divon, M., Osborne, M., Kutter, C., Hadfield, J., Bertone, P., & Caldas, C. 2010. Systematic comparison of microarray profiling, real-time PCR, and next-generation sequencing technologies for measuring differential microRNA expression. *RNA.*, 16, (5) 991-1006 available from: PM:20360395

Gore, A., Li, Z., Fung, H.L., Young, J.E., Agarwal, S., Antosiewicz-Bourget, J., Canto, I., Giorgetti, A., Israel, M.A., Kiskinis, E., Lee, J.H., Loh, Y.H., Manos, P.D., Montserrat, N., Panopoulos, A.D., Ruiz, S., Wilbert, M.L., Yu, J., Kirkness, E.F., Izpisua Belmonte, J.C., Rossi, D.J., Thomson, J.A., Eggan, K., Daley, G.Q., Goldstein, L.S., & Zhang, K. 2011. Somatic coding mutations in human induced pluripotent stem cells. *Nature*, 471, (7336) 63-67 available from: PM:21368825

Gradwohl, G., Dierich, A., LeMeur, M., & Guillemot, F. 2000. neurogenin3 is required for the development of the four endocrine cell lineages of the pancreas. *Proc.Natl.Acad.Sci.U.S.A*, 97, (4) 1607-1611 available from: PM:10677506

Grapin-Botton, A., Majithia, A.R., & Melton, D.A. 2001. Key events of pancreas formation are triggered in gut endoderm by ectopic expression of pancreatic regulatory genes. *Genes Dev.*, 15, (4) 444-454 available from: PM:11230152

Grapin-Botton, A. & Constam, D. 2007. Evolution of the mechanisms and molecular control of endoderm formation. *Mech.Dev.*, 124, (4) 253-278 available from: PM:17307341

Greggio, C., De, F.F., Figueiredo-Larsen, M., Gobaa, S., Ranga, A., Semb, H., Lutolf, M., & Grapin-Botton, A. 2013. Artificial three-dimensional niches deconstruct pancreas development in vitro. *Development*, 140, (21) 4452-4462 available from: PM:24130330

Gregory, P.A., Bracken, C.P., Bert, A.G., & Goodall, G.J. 2008. MicroRNAs as regulators of epithelial-mesenchymal transition. *Cell Cycle*, 7, (20) 3112-3118 available from: PM:18927505

Grinnemo, K.H., Sylven, C., Hovatta, O., Dellgren, G., & Corbascio, M. 2008. Immunogenicity of human embryonic stem cells. *Cell Tissue Res.*, 331, (1) 67-78 available from: PM:17846795

Grose, S. 2007. Critics slam Russian trial to test pig pancreas for diabetics. *Nat.Med*, 13, (4) 390-391 available from: PM:17415358

Grove, J.E., Bruscia, E., & Krause, D.S. 2004. Plasticity of bone marrow-derived stem cells. *Stem Cells*, 22, (4) 487-500 available from: PM:15277695

Gu, G., Dubauskaite, J., & Melton, D.A. 2002. Direct evidence for the pancreatic lineage: NGN3+ cells are islet progenitors and are distinct from duct progenitors. *Development*, 129, (10) 2447-2457 available from: PM:11973276

Gu, G., Brown, J.R., & Melton, D.A. 2003. Direct lineage tracing reveals the ontogeny of pancreatic cell fates during mouse embryogenesis. *Mech.Dev.*, 120, (1) 35-43 available from: PM:12490294

Guenther, M.G., Frampton, G.M., Soldner, F., Hockemeyer, D., Mitalipova, M., Jaenisch, R., & Young, R.A. 2010. Chromatin structure and gene expression programs of human embryonic and induced pluripotent stem cells. *Cell Stem Cell*, 7, (2) 249-257 available from: PM:20682450

Gurdon, J.B., Harger, P., Mitchell, A., & Lemaire, P. 1994. Activin signalling and response to a morphogen gradient. *Nature*, 371, (6497) 487-492 available from: PM:7935761

Gutierrez-Aranda, I., Ramos-Mejia, V., Bueno, C., Munoz-Lopez, M., Real, P.J., Macia, A., Sanchez, L., Ligero, G., Garcia-Parez, J.L., & Menendez, P. 2010. Human induced pluripotent stem cells develop teratoma more efficiently and faster than human embryonic stem cells regardless the site of injection. *Stem Cells*, 28, (9) 1568-1570 available from: PM:20641038

Haase, A., Olmer, R., Schwanke, K., Wunderlich, S., Merkert, S., Hess, C., Zweigerdt, R., Gruh, I., Meyer, J., Wagner, S., Maier, L.S., Han, D.W., Glage, S., Miller, K., Fischer, P., Scholer, H.R., & Martin, U. 2009. Generation of induced pluripotent stem cells from human cord blood. *Cell Stem Cell*, 5, (4) 434-441 available from: PM:19796623

Hanna, J., Wernig, M., Markoulaki, S., Sun, C.W., Meissner, A., Cassady, J.P., Beard, C., Brambrink, T., Wu, L.C., Townes, T.M., & Jaenisch, R. 2007. Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin. *Science*, 318, (5858) 1920-1923 available from: PM:18063756

Hanna, J., Carey, B.W., & Jaenisch, R. 2008. Reprogramming of somatic cell identity. *Cold Spring Harb.Symp.Quant.Biol.*, 73, 147-155 available from: PM:19022760

Hao, E., Lee, S.H., & Levine, F. 2013. Efficient beta-Cell Regeneration by a Combination of Neogenesis and Replication Following beta-Cell Ablation and Reversal of Pancreatic Duct Ligation. *Stem Cells*, 31, (11) 2388-2395 available from: PM:23922239

Harlan, D.M., Kenyon, N.S., Korsgren, O., & Roep, B.O. 2009. Current advances and travails in islet transplantation. *Diabetes*, 58, (10) 2175-2184 available from: PM:19794074

Harrison, K.A., Thaler, J., Pfaff, S.L., Gu, H., & Kehrl, J.H. 1999. Pancreas dorsal lobe agenesis and abnormal islets of Langerhans in Hlxb9-deficient mice. *Nat.Genet.*, 23, (1) 71-75 available from: PM:10471502

Hasegawa, Y., Ogihara, T., Yamada, T., Ishigaki, Y., Imai, J., Uno, K., Gao, J., Kaneko, K., Ishihara, H., Sasano, H., Nakauchi, H., Oka, Y., & Katagiri, H. 2007. Bone marrow (BM) transplantation promotes beta-cell regeneration after acute injury through BM cell mobilization. *Endocrinology*, 148, (5) 2006-2015 available from: PM:17255204

Hawkins, R.D., Hon, G.C., Lee, L.K., Ngo, Q., Lister, R., Pelizzola, M., Edsall, L.E., Kuan, S., Luu, Y., Klugman, S., Antosiewicz-Bourget, J., Ye, Z., Espinoza, C., Agarwahl, S., Shen, L., Ruotti, V., Wang, W., Stewart, R., Thomson, J.A., Ecker, J.R., & Ren, B. 2010. Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell*, 6, (5) 479-491 available from: PM:20452322

Hayek, A., Beattie, G.M., Cirulli, V., Lopez, A.D., Ricordi, C., & Rubin, J.S. 1995. Growth factor/matrix-induced proliferation of human adult beta-cells. *Diabetes*, 44, (12) 1458-1460 available from: PM:7589854

Hebrok, M., Kim, S.K., & Melton, D.A. 1998. Notochord repression of endodermal Sonic hedgehog permits pancreas development. *Genes Dev.*, 12, (11) 1705-1713 available from: PM:9620856

Hedeskov, C.J. 1980. Mechanism of glucose-induced insulin secretion. *Physiol Rev.*, 60, (2) 442-509 available from: PM:6247727

Hennessy, E., Clynes, M., Jeppesen, P.B., & O'Driscoll, L. 2010. Identification of microRNAs with a role in glucose stimulated insulin secretion by expression profiling of MIN6 cells. *Biochem.Biophys.Res.Commun.*, 396, (2) 457-462 available from: PM:20417623

Hering, B.J. & Walawalkar, N. 2009. Pig-to-nonhuman primate islet xenotransplantation. *Transpl.Immunol.*, 21, (2) 81-86 available from: PM:19427901

Hess, D., Li, L., Martin, M., Sakano, S., Hill, D., Strutt, B., Thyssen, S., Gray, D.A., & Bhatia, M. 2003. Bone marrow-derived stem cells initiate pancreatic regeneration. *Nat.Biotechnol.*, 21, (7) 763-770 available from: PM:12819790

Hinton, A., Afrikanova, I., Wilson, M., King, C.C., Maurer, B., Yeo, G.W., Hayek, A., & Pasquinelli, A.E. 2010. A distinct microRNA signature for definitive endoderm derived from human embryonic stem cells. *Stem Cells Dev.*, 19, (6) 797-807 available from: PM:19807270

Hisanaga, E., Park, K.Y., Yamada, S., Hashimoto, H., Takeuchi, T., Mori, M., Seno, M., Umezawa, K., Takei, I., & Kojima, I. 2008. A simple method to induce differentiation of murine bone marrow mesenchymal cells to insulin-producing cells using conophylline and betacellulin-delta4. *Endocr.J*, 55, (3) 535-543 available from: PM:18480554

Hisatomi, Y., Okumura, K., Nakamura, K., Matsumoto, S., Satoh, A., Nagano, K., Yamamoto, T., & Endo, F. 2004. Flow cytometric isolation of endodermal progenitors from mouse salivary gland differentiate into hepatic and pancreatic lineages. *Hepatology*, 39, (3) 667-675 available from: PM:14999685

Hockemeyer, D., Soldner, F., Cook, E.G., Gao, Q., Mitalipova, M., & Jaenisch, R. 2008. A drug-inducible system for direct reprogramming of human somatic cells to pluripotency. *Cell Stem Cell*, 3, (3) 346-353 available from: PM:18786421

Hong, S.G., Winkler, T., Wu, C., Guo, V., Pittaluga, S., Nicolae, A., Donahue, R.E., Metzger, M.E., Price, S.D., Uchida, N., Kuznetsov, S.A., Kilts, T., Li, L., Robey, P.G., & Dunbar, C.E. 2014. Path to the Clinic: Assessment of iPSC-Based Cell Therapies In Vivo in a Nonhuman Primate Model. *Cell Rep.* available from: PM:24835994

Horb, M.E., Shen, C.N., Tosh, D., & Slack, J.M. 2003. Experimental conversion of liver to pancreas. *Curr.Biol.*, 13, (2) 105-115 available from: PM:12546783

Hori, Y., Gu, X., Xie, X., & Kim, S.K. 2005. Differentiation of insulin-producing cells from human neural progenitor cells. *PLoS.Med*, 2, (4) e103 available from: PM:15839736

Hou, P., Li, Y., Zhang, X., Liu, C., Guan, J., Li, H., Zhao, T., Ye, J., Yang, W., Liu, K., Ge, J., Xu, J., Zhang, Q., Zhao, Y., & Deng, H. 2013. Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. *Science*, 341, (6146) 651-654 available from: PM:23868920

Houbaviy, H.B., Murray, M.F., & Sharp, P.A. 2003. Embryonic stem cell-specific MicroRNAs. *Dev.Cell*, 5, (2) 351-358 available from: PM:12919684

Hu, Q., Friedrich, A.M., Johnson, L.V., & Clegg, D.O. 2010. Memory in induced pluripotent stem cells: reprogrammed human retinal-pigmented epithelial cells show tendency for spontaneous redifferentiation. *Stem Cells*, 28, (11) 1981-1991 available from: PM:20882530

Huang, H.P., Chu, K., Nemoz-Gaillard, E., Elberg, D., & Tsai, M.J. 2002. Neogenesis of beta-cells in adult BETA2/NeuroD-deficient mice. *Mol.Endocrinol.*, 16, (3) 541-551 available from: PM:11875114

Huangfu, D., Osafune, K., Maehr, R., Guo, W., Eijkelenboom, A., Chen, S., Muhlestein, W., & Melton, D.A. 2008. Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nat.Biotechnol.*, 26, (11) 1269-1275 available from: PM:18849973

Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A.E., & Melton, D.A. 2008. Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat.Biotechnol.*, 26, (7) 795-797 available from: PM:18568017

Huber, B.C., Ransohoff, J.D., Ransohoff, K.J., Riegler, J., Ebert, A., Kodo, K., Gong, Y., Sanchez-Freire, V., Dey, D., Kooreman, N.G., Diecke, S., Zhang, W.Y., Odegaard, J., Hu, S., Gold, J.D., Robbins, R.C., & Wu, J.C. 2013. Costimulation-adhesion blockade is superior to cyclosporine A and prednisone immunosuppressive therapy for preventing rejection of differentiated human embryonic stem cells following transplantation. *Stem Cells*, 31, (11) 2354-2363 available from: PM:24038578

Huch, M., Bonfanti, P., Boj, S.F., Sato, T., Loomans, C.J., van de Wetering, M., Sojoodi, M., Li, V.S., Schuijers, J., Gracanin, A., Ringnalda, F., Begthel, H., Hamer, K., Mulder, J., van Es, J.H., de, K.E., Vries, R.G., Heimberg, H., & Clevers, H. 2013. Unlimited in vitro expansion of adult bi-potent pancreas progenitors through the Lgr5/R-spondin axis. *EMBO J*, 32, (20) 2708-2721 available from: PM:24045232

Hussein, S.M., Batada, N.N., Vuoristo, S., Ching, R.W., Autio, R., Narva, E., Ng, S., Sourour, M., Hamalainen, R., Olsson, C., Lundin, K., Mikkola, M., Trokovic, R., Peitz, M., Brustle, O., Bazett-Jones, D.P., Alitalo, K., Lahesmaa, R., Nagy, A., & Otonkoski, T. 2011. Copy number variation and selection during reprogramming to pluripotency. *Nature*, 471, (7336) 58-62 available from: PM:21368824

Ianus, A., Holz, G.G., Theise, N.D., & Hussain, M.A. 2003. In vivo derivation of glucose-competent pancreatic endocrine cells from bone marrow without evidence of cell fusion. *J Clin.Invest*, 111, (6) 843-850 available from: PM:12639990

Inada, A., Nienaber, C., Katsuta, H., Fujitani, Y., Levine, J., Morita, R., Sharma, A., & Bonner-Weir, S. 2008. Carbonic anhydrase II-positive pancreatic cells are progenitors for both endocrine and exocrine pancreas after birth. *Proc.Natl.Acad.Sci.U.S.A*, 105, (50) 19915-19919 available from: PM:19052237

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

Ivey, K.N. & Srivastava, D. 2010. MicroRNAs as regulators of differentiation and cell fate decisions. *Cell Stem Cell*, 7, (1) 36-41 available from: PM:20621048

Jackson-Grusby, L., Laird, P.W., Magge, S.N., Moeller, B.J., & Jaenisch, R. 1997. Mutagenicity of 5-aza-2'-deoxycytidine is mediated by the mammalian DNA methyltransferase. *Proc.Natl.Acad.Sci.U.S.A*, 94, (9) 4681-4685 available from: PM:9114051

Jacquemin, P., Durviaux, S.M., Jensen, J., Godfraind, C., Gradwohl, G., Guillemot, F., Madsen, O.D., Carmeliet, P., Dewerchin, M., Collen, D., Rousseau, G.G., & Lemaigre, F.P. 2000. Transcription factor hepatocyte nuclear factor 6 regulates pancreatic endocrine cell differentiation and controls expression of the proendocrine gene ngn3. *Mol.Cell Biol.*, 20, (12) 4445-4454 available from: PM:10825208

Jayawardena, T.M., Egemnazarov, B., Finch, E.A., Zhang, L., Payne, J.A., Pandya, K., Zhang, Z., Rosenberg, P., Mirotsou, M., & Dzau, V.J. 2012. MicroRNA-mediated in vitro and in vivo direct reprogramming of cardiac fibroblasts to cardiomyocytes. *Circ.Res.*, 110, (11) 1465-1473 available from: PM:22539765

Jensen, J., Heller, R.S., Funder-Nielsen, T., Pedersen, E.E., Lindsell, C., Weinmaster, G., Madsen, O.D., & Serup, P. 2000. Independent development of pancreatic alpha- and beta-cells from neurogenin3-expressing precursors: a role for the notch pathway in repression of premature differentiation. *Diabetes*, 49, (2) 163-176 available from: PM:10868931

Jeon, K., Lim, H., Kim, J.H., Thuan, N.V., Park, S.H., Lim, Y.M., Choi, H.Y., Lee, E.R., Kim, J.H., Lee, M.S., & Cho, S.G. 2012. Differentiation and transplantation of functional pancreatic beta cells generated from induced pluripotent stem cells derived from a type 1 diabetes mouse model. *Stem Cells Dev.*, 21, (14) 2642-2655 available from: PM:22512788

Jiang, J., Au, M., Lu, K., Eshpeter, A., Korbutt, G., Fisk, G., & Majumdar, A.S. 2007. Generation of insulin-producing islet-like clusters from human embryonic stem cells. *Stem Cells*, 25, (8) 1940-1953 available from: PM:17510217

Jiang, W., Shi, Y., Zhao, D., Chen, S., Yong, J., Zhang, J., Qing, T., Sun, X., Zhang, P., Ding, M., Li, D., & Deng, H. 2007. In vitro derivation of functional insulin-producing cells from human embryonic stem cells. *Cell Res.*, 17, (4) 333-344 available from: PM:17426693

Jiang, Y., Jahagirdar, B.N., Reinhardt, R.L., Schwartz, R.E., Keene, C.D., Ortiz-Gonzalez, X.R., Reyes, M., Lenvik, T., Lund, T., Blackstad, M., Du, J., Aldrich, S., Lisberg, A., Low, W.C., Largaespada, D.A., & Verfaillie, C.M. 2002. Pluripotency of mesenchymal stem cells derived from adult marrow. *Nature*, 418, (6893) 41-49 available from: PM:12077603

Joglekar, M.V., Parekh, V.S., Mehta, S., Bhonde, R.R., & Hardikar, A.A. 2007. MicroRNA profiling of developing and regenerating pancreas reveal post-transcriptional regulation of neurogenin3. *Dev.Biol.* , 311, (2) 603-612 available from: PM:17936263

Joglekar, M.V., Joglekar, V.M., & Hardikar, A.A. 2009. Expression of islet-specific microRNAs during human pancreatic development. *Gene Expr.Patterns.*, 9, (2) 109-113 available from: PM:18977315

Jones, P.M., Courtney, M.L., Burns, C.J., & Persaud, S.J. 2008. Cell-based treatments for diabetes. *Drug Discov.Today*, 13, (19-20) 888-893 available from: PM:18652911

Jonsson, J., Carlsson, L., Edlund, T., & Edlund, H. 1994. Insulin-promoter-factor 1 is required for pancreas development in mice. *Nature*, 371, (6498) 606-609 available from: PM:7935793

Jorgensen, M.C., Ahnfelt-Ronne, J., Hald, J., Madsen, O.D., Serup, P., & Hecksher-Sorensen, J. 2007. An illustrated review of early pancreas development in the mouse. *Endocr.Rev.*, 28, (6) 685-705 available from: PM:17881611

Juan, A.H., Kumar, R.M., Marx, J.G., Young, R.A., & Sartorelli, V. 2009. Mir-214-dependent regulation of the polycomb protein Ezh2 in skeletal muscle and embryonic stem cells. *Mol.Cell*, 36, (1) 61-74 available from: PM:19818710

Judson, R.L., Babiarz, J.E., Venere, M., & Blelloch, R. 2009. Embryonic stem cell-specific microRNAs promote induced pluripotency. *Nat.Biotechnol.*, 27, (5) 459-461 available from: PM:19363475

Judson, R.L., Greve, T.S., Parchem, R.J., & Blelloch, R. 2013. MicroRNA-based discovery of barriers to dedifferentiation of fibroblasts to pluripotent stem cells. *Nat.Struct.Mol.Biol.*, 20, (10) 1227-1235 available from: PM:24037508

Kahan, B.W., Jacobson, L.M., Hullett, D.A., Ochoada, J.M., Oberley, T.D., Lang, K.M., & Odorico, J.S. 2003. Pancreatic precursors and differentiated islet cell types from murine embryonic stem cells: an in vitro model to study islet differentiation. *Diabetes*, 52, (8) 2016-2024 available from: PM:12882918

Kahn, S.E., Hull, R.L., & Utzschneider, K.M. 2006. Mechanisms linking obesity to insulin resistance and type 2 diabetes. *Nature*, 444, (7121) 840-846 available from: PM:17167471

Kajiwara, M., Aoi, T., Okita, K., Takahashi, R., Inoue, H., Takayama, N., Endo, H., Eto, K., Toguchida, J., Uemoto, S., & Yamanaka, S. 2012. Donor-dependent variations in hepatic differentiation from human-induced pluripotent stem cells. *Proc.Natl.Acad.Sci.U.S.A*, 109, (31) 12538-12543 available from: PM:22802639

Kalis, M., Bolmeson, C., Esguerra, J.L., Gupta, S., Edlund, A., Tormo-Badia, N., Speidel, D., Holmberg, D., Mayans, S., Khoo, N.K., Wendt, A., Eliasson, L., & Cilio, C.M. 2011. Beta-cell specific deletion of Dicer1 leads to defective insulin secretion and diabetes mellitus. *PLoS.One.*, 6, (12) e29166 available from: PM:22216196

Kanai-Azuma, M., Kanai, Y., Gad, J.M., Tajima, Y., Taya, C., Kurohmaru, M., Sanai, Y., Yonekawa, H., Yazaki, K., Tam, P.P., & Hayashi, Y. 2002. Depletion of definitive gut endoderm in Sox17-null mutant mice. *Development*, 129, (10) 2367-2379 available from: PM:11973269

Kanellopoulou, C., Muljo, S.A., Kung, A.L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D.M., & Rajewsky, K. 2005. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev.*, 19, (4) 489-501 available from: PM:15713842

Kaneto, H., Nakatani, Y., Miyatsuka, T., Matsuoka, T.A., Matsuhisa, M., Hori, M., & Yamasaki, Y. 2005. PDX-1/VP16 fusion protein, together with NeuroD or Ngn3, markedly induces insulin gene transcription and ameliorates glucose tolerance. *Diabetes*, 54, (4) 1009-1022 available from: PM:15793239

Karlsson, O., Thor, S., Norberg, T., Ohlsson, H., & Edlund, T. 1990. Insulin gene enhancer binding protein Isl-1 is a member of a novel class of proteins containing both a homeo- and a Cys-His domain. *Nature*, 344, (6269) 879-882 available from: PM:1691825

Karolina, D.S., Armugam, A., Tavintharan, S., Wong, M.T., Lim, S.C., Sum, C.F., & Jeyaseelan, K. 2011. MicroRNA 144 impairs insulin signaling by inhibiting the expression of insulin receptor substrate 1 in type 2 diabetes mellitus. *PLoS.One.*, 6, (8) e22839 available from: PM:21829658

Kaspi, H., Chapnik, E., Levy, M., Beck, G., Hornstein, E., & Soen, Y. 2013. Brief report: miR-290-295 regulate embryonic stem cell differentiation propensities by repressing Pax6. *Stem Cells*, 31, (10) 2266-2272 available from: PM:23843298

Kaufman, D.L., Erlander, M.G., Clare-Salzler, M., Atkinson, M.A., Maclaren, N.K., & Tobin, A.J. 1992. Autoimmunity to two forms of glutamate decarboxylase in insulin-dependent diabetes mellitus. *J.Clin.Invest*, 89, (1) 283-292 available from: PM:1370298

Keller, D.M., McWeeney, S., Arsenlis, A., Drouin, J., Wright, C.V., Wang, H., Wollheim, C.B., White, P., Kaestner, K.H., & Goodman, R.H. 2007. Characterization of pancreatic transcription factor Pdx-1 binding sites using promoter microarray and serial analysis of chromatin occupancy. *J Biol.Chem.*, 282, (44) 32084-32092 available from: PM:17761679

Kelly, O.G., Chan, M.Y., Martinson, L.A., Kadoya, K., Ostertag, T.M., Ross, K.G., Richardson, M., Carpenter, M.K., D'Amour, K.A., Kroon, E., Moorman, M., Baetge, E.E., & Bang, A.G. 2011. Cell-surface markers for the isolation of pancreatic cell types derived from human embryonic stem cells. *Nat.Biotechnol.*, 29, (8) 750-756 available from: PM:21804561

Kelly, W.D., Lillehei, R.C., Merkel, F.K., Idezuki, Y., & Goetz, F.C. 1967. Allotransplantation of the pancreas and duodenum along with the kidney in diabetic nephropathy. *Surgery*, 61, (6) 827-837 available from: PM:5338113

Kim, D., Kim, C.H., Moon, J.I., Chung, Y.G., Chang, M.Y., Han, B.S., Ko, S., Yang, E., Cha, K.Y., Lanza, R., & Kim, K.S. 2009. Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. *Cell Stem Cell*, 4, (6) 472-476 available from: PM:19481515

Kim, J.B., Zaehres, H., Wu, G., Gentile, L., Ko, K., Sebastiano, V., Arauzo-Bravo, M.J., Ruau, D., Han, D.W., Zenke, M., & Scholer, H.R. 2008. Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. *Nature*, 454, (7204) 646-650 available from: PM:18594515

Kim, J.B., Greber, B., Arauzo-Bravo, M.J., Meyer, J., Park, K.I., Zaehres, H., & Scholer, H.R. 2009. Direct reprogramming of human neural stem cells by OCT4. *Nature*, 461, (7264) 649-3 available from: PM:19718018

Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., Kim, J., Aryee, M.J., Ji, H., Ehrlich, L.I., 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., & Daley, G.Q. 2010. Epigenetic memory in induced pluripotent stem cells. *Nature*, 467, (7313) 285-290 available from: PM:20644535

Kim, N., Kim, H., Jung, I., Kim, Y., Kim, D., & Han, Y.M. 2011. Expression profiles of miRNAs in human embryonic stem cells during hepatocyte differentiation. *Hepatol.Res.*, 41, (2) 170-183 available from: PM:21269386

Kim, S., Shin, J.S., Kim, H.J., Fisher, R.C., Lee, M.J., & Kim, C.W. 2007. Streptozotocin-induced diabetes can be reversed by hepatic oval cell activation through hepatic transdifferentiation and pancreatic islet regeneration. *Lab Invest*, 87, (7) 702-712 available from: PM:17483848

Kim, S.K., Hebrok, M., & Melton, D.A. 1997. Notochord to endoderm signaling is required for pancreas development. *Development*, 124, (21) 4243-4252 available from: PM:9334273

Kim, S.K. & Melton, D.A. 1998. Pancreas development is promoted by cyclopamine, a hedgehog signaling inhibitor. *Proc.Natl.Acad.Sci.U.S.A*, 95, (22) 13036-13041 available from: PM:9789036

Kimelman, D. & Griffin, K.J. 2000. Vertebrate mesendoderm induction and patterning. *Curr.Opin.Genet.Dev.*, 10, (4) 350-356 available from: PM:10889062

Kimura, H., Kawasaki, H., & Taira, K. 2004. Mouse microRNA-23b regulates expression of Hes1 gene in P19 cells. *Nucleic Acids Symp.Ser.(Oxf)* (48) 213-214 available from: PM:17150554

Kirk, K., Hao, E., Lahmy, R., & Itkin-Ansari, P. 2014. Human embryonic stem cell derived islet progenitors mature inside an encapsulation device without evidence of increased biomass or cell escape. *Stem Cell Res.*, 12, (3) 807-814 available from: PM:24788136

Klein, D., Misawa, R., Bravo-Egana, V., Vargas, N., Rosero, S., Piroso, J., Ichii, H., Umland, O., Zhijie, J., Tsinoremas, N., Ricordi, C., Inverardi, L., Dominguez-Bendala, J., & Pastori, R.L. 2013. MicroRNA expression in alpha and beta cells of human pancreatic islets. *PLoS.One.* , 8, (1) e55064 available from: PM:23383059

Kloosterman, W.P., Lagendijk, A.K., Ketting, R.F., Moulton, J.D., & Plasterk, R.H. 2007. Targeted inhibition of miRNA maturation with morpholinos reveals a role for miR-375 in pancreatic islet development. *PLoS.Biol.*, 5, (8) e203 available from: PM:17676975

Klymiuk, N., van, B.L., Bahr, A., Offers, M., Kessler, B., Wuensch, A., Kurome, M., Thormann, M., Lochner, K., Nagashima, H., Herbach, N., Wanke, R., Seissler, J., & Wolf, E. 2012. Xenografted islet cell clusters from INSLEA29Y transgenic pigs rescue diabetes and prevent immune rejection in humanized mice. *Diabetes*, 61, (6) 1527-1532 available from: PM:22522620

Kojima, H., Nakamura, T., Fujita, Y., Kishi, A., Fujimiya, M., Yamada, S., Kudo, M., Nishio, Y., Maegawa, H., Haneda, M., Yasuda, H., Kojima, I., Seno, M., Wong, N.C., Kikkawa, R., & Kashiwagi, A. 2002. Combined expression of pancreatic duodenal homeobox 1 and islet factor 1 induces immature enterocytes to produce insulin. *Diabetes*, 51, (5) 1398-1408 available from: PM:11978636

Kojima, H., Fujimiya, M., Matsumura, K., Younan, P., Imaeda, H., Maeda, M., & Chan, L. 2003. NeuroD-betacellulin gene therapy induces islet neogenesis in the liver and reverses diabetes in mice. *Nat.Med*, 9, (5) 596-603 available from: PM:12704384

Kopinke, D. & Murtaugh, L.C. 2010. Exocrine-to-endocrine differentiation is detectable only prior to birth in the uninjured mouse pancreas. *BMC.Dev.Biol.*, 10, 38 available from: PM:20377894

Kopp, J.L., Ormsbee, B.D., Desler, M., & Rizzino, A. 2008. Small increases in the level of Sox2 trigger the differentiation of mouse embryonic stem cells. *Stem Cells*, 26, (4) 903-911 available from: PM:18238855

Kopp, J.L., Dubois, C.L., Schaffer, A.E., Hao, E., Shih, H.P., Seymour, P.A., Ma, J., & Sander, M. 2011. Sox9+ ductal cells are multipotent progenitors throughout development but do not produce new endocrine cells in the normal or injured adult pancreas. *Development*, 138, (4) 653-665 available from: PM:21266405

Korpal, M. & Kang, Y. 2008. The emerging role of miR-200 family of microRNAs in epithelial-mesenchymal transition and cancer metastasis. *RNA.Biol.*, 5, (3) 115-119 available from: PM:19182522

Koyanagi-Aoi, M., Ohnuki, M., Takahashi, K., Okita, K., Noma, H., Sawamura, Y., Teramoto, I., Narita, M., Sato, Y., Ichisaka, T., Amano, N., Watanabe, A., Morizane, A., Yamada, Y., Sato, T., Takahashi, J., & Yamanaka, S. 2013. Differentiation-defective phenotypes revealed by large-scale analyses of human pluripotent stem cells. *Proc.Natl.Acad.Sci.U.S.A* available from: PM:24259714

Kozomara, A. & Griffiths-Jones, S. 2014. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.*, 42, (1) D68-D73 available from: PM:24275495

Kredo-Russo, S., Ness, A., Mandelbaum, A.D., Walker, M.D., & Hornstein, E. 2012. Regulation of pancreatic microRNA-7 expression. *Exp.Diabetes Res.*, 2012, 695214 available from: PM:22675342

Kredo-Russo, S., Mandelbaum, A.D., Ness, A., Alon, I., Lennox, K.A., Behlke, M.A., & Hornstein, E. 2012. Pancreas-enriched miRNA refines endocrine cell differentiation. *Development*, 139, (16) 3021-3031 available from: PM:22764048

Krek, A., Grun, D., Poy, M.N., Wolf, R., Rosenberg, L., Epstein, E.J., MacMenamin, P., da, P., I, Gunsalus, K.C., Stoffel, M., & Rajewsky, N. 2005. Combinatorial microRNA target predictions. *Nat.Genet.* , 37, (5) 495-500 available from: PM:15806104

Kroon, E., Martinson, L.A., Kadoya, K., Bang, A.G., Kelly, O.G., Eliazer, S., Young, H., Richardson, M., Smart, N.G., Cunningham, J., Agulnick, A.D., D'Amour, K.A., Carpenter, M.K., & Baetge, E.E. 2008. Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells in vivo. *Nat.Biotechnol.*, 26, (4) 443-452 available from: PM:18288110

Krupnick, J.G., Damjanov, I., Damjanov, A., Zhu, Z.M., & Fenderson, B.A. 1994. Globo-series carbohydrate antigens are expressed in different forms on human and murine teratocarcinoma-derived cells. *Int.J Cancer*, 59, (5) 692-698 available from: PM:7960243

Krutzfeldt, J., Poy, M.N., & Stoffel, M. 2006. Strategies to determine the biological function of microRNAs. *Nat.Genet.*, 38 Suppl, S14-S19 available from: PM:16736018

Krutzfeldt, J. & Stoffel, M. 2006. MicroRNAs: a new class of regulatory genes affecting metabolism. *Cell Metab*, 4, (1) 9-12 available from: PM:16814728

Kubo, A., Shinozaki, K., Shannon, J.M., Kouskoff, V., Kennedy, M., Woo, S., Fehling, H.J., & Keller, G. 2004. Development of definitive endoderm from embryonic stem cells in culture. *Development*, 131, (7) 1651-1662 available from: PM:14998924

Kunisada, Y., Tsubooka-Yamazoe, N., Shoji, M., & Hosoya, M. 2012. Small molecules induce efficient differentiation into insulin-producing cells from human induced pluripotent stem cells. *Stem Cell Res.* , 8, (2) 274-284 available from: PM:22056147

Kustikova, O., Fehse, B., Modlich, U., Yang, M., Dullmann, J., Kamino, K., von, N.N., Schlegelberger, B., Li, Z., & Baum, C. 2005. Clonal dominance of hematopoietic stem cells triggered by retroviral gene marking. *Science*, 308, (5725) 1171-1174 available from: PM:15905401

Lacy, P.E., Hegre, O.D., Gerasimidi-Vazeou, A., Gentile, F.T., & Dionne, K.E. 1991. Maintenance of normoglycemia in diabetic mice by subcutaneous xenografts of encapsulated islets. *Science*, 254, (5039) 1782-1784 available from: PM:1763328

Lahmy, R., Soleimani, M., Sanati, M.H., Behmanesh, M., Kouhkan, F., & Mobarra, N. 2013. Pancreatic islet differentiation of human embryonic stem cells by microRNA overexpression. *J Tissue Eng Regen.Med* available from: PM:23897763

Lahmy, R., Soleimani, M., Sanati, M.H., Behmanesh, M., Kouhkan, F., & Mobarra, N. 2014. MiRNA-375 promotes beta pancreatic differentiation in human induced pluripotent stem (hiPS) cells. *Mol.Biol.Rep.*, 41, (4) 2055-2066 available from: PM:24469711

Lakshmipathy, U., Davila, J., & Hart, R.P. 2010. miRNA in pluripotent stem cells. *Regen.Med*, 5, (4) 545-555 available from: PM:20632858

Lammert, E., Cleaver, O., & Melton, D. 2001. Induction of pancreatic differentiation by signals from blood vessels. *Science*, 294, (5542) 564-567 available from: PM:11577200

Lan, M.S., Wasserfall, C., Maclaren, N.K., & Notkins, A.L. 1996. IA-2, a transmembrane protein of the protein tyrosine phosphatase family, is a major autoantigen in insulin-dependent diabetes mellitus. *Proc.Natl.Acad.Sci.U.S.A*, 93, (13) 6367-6370 available from: PM:8692821

Langer, R.M. 2010. Islet transplantation: lessons learned since the Edmonton breakthrough. *Transplant.Proc.*, 42, (5) 1421-1424 available from: PM:20620447

Lanzoni, G., Oikawa, T., Wang, Y., Cui, C.B., Carpino, G., Cardinale, V., Gerber, D., Gabriel, M., Dominguez-Bendala, J., Furth, M.E., Gaudio, E., Alvaro, D., Inverardi, L., & Reid, L.M. 2013. Concise

review: Clinical programs of stem cell therapies for liver and pancreas. *Stem Cells*, 31, (10) 2047-2060 available from: PM:23873634

Laurent, L.C., Chen, J., Ulitsky, I., Mueller, F.J., Lu, C., Shamir, R., Fan, J.B., & Loring, J.F. 2008. Comprehensive microRNA profiling reveals a unique human embryonic stem cell signature dominated by a single seed sequence. *Stem Cells*, 26, (6) 1506-1516 available from: PM:18403753

Laurent, L.C., Ulitsky, I., Slavin, I., Tran, H., Schork, A., Morey, R., Lynch, C., Harness, J.V., Lee, S., Barrero, M.J., Ku, S., Martynova, M., Semechkin, R., Galat, V., Gottesfeld, J., Izpisua Belmonte, J.C., Murry, C., Keirstead, H.S., Park, H.S., Schmidt, U., Laslett, A.L., Muller, F.J., Nievergelt, C.M., Shamir, R., & Loring, J.F. 2011. Dynamic changes in the copy number of pluripotency and cell proliferation genes in human ESCs and iPSCs during reprogramming and time in culture. *Cell Stem Cell*, 8, (1) 106-118 available from: PM:21211785

Lavon, N., Yanuka, O., & Benvenisty, N. 2006. The effect of overexpression of Pdx1 and Foxa2 on the differentiation of human embryonic stem cells into pancreatic cells. *Stem Cells*, 24, (8) 1923-1930 available from: PM:16675598

Lawson, K.A., Meneses, J.J., & Pedersen, R.A. 1991. Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. *Development*, 113, (3) 891-911 available from: PM:1821858

Lechner, A., Yang, Y.G., Blacken, R.A., Wang, L., Nolan, A.L., & Habener, J.F. 2004. No evidence for significant transdifferentiation of bone marrow into pancreatic beta-cells in vivo. *Diabetes*, 53, (3) 616-623 available from: PM:14988245

Lechner, A., Nolan, A.L., Blacken, R.A., & Habener, J.F. 2005. Redifferentiation of insulin-secreting cells after in vitro expansion of adult human pancreatic islet tissue. *Biochem.Biophys.Res.Commun.*, 327, (2) 581-588 available from: PM:15629153

Lee, A.S., Tang, C., Rao, M.S., Weissman, I.L., & Wu, J.C. 2013. Tumorigenicity as a clinical hurdle for pluripotent stem cell therapies. *Nat.Med*, 19, (8) 998-1004 available from: PM:23921754

Lee, J.C., Smith, S.B., Watada, H., Lin, J., Scheel, D., Wang, J., Mirmira, R.G., & German, M.S. 2001. Regulation of the pancreatic pro-endocrine gene neurogenin3. *Diabetes*, 50, (5) 928-936 available from: PM:11334435

Lee, M.O., Moon, S.H., Jeong, H.C., Yi, J.Y., Lee, T.H., Shim, S.H., Rhee, Y.H., Lee, S.H., Oh, S.J., Lee, M.Y., Han, M.J., Cho, Y.S., Chung, H.M., Kim, K.S., & Cha, H.J. 2013. Inhibition of pluripotent stem cell-derived teratoma formation by small molecules. *Proc.Natl.Acad.Sci.U.S.A*, 110, (35) E3281-E3290 available from: PM:23918355

Lee, R.C., Feinbaum, R.L., & Ambros, V. 1993. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell*, 75, (5) 843-854 available from: PM:8252621

Lefort, N., Feyeux, M., Bas, C., Feraud, O., Bennaceur-Griscelli, A., Tachdjian, G., Peschanski, M., & Perrier, A.L. 2008. Human embryonic stem cells reveal recurrent genomic instability at 20q11.21. *Nat.Biotechnol.*, 26, (12) 1364-1366 available from: PM:19029913

Leonard, J., Peers, B., Johnson, T., Ferreri, K., Lee, S., & Montminy, M.R. 1993. Characterization of somatostatin transactivating factor-1, a novel homeobox factor that stimulates somatostatin expression in pancreatic islet cells. *Mol.Endocrinol.*, 7, (10) 1275-1283 available from: PM:7505393

Levine, F., Wang, S., Beattie, G.M., Mally, M.I., Cirulli, V., Lopez, A.D., & Hayek, A. 1995. Development of a cell line from the human fetal pancreas. *Transplant.Proc.*, 27, (6) 3410 available from: PM:8540023

Lewis, S.L. & Tam, P.P. 2006. Definitive endoderm of the mouse embryo: formation, cell fates, and morphogenetic function. *Dev.Dyn.*, 235, (9) 2315-2329 available from: PM:16752393

Li, H., Arber, S., Jessell, T.M., & Edlund, H. 1999. Selective agenesis of the dorsal pancreas in mice lacking homeobox gene Hlxb9. *Nat.Genet.*, 23, (1) 67-70 available from: PM:10471501

Li, K., Zhu, S., Russ, H.A., Xu, S., Xu, T., Zhang, Y., Ma, T., Hebrok, M., & Ding, S. 2014. Small molecules facilitate the reprogramming of mouse fibroblasts into pancreatic lineages. *Cell Stem Cell*, 14, (2) 228-236 available from: PM:24506886

Li, W.C., Rukstalis, J.M., Nishimura, W., Tchipashvili, V., Habener, J.F., Sharma, A., & Bonner-Weir, S. 2010. Activation of pancreatic-duct-derived progenitor cells during pancreas regeneration in adult rats. *J Cell Sci.*, 123, (Pt 16) 2792-2802 available from: PM:20663919

Li, Y., Zhang, R., Qiao, H., Zhang, H., Wang, Y., Yuan, H., Liu, Q., Liu, D., Chen, L., & Pei, X. 2007. Generation of insulin-producing cells from PDX-1 gene-modified human mesenchymal stem cells. *J Cell Physiol*, 211, (1) 36-44 available from: PM:17226789

Liao, B., Bao, X., Liu, L., Feng, S., Zovoilis, A., Liu, W., Xue, Y., Cai, J., Guo, X., Qin, B., Zhang, R., Wu, J., Lai, L., Teng, M., Niu, L., Zhang, B., Esteban, M.A., & Pei, D. 2011. MicroRNA cluster 302-367 enhances somatic cell reprogramming by accelerating a mesenchymal-to-epithelial transition. *J Biol.Chem.*, 286, (19) 17359-17364 available from: PM:21454525

Liao, X., Xue, H., Wang, Y.C., Nazor, K.L., Guo, S., Trivedi, N., Peterson, S.E., Liu, Y., Loring, J.F., & Laurent, L.C. 2013. Matched miRNA and mRNA signatures from an hESC-based in vitro model of pancreatic differentiation reveal novel regulatory interactions. *J Cell Sci.*, 126, (Pt 17) 3848-3861 available from: PM:23813959

Liew, C.G., Shah, N.N., Briston, S.J., Shepherd, R.M., Khoo, C.P., Dunne, M.J., Moore, H.D., Cosgrove, K.E., & Andrews, P.W. 2008. PAX4 enhances beta-cell differentiation of human embryonic stem cells. *PLoS.One.*, 3, (3) e1783 available from: PM:18335054

Lin, C.H., Jackson, A.L., Guo, J., Linsley, P.S., & Eisenman, R.N. 2009. Myc-regulated microRNAs attenuate embryonic stem cell differentiation. *EMBO J*, 28, (20) 3157-3170 available from: PM:19745813

Lin, H.T., Kao, C.L., Lee, K.H., Chang, Y.L., Chiou, S.H., Tsai, F.T., Tsai, T.H., Sheu, D.C., Ho, L.L., & Ku, H.H. 2007. Enhancement of insulin-producing cell differentiation from embryonic stem cells using pax4-nucleofection method. *World J Gastroenterol.*, 13, (11) 1672-1679 available from: PM:17461469

Lin, S.L., Chang, D.C., Ying, S.Y., Leu, D., & Wu, D.T. 2010. MicroRNA miR-302 inhibits the tumorigenecity of human pluripotent stem cells by coordinate suppression of the CDK2 and CDK4/6 cell cycle pathways. *Cancer Res.*, 70, (22) 9473-9482 available from: PM:21062975

Lister, R., Pelizzola, M., Kida, Y.S., Hawkins, R.D., Nery, J.R., Hon, G., Antosiewicz-Bourget, J., O'Malley, R., Castanon, R., Klugman, S., Downes, M., Yu, R., Stewart, R., Ren, B., Thomson, J.A., Evans, R.M., & Ecker, J.R. 2011. Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature*, 471, (7336) 68-73 available from: PM:21289626

Liu, H., Zhu, F., Yong, J., Zhang, P., Hou, P., Li, H., Jiang, W., Cai, J., Liu, M., Cui, K., Qu, X., Xiang, T., Lu, D., Chi, X., Gao, G., Ji, W., Ding, M., & Deng, H. 2008. Generation of induced pluripotent stem cells from adult rhesus monkey fibroblasts. *Cell Stem Cell*, 3, (6) 587-590 available from: PM:19041774

Liu, L., Luo, G.Z., Yang, W., Zhao, X., Zheng, Q., Lv, Z., Li, W., Wu, H.J., Wang, L., Wang, X.J., & Zhou, Q. 2010. Activation of the imprinted Dlk1-Dio3 region correlates with pluripotency levels of mouse stem cells. *J Biol.Chem.*, 285, (25) 19483-19490 available from: PM:20382743

Liu, P., Chen, S., Li, X., Qin, L., Huang, K., Wang, L., Huang, W., Li, S., Jia, B., Zhong, M., Pan, G., Cai, J., & Pei, D. 2013. Low immunogenicity of neural progenitor cells differentiated from induced pluripotent stem cells derived from less immunogenic somatic cells. *PLoS.One.* , 8, (7) e69617 available from: PM:23922758

Liu, S.H. & Lee, L.T. 2012. Efficient differentiation of mouse embryonic stem cells into insulin-producing cells. *Exp.Diabetes Res.*, 2012, 201295 available from: PM:22919367

Lizcano, J.M. & Alessi, D.R. 2002. The insulin signalling pathway. *Curr.Biol.*, 12, (7) R236-R238 available from: PM:11937037

Loh, K.M., Ang, L.T., Zhang, J., Kumar, V., Ang, J., Auyeong, J.Q., Lee, K.L., Choo, S.H., Lim, C.Y., Nichane, M., Tan, J., Noghabi, M.S., Azzola, L., Ng, E.S., Durruthy-Durruthy, J., Sebastiano, V., Poellinger, L., Elefanty, A.G., Stanley, E.G., Chen, Q., Prabhakar, S., Weissman, I.L., & Lim, B. 2014. Efficient endoderm induction from human pluripotent stem cells by logically directing signals controlling lineage bifurcations. *Cell Stem Cell*, 14, (2) 237-252 available from: PM:24412311

Loh, Y.H., Hartung, O., Li, H., Guo, C., Sahalie, J.M., Manos, P.D., Urbach, A., Heffner, G.C., Grskovic, M., Vigneault, F., Lensch, M.W., Park, I.H., Agarwal, S., Church, G.M., Collins, J.J., Irion, S., & Daley, G.Q. 2010. Reprogramming of T cells from human peripheral blood. *Cell Stem Cell*, 7, (1) 15-19 available from: PM:20621044

Lohle, M., Hermann, A., Glass, H., Kempe, A., Schwarz, S.C., Kim, J.B., Poulet, C., Ravens, U., Schwarz, J., Scholer, H.R., & Storch, A. 2012. Differentiation efficiency of induced pluripotent stem cells depends on the number of reprogramming factors. *Stem Cells*, 30, (3) 570-579 available from: PM:22213586

Lois, C., Hong, E.J., Pease, S., Brown, E.J., & Baltimore, D. 2002. Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science*, 295, (5556) 868-872 available from: PM:11786607

Lovis, P., Gattesco, S., & Regazzi, R. 2008. Regulation of the expression of components of the exocytotic machinery of insulin-secreting cells by microRNAs. *Biol.Chem.*, 389, (3) 305-312 available from: PM:18177263

Lu, J., Li, Q., Xie, H., Chen, Z.J., Borovitskaya, A.E., Maclaren, N.K., Notkins, A.L., & Lan, M.S. 1996. Identification of a second transmembrane protein tyrosine phosphatase, IA-2beta, as an autoantigen in insulin-dependent diabetes mellitus: precursor of the 37-kDa tryptic fragment. *Proc.Natl.Acad.Sci.U.S.A*, 93, (6) 2307-2311 available from: PM:8637868

Luca, G., Basta, G., Calafiore, R., Rossi, C., Giovagnoli, S., Esposito, E., & Nastruzzi, C. 2003. Multifunctional microcapsules for pancreatic islet cell entrapment: design, preparation and in vitro characterization. *Biomaterials*, 24, (18) 3101-3114 available from: PM:12895583

Lumelsky, N., Blondel, O., Laeng, P., Velasco, I., Ravin, R., & McKay, R. 2001. Differentiation of embryonic stem cells to insulin-secreting structures similar to pancreatic islets. *Science*, 292, (5520) 1389-1394 available from: PM:11326082

Lynn, F.C., Skewes-Cox, P., Kosaka, Y., McManus, M.T., Harfe, B.D., & German, M.S. 2007. MicroRNA expression is required for pancreatic islet cell genesis in the mouse. *Diabetes*, 56, (12) 2938-2945 available from: PM:17804764

Macarthur, C.C., Fontes, A., Ravinder, N., Kuninger, D., Kaur, J., Bailey, M., Taliana, A., Vemuri, M.C., & Lieu, P.T. 2012. Generation of human-induced pluripotent stem cells by a nonintegrating RNA Sendai virus vector in feeder-free or xeno-free conditions. *Stem Cells Int.*, 2012, 564612 available from: PM:22550511

MacFarlane, W.M., Campbell, S.C., Elrick, L.J., Oates, V., Bermano, G., Lindley, K.J., Aynsley-Green, A., Dunne, M.J., James, R.F., & Docherty, K. 2000. Glucose regulates islet amyloid polypeptide gene transcription in a. *J.Biol.Chem.*, 275, (20) 15330-15335 available from: PM:10748090

Maehr, R., Chen, S., Snitow, M., Ludwig, T., Yagasaki, L., Goland, R., Leibel, R.L., & Melton, D.A. 2009. Generation of pluripotent stem cells from patients with type 1 diabetes. *Proc.Natl.Acad.Sci.U.S.A*, 106, (37) 15768-15773 available from: PM:19720998

Maehr, R. 2011. iPS cells in type 1 diabetes research and treatment. *Clin.Pharmacol.Ther.*, 89, (5) 750-753 available from: PM:21430661

Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R., Plath, K., & Hochedlinger, K. 2007. Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell*, 1, (1) 55-70 available from: PM:18371336

Maherali, N., Ahfeldt, T., Rigamonti, A., Utikal, J., Cowan, C., & Hochedlinger, K. 2008. A high-efficiency system for the generation and study of human induced pluripotent stem cells. *Cell Stem Cell*, 3, (3) 340-345 available from: PM:18786420

Maherali, N. & Hochedlinger, K. 2008. Guidelines and techniques for the generation of induced pluripotent stem cells. *Cell Stem Cell*, 3, (6) 595-605 available from: PM:19041776

Maitra, A., Arking, D.E., Shivapurkar, N., Ikeda, M., Stastny, V., Kassauei, K., Sui, G., Cutler, D.J., Liu, Y., Brimble, S.N., Noaksson, K., Hyllner, J., Schulz, T.C., Zeng, X., Freed, W.J., Crook, J., Abraham, S., Colman, A., Sartipy, P., Matsui, S., Carpenter, M., Gazdar, A.F., Rao, M., & Chakravarti, A. 2005. Genomic alterations in cultured human embryonic stem cells. *Nat.Genet.*, 37, (10) 1099-1103 available from: PM:16142235

Mallanna, S.K. & Rizzino, A. 2010. Emerging roles of microRNAs in the control of embryonic stem cells and the generation of induced pluripotent stem cells. *Dev.Biol.*, 344, (1) 16-25 available from: PM:20478297

Mallon, B.S., Chenoweth, J.G., Johnson, K.R., Hamilton, R.S., Tesar, P.J., Yavatkar, A.S., Tyson, L.J., Park, K., Chen, K.G., Fann, Y.C., & McKay, R.D. 2013. StemCellDB: the human pluripotent stem cell database at the National Institutes of Health. *Stem Cell Res.*, 10, (1) 57-66 available from: PM:23117585

Mallon, B.S., Hamilton, R.S., Kozhich, O.A., Johnson, K.R., Fann, Y.C., Rao, M.S., & Robey, P.G. 2014. Comparison of the molecular profiles of human embryonic and induced pluripotent stem cells of isogenic origin. *Stem Cell Res.*, 12, (2) 376-386 available from: PM:24374290

Marchetto, M.C., Yeo, G.W., Kainohana, O., Marsala, M., Gage, F.H., & Muotri, A.R. 2009. Transcriptional signature and memory retention of human-induced pluripotent stem cells. *PLoS.One.*, 4, (9) e7076 available from: PM:19763270

Markoulaki, S., Hanna, J., Beard, C., Carey, B.W., Cheng, A.W., Lengner, C.J., Dausman, J.A., Fu, D., Gao, Q., Wu, S., Cassady, J.P., & Jaenisch, R. 2009. Transgenic mice with defined combinations of drug-inducible reprogramming factors. *Nat.Biotechnol.*, 27, (2) 169-171 available from: PM:19151700

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., & Young, R.A. 2008. Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell*, 134, (3) 521-533 available from: PM:18692474

Martin, G.R. 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc.Natl.Acad.Sci.U.S.A*, 78, (12) 7634-7638 available from: PM:6950406

Martin, M.J., Muotri, A., Gage, F., & Varki, A. 2005. Human embryonic stem cells express an immunogenic nonhuman sialic acid. *Nat.Med*, 11, (2) 228-232 available from: PM:15685172

Martinez, N.J. & Gregory, R.I. 2010. MicroRNA gene regulatory pathways in the establishment and maintenance of ESC identity. *Cell Stem Cell*, 7, (1) 31-35 available from: PM:20621047

Martinez, Y., Dubois-Dauphin, M., & Krause, K.H. 2012. Generation and applications of human pluripotent stem cells induced into neural lineages and neural tissues. *Front Physiol*, 3, 47 available from: PM:22457650

Martini F.H. 2006. *Fundamentals of anatomy and physiology (7th edition)*, 7 ed. Benjamin Cummings.

Martins-Taylor, K., Nisler, B.S., Taapken, S.M., Compton, T., Crandall, L., Montgomery, K.D., Lalande, M., & Xu, R.H. 2011. Recurrent copy number variations in human induced pluripotent stem cells. *Nat.Biotechnol.*, 29, (6) 488-491 available from: PM:21654665

Massumi, M., Hoveizi, E., Baktash, P., Hooti, A., Ghazizadeh, L., Nadri, S., Pourasgari, F., Hajarizadeh, A., Soleimani, M., Nabiuni, M., & Khorramizadeh, M.R. 2014. Efficient programming of human eye conjunctiva-derived induced pluripotent stem (ECiPS) cells into definitive endoderm-like cells. *Exp.Cell Res.*, 322, (1) 51-61 available from: PM:24434355

Matsuoka, T.A., Zhao, L., Artner, I., Jarrett, H.W., Friedman, D., Means, A., & Stein, R. 2003. Members of the large Maf transcription family regulate insulin gene transcription in islet beta cells. *Mol.Cell Biol.*, 23, (17) 6049-6062 available from: PM:12917329

Matsuoka, T.A., Artner, I., Henderson, E., Means, A., Sander, M., & Stein, R. 2004. The MafA transcription factor appears to be responsible for tissue-specific expression of insulin. *Proc.Natl.Acad.Sci.U.S.A*, 101, (9) 2930-2933 available from: PM:14973194

Mayshar, Y., Ben-David, U., Lavon, N., Biancotti, J.C., Yakir, B., Clark, A.T., Plath, K., Lowry, W.E., & Benvenisty, N. 2010. Identification and classification of chromosomal aberrations in human induced pluripotent stem cells. *Cell Stem Cell*, 7, (4) 521-531 available from: PM:20887957

McCarthy, M.I. & Zeggini, E. 2009. Genome-wide association studies in type 2 diabetes. *Curr.Diab.Rep.*, 9, (2) 164-171 available from: PM:19323962

McGrath, K.E., Koniski, A.D., Maltby, K.M., McGann, J.K., & Palis, J. 1999. Embryonic expression and function of the chemokine SDF-1 and its receptor, CXCR4. *Dev.Biol.*, 213, (2) 442-456 available from: PM:10479460

McLean, A.B., D'Amour, K.A., Jones, K.L., Krishnamoorthy, M., Kulik, M.J., Reynolds, D.M., Sheppard, A.M., Liu, H., Xu, Y., Baetge, E.E., & Dalton, S. 2007. Activin a efficiently specifies definitive endoderm from human embryonic stem cells only when phosphatidylinositol 3-kinase signaling is suppressed. *Stem Cells*, 25, (1) 29-38 available from: PM:17204604

Meda, P., Bosco, D., Chanson, M., Giordano, E., Vallar, L., Wollheim, C., & Orci, L. 1990. Rapid and reversible secretion changes during uncoupling of rat insulin-producing cells. *J Clin.Invest*, 86, (3) 759-768 available from: PM:1697604

Meissner, A., Wernig, M., & Jaenisch, R. 2007. Direct reprogramming of genetically unmodified fibroblasts into pluripotent stem cells. *Nat.Biotechnol.*, 25, (10) 1177-1181 available from: PM:17724450

Melkman-Zehavi, T., Oren, R., Kredo-Russo, S., Shapira, T., Mandelbaum, A.D., Rivkin, N., Nir, T., Lennox, K.A., Behlke, M.A., Dor, Y., & Hornstein, E. 2011. miRNAs control insulin content in pancreatic beta-cells via downregulation of transcriptional repressors. *EMBO J*, 30, (5) 835-845 available from: PM:21285947

Melton, C., Judson, R.L., & Blelloch, R. 2010. Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. *Nature*, 463, (7281) 621-626 available from: PM:20054295

Meng, G., Liu, S., Li, X., Krawetz, R., & Rancourt, D.E. 2010. Extracellular matrix isolated from foreskin fibroblasts supports long-term xeno-free human embryonic stem cell culture. *Stem Cells Dev.*, 19, (4) 547-556 available from: PM:19883201

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P., Lee, W., Mendenhall, E., O'Donovan, A., Presser, A., Russ, C., Xie, X., Meissner, A., Wernig, M., Jaenisch, R., Nusbaum, C., Lander, E.S., & Bernstein, B.E. 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*, 448, (7153) 553-560 available from: PM:17603471

Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S., & Meissner, A. 2008. Dissecting direct reprogramming through integrative genomic analysis. *Nature*, 454, (7200) 49-55 available from: PM:18509334

Miller, C.P., McGehee, R.E., Jr., & Habener, J.F. 1994. IDX-1: a new homeodomain transcription factor expressed in rat pancreatic islets and duodenum that transactivates the somatostatin gene. *EMBO J.*, 13, (5) 1145-1156 available from: PM:7907546

Min, H. & Yoon, S. 2010. Got target? Computational methods for microRNA target prediction and their extension. *Exp.Mol.Med*, 42, (4) 233-244 available from: PM:20177143

Minger, S.L., Fisher, L.J., Ray, J., & Gage, F.H. 1996. Long-term survival of transplanted basal forebrain cells following in vitro propagation with fibroblast growth factor-2. *Exp.Neurol.*, 141, (1) 12-24 available from: PM:8797663

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., & Yamanaka, S. 2003. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell*, 113, (5) 631-642 available from: PM:12787504

Miura, K., Okada, Y., Aoi, T., Okada, A., Takahashi, K., Okita, K., Nakagawa, M., Koyanagi, M., Tanabe, K., Ohnuki, M., Ogawa, D., Ikeda, E., Okano, H., & Yamanaka, S. 2009. Variation in the safety of induced pluripotent stem cell lines. *Nat.Biotechnol.*, 27, (8) 743-745 available from: PM:19590502

Miyatsuka, T., Kaneto, H., Shiraiwa, T., Matsuoka, T.A., Yamamoto, K., Kato, K., Nakamura, Y., Akira, S., Takeda, K., Kajimoto, Y., Yamasaki, Y., Sandgren, E.P., Kawaguchi, Y., Wright, C.V., & Fujitani, Y. 2006. Persistent expression of PDX-1 in the pancreas causes acinar-to-ductal metaplasia through Stat3 activation. *Genes Dev.*, 20, (11) 1435-1440 available from: PM:16751181

Miyazaki, J., Araki, K., Yamato, E., Ikegami, H., Asano, T., Shibasaki, Y., Oka, Y., & Yamamura, K. 1990. Establishment of a pancreatic beta cell line that retains glucose-inducible insulin secretion: special reference to expression of glucose transporter isoforms. *Endocrinology*, 127, (1) 126-132 available from: PM:2163307

Miyazaki, S., Yamato, E., & Miyazaki, J. 2004. Regulated expression of pdx-1 promotes in vitro differentiation of insulin-producing cells from embryonic stem cells. *Diabetes*, 53, (4) 1030-1037 available from: PM:15047618

Miyoshi, N., Ishii, H., Nagano, H., Haraguchi, N., Dewi, D.L., Kano, Y., Nishikawa, S., Tanemura, M., Mimori, K., Tanaka, F., Saito, T., Nishimura, J., Takemasa, I., Mizushima, T., Ikeda, M., Yamamoto, H., Sekimoto, M., Doki, Y., & Mori, M. 2011. Reprogramming of mouse and human cells to pluripotency using mature microRNAs. *Cell Stem Cell*, 8, (6) 633-638 available from: PM:21620789

Monaghan, A.P., Kaestner, K.H., Grau, E., & Schutz, G. 1993. Postimplantation expression patterns indicate a role for the mouse forkhead/HNF-3 alpha, beta and gamma genes in determination of the definitive endoderm, chordamesoderm and neuroectoderm. *Development*, 119, (3) 567-578 available from: PM:8187630

Monaghan, A.P., Kaestner, K.H., Grau, E., & Schutz, G. 1993. Postimplantation expression patterns indicate a role for the mouse forkhead/HNF-3 alpha, beta and gamma genes in determination of the definitive endoderm, chordamesoderm and neuroectoderm. *Development*, 119, (3) 567-578 available from: PM:8187630

Morin, R.D., O'Connor, M.D., Griffith, M., Kuchenbauer, F., Delaney, A., Prabhu, A.L., Zhao, Y., McDonald, H., Zeng, T., Hirst, M., Eaves, C.J., & Marra, M.A. 2008. Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res.*, 18, (4) 610-621 available from: PM:18285502

Morita, S., Kojima, T., & Kitamura, T. 2000. Plat-E: an efficient and stable system for transient packaging of retroviruses. *Gene Ther.*, 7, (12) 1063-1066 available from: PM:10871756

Morita, S., Hara, A., Kojima, I., Horii, T., Kimura, M., Kitamura, T., Ochiya, T., Nakanishi, K., Matoba, R., Matsubara, K., & Hatada, I. 2009. Dicer is required for maintaining adult pancreas. *PLoS.One.*, 4, (1) e4212 available from: PM:19148298

Morrison, G.M., Oikonomopoulou, I., Migueles, R.P., Soneji, S., Livigni, A., Enver, T., & Brickman, J.M. 2008. Anterior definitive endoderm from ESCs reveals a role for FGF signaling. *Cell Stem Cell*, 3, (4) 402-415 available from: PM:18940732

Muir, K.R., Lima, M.J., Docherty, H.M., & Docherty, K. 2014. Cell Therapy for Type 1 Diabetes. *QJM.* available from: PM:24481872

Muller, F.J., Brandl, B., & Loring, J.F. 2008. Assessment of human pluripotent stem cells with PluriTest. available from: PM:23658970

Muller, F.J., Laurent, L.C., Kostka, D., Ulitsky, I., Williams, R., Lu, C., Park, I.H., Rao, M.S., Shamir, R., Schwartz, P.H., Schmidt, N.O., & Loring, J.F. 2008. Regulatory networks define phenotypic classes of human stem cell lines. *Nature*, 455, (7211) 401-405 available from: PM:18724358

Muller, L.U., Schlaeger, T.M., DeVine, A.L., & Williams, D.A. 2012. Induced pluripotent stem cells as a tool for gaining new insights into Fanconi anemia. *Cell Cycle*, 11, (16) 2985-2990 available from: PM:22825249

Murry, C.E. & Keller, G. 2008. Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. *Cell*, 132, (4) 661-680 available from: PM:18295582

Nair, S. & Schilling, T.F. 2008. Chemokine signaling controls endodermal migration during zebrafish gastrulation. *Science* , 322, (5898) 89-92 available from: PM:18719251

Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y., Takizawa, N., & Yamanaka, S. 2008. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat.Biotechnol.*, 26, (1) 101-106 available from: PM:18059259

Nakajima-Nagata, N., Sakurai, T., Mitaka, T., Katakai, T., Yamato, E., Miyazaki, J., Tabata, Y., Sugai, M., & Shimizu, A. 2004. In vitro induction of adult hepatic progenitor cells into insulin-producing cells. *Biochem.Biophys.Res.Commun.*, 318, (3) 625-630 available from: PM:15144883

Narsinh, K.H., Sun, N., Sanchez-Freire, V., Lee, A.S., Almeida, P., Hu, S., Jan, T., Wilson, K.D., Leong, D., Rosenberg, J., Yao, M., Robbins, R.C., & Wu, J.C. 2011. Single cell transcriptional profiling reveals heterogeneity of human induced pluripotent stem cells. *J Clin.Invest*, 121, (3) 1217-1221 available from: PM:21317531

Narushima, M., Kobayashi, N., Okitsu, T., Tanaka, Y., Li, S.A., Chen, Y., Miki, A., Tanaka, K., Nakaji, S., Takei, K., Gutierrez, A.S., Rivas-Carrillo, J.D., Navarro-Alvarez, N., Jun, H.S., Westerman, K.A., Noguchi, H., Lakey, J.R., Leboulch, P., Tanaka, N., & Yoon, J.W. 2005. A human beta-cell line for transplantation therapy to control type 1 diabetes. *Nat.Biotechnol.*, 23, (10) 1274-1282 available from: PM:16186810

Narva, E., Autio, R., Rahkonen, N., Kong, L., Harrison, N., Kitsberg, D., Borghese, L., Itskovitz-Eldor, J., Rasool, O., Dvorak, P., Hovatta, O., Otonkoski, T., Tuuri, T., Cui, W., Brustle, O., Baker, D., Maltby, E., Moore, H.D., Benvenisty, N., Andrews, P.W., Yli-Harja, O., & Lahesmaa, R. 2010. High-resolution DNA

analysis of human embryonic stem cell lines reveals culture-induced copy number changes and loss of heterozygosity. *Nat.Biotechnol.*, 28, (4) 371-377 available from: PM:20351689

National Diabetes Information Clearinghouse. DCCT and EDIC: The Diabetes Control and Complications Trial and Follow-up Study. 21113.
Ref Type: Online Source

Naujok, O., Francini, F., Picton, S., Bailey, C.J., Lenzen, S., & Jorns, A. 2009. Changes in gene expression and morphology of mouse embryonic stem cells on differentiation into insulin-producing cells in vitro and in vivo. *Diabetes Metab Res.Rev.*, 25, (5) 464-476 available from: PM:19425055

Naya, F.J., Stellrecht, C.M., & Tsai, M.J. 1995. Tissue-specific regulation of the insulin gene by a novel basic helix-loop-helix transcription factor. *Genes Dev.*, 9, (8) 1009-1019 available from: PM:7774807

Neveu, P., Kye, M.J., Qi, S., Buchholz, D.E., Clegg, D.O., Sahin, M., Park, I.H., Kim, K.S., Daley, G.Q., Kornblum, H.I., Shraiman, B.I., & Kosik, K.S. 2010. MicroRNA profiling reveals two distinct p53-related human pluripotent stem cell states. *Cell Stem Cell*, 7, (6) 671-681 available from: PM:21112562

Newman, A.M. & Cooper, J.B. 2010. Lab-specific gene expression signatures in pluripotent stem cells. *Cell Stem Cell*, 7, (2) 258-262 available from: PM:20682451

Nieto, M., Hevia, P., Garcia, E., Klein, D., Alvarez-Cubela, S., Bravo-Egana, V., Rosero, S., Damaris, M.R., Vargas, N., Ricordi, C., Pileggi, A., Diez, J., Dominguez-Bendala, J., & Pastori, R.L. 2012. Antisense miR-7 impairs insulin expression in developing pancreas and in cultured pancreatic buds. *Cell Transplant.*, 21, (8) 1761-1774 available from: PM:22186137

NIH. Stem Cell Information. 2001.
Ref Type: Online Source

Nir, T., Melton, D.A., & Dor, Y. 2007. Recovery from diabetes in mice by beta cell regeneration. *J Clin.Invest*, 117, (9) 2553-2561 available from: PM:17786244

Nishio, J., Gaglia, J.L., Turvey, S.E., Campbell, C., Benoist, C., & Mathis, D. 2006. Islet recovery and reversal of murine type 1 diabetes in the absence of any infused spleen cell contribution. *Science*, 311, (5768) 1775-1778 available from: PM:16556845

Niwa, H., Miyazaki, J., & Smith, A.G. 2000. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat.Genet.*, 24, (4) 372-376 available from: PM:10742100

Norrman, K., Strombeck, A., Semb, H., & Stahlberg, A. 2013. Distinct gene expression signatures in human embryonic stem cells differentiated towards definitive endoderm at single-cell level. *Methods*, 59, (1) 59-70 available from: PM:22503774

Nostro, M.C., Sarangi, F., Ogawa, S., Holtzinger, A., Corneo, B., Li, X., Micallef, S.J., Park, I.H., Basford, C., Wheeler, M.B., Daley, G.Q., Elefanty, A.G., Stanley, E.G., & Keller, G. 2011. Stage-specific signaling through TGFbeta family members and WNT regulates patterning and pancreatic specification of human pluripotent stem cells. *Development*, 138, (5) 861-871 available from: PM:21270052

Odom, D.T., Zizlsperger, N., Gordon, D.B., Bell, G.W., Rinaldi, N.J., Murray, H.L., Volkert, T.L., Schreiber, J., Rolfe, P.A., Gifford, D.K., Fraenkel, E., Bell, G.I., & Young, R.A. 2004. Control of pancreas and liver gene expression by HNF transcription factors. *Science*, 303, (5662) 1378-1381 available from: PM:14988562

Offield, M.F., Jetton, T.L., Labosky, P.A., Ray, M., Stein, R.W., Magnuson, M.A., Hogan, B.L., & Wright, C.V. 1996. PDX-1 is required for pancreatic outgrowth and differentiation of the rostral duodenum. *Development*, 122, (3) 983-995 available from: PM:8631275

Ohi, Y., Qin, H., Hong, C., Blouin, L., Polo, J.M., Guo, T., Qi, Z., Downey, S.L., Manos, P.D., Rossi, D.J., Yu, J., Hebrok, M., Hochedlinger, K., Costello, J.F., Song, J.S., & Ramalho-Santos, M. 2011. Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. *Nat.Cell Biol.*, 13, (5) 541-549 available from: PM:21499256

Ohlsson, H., Karlsson, K., & Edlund, T. 1993. IPF1, a homeodomain-containing transactivator of the insulin gene. *EMBO J.*, 12, (11) 4251-4259 available from: PM:7901001

Okita, K., Ichisaka, T., & Yamanaka, S. 2007. Generation of germline-competent induced pluripotent stem cells. *Nature*, 448, (7151) 313-317 available from: PM:17554338

Oliver, J.A. 2004. Adult renal stem cells and renal repair. *Curr.Opin.Nephrol.Hypertens.*, 13, (1) 17-22 available from: PM:15090855

Osafune, K., Caron, L., Borowiak, M., Martinez, R.J., Fitz-Gerald, C.S., Sato, Y., Cowan, C.A., Chien, K.R., & Melton, D.A. 2008. Marked differences in differentiation propensity among human embryonic stem cell lines. *Nat.Biotechnol.*, 26, (3) 313-315 available from: PM:18278034

Oster, A., Jensen, J., Edlund, H., & Larsson, L.I. 1998. Homeobox gene product Nkx 6.1 immunoreactivity in nuclei of endocrine cells of rat and mouse stomach. *J.Histochem.Cytochem.*, 46, (6) 717-721 available from: PM:9603782

Ott, M.O., Rey-Campos, J., Cereghini, S., & Yaniv, M. 1991. vHNF1 is expressed in epithelial cells of distinct embryonic origin during development and precedes HNF1 expression. *Mech.Dev.*, 36, (1-2) 47-58 available from: PM:1685890

Ouziel-Yahalom, L., Zalzman, M., Anker-Kitai, L., Knoller, S., Bar, Y., Glandt, M., Herold, K., & Efrat, S. 2006. Expansion and redifferentiation of adult human pancreatic islet cells. *Biochem.Biophys.Res.Commun.*, 341, (2) 291-298 available from: PM:16446152

Pan, C., Hicks, A., Guan, X., Chen, H., & Bishop, C.E. 2010. SNL fibroblast feeder layers support derivation and maintenance of human induced pluripotent stem cells. *J Genet.Genomics*, 37, (4) 241-248 available from: PM:20439100

Patterson, M., Chan, D.N., Ha, I., Case, D., Cui, Y., Van, H.B., Mikkola, H.K., & Lowry, W.E. 2012. Defining the nature of human pluripotent stem cell progeny. *Cell Res.*, 22, (1) 178-193 available from: PM:21844894

Pauklin, S. & Vallier, L. 2013. The cell-cycle state of stem cells determines cell fate propensity. *Cell*, 155, (1) 135-147 available from: PM:24074866

Pennarossa, G., Maffei, S., Campagnol, M., Tarantini, L., Gandolfi, F., & Brevini, T.A. 2013. Brief demethylation step allows the conversion of adult human skin fibroblasts into insulin-secreting cells. *Proc.Natl.Acad.Sci.U.S.A*, 110, (22) 8948-8953 available from: PM:23696663

Perez-Cunningham, J., Ames, E., Smith, R.C., Peter, A.K., Naidu, R., Nolta, J.A., & Murphy, W.J. 2014. Natural Killer Cell Subsets Differentially Reject Embryonic Stem Cells Based on Licensing. *Transplantation* available from: PM:24704665

Pessina, A., Eletti, B., Croera, C., Savalli, N., Diodovich, C., & Gribaldo, L. 2004. Pancreas developing markers expressed on human mononucleated umbilical cord blood cells. *Biochem.Biophys.Res.Commun.*, 323, (1) 315-322 available from: PM:15351739

Pick, M., Stelzer, Y., Bar-Nur, O., Mayshar, Y., Eden, A., & Benvenisty, N. 2009. Clone- and gene-specific aberrations of parental imprinting in human induced pluripotent stem cells. *Stem Cells*, 27, (11) 2686-2690 available from: PM:19711451

Pictet, R.L., Clark, W.R., Williams, R.H., & Rutter, W.J. 1972. An ultrastructural analysis of the developing embryonic pancreas. *Dev.Biol.*, 29, (4) 436-467 available from: PM:4570759

Place, R.F., Li, L.C., Pookot, D., Noonan, E.J., & Dahiya, R. 2008. MicroRNA-373 induces expression of genes with complementary promoter sequences. *Proc.Natl.Acad.Sci.U.S.A*, 105, (5) 1608-1613 available from: PM:18227514

Plaisance, V., Abderrahmani, A., Perret-Menoud, V., Jacquemin, P., Lemaigre, F., & Regazzi, R. 2006. MicroRNA-9 controls the expression of Granuphilin/Slp4 and the secretory response of insulin-producing cells. *J Biol.Chem.*, 281, (37) 26932-26942 available from: PM:16831872

Polanco, J.C., Ho, M.S., Wang, B., Zhou, Q., Wolvetang, E., Mason, E., Wells, C.A., Kolle, G., Grimmond, S.M., Bertoncello, I., O'Brien, C., & Laslett, A.L. 2013. Identification of unsafe human induced pluripotent stem cell lines using a robust surrogate assay for pluripotency. *Stem Cells*, 31, (8) 1498-1510 available from: PM:23728894

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., & Hochedlinger, K. 2010. Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nat.Biotechnol.*, 28, (8) 848-855 available from: PM:20644536

Porciuncula, A., Zapata, N., Guruceaga, E., Agirre, X., Barajas, M., & Prosper, F. 2013. MicroRNA signatures of iPSCs and endoderm-derived tissues. *Gene Expr.Patterns.*, 13, (1-2) 12-20 available from: PM:22982176

Potten, C.S. 1998. Stem cells in gastrointestinal epithelium: numbers, characteristics and death. *Philos.Trans.R.Soc.Lond B Biol.Sci.*, 353, (1370) 821-830 available from: PM:9684279

Poy, M.N., Eliasson, L., Krutzfeldt, J., Kuwajima, S., Ma, X., Macdonald, P.E., Pfeffer, S., Tuschl, T., Rajewsky, N., Rorsman, P., & Stoffel, M. 2004. A pancreatic islet-specific microRNA regulates insulin secretion. *Nature*, 432, (7014) 226-230 available from: PM:15538371

Poy, M.N., Hausser, J., Trajkovski, M., Braun, M., Collins, S., Rorsman, P., Zavolan, M., & Stoffel, M. 2009. miR-375 maintains normal pancreatic alpha- and beta-cell mass. *Proc.Natl.Acad.Sci.U.S.A*, 106, (14) 5813-5818 available from: PM:19289822

Pullen, T.J., da, S., X, Kelsey, G., & Rutter, G.A. 2011. miR-29a and miR-29b contribute to pancreatic beta-cell-specific silencing of monocarboxylate transporter 1 (Mct1). *Mol.Cell Biol.*, 31, (15) 3182-3194 available from: PM:21646425

Qiu, Y., Guo, M., Huang, S., & Stein, R. 2002. Insulin gene transcription is mediated by interactions between the p300 coactivator and PDX-1, BETA2, and E47. *Mol.Cell Biol.*, 22, (2) 412-420 available from: PM:11756538

Qu, X.B., Pan, J., Zhang, C., & Huang, S.Y. 2008. Sox17 facilitates the differentiation of mouse embryonic stem cells into primitive and definitive endoderm in vitro. *Dev.Growth Differ.*, 50, (7) 585-593 available from: PM:19238729

Rajagopal, J., Anderson, W.J., Kume, S., Martinez, O.I., & Melton, D.A. 2003. Insulin staining of ES cell progeny from insulin uptake. *Science*, 299, (5605) 363 available from: PM:12532008

Rajala, K., Lindroos, B., Hussein, S.M., Lappalainen, R.S., Pekkanen-Mattila, M., Inzunza, J., Rozell, B., Miettinen, S., Narkilahti, S., Kerkela, E., Aalto-Setala, K., Otonkoski, T., Suuronen, R., Hovatta, O., & Skottman, H. 2010. A defined and xeno-free culture method enabling the establishment of clinical-grade human embryonic, induced pluripotent and adipose stem cells. *PLoS.One.*, 5, (4) e10246 available from: PM:20419109

Rajewsky, N. 2006. microRNA target predictions in animals. *Nat.Genet.*, 38 Suppl, S8-13 available from: PM:16736023

Ramachandran, D., Roy, U., Garg, S., Ghosh, S., Pathak, S., & Kolthur-Seetharam, U. 2011. Sirt1 and mir-9 expression is regulated during glucose-stimulated insulin secretion in pancreatic beta-islets. *FEBS J*, 278, (7) 1167-1174 available from: PM:21288303

Rankin, M.M., Wilbur, C.J., Rak, K., Shields, E.J., Granger, A., & Kushner, J.A. 2013. beta-Cells are not generated in pancreatic duct ligation-induced injury in adult mice. *Diabetes*, 62, (5) 1634-1645 available from: PM:23349489

Rausa, F., Samadani, U., Ye, H., Lim, L., Fletcher, C.F., Jenkins, N.A., Copeland, N.G., & Costa, R.H. 1997. The cut-homeodomain transcriptional activator HNF-6 is coexpressed with its target gene HNF-3 beta in the developing murine liver and pancreas. *Dev.Biol.*, 192, (2) 228-246 available from: PM:9441664

Ravassard, P., Hazhouz, Y., Pechberty, S., Bricout-Neveu, E., Armanet, M., Czernichow, P., & Scharfmann, R. 2011. A genetically engineered human pancreatic beta cell line exhibiting glucose-inducible insulin secretion. *J Clin.Invest*, 121, (9) 3589-3597 available from: PM:21865645

Razak, S.R., Ueno, K., Takayama, N., Nariai, N., Nagasaki, M., Saito, R., Koso, H., Lai, C.Y., Murakami, M., Tsuji, K., Michiue, T., Nakauchi, H., Otsu, M., & Watanabe, S. 2013. Profiling of microRNA in human and mouse ES and iPS cells reveals overlapping but distinct microRNA expression patterns. *PLoS.One.*, 8, (9) e73532 available from: PM:24086284

Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R., & Ruvkun, G. 2000. The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature*, 403, (6772) 901-906 available from: PM:10706289

Ren, J., Jin, P., Wang, E., Marincola, F.M., & Stroncek, D.F. 2009. MicroRNA and gene expression patterns in the differentiation of human embryonic stem cells. *J Transl.Med*, 7, 20 available from: PM:19309508

Reubinoff, B.E., Pera, M.F., Fong, C.Y., Trounson, A., & Bongso, A. 2000. Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. *Nat.Biotechnol.*, 18, (4) 399-404 available from: PM:10748519

Rodaway, A. & Patient, R. 2001. Mesendoderm. an ancient germ layer? *Cell*, 105, (2) 169-172 available from: PM:11336666

Rogers, M.B., Hosler, B.A., & Gudas, L.J. 1991. Specific expression of a retinoic acid-regulated, zinc-finger gene, Rex-1, in preimplantation embryos, trophoblast and spermatocytes. *Development*, 113, (3) 815-824 available from: PM:1821852

Roggli, E., Gattesco, S., Caille, D., Briet, C., Boitard, C., Meda, P., & Regazzi, R. 2012. Changes in microRNA expression contribute to pancreatic beta-cell dysfunction in prediabetic NOD mice. *Diabetes*, 61, (7) 1742-1751 available from: PM:22537941

Rohlfing, C.L., Wiedmeyer, H.M., Little, R.R., England, J.D., Tennill, A., & Goldstein, D.E. 2002. Defining the relationship between plasma glucose and HbA(1c): analysis of glucose profiles and HbA(1c) in the Diabetes Control and Complications Trial. *Diabetes Care*, 25, (2) 275-278 available from: PM:11815495

Rosa, A. & Brivanlou, A.H. 2011. A regulatory circuitry comprised of miR-302 and the transcription factors OCT4 and NR2F2 regulates human embryonic stem cell differentiation. *EMBO J*, 30, (2) 237-248 available from: PM:21151097

Rosner, M.H., Vigano, M.A., Ozato, K., Timmons, P.M., Poirier, F., Rigby, P.W., & 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 available from: PM:1972777

Rossant, J. & Tam, P.P. 2004. Emerging asymmetry and embryonic patterning in early mouse development. *Dev.Cell*, 7, (2) 155-164 available from: PM:15296713

Ruhnke, M., Ungefroren, H., Nussler, A., Martin, F., Brulport, M., Schormann, W., Hengstler, J.G., Klapper, W., Ulrichs, K., Hutchinson, J.A., Soria, B., Parwaresch, R.M., Heeckt, P., Kremer, B., & Fandrich, F. 2005. Differentiation of in vitro-modified human peripheral blood monocytes into hepatocyte-like and pancreatic islet-like cells. *Gastroenterology*, 128, (7) 1774-1786 available from: PM:15940611

Ruiz, S., Diep, D., Gore, A., Panopoulos, A.D., Montserrat, N., Plongthongkum, N., Kumar, S., Fung, H.L., Giorgetti, A., Bilic, J., Batchelder, E.M., Zaehres, H., Kan, N.G., Scholer, H.R., Mercola, M., Zhang, K., & Izpisua Belmonte, J.C. 2012. Identification of a specific reprogramming-associated epigenetic signature in human induced pluripotent stem cells. *Proc.Natl.Acad.Sci.U.S.A*, 109, (40) 16196-16201 available from: PM:22991473

Saito, H., Takeuchi, M., Chida, K., & Miyajima, A. 2011. Generation of glucose-responsive functional islets with a three-dimensional structure from mouse fetal pancreatic cells and iPS cells in vitro. *PLoS.One.*, 6, (12) e28209 available from: PM:22145030

Saito, T. & Saetrom, P. 2010. MicroRNAs--targeting and target prediction. *N.Biotechnol.*, 27, (3) 243-249 available from: PM:20219708

Sander, M., Sussel, L., Conners, J., Scheel, D., Kalamaras, J., Dela, C.F., Schwitzgebel, V., Hayes-Jordan, A., & German, M. 2000. Homeobox gene Nkx6.1 lies downstream of Nkx2.2 in the major pathway of beta-cell formation in the pancreas. *Development*, 127, (24) 5533-5540 available from: PM:11076772

Santerre, R.F., Cook, R.A., Crisel, R.M., Sharp, J.D., Schmidt, R.J., Williams, D.C., & Wilson, C.P. 1981. Insulin synthesis in a clonal cell line of simian virus 40-transformed hamster pancreatic beta cells. *Proc.Natl.Acad.Sci.U.S.A*, 78, (7) 4339-4343 available from: PM:6270673

Sapir, T., Shternhall, K., Meivar-Levy, I., Blumenfeld, T., Cohen, H., Skutelsky, E., Eventov-Friedman, S., Barshack, I., Goldberg, I., Pri-Chen, S., Ben-Dor, L., Polak-Charcon, S., Karasik, A., Shimon, I., Mor, E., & Ferber, S. 2005. Cell-replacement therapy for diabetes: Generating functional insulin-producing tissue from adult human liver cells. *Proc.Natl.Acad.Sci.U.S.A*, 102, (22) 7964-7969 available from: PM:15899968

Sasaki, H. & Hogan, B.L. 1993. Differential expression of multiple fork head related genes during gastrulation and axial pattern formation in the mouse embryo. *Development*, 118, (1) 47-59 available from: PM:8375339

Schott, M., Schatz, D., Atkinson, M., Krischer, J., Mehta, H., Vold, B., & Maclaren, N. 1994. GAD65 autoantibodies increase the predictability but not the sensitivity of islet cell and insulin autoantibodies for developing insulin dependent diabetes mellitus. *J.Autoimmun.*, 7, (6) 865-872 available from: PM:7888042

Schuldiner, M., Itskovitz-Eldor, J., & Benvenisty, N. 2003. Selective ablation of human embryonic stem cells expressing a "suicide" gene. *Stem Cells*, 21, (3) 257-265 available from: PM:12743320

Schwitzgebel, V.M., Scheel, D.W., Conners, J.R., Kalamaras, J., Lee, J.E., Anderson, D.J., Sussel, L., Johnson, J.D., & German, M.S. 2000. Expression of neurogenin3 reveals an islet cell precursor population in the pancreas. *Development*, 127, (16) 3533-3542 available from: PM:10903178

Seguin, C.A., Draper, J.S., Nagy, A., & Rossant, J. 2008. Establishment of endoderm progenitors by SOX transcription factor expression in human embryonic stem cells. *Cell Stem Cell*, 3, (2) 182-195 available from: PM:18682240

Seki, T., Yuasa, S., Oda, M., Egashira, T., Yae, K., Kusumoto, D., Nakata, H., Tohyama, S., Hashimoto, H., Kodaira, M., Okada, Y., Seimiya, H., Fusaki, N., Hasegawa, M., & Fukuda, K. 2010. Generation of induced pluripotent stem cells from human terminally differentiated circulating T cells. *Cell Stem Cell*, 7, (1) 11-14 available from: PM:20621043

Sekine, K., Takebe, T., Suzuki, Y., Kamiya, A., Nakauchi, H., & Taniguchi, H. 2012. Highly efficient generation of definitive endoderm lineage from human induced pluripotent stem cells. *Transplant.Proc.*, 44, (4) 1127-1129 available from: PM:22564643

Sengupta, S., Nie, J., Wagner, R.J., Yang, C., Stewart, R., & Thomson, J.A. 2009. MicroRNA 92b controls the G1/S checkpoint gene p57 in human embryonic stem cells. *Stem Cells*, 27, (7) 1524-1528 available from: PM:19544458

Serup, P., Madsen, O.D., & Mandrup-Poulsen, T. 2001. Islet and stem cell transplantation for treating diabetes. *BMJ*, 322, (7277) 29-32 available from: PM:11141151

Sethupathy, P., Megraw, M., & Hatzigeorgiou, A.G. 2006. A guide through present computational approaches for the identification of mammalian microRNA targets. *Nat.Methods*, 3, (11) 881-886 available from: PM:17060911

Shafa, M., Sjonnesen, K., Yamashita, A., Liu, S., Michalak, M., Kallos, M.S., & Rancourt, D.E. 2012. Expansion and long-term maintenance of induced pluripotent stem cells in stirred suspension bioreactors. *J Tissue Eng Regen.Med*, 6, (6) 462-472 available from: PM:21761573

Shao, K., Koch, C., Gupta, M.K., Lin, Q., Lenz, M., Laufs, S., Denecke, B., Schmidt, M., Linke, M., Hennies, H.C., Hescheler, J., Zenke, M., Zechner, U., Saric, T., & Wagner, W. 2013. Induced pluripotent mesenchymal stromal cell clones retain donor-derived differences in DNA methylation profiles. *Mol.Ther.*, 21, (1) 240-250 available from: PM:23032973

Shapiro, A.M., Lakey, J.R., Ryan, E.A., Korbutt, G.S., Toth, E., Warnock, G.L., Kneteman, N.M., & Rajotte, R.V. 2000. Islet transplantation in seven patients with type 1 diabetes mellitus using a glucocorticoid-free immunosuppressive regimen. *N.Engl.J Med*, 343, (4) 230-238 available from: PM:10911004

Shen, W., Tremblay, M.S., Deshmukh, V.A., Wang, W., Filippi, C.M., Harb, G., Zhang, Y.Q., Kamireddy, A., Baaten, J.E., Jin, Q., Wu, T., Swoboda, J.G., Cho, C.Y., Li, J., Laffitte, B.A., McNamara, P., Glynne, R., Wu, X., Herman, A.E., & Schultz, P.G. 2013. Small-molecule inducer of beta cell proliferation identified by high-throughput screening. *J Am.Chem.Soc.* , 135, (5) 1669-1672 available from: PM:23330637

Shi, C.M. & Cheng, T.M. 2004. Differentiation of dermis-derived multipotent cells into insulin-producing pancreatic cells in vitro. *World J Gastroenterol.*, 10, (17) 2550-2552 available from: PM:15300903

Shi, Y., Desponts, C., Do, J.T., Hahm, H.S., Scholer, H.R., & Ding, S. 2008. Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. *Cell Stem Cell*, 3, (5) 568-574 available from: PM:18983970

Shim, J.H., Kim, S.E., Woo, D.H., Kim, S.K., Oh, C.H., McKay, R., & Kim, J.H. 2007. Directed differentiation of human embryonic stem cells towards a pancreatic cell fate. *Diabetologia*, 50, (6) 1228-1238 available from: PM:17457565

Shiraki, N., Yoshida, T., Araki, K., Umezawa, A., Higuchi, Y., Goto, H., Kume, K., & Kume, S. 2008. Guided differentiation of embryonic stem cells into Pdx1-expressing regional-specific definitive endoderm. *Stem Cells*, 26, (4) 874-885 available from: PM:18238854

Shternhall-Ron, K., Quintana, F.J., Perl, S., Meivar-Levy, I., Barshack, I., Cohen, I.R., & Ferber, S. 2007. Ectopic PDX-1 expression in liver ameliorates type 1 diabetes. *J Autoimmun.*, 28, (2-3) 134-142 available from: PM:17383157

Si-Tayeb, K., Noto, F.K., Sepac, A., Sedlic, F., Bosnjak, Z.J., Lough, J.W., & Duncan, S.A. 2010. Generation of human induced pluripotent stem cells by simple transient transfection of plasmid DNA encoding reprogramming factors. *BMC.Dev.Biol.*, 10, 81 available from: PM:20682060

Simion, A., Laudadio, I., Prevot, P.P., Raynaud, P., Lemaigre, F.P., & Jacquemin, P. 2010. MiR-495 and miR-218 regulate the expression of the Onecut transcription factors HNF-6 and OC-2. *Biochem.Biophys.Res.Commun.*, 391, (1) 293-298 available from: PM:19913497

Singh, A., Suri, S., Lee, T., Chilton, J.M., Cooke, M.T., Chen, W., Fu, J., Stice, S.L., Lu, H., McDevitt, T.C., & Garcia, A.J. 2013. Adhesion strength-based, label-free isolation of human pluripotent stem cells. *Nat.Methods*, 10, (5) 438-444 available from: PM:23563795

Slack, F.J. 2010. Stem cells: Big roles for small RNAs. *Nature*, 463, (7281) 616 available from: PM:20130638

Slack, J.M. 1985. Homoeotic transformations in man: implications for the mechanism of embryonic development and for the organization of epithelia. *J.Theor.Biol.*, 114, (3) 463-490 available from: PM:3927066

Slack, J.M. 1995. Developmental biology of the pancreas. *Development*, 121, (6) 1569-1580 available from: PM:7600975

Smith, J.C., Price, B.M., Van, N.K., & Huylebroeck, D. 1990. Identification of a potent Xenopus mesoderm-inducing factor as a homologue of activin A. *Nature*, 345, (6277) 729-731 available from: PM:2113615

Smith, K.P., Luong, M.X., & Stein, G.S. 2009. Pluripotency: toward a gold standard for human ES and iPS cells. *J Cell Physiol*, 220, (1) 21-29 available from: PM:19326392

Smith, S.B., Ee, H.C., Conners, J.R., & German, M.S. 1999. Paired-homeodomain transcription factor PAX4 acts as a transcriptional repressor in early pancreatic development. *Mol.Cell Biol.*, 19, (12) 8272-8280 available from: PM:10567552

Sneddon, J.B., Borowiak, M., & Melton, D.A. 2012. Self-renewal of embryonic-stem-cell-derived progenitors by organ-matched mesenchyme. *Nature*, 491, (7426) 765-768 available from: PM:23041930

Solar, M., Cardalda, C., Houbracken, I., Martin, M., Maestro, M.A., De, M.N., Xu, X., Grau, V., Heimberg, H., Bouwens, L., & Ferrer, J. 2009. Pancreatic exocrine duct cells give rise to insulin-producing beta cells during embryogenesis but not after birth. *Dev.Cell*, 17, (6) 849-860 available from: PM:20059954

Soldner, F., Hockemeyer, D., Beard, C., Gao, Q., Bell, G.W., Cook, E.G., Hargus, G., Blak, A., Cooper, O., Mitalipova, M., Isacson, O., & Jaenisch, R. 2009. Parkinson's disease patient-derived induced pluripotent stem cells free of viral reprogramming factors. *Cell*, 136, (5) 964-977 available from: PM:19269371

Sommer, C.A., Stadtfeld, M., Murphy, G.J., Hochedlinger, K., Kotton, D.N., & Mostoslavsky, G. 2009. Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. *Stem Cells*, 27, (3) 543-549 available from: PM:19096035

Song, P.P., Hu, Y., Liu, C.M., Yan, M.J., Song, G., Cui, Y., Xia, H.F., & Ma, X. 2011. Embryonic ectoderm development protein is regulated by microRNAs in human neural tube defects. *Am.J Obstet.Gynecol.*, 204, (6) 544-17 available from: PM:21497788

Sorenson, R.L. & Brelje, T.C. 1997. Adaptation of islets of Langerhans to pregnancy: beta-cell growth, enhanced insulin secretion and the role of lactogenic hormones. *Horm.Metab Res.*, 29, (6) 301-307 available from: PM:9230352

Soria, B., Roche, E., Berna, G., Leon-Quinto, T., Reig, J.A., & Martin, F. 2000. Insulin-secreting cells derived from embryonic stem cells normalize glycemia in streptozotocin-induced diabetic mice. *Diabetes*, 49, (2) 157-162 available from: PM:10868930

Spits, C., Mateizel, I., Geens, M., Mertzanidou, A., Staessen, C., Vandeskelde, Y., Van der Elst, J., Liebaers, I., & Sermon, K. 2008. Recurrent chromosomal abnormalities in human embryonic stem cells. *Nat.Biotechnol.*, 26, (12) 1361-1363 available from: PM:19029912

St-Onge, L., Sosa-Pineda, B., Chowdhury, K., Mansouri, A., & Gruss, P. 1997. Pax6 is required for differentiation of glucagon-producing alpha-cells in mouse pancreas. *Nature*, 387, (6631) 406-409 available from: PM:9163426

Stadtfeld, M., Brennand, K., & Hochedlinger, K. 2008. Reprogramming of pancreatic beta cells into induced pluripotent stem cells. *Curr.Biol.*, 18, (12) 890-894 available from: PM:18501604

Stadtfeld, M., Maherali, N., Breault, D.T., & Hochedlinger, K. 2008. Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. *Cell Stem Cell*, 2, (3) 230-240 available from: PM:18371448

Stadtfeld, M., Apostolou, E., Akutsu, H., Fukuda, A., Follett, P., Natesan, S., Kono, T., Shioda, T., & Hochedlinger, K. 2010. Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature*, 465, (7295) 175-181 available from: PM:20418860

Stadtfeld, M., Apostolou, E., Ferrari, F., Choi, J., Walsh, R.M., Chen, T., Ooi, S.S., Kim, S.Y., Bestor, T.H., Shioda, T., Park, P.J., & Hochedlinger, K. 2012. Ascorbic acid prevents loss of Dlk1-Dio3 imprinting and facilitates generation of all-iPS cell mice from terminally differentiated B cells. *Nat.Genet.*, 44, (4) 398-2 available from: PM:22387999

Staerk, J., Dawlaty, M.M., Gao, Q., Maetzel, D., Hanna, J., Sommer, C.A., Mostoslavsky, G., & Jaenisch, R. 2010. Reprogramming of human peripheral blood cells to induced pluripotent stem cells. *Cell Stem Cell*, 7, (1) 20-24 available from: PM:20621045

Stafford, D., Hornbruch, A., Mueller, P.R., & Prince, V.E. 2004. A conserved role for retinoid signaling in vertebrate pancreas development. *Dev.Genes Evol.*, 214, (9) 432-441 available from: PM:15322880

Steiner, D.F., Chan, S.J., Welsh, J.M., & Kwok, S.C. 1985. Structure and evolution of the insulin gene. *Annu.Rev.Genet.*, 19, 463-484 available from: PM:3909947

Strauss, W.M., Chen, C., Lee, C.T., & Ridzon, D. 2006. Nonrestrictive developmental regulation of microRNA gene expression. *Mamm.Genome*, 17, (8) 833-840 available from: PM:16897339

Subramanyam, D., Lamouille, S., Judson, R.L., Liu, J.Y., Bucay, N., Derynck, R., & Blelloch, R. 2011. Multiple targets of miR-302 and miR-372 promote reprogramming of human fibroblasts to induced pluripotent stem cells. *Nat.Biotechnol.*, 29, (5) 443-448 available from: PM:21490602

Suh, M.R., Lee, Y., Kim, J.Y., Kim, S.K., Moon, S.H., Lee, J.Y., Cha, K.Y., Chung, H.M., Yoon, H.S., Moon, S.Y., Kim, V.N., & Kim, K.S. 2004. Human embryonic stem cells express a unique set of microRNAs. *Dev.Biol.*, 270, (2) 488-498 available from: PM:15183728

Sui, L., Mfopou, J.K., Geens, M., Sermon, K., & Bouwens, L. 2012. FGF signaling via MAPK is required early and improves Activin A-induced definitive endoderm formation from human embryonic stem cells. *Biochem.Biophys.Res.Commun.*, 426, (3) 380-385 available from: PM:22960178

Sui, L., Geens, M., Sermon, K., Bouwens, L., & Mfopou, J.K. 2013. Role of BMP signaling in pancreatic progenitor differentiation from human embryonic stem cells. *Stem Cell Rev.*, 9, (5) 569-577 available from: PM:23468018

Sulzbacher, S., Schroeder, I.S., Truong, T.T., & Wobus, A.M. 2009. Activin A-induced differentiation of embryonic stem cells into endoderm and pancreatic progenitors-the influence of differentiation factors and culture conditions. *Stem Cell Rev.*, 5, (2) 159-173 available from: PM:19263252

Sun, L.L., Jiang, B.G., Li, W.T., Zou, J.J., Shi, Y.Q., & Liu, Z.M. 2011. MicroRNA-15a positively regulates insulin synthesis by inhibiting uncoupling protein-2 expression. *Diabetes Res.Clin.Pract.*, 91, (1) 94-100 available from: PM:21146880

Sun, N., Panetta, N.J., Gupta, D.M., Wilson, K.D., Lee, A., Jia, F., Hu, S., Cherry, A.M., Robbins, R.C., Longaker, M.T., & Wu, J.C. 2009. Feeder-free derivation of induced pluripotent stem cells from adult human adipose stem cells. *Proc.Natl.Acad.Sci.U.S.A*, 106, (37) 15720-15725 available from: PM:19805220

Sun, Y., Chen, L., Hou, X.G., Hou, W.K., Dong, J.J., Sun, L., Tang, K.X., Wang, B., Song, J., Li, H., & Wang, K.X. 2007. Differentiation of bone marrow-derived mesenchymal stem cells from diabetic patients into insulin-producing cells in vitro. *Chin Med J (Engl.)*, 120, (9) 771-776 available from: PM:17531117

Suri, A., Calderon, B., Esparza, T.J., Frederick, K., Bittner, P., & Unanue, E.R. 2006. Immunological reversal of autoimmune diabetes without hematopoietic replacement of beta cells. *Science*, 311, (5768) 1778-1780 available from: PM:16556846

Sussel, L., Kalamaras, J., Hartigan-O'Connor, D.J., Meneses, J.J., Pedersen, R.A., Rubenstein, J.L., & German, M.S. 1998. Mice lacking the homeodomain transcription factor Nkx2.2 have diabetes due to arrested differentiation of pancreatic beta cells. *Development*, 125, (12) 2213-2221 available from: PM:9584121

Suzuki, A., Zheng, Y.W., Kaneko, S., Onodera, M., Fukao, K., Nakauchi, H., & Taniguchi, H. 2002. Clonal identification and characterization of self-renewing pluripotent stem cells in the developing liver. *J Cell Biol.*, 156, (1) 173-184 available from: PM:11781341

Suzuki, A., Nakauchi, H., & Taniguchi, H. 2003. Glucagon-like peptide 1 (1-37) converts intestinal epithelial cells into insulin-producing cells. *Proc.Natl.Acad.Sci.U.S.A*, 100, (9) 5034-5039 available from: PM:12702762

Suzuki, N., Rohdewohld, H., Neuman, T., Gruss, P., & Scholer, H.R. 1990. Oct-6: a POU transcription factor expressed in embryonal stem cells and in the developing brain. *EMBO J*, 9, (11) 3723-3732 available from: PM:1976514

Sykes, M. & Cozzi, E. 2006. Xenotransplantation of pig islets into Mexican children: were the fundamental ethical requirements to proceed with such a study really met? *Eur.J Endocrinol.*, 154, (6) 921-922 available from: PM:16728555

Taapken, S.M., Nisler, B.S., Newton, M.A., Sampsell-Barron, T.L., Leonhard, K.A., McIntire, E.M., & Montgomery, K.D. 2011. Karotypic abnormalities in human induced pluripotent stem cells and embryonic stem cells. *Nat.Biotechnol.*, 29, (4) 313-314 available from: PM:21478842

Tachibana, M., Amato, P., Sparman, M., Gutierrez, N.M., Tippner-Hedges, R., Ma, H., Kang, E., Fulati, A., Lee, H.S., Sritanaudomchai, H., Masterson, K., Larson, J., Eaton, D., Sadler-Fredd, K., Battaglia, D., Lee, D., Wu, D., Jensen, J., Patton, P., Gokhale, S., Stouffer, R.L., Wolf, D., & Mitalipov, S. 2013. Human embryonic stem cells derived by somatic cell nuclear transfer. *Cell*, 153, (6) 1228-1238 available from: PM:23683578

Tada, S., Era, T., Furusawa, C., Sakurai, H., Nishikawa, S., Kinoshita, M., Nakao, K., Chiba, T., & Nishikawa, S. 2005. Characterization of mesendoderm: a diverging point of the definitive endoderm and mesoderm in embryonic stem cell differentiation culture. *Development*, 132, (19) 4363-4374 available from: PM:16141227

Tahamtani, Y., Azarnia, M., Farrokhi, A., Sharifi-Zarchi, A., Aghdami, N., & Baharvand, H. 2013. Treatment of human embryonic stem cells with different combinations of priming and inducing factors toward definitive endoderm. *Stem Cells Dev.*, 22, (9) 1419-1432 available from: PM:23249309

Takahashi, K. & Yamanaka, S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, 126, (4) 663-676 available from: PM:16904174

Takahashi, K., Okita, K., Nakagawa, M., & Yamanaka, S. 2007. Induction of pluripotent stem cells from fibroblast cultures. *Nat.Protoc.*, 2, (12) 3081-3089 available from: PM:18079707

Takahashi, K., Narita, M., Yokura, M., Ichisaka, T., & Yamanaka, S. 2009. Human induced pluripotent stem cells on autologous feeders. *PLoS.One.*, 4, (12) e8067 available from: PM:19956543

Talchai, C., Xuan, S., Kitamura, T., DePinho, R.A., & Accili, D. 2012. Generation of functional insulin-producing cells in the gut by Foxo1 ablation. *Nat.Genet.*, 44, (4) 406-12, S1 available from: PM:22406641

Tam, P.P., Kanai-Azuma, M., & Kanai, Y. 2003. Early endoderm development in vertebrates: lineage differentiation and morphogenetic function. *Curr.Opin.Genet.Dev.*, 13, (4) 393-400 available from: PM:12888013

Tan, H.L., Fong, W.J., Lee, E.H., Yap, M., & Choo, A. 2009. mAb 84, a cytotoxic antibody that kills undifferentiated human embryonic stem cells via oncosis. *Stem Cells*, 27, (8) 1792-1801 available from: PM:19544435

Taneera, J., Rosengren, A., Renstrom, E., Nygren, J.M., Serup, P., Rorsman, P., & Jacobsen, S.E. 2006. Failure of transplanted bone marrow cells to adopt a pancreatic beta-cell fate. *Diabetes*, 55, (2) 290-296 available from: PM:16443759

Tang, C., Lee, A.S., Volkmer, J.P., Sahoo, D., Nag, D., Mosley, A.R., Inlay, M.A., Ardehali, R., Chavez, S.L., Pera, R.R., Behr, B., Wu, J.C., Weissman, I.L., & Drukker, M. 2011. An antibody against SSEA-5 glycan on human pluripotent stem cells enables removal of teratoma-forming cells. *Nat.Biotechnol.*, 29, (9) 829-834 available from: PM:21841799

Tang, D.Q., Cao, L.Z., Burkhardt, B.R., Xia, C.Q., Litherland, S.A., Atkinson, M.A., & Yang, L.J. 2004. In vivo and in vitro characterization of insulin-producing cells obtained from murine bone marrow. *Diabetes*, 53, (7) 1721-1732 available from: PM:15220196

Tang, X., Muniappan, L., Tang, G., & Ozcan, S. 2009. Identification of glucose-regulated miRNAs from pancreatic {beta} cells reveals a role for miR-30d in insulin transcription. *RNA.*, 15, (2) 287-293 available from: PM:19096044

Taniguchi, H., Yamato, E., Tashiro, F., Ikegami, H., Ogihara, T., & Miyazaki, J. 2003. beta-cell neogenesis induced by adenovirus-mediated gene delivery of transcription factor pdx-1 into mouse pancreas. *Gene Ther.*, 10, (1) 15-23 available from: PM:12525833

Tannenbaum, S.E., Turetsky, T.T., Singer, O., Aizenman, E., Kirshberg, S., Ilouz, N., Gil, Y., Berman-Zaken, Y., Perlman, T.S., Geva, N., Levy, O., Arbell, D., Simon, A., Ben-Meir, A., Shufaro, Y., Laufer, N., & Reubinoff, B.E. 2012. Derivation of xeno-free and GMP-grade human embryonic stem cells--platforms for future clinical applications. *PLoS.One.*, 7, (6) e35325 available from: PM:22745653

Taraviras, S., Monaghan, A.P., Schutz, G., & Kelsey, G. 1994. Characterization of the mouse HNF-4 gene and its expression during mouse embryogenesis. *Mech.Dev.*, 48, (2) 67-79 available from: PM:7873404

Tateishi, K., He, J., Taranova, O., Liang, G., D'Alessio, A.C., & Zhang, Y. 2008. Generation of insulin-secreting islet-like clusters from human skin fibroblasts. *J Biol.Chem.*, 283, (46) 31601-31607 available from: PM:18782754

Tay, Y., Zhang, J., Thomson, A.M., Lim, B., & Rigoutsos, I. 2008. MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. *Nature*, 455, (7216) 1124-1128 available from: PM:18806776

Taylor-Weiner, H., Schwarzbauer, J.E., & Engler, A.J. 2013. Defined extracellular matrix components are necessary for definitive endoderm induction. *Stem Cells*, 31, (10) 2084-2094 available from: PM:23766144

Taylor, C.J., Bolton, E.M., Pocock, S., Sharples, L.D., Pedersen, R.A., & Bradley, J.A. 2005. Banking on human embryonic stem cells: estimating the number of donor cell lines needed for HLA matching. *Lancet*, 366, (9502) 2019-2025 available from: PM:16338451

Teo, A.K., Arnold, S.J., Trotter, M.W., Brown, S., Ang, L.T., Chng, Z., Robertson, E.J., Dunn, N.R., & Vallier, L. 2011. Pluripotency factors regulate definitive endoderm specification through eomesodermin. *Genes Dev.*, 25, (3) 238-250 available from: PM:21245162

Teta, M., Rankin, M.M., Long, S.Y., Stein, G.M., & Kushner, J.A. 2007. Growth and regeneration of adult beta cells does not involve specialized progenitors. *Dev.Cell*, 12, (5) 817-826 available from: PM:17488631

Thatava, T., Kudva, Y.C., Edukulla, R., Squillace, K., De Lamo, J.G., Khan, Y.K., Sakuma, T., Ohmine, S., Terzic, A., & Ikeda, Y. 2013. Intrapatient variations in type 1 diabetes-specific iPS cell differentiation into insulin-producing cells. *Mol.Ther.*, 21, (1) 228-239 available from: PM:23183535

Theise, N.D., Saxena, R., Portmann, B.C., Thung, S.N., Yee, H., Chiriboga, L., Kumar, A., & Crawford, J.M. 1999. The canals of Hering and hepatic stem cells in humans. *Hepatology*, 30, (6) 1425-1433 available from: PM:10573521

Thomas, M., Lieberman, J., & Lal, A. 2010. Desperately seeking microRNA targets. *Nat.Struct.Mol.Biol.*, 17, (10) 1169-1174 available from: PM:20924405

Thomson, D.W., Bracken, C.P., & Goodall, G.J. 2011. Experimental strategies for microRNA target identification. *Nucleic Acids Res.*, 39, (16) 6845-6853 available from: PM:21652644

Thomson, J.A., Kalishman, J., Golos, T.G., Durning, M., Harris, C.P., Becker, R.A., & Hearn, J.P. 1995. Isolation of a primate embryonic stem cell line. *Proc.Natl.Acad.Sci.U.S.A*, 92, (17) 7844-7848 available from: PM:7544005

Thomson, J.A., Kalishman, J., Golos, T.G., Durning, M., Harris, C.P., & Hearn, J.P. 1996. Pluripotent cell lines derived from common marmoset (Callithrix jacchus) blastocysts. *Biol.Reprod.*, 55, (2) 254-259 available from: PM:8828827

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., & Jones, J.M. 1998. Embryonic stem cell lines derived from human blastocysts. *Science*, 282, (5391) 1145-1147 available from: PM:9804556

Thor, S., Ericson, J., Brannstrom, T., & Edlund, T. 1991. The homeodomain LIM protein Isl-1 is expressed in subsets of neurons and endocrine cells in the adult rat. *Neuron*, 7, (6) 881-889 available from: PM:1764243

Thorel, F., Nepote, V., Avril, I., Kohno, K., Desgraz, R., Chera, S., & Herrera, P.L. 2010. Conversion of adult pancreatic alpha-cells to beta-cells after extreme beta-cell loss. *Nature*, 464, (7292) 1149-1154 available from: PM:20364121

Tiscornia, G. & Izpisua Belmonte, J.C. 2010. MicroRNAs in embryonic stem cell function and fate. *Genes Dev.*, 24, (24) 2732-2741 available from: PM:21159814

Tropepe, V., Coles, B.L., Chiasson, B.J., Horsford, D.J., Elia, A.J., McInnes, R.R., & van der Kooy, D. 2000. Retinal stem cells in the adult mammalian eye. *Science*, 287, (5460) 2032-2036 available from: PM:10720333

Trucco, M., Casu, A., & Bottino, R. 2007. The pig as the donor of pancreatic islets for men. *Vet.Res.Commun.*, 31 Suppl 1, 27-33 available from: PM:17682843

Tsai, Z.Y., Singh, S., Yu, S.L., Kao, L.P., Chen, B.Z., Ho, B.C., Yang, P.C., & Li, S.S. 2010. Identification of microRNAs regulated by activin A in human embryonic stem cells. *J Cell Biochem.*, 109, (1) 93-102 available from: PM:19885849

Tzur, G., Levy, A., Meiri, E., Barad, O., Spector, Y., Bentwich, Z., Mizrahi, L., Katzenellenbogen, M., Ben-Shushan, E., Reubinoff, B.E., & Galun, E. 2008. MicroRNA expression patterns and function in endodermal differentiation of human embryonic stem cells. *PLoS.One.*, 3, (11) e3726 available from: PM:19015728

Ungrin, M.D., Clarke, G., Yin, T., Niebrugge, S., Nostro, M.C., Sarangi, F., Wood, G., Keller, G., & Zandstra, P.W. 2012. Rational bioprocess design for human pluripotent stem cell expansion and endoderm differentiation based on cellular dynamics. *Biotechnol.Bioeng.*, 109, (4) 853-866 available from: PM:22139975

Urbach, A., Bar-Nur, O., Daley, G.Q., & Benvenisty, N. 2010. Differential modeling of fragile X syndrome by human embryonic stem cells and induced pluripotent stem cells. *Cell Stem Cell*, 6, (5) 407-411 available from: PM:20452313

Urban, V.S., Kiss, J., Kovacs, J., Gocza, E., Vas, V., Monostori, E., & Uher, F. 2008. Mesenchymal stem cells cooperate with bone marrow cells in therapy of diabetes. *Stem Cells*, 26, (1) 244-253 available from: PM:17932424

Utikal, J., Maherali, N., Kulalert, W., & Hochedlinger, K. 2009. Sox2 is dispensable for the reprogramming of melanocytes and melanoma cells into induced pluripotent stem cells. *J Cell Sci.*, 122, (Pt 19) 3502-3510 available from: PM:19723802

Valdes-Gonzalez, R.A., White, D.J., Dorantes, L.M., Teran, L., Garibay-Nieto, G.N., Bracho-Blanchet, E., Davila-Perez, R., Evia-Viscarra, L., Ormsby, C.E., Ayala-Sumuano, J.T., Silva-Torres, M.L., & Ramirez-Gonzalez, B. 2007. Three-yr follow-up of a type 1 diabetes mellitus patient with an islet xenotransplant. *Clin.Transplant.*, 21, (3) 352-357 available from: PM:17488384

Vallier, L., Alexander, M., & Pedersen, R.A. 2005. Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *J Cell Sci.*, 118, (Pt 19) 4495-4509 available from: PM:16179608

van der Laan, L.J., Lockey, C., Griffeth, B.C., Frasier, F.S., Wilson, C.A., Onions, D.E., Hering, B.J., Long, Z., Otto, E., Torbett, B.E., & Salomon, D.R. 2000. Infection by porcine endogenous retrovirus after islet xenotransplantation in SCID mice. *Nature*, 407, (6800) 90-94 available from: PM:10993079

Vander, A., Sherman, J., & Luciano, D. 1998. *Human Physiology: the Mechanisms of Body Function*, 7th Edition ed. WCB/McGraw-Hill.

VanHoof, D., D'Amour, K.A., & German, M.S. 2009. Derivation of insulin-producing cells from human embryonic stem cells. *Stem Cell Res.*, 3, (2-3) 73-87 available from: PM:19766074

VanHoof, D., Mendelsohn, A.D., Seerke, R., Desai, T.A., & German, M.S. 2011. Differentiation of human embryonic stem cells into pancreatic endoderm in patterned size-controlled clusters. *Stem Cell Res.* , 6, (3) 276-285 available from: PM:21513906

Vasudevan, S., Tong, Y., & Steitz, J.A. 2007. Switching from repression to activation: microRNAs can up-regulate translation. *Science*, 318, (5858) 1931-1934 available from: PM:18048652

Vazquez-Martin, A., Cufi, S., Lopez-Bonet, E., Corominas-Faja, B., Oliveras-Ferraros, C., Martin-Castillo, B., & Menendez, J.A. 2012. Metformin limits the tumourigenicity of iPS cells without affecting their pluripotency. *Sci.Rep.*, 2, 964 available from: PM:23236586

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., & 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 available from: PM:18329372

Villa-Diaz, L.G., Brown, S.E., Liu, Y., Ross, A.M., Lahann, J., Parent, J.M., & Krebsbach, P.H. 2012. Derivation of mesenchymal stem cells from human induced pluripotent stem cells cultured on synthetic substrates. *Stem Cells*, 30, (6) 1174-1181 available from: PM:22415987

Voltarelli, J.C., Couri, C.E., Stracieri, A.B., Oliveira, M.C., Moraes, D.A., Pieroni, F., Coutinho, M., Malmegrim, K.C., Foss-Freitas, M.C., Simoes, B.P., Foss, M.C., Squiers, E., & Burt, R.K. 2007. Autologous nonmyeloablative hematopoietic stem cell transplantation in newly diagnosed type 1 diabetes mellitus. *JAMA*, 297, (14) 1568-1576 available from: PM:17426276

Wahlestedt, M., Ameur, A., Moraghebi, R., Norddahl, G.L., Sten, G., Woods, N.B., & Bryder, D. 2014. Somatic cells with a heavy mitochondrial DNA mutational load render iPS cells with distinct differentiation defects. *Stem Cells* available from: PM:24446123

Wang, C., Yao, N., Lu, C.L., Li, D., & Ma, X. 2010. Mouse microRNA-124 regulates the expression of Hes1 in P19 cells. *Front Biosci.(Elite.Ed)*, 2, 127-132 available from: PM:20036862

Wang, J., Cao, N., Yuan, M., Cui, H., Tang, Y., Qin, L., Huang, X., Shen, N., & Yang, H.T. 2012. MicroRNA-125b/Lin28 pathway contributes to the mesendodermal fate decision of embryonic stem cells. *Stem Cells Dev.*, 21, (9) 1524-1537 available from: PM:22277001

Wang, L., Xue, Y., Shen, Y., Li, W., Cheng, Y., Yan, X., Shi, W., Wang, J., Gong, Z., Yang, G., Guo, C., Zhou, Y., Wang, X., Zhou, Q., & Zeng, F. 2012. Claudin 6: a novel surface marker for characterizing mouse pluripotent stem cells. *Cell Res.*, 22, (6) 1082-1085 available from: PM:22565286

Wang, R.N., Kloppel, G., & Bouwens, L. 1995. Duct- to islet-cell differentiation and islet growth in the pancreas of duct-ligated adult rats. *Diabetologia*, 38, (12) 1405-1411 available from: PM:8786013

Wang, X., Metzger, D.L., Meloche, M., Hao, J., Ao, Z., & Warnock, G.L. 2012. Generation of transplantable Beta cells for patient-specific cell therapy. *Int.J Endocrinol.*, 2012, 414812 available from: PM:22611393

Wang, Y., Medvid, R., Melton, C., Jaenisch, R., & Blelloch, R. 2007. DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. *Nat.Genet.*, 39, (3) 380-385 available from: PM:17259983

Wang, Y., Liu, J., Liu, C., Naji, A., & Stoffers, D.A. 2013. MicroRNA-7 regulates the mTOR pathway and proliferation in adult pancreatic beta-cells. *Diabetes*, 62, (3) 887-895 available from: PM:23223022

Wang, Y.C., Nakagawa, M., Garitaonandia, I., Slavin, I., Altun, G., Lacharite, R.M., Nazor, K.L., Tran, H.T., Lynch, C.L., Leonardo, T.R., Liu, Y., Peterson, S.E., Laurent, L.C., Yamanaka, S., & Loring, J.F. 2011. Specific lectin biomarkers for isolation of human pluripotent stem cells identified through array-based glycomic analysis. *Cell Res.*, 21, (11) 1551-1563 available from: PM:21894191

Warren, L., Manos, P.D., Ahfeldt, T., Loh, Y.H., Li, H., Lau, F., Ebina, W., Mandal, P.K., Smith, Z.D., Meissner, A., Daley, G.Q., Brack, A.S., Collins, J.J., Cowan, C., Schlaeger, T.M., & Rossi, D.J. 2010. Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell*, 7, (5) 618-630 available from: PM:20888316

Warren, L., Ni, Y., Wang, J., & Guo, X. 2012. Feeder-free derivation of human induced pluripotent stem cells with messenger RNA. *Sci.Rep.*, 2, 657 available from: PM:22984641

Watada, H., Kajimoto, Y., Miyagawa, J., Hanafusa, T., Hamaguchi, K., Matsuoka, T., Yamamoto, K., Matsuzawa, Y., Kawamori, R., & Yamasaki, Y. 1996. PDX-1 induces insulin and glucokinase gene

expressions in alphaTC1 clone 6 cells in the presence of betacellulin. *Diabetes*, 45, (12) 1826-1831 available from: PM:8922372

Watada, H., Kajimoto, Y., Umayahara, Y., Matsuoka, T., Kaneto, H., Fujitani, Y., Kamada, T., Kawamori, R., & Yamasaki, Y. 1996. The human glucokinase gene beta-cell-type promoter: an essential role of insulin promoter factor 1/PDX-1 in its activation in HIT-T15 cells. *Diabetes*, 45, (11) 1478-1488 available from: PM:8866550

Watanabe, K., Ueno, M., Kamiya, D., Nishiyama, A., Matsumura, M., Wataya, T., Takahashi, J.B., Nishikawa, S., Nishikawa, S., Muguruma, K., & Sasai, Y. 2007. A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nat.Biotechnol.*, 25, (6) 681-686 available from: PM:17529971

Wei, R., Yang, J., Liu, G.Q., Gao, M.J., Hou, W.F., Zhang, L., Gao, H.W., Liu, Y., Chen, G.A., & Hong, T.P. 2013. Dynamic expression of microRNAs during the differentiation of human embryonic stem cells into insulin-producing cells. *Gene*, 518, (2) 246-255 available from: PM:23370336

Weinstein, D.C., Altaba, A., Chen, W.S., Hoodless, P., Prezioso, V.R., Jessell, T.M., & Darnell, J.E., Jr. 1994. The winged-helix transcription factor HNF-3 beta is required for notochord development in the mouse embryo. *Cell*, 78, (4) 575-588 available from: PM:8069910

Weinzimer, S.A., Swan, K.L., Sikes, K.A., & Ahern, J.H. 2006. Emerging evidence for the use of insulin pump therapy in infants, toddlers, and preschool-aged children with type 1 diabetes. *Pediatr.Diabetes*, 7 Suppl 4, 15-19 available from: PM:16774613

Wells, J.M. & Melton, D.A. 1999. Vertebrate endoderm development. *Annu.Rev.Cell Dev.Biol.*, 15, 393-410 available from: PM:10611967

Wells, J.M. & Melton, D.A. 2000. Early mouse endoderm is patterned by soluble factors from adjacent germ layers. *Development*, 127, (8) 1563-1572 available from: PM:10725233

Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B.E., & Jaenisch, R. 2007. In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* , 448, (7151) 318-324 available from: PM:17554336

Wernig, M., Lengner, C.J., Hanna, J., Lodato, M.A., Steine, E., Foreman, R., Staerk, J., Markoulaki, S., & Jaenisch, R. 2008. A drug-inducible transgenic system for direct reprogramming of multiple somatic cell types. *Nat.Biotechnol.*, 26, (8) 916-924 available from: PM:18594521

Wernig, M., Meissner, A., Cassady, J.P., & Jaenisch, R. 2008. c-Myc is dispensable for direct reprogramming of mouse fibroblasts. *Cell Stem Cell*, 2, (1) 10-12 available from: PM:18371415

Wessels, N.K. & Cohen, J.H. 1967. Early pancreas organogenesis: morphogenesis, tissue interactions and mass effects. *Dev.Biol.*, 15, 237-270

Wienholds, E., Kloosterman, W.P., Miska, E., Alvarez-Saavedra, E., Berezikov, E., de, B.E., Horvitz, H.R., Kauppinen, S., & Plasterk, R.H. 2005. MicroRNA expression in zebrafish embryonic development. *Science*, 309, (5732) 310-311 available from: PM:15919954

Wightman, B., Ha, I., & Ruvkun, G. 1993. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell*, 75, (5) 855-862 available from: PM:8252622

Wijesekara, N., Zhang, L.H., Kang, M.H., Abraham, T., Bhattacharjee, A., Warnock, G.L., Verchere, C.B., & Hayden, M.R. 2012. miR-33a modulates ABCA1 expression, cholesterol accumulation, and insulin secretion in pancreatic islets. *Diabetes*, 61, (3) 653-658 available from: PM:22315319

Wilkinson, D.G., Bhatt, S., & Herrmann, B.G. 1990. Expression pattern of the mouse T gene and its role in mesoderm formation. *Nature*, 343, (6259) 657-659 available from: PM:1689462

Wilson, K.D., Venkatasubrahmanyam, S., Jia, F., Sun, N., Butte, A.J., & Wu, J.C. 2009. MicroRNA profiling of human-induced pluripotent stem cells. *Stem Cells Dev.*, 18, (5) 749-758 available from: PM:19284351

Wilson, M.E., Scheel, D., & German, M.S. 2003. Gene expression cascades in pancreatic development. *Mech.Dev.*, 120, (1) 65-80 available from: PM:12490297

Witkos, T.M., Koscianska, E., & Krzyzosiak, W.J. 2011. Practical Aspects of microRNA Target Prediction. *Curr.Mol.Med*, 11, (2) 93-109 available from: PM:21342132

Woltjen, K., Michael, I.P., Mohseni, P., Desai, R., Mileikovsky, M., Hamalainen, R., Cowling, R., Wang, W., Liu, P., Gertsenstein, M., Kaji, K., Sung, H.K., & Nagy, A. 2009. piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature*, 458, (7239) 766-770 available from: PM:19252478

World Health Organisation 2006, *Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia*, WHO document production services, Geneva, Switzerland.

Xiao, X., Chen, Z., Shiota, C., Prasadan, K., Guo, P., El-Gohary, Y., Paredes, J., Welsh, C., Wiersch, J., & Gittes, G.K. 2013. No evidence for beta cell neogenesis in murine adult pancreas. *J Clin.Invest*, 123, (5) 2207-2217 available from: PM:23619362

Xie, R., Everett, L.J., Lim, H.W., Patel, N.A., Schug, J., Kroon, E., Kelly, O.G., Wang, A., D'Amour, K.A., Robins, A.J., Won, K.J., Kaestner, K.H., & Sander, M. 2013. Dynamic chromatin remodeling mediated by polycomb proteins orchestrates pancreatic differentiation of human embryonic stem cells. *Cell Stem Cell*, 12, (2) 224-237 available from: PM:23318056

Xu, H., Tsang, K.S., Chan, J.C., Yuan, P., Fan, R., Kaneto, H., & Xu, G. 2013. The combined expression of Pdx1 and MafA with either Ngn3 or NeuroD improves the differentiation efficiency of mouse embryonic stem cells into insulin-producing cells. *Cell Transplant.*, 22, (1) 147-158 available from: PM:22776709

Xu, N., Papagiannakopoulos, T., Pan, G., Thomson, J.A., & Kosik, K.S. 2009. MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. *Cell*, 137, (4) 647-658 available from: PM:19409607

Xu, X., Kahan, B., Forgianni, A., Jing, P., Jacobson, L., Browning, V., Treff, N., & Odorico, J. 2006. Endoderm and pancreatic islet lineage differentiation from human embryonic stem cells. *Cloning Stem Cells*, 8, (2) 96-107 available from: PM:16776601

Xu, X., D'Hoker, J., Stange, G., Bonne, S., De, L.N., Xiao, X., Van de Casteele, M., Mellitzer, G., Ling, Z., Pipeleers, D., Bouwens, L., Scharfmann, R., Gradwohl, G., & Heimberg, H. 2008. Beta cells can be generated from endogenous progenitors in injured adult mouse pancreas. *Cell*, 132, (2) 197-207 available from: PM:18243096

Xu, X., Browning, V.L., & Odorico, J.S. 2011. Activin, BMP and FGF pathways cooperate to promote endoderm and pancreatic lineage cell differentiation from human embryonic stem cells. *Mech.Dev.*, 128, (7-10) 412-427 available from: PM:21855631

Xue, Y., Ouyang, K., Huang, J., Zhou, Y., Ouyang, H., Li, H., Wang, G., Wu, Q., Wei, C., Bi, Y., Jiang, L., Cai, Z., Sun, H., Zhang, K., Zhang, Y., Chen, J., & Fu, X.D. 2013. Direct conversion of fibroblasts to neurons by reprogramming PTB-regulated microRNA circuits. *Cell*, 152, (1-2) 82-96 available from: PM:23313552

Yakhnenko, I., Wong, W.K., Katkov, I.I., & Itkin-Ansari, P. 2012. Cryopreservation of human insulin expressing cells macro-encapsulated in a durable therapeutic immunoisolating device theracyte. *Cryo.Letters.*, 33, (6) 518-531 available from: PM:23250411

Yamanaka, S. 2007. Strategies and new developments in the generation of patient-specific pluripotent stem cells. *Cell Stem Cell*, 1, (1) 39-49 available from: PM:18371333

Yang, L., Li, S., Hatch, H., Ahrens, K., Cornelius, J.G., Petersen, B.E., & Peck, A.B. 2002. In vitro trans-differentiation of adult hepatic stem cells into pancreatic endocrine hormone-producing cells. *Proc.Natl.Acad.Sci.U.S.A*, 99, (12) 8078-8083 available from: PM:12048252

Yang, X.J., Kow, L.M., Funabashi, T., & Mobbs, C.V. 1999. Hypothalamic glucose sensor: similarities to and differences from pancreatic beta-cell mechanisms. *Diabetes*, 48, (9) 1763-1772 available from: PM:10480606

Yasunaga, M., Tada, S., Torikai-Nishikawa, S., Nakano, Y., Okada, M., Jakt, L.M., Nishikawa, S., Chiba, T., Era, T., & Nishikawa, S. 2005. Induction and monitoring of definitive and visceral endoderm differentiation of mouse ES cells. *Nat.Biotechnol.*, 23, (12) 1542-1550 available from: PM:16311587

Yatoh, S., Akashi, T., Chan, P.P., Kaneto, H., Sharma, A., Bonner-Weir, S., & Weir, G.C. 2007. NeuroD and reaggregation induce beta-cell specific gene expression in cultured hepatocytes. *Diabetes Metab Res.Rev.*, 23, (3) 239-249 available from: PM:16921545

Yechoor, V., Liu, V., Paul, A., Lee, J., Buras, E., Ozer, K., Samson, S., & Chan, L. 2009. Gene therapy with neurogenin 3 and betacellulin reverses major metabolic problems in insulin-deficient diabetic mice. *Endocrinology*, 150, (11) 4863-4873 available from: PM:19819964

Yoo, A.S., Sun, A.X., Li, L., Shcheglovitov, A., Portmann, T., Li, Y., Lee-Messer, C., Dolmetsch, R.E., Tsien, R.W., & Crabtree, G.R. 2011. MicroRNA-mediated conversion of human fibroblasts to neurons. *Nature*, 476, (7359) 228-231 available from: PM:21753754

Yoshida, S., Kajimoto, Y., Yasuda, T., Watada, H., Fujitani, Y., Kosaka, H., Gotow, T., Miyatsuka, T., Umayahara, Y., Yamasaki, Y., & Hori, M. 2002. PDX-1 induces differentiation of intestinal epithelioid IEC-6 into insulin-producing cells. *Diabetes*, 51, (8) 2505-2513 available from: PM:12145164

Yoshida, Y., Takahashi, K., Okita, K., Ichisaka, T., & Yamanaka, S. 2009. Hypoxia enhances the generation of induced pluripotent stem cells. *Cell Stem Cell*, 5, (3) 237-241 available from: PM:19716359

Yoshioka, N., Gros, E., Li, H.R., Kumar, S., Deacon, D.C., Maron, C., Muotri, A.R., Chi, N.C., Fu, X.D., Yu, B.D., & Dowdy, S.F. 2013. Efficient generation of human iPSCs by a synthetic self-replicative RNA. *Cell Stem Cell*, 13, (2) 246-254 available from: PM:23910086

Young, M.A., Larson, D.E., Sun, C.W., George, D.R., Ding, L., Miller, C.A., Lin, L., Pawlik, K.M., Chen, K., Fan, X., Schmidt, H., Kalicki-Veizer, J., Cook, L.L., Swift, G.W., Demeter, R.T., Wendl, M.C., Sands, M.S., Mardis, E.R., Wilson, R.K., Townes, T.M., & Ley, T.J. 2012. Background mutations in parental cells account for most of the genetic heterogeneity of induced pluripotent stem cells. *Cell Stem Cell*, 10, (5) 570-582 available from: PM:22542160

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., Slukvin, I.I., & Thomson, J.A. 2007. Induced pluripotent stem cell lines derived from human somatic cells. *Science*, 318, (5858) 1917-1920 available from: PM:18029452

Yu, J., Hu, K., Smuga-Otto, K., Tian, S., Stewart, R., Slukvin, I.I., & Thomson, J.A. 2009. Human induced pluripotent stem cells free of vector and transgene sequences. *Science*, 324, (5928) 797-801 available from: PM:19325077

Yu, J., Hu, K., Smuga-Otto, K., Tian, S., Stewart, R., Slukvin, I.I., & Thomson, J.A. 2009. Human induced pluripotent stem cells free of vector and transgene sequences. *Science*, 324, (5928) 797-801 available from: PM:19325077

Yuan, H., Corbi, N., Basilico, C., & Dailey, L. 1995. Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. *Genes Dev.*, 9, (21) 2635-2645 available from: PM:7590241

Zalzman, M., Gupta, S., Giri, R.K., Berkovich, I., Sappal, B.S., Karnieli, O., Zern, M.A., Fleischer, N., & Efrat, S. 2003. Reversal of hyperglycemia in mice by using human expandable insulin-producing cells differentiated from fetal liver progenitor cells. *Proc.Natl.Acad.Sci.U.S.A*, 100, (12) 7253-7258 available from: PM:12756298

Zhang, C., Moriguchi, T., Kajihara, M., Esaki, R., Harada, A., Shimohata, H., Oishi, H., Hamada, M., Morito, N., Hasegawa, K., Kudo, T., Engel, J.D., Yamamoto, M., & Takahashi, S. 2005. MafA is a key regulator of glucose-stimulated insulin secretion. *Mol.Cell Biol.*, 25, (12) 4969-4976 available from: PM:15923615

Zhang, D., Jiang, W., Liu, M., Sui, X., Yin, X., Chen, S., Shi, Y., & Deng, H. 2009. Highly efficient differentiation of human ES cells and iPS cells into mature pancreatic insulin-producing cells. *Cell Res.*, 19, (4) 429-438 available from: PM:19255591

Zhang, Z.W., Zhang, L.Q., Ding, L., Wang, F., Sun, Y.J., An, Y., Zhao, Y., Li, Y.H., & Teng, C.B. 2011. MicroRNA-19b downregulates insulin 1 through targeting transcription factor NeuroD1. *FEBS Lett.*, 585, (16) 2592-2598 available from: PM:21781967

Zhao, L., Guo, M., Matsuoka, T.A., Hagman, D.K., Parazzoli, S.D., Poitout, V., & Stein, R. 2005. The islet beta cell-enriched MafA activator is a key regulator of insulin gene transcription. *J.Biol.Chem.*, 280, (12) 11887-11894 available from: PM:15665000

Zhao, X., Mohan, R., Ozcan, S., & Tang, X. 2012. MicroRNA-30d induces insulin transcription factor MafA and insulin production by targeting mitogen-activated protein 4 kinase 4 (MAP4K4) in pancreatic beta-cells. *J Biol.Chem.*, 287, (37) 31155-31164 available from: PM:22733810

Zhao, Y., Yin, X., Qin, H., Zhu, F., Liu, H., Yang, W., Zhang, Q., Xiang, C., Hou, P., Song, Z., Liu, Y., Yong, J., Zhang, P., Cai, J., Liu, M., Li, H., Li, Y., Qu, X., Cui, K., Zhang, W., Xiang, T., Wu, Y., Zhao, Y., Liu, C., Yu, C., Yuan, K., Lou, J., Ding, M., & Deng, H. 2008. Two supporting factors greatly improve the efficiency of human iPSC generation. *Cell Stem Cell*, 3, (5) 475-479 available from: PM:18983962

Zhou, H., Wu, S., Joo, J.Y., Zhu, S., Han, D.W., Lin, T., Trauger, S., Bien, G., Yao, S., Zhu, Y., Siuzdak, G., Scholer, H.R., Duan, L., & Ding, S. 2009. Generation of induced pluripotent stem cells using recombinant proteins. *Cell Stem Cell*, 4, (5) 381-384 available from: PM:19398399

Zhou, Q., Brown, J., Kanarek, A., Rajagopal, J., & Melton, D.A. 2008. In vivo reprogramming of adult pancreatic exocrine cells to beta-cells. *Nature*, 455, (7213) 627-632 available from: PM:18754011

Zhou, Q. & Melton, D.A. 2008. Extreme makeover: converting one cell into another. *Cell Stem Cell*, 3, (4) 382-388 available from: PM:18940730

Zhou, W. & Freed, C.R. 2009. Adenoviral gene delivery can reprogram human fibroblasts to induced pluripotent stem cells. *Stem Cells*, 27, (11) 2667-2674 available from: PM:19697349

Zhu, F.F., Zhang, P.B., Zhang, D.H., Sui, X., Yin, M., Xiang, T.T., Shi, Y., Ding, M.X., & Deng, H. 2011. Generation of pancreatic insulin-producing cells from rhesus monkey induced pluripotent stem cells. *Diabetologia*, 54, (9) 2325-2336 available from: PM:21755313

Zhu, H., Shyh-Chang, N., Segre, A.V., Shinoda, G., Shah, S.P., Einhorn, W.S., Takeuchi, A., Engreitz, J.M., Hagan, J.P., Kharas, M.G., Urbach, A., Thornton, J.E., Triboulet, R., Gregory, R.I., Altshuler, D., & Daley, G.Q. 2011. The Lin28/let-7 axis regulates glucose metabolism. *Cell*, 147, (1) 81-94 available from: PM:21962509

Zhu, S., Wurdak, H., Wang, J., Lyssiotis, C.A., Peters, E.C., Cho, C.Y., Wu, X., & Schultz, P.G. 2009. A small molecule primes embryonic stem cells for differentiation. *Cell Stem Cell*, 4, (5) 416-426 available from: PM:19427291

Zhu, S., Li, W., Zhou, H., Wei, W., Ambasudhan, R., Lin, T., Kim, J., Zhang, K., & Ding, S. 2010. Reprogramming of human primary somatic cells by OCT4 and chemical compounds. *Cell Stem Cell*, 7, (6) 651-655 available from: PM:21112560

Zweigerdt, R., Olmer, R., Singh, H., Haverich, A., & Martin, U. 2011. Scalable expansion of human pluripotent stem cells in suspension culture. *Nat.Protoc.*, 6, (5) 689-700 available from: PM:21527925

Appendix 1: Supplementary miRNA Tables

Supplementary Table 1
ESCs: Differentiated vs. Undifferentiated

miRNA	*p-value*	*logFC*
hsa-miR-124-5p	0.015	-0.207
hsa-miR-99b-5p	0.019	1.579
hsa-miR-483-5p	0.019	0.729
hsa-miR-375	0.019	0.502
hsa-miR-5572	0.019	0.398
hsa-miR-30d-5p	0.019	0.442
hsa-miR-4530	0.019	0.687
hsa-miR-4795-5p	0.021	-0.273
hsa-miR-26b-5p	0.021	0.574
hsa-miR-181d	0.021	2.007
hsa-miR-378a-3p	0.021	-0.309
hsa-miR-185-5p	0.023	0.251
hsa-miR-191-5p	0.023	0.453
hsa-miR-181a-5p	0.023	2.261
hsa-miR-3182	0.026	0.850
hsa-miR-484	0.027	0.129
hsa-miR-3202	0.028	0.571
hsa-miR-342-3p	0.028	0.680
hsa-miR-181b-5p	0.029	1.497
hsa-miR-4306	0.029	0.519
hsa-miR-4472	0.029	0.247
hsa-miR-744-5p	0.029	0.471
hsa-miR-3149	0.029	-0.311
hsa-miR-197-3p	0.029	0.256
hsa-miR-135b-5p	0.029	1.366
hsa-miR-4792	0.030	0.321
hsa-miR-4516	0.035	1.078
hsa-miR-4451	0.035	-0.351
hsa-miR-27b-3p	0.036	0.846
hsa-miR-1247-5p	0.038	0.127
hsa-miR-516b-5p	0.039	-0.232
hsa-miR-4800-3p	0.042	0.822
hsa-miR-1587	0.044	0.319
hsa-miR-4285	0.045	-0.555
hsa-miR-15a-5p	0.045	0.449
hsa-miR-5581-3p	0.046	0.204
hsa-miR-24-3p	0.050	0.943
hsa-miR-125a-5p	0.050	1.373
hsa-miR-708-5p	0.050	1.066
hsa-miR-151a-5p	0.050	0.764
hsa-miR-3935	0.050	-0.236
hsa-miR-30b-5p	0.050	0.699
hsa-miR-218-5p	0.050	0.778
hsa-miR-4297	0.051	-0.150
hsa-miR-454-3p	0.054	0.510
hsa-miR-16-5p	0.055	0.396
hsa-miR-320a	0.056	0.737
hsa-miR-320c	0.058	0.852
hsa-miR-519d	0.061	-0.244
hsa-miR-99b-3p	0.065	0.211
hsa-miR-5002-5p	0.065	-0.475
hsa-miR-4301	0.065	0.632
hsa-miR-3941	0.065	-0.480
hsa-miR-193b-3p	0.066	0.525
hsa-miR-151a-5p	0.066	0.507
hsa-miR-4639-3p	0.066	-0.439
hsa-miR-3148	0.066	-0.302
hsa-miR-3136-3p	0.066	-0.274
hsa-miR-150-5p	0.066	-0.146
hsa-miR-30e-3p	0.066	0.556
hsa-miR-378c	0.066	-0.147
hsa-miR-301a-3p	0.066	0.888
hsa-miR-149-5p	0.066	0.503
hsa-miR-877-5p	0.070	0.314
hsa-miR-331-3p	0.072	0.761
hsa-miR-340-5p	0.073	0.391
hsa-miR-1825	0.073	-0.123
hsa-miR-374a-5p	0.075	0.630
hsa-miR-4644	0.077	0.295
hsa-miR-23a-3p	0.077	1.047
hsa-miR-320b	0.077	0.739
hsa-miR-1249	0.077	-0.265
hsa-miR-3178	0.077	-0.454
hsa-miR-27a-3p	0.078	0.716
hsa-miR-149-3p	0.079	0.530
hsa-miR-4284	0.084	-0.427
hsa-miR-4292	0.088	0.152
hsa-miR-23b-3p	0.090	1.181
hsa-miR-151a-3p	0.094	0.406
hsa-miR-9-5p	0.096	2.705
hsa-miR-943	0.099	0.454
hsa-miR-30c-5p	0.100	0.684

Supplementary Table 2
iPSCs: Differentiated vs. Undifferentiated

miRNA	*p-value*	*logFC*
hsa-miR-708-5p	0.000	1.702
hsa-miR-4732-3p	0.005	-0.719
hsa-miR-32-3p	0.005	-0.719
hsa-miR-106a-5p	0.008	0.569
hsa-miR-302c-5p	0.009	0.420
hsa-miR-107	0.010	0.372
hsa-miR-93-5p	0.011	0.631
hsa-miR-18a-5p	0.011	0.209
hsa-miR-17-5p	0.011	0.544
hsa-miR-92a-3p	0.011	0.464
hsa-miR-9-3p	0.011	-0.331
hsa-miR-659-3p	0.011	-0.331
hsa-miR-103a-3p	0.013	0.633
hsa-miR-4289	0.017	0.236
hsa-miR-106b-5p	0.019	0.438
hsa-miR-151a-3p	0.019	0.188
hsa-miR-151a-5p	0.020	0.181
hsa-miR-4791	0.020	0.373
hsa-miR-302c-3p	0.020	0.558
hsa-miR-302f	0.020	0.339
hsa-miR-205-3p	0.020	-0.306
hsa-miR-210	0.024	0.222
hsa-miR-365a-3p/hsa-miR-365b-3p	0.035	0.229
hsa-miR-3591-5p	0.035	-0.406
hsa-miR-4436b-5p	0.037	-0.371
hsa-miR-4532	0.037	0.589
hsa-miR-18a-3p	0.037	0.154
hsa-miR-744-5p	0.038	0.655
hsa-miR-302a-5p	0.039	0.350
hsa-miR-3201	0.039	0.405
hsa-miR-3148	0.041	-0.216
hsa-miR-20a-5p	0.041	0.487
hsa-miR-375	0.042	0.876
hsa-miR-3941	0.042	-0.376
hsa-miR-148b-3p	0.042	0.209
hsa-miR-4505	0.042	0.755
hsa-miR-361-5p	0.042	0.239
hsa-miR-25-3p	0.042	0.250
hsa-miR-4484	0.042	0.803
hsa-miR-363-3p	0.042	0.477
hsa-miR-199a-5p	0.042	0.509
hsa-miR-491-3p	0.043	-0.679
hsa-miR-3611	0.044	0.634
hsa-miR-2116-5p	0.044	-0.242
hsa-miR-92b-3p	0.050	0.145
hsa-miR-4530	0.051	0.929
hsa-miR-1587	0.058	0.679
hsa-miR-149-3p	0.059	0.960
hsa-miR-4508	0.059	0.858
hsa-miR-3183	0.059	-0.214
hsa-miR-124-5p	0.063	-0.148
hsa-miR-4507	0.063	0.508
hsa-miR-3656	0.063	0.439
hsa-miR-26b-5p	0.066	0.356
hsa-miR-4285	0.066	-0.578
hsa-miR-130a-3p	0.066	0.342
hsa-miR-18b-5p	0.066	0.335
hsa-miR-30b-5p	0.068	0.252
hsa-miR-711	0.068	0.620
hsa-miR-620	0.068	-0.101
hsa-miR-155-5p	0.068	-0.450
hsa-miR-4646-3p	0.068	-0.266
hsa-miR-27b-3p	0.070	0.307
hsa-miR-320a	0.071	0.225
hsa-miR-1285-3p	0.071	-0.259
hsa-miR-638	0.071	0.364
hsa-miR-30c-5p	0.074	0.250
hsa-miR-191-5p	0.081	0.304
hsa-miR-423-3p	0.081	0.190
hsa-miR-2355-3p	0.081	-0.604
hsa-miR-361-3p	0.081	-0.272
hsa-miR-675-5p	0.081	-0.287
hsa-miR-302e	0.082	0.445
hsa-miR-20b-5p	0.089	0.404
hsa-miR-3676-3p	0.090	0.630
hsa-miR-302d-3p	0.091	0.426
hsa-miR-1976	0.092	-0.114
hsa-miR-3924	0.098	-0.485
hsa-miR-187-3p	0.098	0.124
hsa-miR-3651	0.098	0.306
hsa-miR-1255a	0.098	-0.265
hsa-miR-5701	0.098	0.380
hsa-miR-19b-3p	0.099	0.352
hsa-miR-320b	0.099	0.207

Supplementary Table 3
MRC5I: Differentiated vs. Undifferentiated

miRNA	*p-value*	*logFC*
hsa-miR-3676-3p	0.004	1.357
hsa-miR-4707-5p	0.004	0.776
hsa-miR-1469	0.009	0.761
hsa-miR-585	0.046	0.829
hsa-miR-4488	0.046	1.029
hsa-miR-4463	0.046	0.812
hsa-miR-708-5p	0.048	1.464
hsa-miR-4792	0.048	0.975
hsa-miR-638	0.048	0.707
hsa-miR-664-3p	0.048	0.193
hsa-miR-1909-3p	0.048	0.524
hsa-miR-652-5p	0.049	0.662
hsa-miR-3195	0.052	0.925
hsa-miR-1184	0.056	0.841
hsa-miR-4436b-5p	0.056	-0.439
hsa-miR-1290	0.056	0.786
hsa-miR-3941	0.059	-0.646
hsa-miR-365a-3p/hsa-miR-365b-3p	0.063	0.200
hsa-miR-1321	0.063	-0.169
hsa-miR-4791	0.072	0.141
hsa-miR-4449	0.074	0.617
hsa-miR-744-5p	0.085	0.713
hsa-miR-9-3p	0.085	-0.494
hsa-miR-659-3p	0.085	-0.494
hsa-miR-5681b	0.087	0.561
hsa-miR-4726-5p	0.089	0.218
hsa-miR-204-3p	0.089	0.993
hsa-miR-492	0.089	0.699
hsa-miR-634	0.089	0.536
hsa-miR-4674	0.089	0.757
hsa-miR-4532	0.094	0.766
hsa-miR-25-5p	0.094	1.321
hsa-miR-149-3p	0.095	1.275

Supplementary Table 4
MRC9G: Differentiated vs. Undifferentiated

miRNA	*p-value*	*logFC*
hsa-miR-708-5p	0.020	1.835
hsa-miR-4791	0.079	0.512
hsa-miR-361-5p	0.079	0.335
hsa-miR-106a-5p	0.079	0.546
hsa-miR-93-5p	0.079	0.667
hsa-miR-3183	0.079	-0.318
hsa-miR-4732-3p	0.079	-0.789
hsa-miR-32-3p	0.079	-0.789
hsa-miR-103a-3p	0.079	0.678
hsa-miR-106b-5p	0.079	0.415
hsa-miR-4456	0.079	-0.572
hsa-miR-4417	0.079	-0.404
hsa-miR-302c-3p	0.079	0.674
hsa-miR-4768-5p	0.079	-0.661
hsa-miR-20a-5p	0.079	0.529
hsa-miR-518e-5p/hsa-miR-519a-5p/hsa-miR-519b-5p/hsa-miR-519c-5p/hsa-miR-522-5p/hsa-miR-523-5p	0.079	-0.385
hsa-miR-302a-5p	0.079	0.415
hsa-miR-302f	0.079	0.467
hsa-miR-302e	0.079	0.688
hsa-miR-302d-3p	0.079	0.648
hsa-miR-107	0.084	0.433
hsa-miR-302c-5p	0.084	0.388
hsa-miR-27b-3p	0.085	0.308
hsa-miR-3158-5p	0.085	-1.273
hsa-miR-665	0.091	-0.551
hsa-miR-17-5p	0.092	0.505
hsa-miR-375	0.092	0.479
hsa-miR-18a-5p	0.092	0.200
hsa-miR-3915	0.092	0.422
hsa-miR-210	0.092	0.257
hsa-miR-18b-5p	0.092	0.477
hsa-miR-26b-5p	0.092	0.499
hsa-miR-320a	0.092	0.344
hsa-miR-361-3p	0.092	-0.462

Supplementary Table 5
H9: Differentiated vs. Undifferentiated

miRNA	*p-value*	*logFC*
hsa-miR-4800-3p	0.008	1.007
hsa-miR-302a-3p	0.015	-0.903
hsa-miR-4451	0.015	-0.442
hsa-miR-302b-3p	0.015	-0.727
hsa-miR-4516	0.015	1.423
hsa-miR-99b-5p	0.015	2.187
hsa-miR-3162-3p	0.015	0.613
hsa-miR-4532	0.015	0.670
hsa-miR-125a-5p	0.015	2.040
hsa-miR-181d	0.015	2.631
hsa-miR-125b-5p	0.015	3.228
hsa-miR-181b-5p	0.016	2.074
hsa-miR-483-5p	0.017	0.702
hsa-miR-520c-3p	0.017	-0.778
hsa-miR-181a-5p	0.017	3.151
hsa-miR-375	0.017	0.620
hsa-miR-4484	0.017	1.043
hsa-miR-943	0.017	0.676
hsa-miR-218-5p	0.017	1.227
hsa-miR-3202	0.017	0.658
hsa-miR-3687	0.017	1.083
hsa-miR-30c-1-3p	0.017	0.386
hsa-miR-99b-3p	0.017	0.304
hsa-miR-542-5p	0.017	0.460
hsa-miR-4324	0.019	1.349
hsa-miR-3651	0.019	-0.338
hsa-miR-9-5p	0.019	4.018
hsa-miR-518b	0.019	-0.280
hsa-miR-512-3p	0.020	-0.966
hsa-miR-4508	0.020	0.895
hsa-miR-4780	0.020	-0.354
hsa-miR-4534	0.020	0.637
hsa-miR-708-5p	0.021	1.082
hsa-miR-302c-3p	0.021	-0.654
hsa-miR-4530	0.025	0.742
hsa-miR-1469	0.025	0.341
hsa-miR-675-5p	0.026	0.599
hsa-miR-4792	0.026	0.310
hsa-miR-4306	0.026	0.579
hsa-miR-491-3p	0.029	0.731
hsa-miR-3656	0.029	0.360
hsa-miR-4714-5p	0.029	-0.555
hsa-miR-5572	0.029	0.460
hsa-miR-135a-5p	0.029	1.653
hsa-miR-4725-3p	0.029	0.316
hsa-miR-4791	0.030	-0.419
hsa-miR-5701	0.030	-0.395
hsa-miR-23b-3p	0.030	1.200
hsa-miR-301a-3p	0.030	1.053
hsa-miR-23a-3p	0.031	1.089
hsa-miR-454-3p	0.031	0.556
hsa-miR-4728-3p	0.031	-0.209
hsa-miR-5002-5p	0.032	-0.347
hsa-miR-340-3p	0.032	0.226
hsa-miR-4459	0.033	0.190
hsa-miR-19b-3p	0.037	-0.774
hsa-miR-519b-3p	0.037	-0.774
hsa-miR-135b-5p	0.037	1.584
hsa-miR-24-3p	0.037	1.095
hsa-miR-1825	0.037	-0.197
hsa-miR-1247-5p	0.037	0.207
hsa-miR-4472	0.039	0.287
hsa-miR-1184	0.039	0.317
hsa-miRPlus-G1246-3p	0.039	-0.479
hsa-miR-485-3p	0.040	0.141
hsa-miR-652-5p	0.041	0.165
hsa-miR-4301	0.042	0.871
hsa-miR-331-3p	0.043	0.702
hsa-miR-4726-5p	0.044	0.428
hsa-miR-197-3p	0.044	0.245
hsa-miR-335-5p	0.044	0.887
hsa-miR-4732-5p	0.045	0.687
hsa-miR-96-5p	0.045	-0.332
hsa-miR-3621	0.050	-0.392
hsa-miR-302e	0.051	-0.449
hsa-miR-371a-5p	0.051	-0.474
hsa-miR-320c	0.051	0.700
hsa-miR-4531	0.051	0.373
hsa-miR-22-3p	0.052	0.677
hsa-miR-4289	0.053	0.570
hsa-let-7a-5p	0.053	0.599
hsa-miR-483-3p	0.053	0.537
hsa-miR-4429	0.055	0.593
hsa-let-7e-5p	0.056	0.690
hsa-miR-302f	0.057	-0.554
hsa-miR-638	0.057	0.344
hsa-miR-204-3p	0.057	0.339
hsa-miR-30e-3p	0.057	0.608
hsa-miR-4657	0.057	-0.415

hsa-miR-519d	0.057	-0.440
hsa-miR-3607-3p	0.057	-0.692
hsa-miR-484	0.057	0.138
hsa-miR-193a-3p	0.057	0.327
hsa-miR-340-5p	0.060	0.579
hsa-miR-30d-5p	0.060	0.424
hsa-miR-3611	0.060	0.740
hsa-miR-320a	0.062	0.614
hsa-miR-27a-3p	0.062	0.536
hsa-miR-15a-5p	0.062	0.706
hsa-miR-374a-5p	0.063	0.580
hsa-miR-148b-3p	0.064	0.369
hsa-miR-711	0.064	0.144
hsa-miR-1275	0.064	0.470
hsa-miR-744-5p	0.064	0.548
hsa-miR-342-3p	0.064	0.809
hsa-miR-585	0.064	0.212
hsa-miR-3607-5p	0.064	-0.530
hsa-miR-874	0.064	-0.268
hsa-miR-200b-3p	0.064	-0.496
hsa-miR-26b-5p	0.065	0.569
hsa-miR-302a-5p	0.065	-0.606
hsa-miR-149-3p	0.065	0.482
hsa-miR-30b-5p	0.065	0.719
hsa-miR-30c-5p	0.065	0.835
hsa-miR-548ap-5p/hsa-miR-548j	0.065	0.232
hsa-miR-877-5p	0.065	0.357
hsa-miR-185-5p	0.066	0.184
hsa-miR-93-5p	0.066	0.829
hsa-miR-9-3p	0.066	0.257
hsa-miR-659-3p	0.066	0.257
hsa-miR-551b-3p	0.066	0.831
hsa-miR-4417	0.066	-0.315
hsa-miR-1915-3p	0.067	-0.912
hsa-miR-4644	0.069	0.474
hsa-miR-516b-5p	0.072	-0.276
hsa-miR-106b-5p	0.073	0.852
hsa-miR-4739	0.073	0.302
hsa-miR-25-3p	0.073	0.560
hsa-miR-3182	0.073	1.413
hsa-miR-4497	0.073	-0.276
hsa-miR-182-5p	0.073	-0.393
hsa-miR-4284	0.076	-0.616
hsa-miR-149-5p	0.077	0.552
hsa-miR-27b-3p	0.077	0.550
hsa-miR-4795-5p	0.078	-0.202
hsa-miR-3178	0.082	-0.291
hsa-let-7i-5p	0.083	-0.443
hsa-miR-4776-3p	0.086	-0.501
hsa-miR-378a-3p	0.086	-0.356
hsa-miR-5100	0.090	0.522
hsa-miR-373-3p	0.092	-0.385
hsa-miR-664-3p	0.092	-0.098
hsa-miR-1587	0.092	0.401
hsa-miR-200c-3p	0.092	-0.295
hsa-miR-3124-3p	0.092	0.358
hsa-miR-4698	0.093	0.383
hsa-miR-4708-3p	0.094	0.355
hsa-miR-16-5p	0.096	0.643
hsa-miR-4421	0.096	0.341
hsa-miR-4787-5p	0.096	0.249
hsa-miR-29a-3p	0.096	-0.326
hsa-miR-4292	0.096	0.232
hsa-miR-5581-3p	0.097	0.189

Supplementary Table 6
H1: Differentiated vs. Undifferentiated

miRNA	*p-value*	*logFC*
hsa-miR-3158-5p	0.048	-0.849
hsa-miR-375	0.048	0.404
hsa-miR-4639-3p	0.063	-0.443
hsa-miR-107	0.075	0.807
hsa-miR-3621	0.075	-0.799
hsa-miR-125a-5p	0.075	0.706
hsa-miR-378a-3p	0.075	-0.184
hsa-miR-181a-5p	0.075	1.216
hsa-miR-193b-3p	0.075	1.116
hsa-miR-3941	0.077	-0.462
hsa-miR-708-5p	0.083	1.095
hsa-miR-93-5p	0.083	0.538
hsa-miR-4732-3p	0.083	-0.864
hsa-miR-320a	0.083	0.859
hsa-miR-23a-3p	0.083	1.284
hsa-miR-181d	0.083	1.284
hsa-miR-23b-3p	0.083	1.370
hsa-miR-4436b-5p	0.083	-0.446
hsa-miR-24-3p	0.083	0.984
hsa-miR-4285	0.083	-0.868
hsa-miR-99b-5p	0.083	0.973
hsa-miR-331-3p	0.083	0.861
hsa-miR-877-5p	0.083	0.237
hsa-miR-103a-3p	0.084	1.001
hsa-miR-20a-5p	0.084	0.759
hsa-miR-27b-3p	0.084	1.400
hsa-miR-17-5p	0.084	0.807
hsa-miR-18b-5p	0.084	0.783
hsa-miR-26b-5p	0.084	0.659
hsa-miR-27a-3p	0.084	1.278
hsa-miR-4289	0.084	0.635
hsa-miR-151a-3p	0.084	0.780
hsa-miR-302b-3p	0.084	-0.278
hsa-miR-9-3p	0.084	0.933
hsa-miR-483-5p	0.084	0.896
hsa-miR-151a-5p	0.084	0.824
hsa-miR-135b-5p	0.084	0.946
hsa-miR-181b-5p	0.084	0.774
hsa-miR-30d-5p	0.084	0.576
hsa-miR-148b-3p	0.084	0.544
hsa-miR-365a-3p/hsa-miR-365b-3p	0.084	0.618
hsa-miR-744-5p	0.084	0.430
hsa-miR-423-3p	0.084	0.643
hsa-miR-191-5p	0.084	0.511
hsa-miR-130a-3p	0.084	0.571
hsa-miR-5701	0.084	0.640
hsa-miR-320c	0.084	1.068
hsa-miR-5572	0.084	0.367
hsa-miR-124-5p	0.084	-0.438
hsa-miR-1255a	0.084	-0.347
hsa-miR-664-3p	0.084	-0.083
hsa-miR-151a-5p/hsa-miR-151b	0.084	1.147
hsa-miR-320b	0.086	0.890
hsa-miR-92a-3p	0.088	0.624
hsa-miR-5002-5p	0.088	-0.454
hsa-miR-374a-5p	0.088	0.768
hsa-miR-200c-3p	0.088	1.019
hsa-miR-516b-5p	0.088	-0.190
hsa-miR-4306	0.088	0.501
hsa-miR-4429	0.088	0.318
hsa-miR-106a-5p	0.088	0.804
hsa-miR-149-3p	0.088	0.581
hsa-miR-3651	0.088	0.560
hsa-miR-3935	0.089	-0.336
hsa-miR-361-5p	0.090	0.313
hsa-miR-32-3p	0.090	-0.361
hsa-miR-4472	0.098	0.358
hsa-miR-342-3p	0.098	0.477

Supplementary Table 7
H1 Undifferentiated vs. iPSC Undifferentiated

miRNA	LogFC	p-value
hsa-miR-363-3p	-0.7661	0.0288
hsa-miR-371a-3p	1.9052	0.0288
hsa-miR-3183	-0.287	0.0288
hsa-miR-4417	1.1804	0.0288
hsa-miR-373-5p	0.9553	0.0288

Supplementary Table 8

Undifferentiated H9 vs. iPSC Undifferentiated

miRNA	LogFC	p-value
hsa-miR-4726-5p	-0.7851	0
hsa-miR-4534	-1.5359	0
hsa-miR-92a-3p	1.1306	0
hsa-miR-302a-3p	1.3641	0
hsa-miR-29a-3p	-2.866	0
hsa-miR-5701	0.8753	0
hsa-miR-302a-5p	1.2003	0
hsa-miR-200c-3p	-2.1629	0
hsa-miR-423-3p	1.2172	0
hsa-miR-4417	2.492	0
hsa-miR-92b-3p	1.3856	0
hsa-miR-17-5p	1.4209	0
hsa-miR-675-5p	-1.067	0
hsa-miR-302f	1.2158	0.0001
hsa-miR-106a-5p	1.4449	0.0001
hsa-miR-4714-5p	1.1868	0.0001
hsa-miR-130b-3p	1.4614	0.0001
hsa-miR-363-3p	1.7939	0.0001
hsa-miR-18b-5p	1.308	0.0001
hsa-miR-20a-5p	1.2902	0.0001
hsa-miR-22-3p	-2.0891	0.0001
hsa-miR-29b-3p	-2.0441	0.0001
hsa-miR-302c-3p	1.4419	0.0001
hsa-miR-4289	1.1549	0.0001
hsa-miR-3651	0.8524	0.0001
hsa-miR-659-3p	-0.9603	0.0001
hsa-miR-4501	-2.7858	0.0001
hsa-miR-30c-1-3p	-0.7297	0.0001
hsa-miR-4780	1.0119	0.0001
hsa-miR-425-5p	0.7575	0.0001
hsa-miR-125b-5p	-2.8101	0.0001
hsa-miR-361-5p	0.8992	0.0001
hsa-miR-17-3p	1.1663	0.0001
hsa-miR-4451	0.4071	0.0001
hsa-miR-4657	0.9052	0.0002
hsa-miR-103a-3p	1.2669	0.0002
hsa-miR-3175	0.7094	0.0002
hsa-miR-4508	-1.1411	0.0002
hsa-miR-20b-5p	1.3988	0.0002
hsa-miR-302d-5p	0.6607	0.0002
hsa-miR-199a-3p	-1.8071	0.0002
hsa-miR-3156-3p	0.3042	0.0002
hsa-miR-3654	0.628	0.0002
hsa-miR-4791	0.7508	0.0002
hsa-miR-106b-5p	0.95	0.0002
hsa-miR-3607-3p	1.0385	0.0003
hsa-miR-199a-5p	-1.5709	0.0003
hsa-miR-4532	-0.891	0.0003
hsa-miR-19b-3p	1.2728	0.0003
hsa-miR-151a-5p	0.7496	0.0003
hsa-miR-107	1.1011	0.0003
hsa-miR-151a-3p	0.5477	0.0003
hsa-miR-4484	-1.7824	0.0003
hsa-miR-767-5p	0.5764	0.0003
hsa-miR-5684	0.5542	0.0004
hsa-miR-302b-3p	1.209	0.0005
hsa-miR-4419b	-0.4261	0.0005
hsa-miR-3607-5p	0.8728	0.0006
hsa-miR-4747-5p	-1.325	0.0006
hsa-miR-1275	-0.763	0.0006
hsa-miR-31-5p	-1.2817	0.0007
hsa-miR-936	-1.4002	0.0007
hsa-miR-491-3p	-1.1362	0.0007
hsa-miR-302e	1.1314	0.0007
hsa-miR-130a-3p	0.878	0.0008
hsa-miR-708-5p	0.9549	0.0009
hsa-miR-151a-5p	0.9931	0.001
hsa-miR-302d-3p	0.903	0.001
hsa-miR-30c-5p	0.6628	0.001
hsa-miR-214-3p	-1.3979	0.0011
hsa-miR-141-3p	-1.197	0.0011
hsa-miR-148b-3p	0.5392	0.0011
hsa-miR-30b-5p	0.7513	0.0012
hsa-miR-302b-5p	0.8361	0.0013
hsa-let-7g-5p	-1.345	0.0013
hsa-miR-4324	-1.3979	0.0014
hsa-miR-25-3p	0.4876	0.0014
hsa-miR-18a-3p	0.6839	0.0014
hsa-let-7d-5p	-1.3636	0.0014
hsa-miR-493-5p	0.8346	0.0014
hsa-miR-320c	0.6394	0.0014
hsa-miR-520c-3p	1.0857	0.0014
hsa-let-7a-5p	-1.1651	0.0014
hsa-miR-3133	0.5877	0.0014
hsa-miR-135b-5p	1.34	0.0016
hsa-miR-200b-3p	0.8383	0.0017
hsa-miR-93-5p	0.8043	0.0017
hsa-miR-135a-5p	0.7005	0.0017
hsa-miR-3201	-0.7565	0.0017

hsa-miR-1827	-1.5736	0.0018		hsa-miR-3676-3p	0.9271	0.0081
hsa-miR-320a	0.4886	0.0021		hsa-miR-5100	-0.6512	0.0081
hsa-miR-3611	-0.9106	0.0023		hsa-miR-483-3p	-0.7586	0.0081
hsa-miR-19a-3p	0.7899	0.0024		hsa-miR-4787-5p	-0.4608	0.0087
hsa-miR-1323	1.1581	0.0027		hsa-miR-4454	-0.6105	0.0089
hsa-miR-124-3p	0.9673	0.0028		hsa-miR-711	-0.6411	0.0093
hsa-miR-4299	-0.8917	0.0029		hsa-miR-331-3p	0.4765	0.01
hsa-miR-34c-5p	-0.8611	0.0031		hsa-miR-518b	0.3005	0.0112
hsa-miR-4279	-0.4803	0.0032		hsa-miR-3924	-0.9507	0.0118
hsa-miR-1915-3p	1.53	0.0032		hsa-miR-143-3p	-0.8403	0.0119
hsa-miR-3124-3p	-0.5258	0.0032		hsa-miR-1264	-0.1717	0.0122
hsa-miR-320b	0.5851	0.0032		hsa-let-7c	-0.8383	0.0123
hsa-miR-4698	-0.4379	0.0033		hsa-miR-197-3p	0.239	0.0123
hsa-miR-4732-3p	-0.7974	0.0033		hsa-miR-30e-5p	0.4951	0.0144
hsa-let-7e-5p	-0.5298	0.0034		hsa-miR-516b-5p	0.2407	0.0144
hsa-miR-21-5p	-0.9867	0.0034		hsa-miR-1273g-3p	0.361	0.0144
hsa-miR-3653	0.6117	0.0034		hsa-miR-4521	0.6734	0.0144
hsa-miR-155-5p	-1.0108	0.0034		hsa-miR-2355-3p	-1.0654	0.0158
hsa-miR-335-3p	0.3802	0.0037		hsa-miR-15b-5p	0.5924	0.0158
hsa-miR-4800-5p	-1.1052	0.0038		hsa-miR-183-5p	0.679	0.016
hsa-miR-106b-3p	0.6589	0.0041		hsa-miR-642b-5p	-0.4632	0.0174
hsa-miR-4284	0.7562	0.0044		hsa-miR-1285-3p	0.3901	0.018
hsa-miR-340-5p	0.3042	0.0044		hsa-miR-585	0.6848	0.0184
hsa-miR-342-3p	0.7927	0.0045		hsa-miR-874	1.0185	0.0198
hsa-miR-4467	-0.5179	0.0046		hsa-miR-4708-3p	-0.438	0.0215
hsa-miR-3960	-0.4991	0.0046		hsa-miR-27b-3p	0.5597	0.0226
hsa-miR-3976	-0.2943	0.0049		hsa-miR-96-5p	0.2815	0.0248
hsa-miR-373-3p	-1.9053	0.005		hsa-miR-140-3p	0.6235	0.027
hsa-miR-4473	-0.7435	0.0052		hsa-miR-27a-3p	-0.4178	0.0274
hsa-miR-4459	-0.3123	0.0052		hsa-miR-15a-5p	0.4883	0.0274
hsa-miR-193a-3p	-0.6587	0.0058		hsa-miR-548an	0.4548	0.0277
hsa-miR-191-5p	0.7436	0.006		hsa-miR-638	0.3622	0.0278
hsa-miR-454-3p	0.5659	0.006		hsa-miR-3656	0.5232	0.0278
hsa-miR-34a-5p	0.6457	0.006		hsa-miR-4531	-0.2953	0.0283
hsa-let-7i-5p	-0.9706	0.006		hsa-miR-664-3p	0.2758	0.0285
hsa-miR-367-3p	0.7045	0.0061		hsa-miR-5193	0.3082	0.0302
hsa-miR-4421	-0.3504	0.0062		hsa-miR-16-5p	0.5091	0.0308
hsa-miR-371a-3p	-1.4247	0.0064		hsa-miR-4712-3p	0.2608	0.0311
hsa-miR-4285	-0.6305	0.0064		hsa-miR-4716-5p	0.1746	0.0323
hsa-miR-302c-5p	0.4494	0.0065		hsa-miR-548ap-5p	0.5216	0.0333
hsa-miR-4436b-5p	-0.4399	0.0069		hsa-miR-374b-5p	0.8743	0.0337
hsa-miR-3687	0.4422	0.0072		hsa-miR-4784	-0.3031	0.0358
hsa-miR-23a-3p	-0.9405	0.0074		hsa-miR-4500	-0.4531	0.0363
hsa-miR-4429	-0.6818	0.0075		hsa-miR-24-3p	-0.6587	0.0369
hsa-miR-424-5p	-0.4962	0.0075		hsa-miR-1587	-0.4785	0.0382
hsa-miR-3195	1.0117	0.0081		hsa-miR-205-3p	-0.3022	0.0383

hsa-miR-4530	-0.5563	0.0416
hsa-miR-374c-5p	0.7103	0.0437
hsa-miR-99b-5p	-0.5374	0.0464
hsa-miR-371a-5p	-1.1702	0.0466
hsa-let-7b-5p	-0.6981	0.0466

Appendix 2: Results of miRNA Target Prediction Algorithms

miR-151a-5p

miRanda (www.microrna.org):

Gene	miRanda Score
LOC541473	-1.67
LGI1	-1.35
FANCA	-1.25
XPO7	-1.2
UTY	-1.2
SEZ6L	-1.19
RALGAPA1	-1.16
CYP19A1	-1.16
IMPDH2	-1.14
PMFBP1	-1.11
SOX17	**-1.11**
TOM1L1	-1.1
CAPN10	-1.1
NDE1	-1.09
APH1A	-1.08
NXF2B	-1.07
HAS1	-1.05
FKBP1A	-1.04
SH2D4B	-1.04
KIAA1328	-1.03
PLP2	-1.03
RGS17	-1.03
CNTNAP4	-1.02
FLJ35776	-1.02
PATZ1	-1.02
C7orf63	-1
NRP2	-1
LY9	-0.99
KCNK1	-0.99
SLC44A5	-0.96
SUCLG2	-0.96
GALP	-0.95
HS3ST1	-0.95
DDX20	-0.95
DNMBP	-0.94
SNAP25	-0.94
PILRB	-0.94
KBTBD10	-0.94
SIGLEC1	-0.93
TOMM40	-0.92
CACNA1G	-0.91
WDR6	-0.91
C6orf136	-0.89
MTRR	-0.89
CEP72	-0.89
ARL4D	-0.88
NXF2	-0.88
CPSF3L	-0.88
TXNDC12	-0.87
KCNJ10	-0.87
FKBP6	-0.86
IQGAP1	-0.82
SLC25A12	-0.82
PABPN1	-0.82
PCSK6	-0.81
SLC25A24	-0.81
PRMT1	-0.8
AP3B1	-0.8
LHCGR	-0.8
CDC7	-0.8
TMEM98	-0.79
CHCHD3	-0.78
LOC644538	-0.78
ELAVL3	-0.78
VSTM2L	-0.77
PDIA5	-0.76
CALCOCO1	-0.76
FLJ11292	-0.75
ZNF638	-0.75
FAM45B	-0.74
CACHD1	-0.74
NAIP	-0.74
FOXP2	-0.73
EPHA6	-0.73
MEPCE	-0.73
RMST	-0.73
STX10	-0.72
PPP2R2A	-0.72
ERAP1	-0.72
BEST2	-0.71
RUNDC2B	-0.71
LOC442308	-0.69
MAX	-0.68
PPP1R3B	-0.68
SRRM1	-0.67
TP53TG3	-0.66
RASGRF2	-0.65
RNF220	-0.65
CLDN2	-0.64
ADAM18	-0.64
KALRN	-0.63
CROT	-0.63
PRODH	-0.63
SYNPR	-0.61
TBXAS1	-0.6
C1orf81	-0.6
TRYX3	-0.59
FAM45A	-0.5
EPHB1	-0.28
ERI1	-0.21

Note: Top 100 predictions only shown.

TargetScan (www.targetscan.org):

Gene	TargetScan Score
SOX17	**-0.75**
PQLC1	-0.75
IGSF9B	-0.69
SLC25A24	-0.64
AK2	-0.62
WNT1	-0.59
EFNB2	-0.59
LYPD3	-0.56
FANCA	-0.56
LPPR5	-0.56
BEST2	-0.55
SEZ6L	-0.55
DNMBP	-0.54
NTRK2	-0.54
SLC25A12	-0.53
GALP	-0.52
GLT1D1	-0.51
ERAP1	-0.51
DDX20	-0.51
GDI1	-0.5
IQGAP1	-0.5
PRODH	-0.49
CYB5B	-0.48
UTY	-0.48
SIGLEC1	-0.48
GPR44	-0.47
EVX1	-0.47
SLC44A5	-0.46
SLC7A5	-0.46
SIM2	-0.46
CARHSP1	-0.46
TRIM58	-0.45
APH1A	-0.45
N4BP1	-0.45
MPI	-0.45
CLEC4F	-0.44
TRPM2	-0.44
HOXA3	-0.43
LARP4B	-0.42
SERPINF2	-0.42
SFT2D3	-0.42
ERI1	-0.42
METTL19	-0.42
KDM2A	-0.41
MEST	-0.41
C9orf47	-0.41
SH3BP4	-0.41
CACNG7	-0.41
ZNF532	-0.41
BSN	-0.39
LHCGR	-0.39
OLFM4	-0.39
SH2D4B	-0.39
GPR173	-0.39
ZNF865	-0.38
CORO1B	-0.38
LRFN1	-0.38
ITGA10	-0.38
GNLY	-0.38
CHMP1A	-0.38
MAPK11	-0.37
ULK1	-0.37
PRLR	-0.37
CLDN2	-0.37
HAS1	-0.37
RNF40	-0.37
TAPBP	-0.37
CHRAC1	-0.36
RAB40C	-0.36
LRCH1	-0.36
SLC35E2B	-0.35
AFF4	-0.35
INPP5D	-0.35
ZNF282	-0.35
TMEM30B	-0.35
SUCLG2	-0.35
NRP2	-0.35
MBD3	-0.34
FXR2	-0.34
WHSC1	-0.34
HEYL	-0.34
TOMM40	-0.34
CDK11A	-0.34
DENND1A	-0.34
SNTA1	-0.34
RAB35	-0.34
WFIKKN2	-0.34
LRRC19	-0.34
RFX2	-0.34
TAL1	-0.34
C17orf28	-0.33
WDR5	-0.33
PILRB	-0.33
CACNA1G	-0.32
PDLIM4	-0.32
TXNDC12	-0.32
VSTM2L	-0.32
CALCOCO1	-0.32
FKBP1A	-0.32
CROT	-0.32

Note: Top 100 predictions only shown.

PicTar (pictar.mdc-berlin.de):

Gene	PicTar score
CACNB4	4.17
ZNF364	4.15
DBCCR1L	3.8
NELL1	3.45
VDAC3	3.4
PTX1	3.36
RBM5	3.04
TWIST1	2.91
FLJ38608	2.8
RGAG1	2.73
RAN	2.42
RDH10	2.31
ARID1B	2.21
ARID1B	2.21
ARID1B	2.21
INPP5A	2.2
CRK	2.19
HPCAL4	2.14
PLEKHM1	2.13
CRK	2.08
ROBO1	2.03
ROBO1	2.03
HNT	1.96
AMD1	1.7
SULF1	1.68
EPN2	1.62
EPN2	1.62
APRIN	1.55
SCC-112	1.51
DKFZp667B0210	1.43
CAMTA1	1.36
CAST	1.29
SEP11	1.14
SRGAP3	0.98
KLF12	0.61
OGT	0.36

miR-151a-3p

miRanda (www.microrna.org):

Gene	miRanda Score
RUNDC2C	-2.4
ZMAT1	-2.37
RERG	-2.16
EIF2C2	-2.1
F11	-2.05
RPS6KA5	-1.98
CNOT10	-1.95
FBXL3	-1.94
IKZF3	-1.87
SOCS5	-1.84
EPGN	-1.71
KCNH8	-1.63
C12orf40	-1.6
PKN2	-1.6
LOC200726	-1.57
SHC4	-1.56
KIAA1217	-1.51
GLS	-1.47
ZNF750	-1.45
PCBP2	-1.44
RNF6	-1.43
LOC387646	-1.43
EIF4G2	-1.42
ZFPM2	-1.42
KDELC2	-1.4
HDAC8	-1.39
PTPN12	-1.36
GABRA6	-1.34
UPP2	-1.33
PANK2	-1.33
PFN2	-1.32
GFM2	-1.32
LRRIQ1	-1.31
PRELID2	-1.31
PLCL2	-1.31
CSNK2A1	-1.31
RUNDC2B	-1.3
PLEKHA5	-1.29
GOSR2	-1.29
ARMC8	-1.29
ME1	-1.28
FAM120AOS	-1.28
PSMA5	-1.28
LUC7L	-1.27
CASD1	-1.27
IL26	-1.27
TIAM1	-1.26
YSK4	-1.26
SPIN3	-1.26
SIX1	-1.26
FXR1	-1.26
CLASP2	-1.26
CAPZA2	-1.25
FMN2	-1.24
LEPR	-1.24
TTL	-1.24
AP4E1	-1.24
GRM3	-1.23
LRP1	-1.23
CCDC132	-1.23
RAD54B	-1.23
PRKACB	-1.22
ZNF326	-1.22
TUBGCP4	-1.22
PSMA1	-1.22
COL24A1	-1.21
HTR1F	-1.21
NEUROD4	-1.21
TRA2B	-1.2
C16orf87	-1.2
PANX3	-1.2
CYTIP	-1.2
PITPNA	-1.2
CSNK2A1P	-1.2
FBXO4	-1.19
C15orf41	-1.19
HIF1A	-1.19
CCDC67	-1.18
LRRC37A2	-1.18
CHL1	-1.18
CLK1	-1.18
EXOC4	-1.17
KBTBD2	-1.17
FZD6	-1.17
ALPK1	-1.17
ACTR2	-1.17
SMARCAD1	-1.16
FAM13A	-1.16
PCDHB7	-1.16
PGBD5	-1.16
GABPA	-1.16
CCDC45	-1.14
SFRS2	-1.13
RANBP17	-1.1
NIPBL	-0.44
PMS2L5	-0.02

Note: Top 100 predictions only shown.

TargetScan (www.targetscan.org):

Gene	TargetScan Score
RPS6KA5	-0.68
KCNH8	-0.64
ZMAT1	-0.59
STXBP4	-0.58
EIF2C2	-0.56
ZNF326	-0.56
ZNF345	-0.54
KIAA1217	-0.52
GFM2	-0.51
FBXL3	-0.51
RPE	-0.49
SETBP1	-0.49
PITPNA	-0.48
ACTR2	-0.47
CLK1	-0.47
RGS6	-0.47
MS4A12	-0.46
ZFAND5	-0.46
AKT3	-0.46
ME1	-0.46
UPP2	-0.46
ARMC8	-0.46
MANEA	-0.45
NEUROD4	-0.45
GLS	-0.45
PANK2	-0.45
PCBP2	-0.45
LIG4	-0.44
CEACAM5	-0.43
LOC200726	-0.43
SOCS5	-0.43
A2ML1	-0.43
PAPOLG	-0.42
AQP4	-0.42
KLHL4	-0.42
PFN2	-0.42
ADIPOQ	-0.42
TRA2B	-0.41
ANKRD44	-0.41
PTPN12	-0.41
CLASP2	-0.41
DISC1	-0.4
EXOC4	-0.4
YSK4	-0.4
CP	-0.4
TRDN	-0.39
CHL1	-0.39
LUC7L	-0.39
NKAIN3	-0.39
CEP95	-0.39
C14orf119	-0.39
GABRA6	-0.38
ZFPM2	-0.38
FXR1	-0.38
ZNF785	-0.38
RYBP	-0.38
C1orf9	-0.38
KRTAP2-2	-0.38
SEC22C	-0.38
FAM170A	-0.37
HTR1F	-0.37
CASD1	-0.37
KNG1	-0.37
LRCH1	-0.37
PRKACB	-0.37
CYP7B1	-0.36
NIPAL2	-0.36
SERPINA1	-0.36
GLYCTK	-0.36
ATP2A2	-0.36
RAB3GAP1	-0.36
NEURL1B	-0.36
FAM76B	-0.36
DTX3L	-0.35
C17orf76	-0.35
PURB	-0.35
OSBPL3	-0.35
RBM7	-0.35
RERG	-0.35
C6orf106	-0.35
FAM199X	-0.35
KDELC2	-0.35
ATF1	-0.34
MFSD4	-0.34
TNPO1	-0.34
ASXL2	-0.34
RBM27	-0.34
CAPZA2	-0.34
ZNF24	-0.34
ENPP1	-0.34
TPM1	-0.34
PGAM1	-0.34
PPP2CB	-0.34
PTGER3	-0.34
COL6A6	-0.34
FEM1C	-0.34
COL24A1	-0.33
CYTIP	-0.33
CCDC22	-0.33

Note: Top 100 predictions only shown.

PicTar (pictar.mdc-berlin.de):

Gene	PicTar score
CACNB4	4.17
ZNF364	4.15
DBCCR1L	3.8
NELL1	3.45
VDAC3	3.4
PTX1	3.36
RBM5	3.04
TWIST1	2.91
FLJ38608	2.8
RGAG1	2.73
RAN	2.42
RDH10	2.31
ARID1B	2.21
INPP5A	2.2
CRK	2.19
HPCAL4	2.14
PLEKHM1	2.13
CRK	2.08
ROBO1	2.03
HNT	1.96
AMD1	1.7
SULF1	1.68
EPN2	1.62
APRIN	1.55
SCC-112	1.51
DKFZp667B0210	1.43
CAMTA1	1.36
CAST	1.29
SEP11	1.14
SRGAP3	0.98
KLF12	0.61
OGT	0.36

Appendix 3: List of Reagents

Reagent	Supplier
100% Ethanol	Sigma
100bp DNA Ladder	Promega
1kbp DNA Ladder	Promega
2x Sensimix Plus	Bioline
5x PCR Reaction Buffer	Promega
5x Reverse Transcription Reaction Buffer	Promega
6x Blue/Orange Loading Dye	Promega
Accutase™	Stem Cell Technologies
Activin A	R & D Systems
Agarose	Sigma
Alkaline Phosphatase Detection Kit	Millipore
AllPrep DNA/RNA/Protein mini kit	Qiagen
B27® Supplement	Invitrogen
bFGF	Invitrogen
Blasticidin	Sigma
BMP4	R & D Systems
Chloroform	Sigma
DAPT	R & D Systems
Dispase	Sigma
DMEM	Invitrogen
DMSO	Sigma
dNTPs	Promega
DTT	Promega
EDTA	Sigma
Ethidium Bromide	Sigma
Ethylene Glycol	Sigma
Exendin4	R & D Systems
FuGENE® 6 Transfection Reagent	Promega
Geneticin	Invitrogen
Glycerol Stocks Of E. Coli Containing pMX Plasmids	Addgene
Goat Anti-Human Oct4 IgG	Abcam
Goat Anti-Human Sox17 IgG	Abcam
Goat Anti-Mouse IgG + IgM (FITC)	Abcam
Goat IgG	Abcam
GoTaq Polymerase	Promega

HEPES Buffer	Invitrogen
HGF	R & D Systems
Hoescht DNA Stain	Sigma
Hygromycin	Sigma
Hyperladder	Bioline
IDE1	StemGent
IGF-1	R & D Systems
KAAD-Cyclopamine	Millipore
KGF	R & D Systems
Knockout Serum Replacement	Invitrogen
L-Glutamine	Invitrogen
Lipofectamine 2000 ®	Invitrogen
Matrigel™	BD
$MgCl_2$	Promega
miArrest™/miExpress™ miRNA Lentiviral Vectors	Genecopoeia
miRNeasy™ Mini Kit	Qiagen
miScript™ Reverse Transcription Kit	Qiagen
miScript™ SYBR Green PCR Kit	Qiagen
miTarget ™ 3'UTR-Luciferase Reporter Constructs	Genecopoeia
Mitomycin C	Sigma
MMLV Reverse Transcriptase	Promega
Mouse Anti-Human TRA-1-81 IgM	UK Stem Cell Bank
Mouse IgG	Abcam
mTeSR-1™	Stem Cell Technologies
Nanog Antibody	Abcam
Noggin	R & D Systems
Non-Essential Amino Acids	Invitrogen
$Oligo(DT)_{15}$ Primers	Promega
Opti-MEM	Invitrogen
PCR-Grade Water	Sigma
Penicillin-Streptomycin	Invitrogen
pLenti6-mSlc7a1 Plasmid	AddGene
Polybrene	Sigma
Puromycin	Sigma
QIAquick Gel Extraction Kit	Qiagen
Rabbit Anti-Goat Alexafluor® 488	Abcam
Random Primers	Promega
Retinoic Acid	Sigma
RNA 6000 Nano Bioanalyzer Chips	Agilent
RNA 6000 Nano Bioanalyzer Reagents	Agilent

RNase-Free DNase Kit	Qiagen
RNAsin	Promega
RNeasy Mini Kit	Qiagen
RPMI	Invitrogen
SecretePair™ Dual Luminescence Assay Kit	Genecopoeia
SNAP Midiprep Kit	Invitrogen
Sodium Pyruvate	Invitrogen
TRA-1-60 StainAlive Dye	Stemgent
Trypan Blue	Sigma
Trypsin/EDTA	Sigma
Valproic Acid	Sigma
VEGF	R & D Systems
Virapower Lentiviral Expression System	Invitrogen
Wnt3A	R & D Systems
Y-27632	Sigma
β –Mercaptoethanol (50mM)	Invitrogen
β-Mercaptoethanol	Sigma