CHAPTER 4 DISCUSSION

4.1. DE and pancreatic differentiation of hPSCs *in vitro* mimics developmental events during pancreatic formation in humans

Human PSCs offer a unique opportunity to study disease phenotypes not reproduced in model organisms such as the mouse. This is particularly relevant for my project where a discrepancy in genetics and the subsequent disease phenotype has been observed between mice and human. In this aspect, efficient generation of pancreatic cell types *in vitro* using hPSCs presents the first step toward successful disease modelling to potentially provide insights into the molecular mechanisms underlying pancreatic agenesis. Indeed, my results show that hPSCs can be efficiently differentiated into near homogenous populations of DE and pancreatic progenitor cells using several established defined culture systems. Importantly, the DE and pancreatic progenitor cells generated across these protocols follow a normal path of development, with the initial down-regulation of pluripotency genes such as *NANOG*, *OCT4* and *SOX2* followed by the up-regulation of DE markers *CXCR4* and *SOX17*, and the subsequent up-regulation of key pancreatic-specific genes such as *PDX1* and *NKX6-1*. In addition, GO analyses derived from RNA-seq show endoderm (Figure 106) and pancreatic (Figure 126) development among the top enriched pathways, further reinforcing these observations.

It has been well documented that hESCs and hiPSCs share many similar properties such as morphology, proliferation, gene expression, and the ability to differentiate into various cell types etc. (Takahashi et al., 2007, Takahashi and Yamanaka, 2006, Evans, 2011). However, variation in the efficiencies of differentiation has been reported between different hPSC lines (Osafune et al., 2008, Chin et al., 2009). Hence, it was not surprising to observe a difference in β-cell specification efficiency between the H9 and FSPS13.B cell lines despite their similar pancreatic progenitor specification efficiencies. The process of β-cell differentiation is controlled by a complex network involving tight regulation of genes required for the development of the pancreas. Naturally, culture conditions play a critical role in

determining the successful generation of pancreatic insulin-secreting β cells from hPSCs (Rostovskaya et al., 2015). Since the difference in β-cell specification efficiency between the H9 and FSPS13.B cell lines was observed using the same lab-derived protocol, variation in the efficiencies of differentiation is most likely explained by the different genetic backgrounds of the H9 and FSPS13.B cell lines.

The goal of deriving functional β-cells from hPSCs still remains a major challenge. Although substantial improvement has been made to differentiate hPSCs toward functional pancreatic β-cells, existing protocols for *in vitro* differentiation produce immature pancreatic β-cells that are not highly responsive to glucose stimulation. Pancreatic β-cell maturation is characterised by the ability of the differentiated β-cells to perform glucose-stimulated insulin secretion (GSIS). This challenge presents a hindrance to the use of hPSCs in applications such as disease modelling, where differentiation of hPSCs into mature, glucose responsive β-cells is required for establishing the disease phenotypes *in vitro*, and to understanding the molecular mechanisms underlying different forms of diabetes. While previously published protocols have shown an improvement in producing glucose-responsive insulin-secreting β-cells *in vivo*, the GSIS of the β-cells *in vitro* still remains limited, indicating an immature nature of these cells (Maehr et al., 2009, Zhang et al., 2009, Jiang et al., 2007a, D'Amour et al., 2006). Furthermore, a recent study has reported that insulin-secreting β cells differentiated from hPSCs are highly similar to human fetal pancreatic β-cells and do not resemble adult β-cells (Hrvatin et al., 2014). This could be one of the factors responsible for the inverse GSIS response that I observed with the H9 cells, where the immature nature of the cells impedes the cells from proper function.

Another factor that could have hindered the success of eliciting a GSIS response from β-cells *in vitro* is the low efficiency in producing insulin-secreting β-cells. With only approximately 10% and 6% of insulin-secreting β-cells generated from H9 or FSPS13.B cells respectively, the current lab-derived protocol is most likely lacking critical signals required for efficient generation of insulin-secreting β cells. Perhaps one solution to circumvent these problems is to adopt more recently

published protocols that report an increased efficiency in generating an average of 30-50% of insulin-producing cells (Rezania et al., 2014, Pagliuca et al., 2014).

4.2. *GATA6* and *GATA4* expression patterns during human pancreatic development

The expression pattern of *GATA6* and *GATA4* during human pancreas development has not been well characterised to date. My work has revealed the precise expression kinetics of *GATA6* and *GATA4* during *in vitro* differentiation into the pancreatic lineage.

That *GATA6* is expressed from the DE stage and remains expressed throughout pancreatic development suggests an important role of this transcription factor during human pancreas development. Interestingly, the expression pattern of *GATA4* is highly similar to that of *GATA6* in that it is not expressed in pluripotent cells and its expression is first observed in the DE stage and remains expressed throughout pancreatic development. This indicates a similar and possibly redundant role of both transcription factors. These findings confirm previous studies where both *GATA6* and *GATA4* have been reported to be expressed in DE cells in hPSC differentiation cultures (McLean et al., 2007, Vallier et al., 2009b). Similarly in mice, *Gata6* and *Gata4* are both expressed in the DE and its derivatives, including the pancreas (Decker et al., 2006, Watt et al., 2007).

My results show that *GATA4* levels are consistently more highly expressed than *GATA6* at the later stages of pancreatic development (D9 onwards), suggesting a critical role of *GATA4* in the development of the human pancreas. This is consistent with a previous report describing GATA4 expression at the onset of pancreatic development in human embryos, although it was unknown from the study whether GATA4 is also expressed in DE cells prior to pancreas formation (Jennings et al., 2013). It has also been reported that *GATA4* mutations are a rare cause of NDM and pancreatic agenesis in five patients harbouring *GATA4* mutations, confirming a role for *GATA4* in the development of the human pancreas (Shaw-Smith et al., 2014).

4.3. TALEN as a genome editing tool for disease modelling

The successful generation of TALEN-derived *GATA6* mutant hESC and hiPSC lines via both NHEJ and HR shows the versatility of TALENs as a genome editing tool for disease modelling. Interestingly, using the same TALEN cut sites, the cutting efficiency in H9 cells was observed to be higher than in FSPS13.B cells, possibly due to differences in nuclease cleavage efficiencies and/or intrinsic differences in activities of DNA repair pathways.

In addition to generating mutant lines, TALENs can also be used to correct mutations via homologous recombination. In the context of my project, it would be useful to correct the missense *GATA6* mutation in Patient A and Patient B to derive isogenic wild-type control cell lines as this would eliminate differences arising from different genetic backgrounds. However, this was not performed due to time constraints and was not prioritised because the patient phenotypes were similar to the TALEN-generated mutants which indicated the suitability of the TALEN-generated mutants as a disease modelling platform.

The similar differentiation efficiencies between the untargeted hPSCs (H9 and FPSP13.B) and their respective isogenic controls, which are targeted hPSCs but harbour no observable mutation around the target site (H9* and FPSP13.B*), indicated that off-target effects (if any) did not affect pancreatic specification. The normal karyotype displayed by all the targeted wild-type and mutant hPSC lines over many passages also suggested that the TALEN targeting did not introduce any gross chromosomal rearrangements and abnormalities such as deletions, inversions or translocations in addition to local mutations that can occur when imprecise repair of on- and off-target DNA cleavages take place (Lee et al., 2010, Lee et al., 2012, Park et al., 2014, Brunet et al., 2009, Cho et al., 2014).

The ability of the TALEN-derived mutant hPSC lines to retain pluripotency similarly to wild-type and untargeted hPSCs indicate that genome editing via TALENs did not affect the pluripotency status of hPSCs. As *GATA6* is not expressed in

undifferentiated cells, it is unlikely that this transcription factor plays an important role in pluripotency, which is evident from the non-effect that loss of *GATA6* has on pluripotency.

4.4. *GATA6* is required for DE specification

Using the TALEN-generated *GATA6* homozygous knockout H9 (*GATA6* $^{\Delta4/\Delta4}$ and *GATA6* $^{GFP/GFP}$) and FSPS13.B (*GATA6* $^{\Delta14/\Delta11}$) mutant cell lines, I have shown that *GATA6* is essential for the formation of the DE in humans. Although truncated proteins were detected in H9-*GATA6* $^{\Delta4/\Delta4}$ and FSPS13.B-*GATA6* $^{\Delta14/\Delta11}$ mutant cell lines, their phenotypes were indistinguishable from the H9-*GATA6* $^{GFP/GFP}$ mutant cell line that had no detectable GATA6 protein. This indicates that, consistent with its known biochemical characteristics (Bates et al., 2008, Molkentin, 2000), the truncated GATA6 protein lacking the C-terminal DNA-binding zinc-finger domains is non-functional. Thus, it can be inferred that in the absence of functional GATA6 proteins, a human embryo would likely fail to form the pancreas as a consequence to a primary defect in definitive endoderm formation. This finding, coupled with the rapid and strong up-regulation of *GATA6* at the DE stage, as well as its co-localisation with key DE marker *SOX17* strongly suggests that *GATA6* does indeed play an important role in DE formation.

Using the TALEN-generated *GATA6* heterozygous knockout H9 (*GATA6* $^{4ins/+}$ and *GATA6* $^{GFP/+}$) and FSPS13.B (*GATA6* $^{1ins/+}$, *GATA6* $^{\Delta14/+}$, *GATA6* $^{\Delta21\_8ins/+}$ and *GATA6* $^{GFP/+}$) mutant cell lines, as well as the patient-derived *GATA6* heterozygous mutant lines Patient A and Patient B, an impairment in DE formation was observed using the lab-derived protocol and STEMdiff pancreatic progenitor kit from SCT, with FSPS13.B-*GATA6* $^{GFP/+}$ mutant cells being the exception. However, a discrepancy was observed in cells differentiated via the PSC Definitive Endoderm Induction Kit from Life Technologies, where the DE formed as efficiently as the wild-type controls. Hence, it is possible that the DE phenotypes observed in the heterozygous mutants differentiated via the lab-derived protocol and STEMdiff pancreatic progenitor kit from SCT are a consequence of the DE differentiation protocol used. Unfortunately, I

am unable to verify this hypothesis through detailed comparison of the protocols by eliminating/adding certain growth factors due to the restrictions imposed by Life Technologies, where I was unable to obtain more information on the media formulation.

Furthermore, the fundamental differences in the differentiation protocols may underlie (or contribute to) the results I obtained and those recently published by Shi et al. and Tiyaboonchai et al. (Shi et al., 2017, Tiyaboonchai et al., 2017). For example, the growth factor and small molecule components as well as medium formulations differ substantially for the first three days of DE differentiation among the three studies. This was further evidenced in the study led by Tiyaboonchai where the group showed that a *GATA6* heterozygous iPSC line derived from an agenesis patient unexpectedly produced beta-like cells *in vitro* by simply reducing the concentration of retinoic acid 80-fold (Tiyaboonchai et al., 2017). This change led to statistically significantly fewer $PDX1^{+}$ cells from the patient line when compared to a wild-type iPSC line that showed negligible sensitivity to the same culture regime.

Comparing across the patient-derived and TALEN-derived H9 and FSPS13.B heterozygous mutants, a spectrum of DE phenotypes was observed using the lab-derived and SCT protocols; FSPS13.B-*GATA6* $^{GFP/+}$ mutant cells displayed no defect in DE formation, and FSPS13.B NHEJ-generated mutants as well as Patients A and B displayed a similar but a slightly more severe DE defected compared to H9-*GATA6* $^{4ins/+}$ and *GATA6* $^{GFP/+}$ mutant cells. This observation, however, was unsurprising as recent studies have reported that *GATA6* mutations can cause diverse diabetic phenotypes, ranging from pancreatic agenesis to adult-onset diabetes where most, but not all, patients display exocrine insufficiency requiring insulin treatment and enzyme replacement therapy, and other extrapancreatic features (De Franco et al., 2013, Bonnefond et al., 2012).

In the report by Bonnefond et al., two French sisters were described with the same *GATA6* allele (c.1504_1505del; p.Lys502Aspfs*5) but presented strikingly different clinical manifestations—one with permanent neonatal diabetes and the other without (Bonnefond et al., 2012). Similarly, Shi et al. (2017) engineered using CRISPR/Cas9 the common *GATA6* agenesis mutation c.1366C>T (p.Arg456Cys) in HUES8 cells—the same allele present in the patient A-derived iPSC line (*GATA6*$^{R456C/+}$)—and observe no heterozygous phenotype at the DE or pancreatic progenitor (PDX1^{+}) stages (Shi et al., 2017), whereas I do, both at the DE stage and beyond. A more recent publication by Yau et al. (2017) describes the inheritance of a novel *GATA6* frame-shift mutation (c.635_660del; p.Pro212fs) in three children from a *GATA6* mosaic mother, and each child (one is an abortus) presents a different phenotype (Yau et al., 2017).

By analogy, it is entirely possible that hiPSC derived from the patients in the Yau et al. (2017) and Bonnefond et al. (2012) studies would each behave entirely differently when differentiated *in vitro*. The simplest explanation for the existence of such "resilient individuals" who are not impacted by deleterious *GATA6* alleles is the influence of modifier genes and rare variants attributable to individual genetic backgrounds (Lek et al., 2016, Chen et al., 2016). *GATA4* is an obvious choice for a genetic modifier, given its expression in the DE, genetic interaction with *Gata6* in mice, the identification of rare *GATA4* heterozygous patients with pancreatic agenesis as well as our finding that *GATA4* is bound and regulated by GATA6 *in vitro* (Figure 113) (Morrisey et al., 1996b, Freyer et al., 2015, Shaw-Smith et al., 2014, D'Amato et al., 2010). Indeed, Shi et al. (2017) elegantly show dosage-dependent effects of *GATA4* alleles on phenotypes associated with *GATA6* heterozygosity during *in vitro* differentiation.

The observation of extrapancreatic abnormalities in *GATA6* patients, which include malformations in endodermal-derived organs such as congenital heart defects (Kodo et al., 2009, Lin et al., 2010), hepatobiliary malformations, gall bladder agenesis, and gut herniation (Allen et al., 2012), further provide evidence that *GATA6* plays an important role in the development of the DE. In addition, the two patients in the family studied in Yau et al. (2017) present defects in a number of endoderm-derived organs, further supporting that diminished GATA6 levels during DE formation underlie a constellation of clinical endodermal phenotypes (Yau et al., 2017). Indeed, when I ran pilot differentiations to specify H9* and H9-*GATA6* $^{4ins/+}$ mutant cells into the hepatic lineage, I observed decreased differentiation efficiencies for H9-*GATA6* $^{4ins/+}$ mutant cells to differentiate into hepatic progenitors, the precursors of hepatic cells (data not shown). Unfortunately, I was unable to successfully differentiate the wild-type H9 cells into mature hepatocytes as the hepatic differentiation protocol was still being optimised in the lab when the experiments were performed. As such, I was unable to perform phenotypic comparisons between H9* and H9-*GATA6* $^{4ins/+}$ mutant cells at a later hepatic development stage.

4.5. *GATA6* is required for pancreatic progenitor specification

My work has demonstrated that *GATA6* is required for pancreatic specification. Surprisingly, despite the broad spectrum of phenotypes observed at the DE stage, a less variable phenotype of 50-80% down-regulation of *PDX1* across all heterozygous mutants was seen at the pancreatic stage and this was consistently observed using all three DE specification protocols. This finding strongly indicates that the pancreatic phenotype seen in the *GATA6* heterozygous mutant cell lines is most likely a true effect of *GATA6* haploinsufficiency, thus establishing a human *in vitro* hPSC model system to study the role of *GATA6* in the development of the human pancreas. However, this present system has its limitations. As heterozygous *GATA6* mutations have been reported to have incomplete penetrance as displayed by different phenotypes in family members having identical mutations (Bonnefond et al., 2012), it is possible that my *in vitro* PSC model system is lacking intrinsic signalling pathways or factors present *in vivo* that may mitigate the negative effects of *GATA6* haploinsufficiency, thus driving a more severe phenotype.

The discordant phenotypes between mice and human models, especially for haploinsufficient disease genes, have been observed and widely discussed (Seidman and Seidman, 2002). In my hPSC model system, deleting one allele of *GATA6* impaired pancreatic formation as seen from the reduction of PDX1+ cells across all genetic backgrounds of the hPSCs used in this study. This finding demonstrates phenotypes not previously reported in mice (Morrisey et al., 1998, Carrasco et al., 2012, Xuan et al., 2012). For instance, pancreatic defects were not observed in *Gata6*$^{-/+}$ or *Gata4*$^{-/+}$ mouse embryos or adults. This suggests distinct gene dosage sensitivities between both species. Furthermore, the genetic background of *GATA6* patients is much more diverse than the inbred mouse strains. Thus, extreme phenotypes such as pancreatic agenesis may be seen in some, but not all, *GATA6* heterozygous patients. Lastly, the different timing where *GATA6* is deleted between the mice and human model systems could also contribute toward the discordant phenotypes between both systems. Due to early embryonic lethality of *Gata6* and *Gata4* embryos caused by extraembryonic defects (Morrisey et al., 1998), it was

necessary for *Gata6/4* to be conditionally inactivated using the Cre-LoxP system in the early pancreatic progenitors or the gut endoderm stages prior to pancreatic specification (Carrasco et al., 2012, Xuan et al., 2012). In contrast, in my *in vitro* hPSC model system, *GATA6* is deleted in pluripotent cells, before the initiation of differentiation. An inducible knockout system, such as a tetracycline-inducible shRNA or doxycycline-inducible CRISPR interference system, where *GATA6* can be inactivated at specific stages such as the gut endoderm could be a suitable method to replicate the mice model more closely (Bertero et al., 2016, Mandegar et al., 2016). Despite these differences, *GATA6/4* interactions are observed in both mice and hPSC model systems, supporting the use of both systems for investigating genetic and environmental modifiers in *GATA6*-linked human disease.

4.6. *GATA6* is a key regulator of DE and pancreatic progenitor specification

The molecular mechanism by which *GATA6* controls DE and pancreatic specification in humans was not known prior to my study. My results using genome-wide transcriptional analyses from RNA-seq revealed that *GATA6* transcriptionally activates the expression of endoderm markers in human. Interestingly, loss of both alleles of *GATA6* leads to an increase in mesoderm development. This suggests a possible role of *GATA6* not only to enable endoderm formation, but also to suppress mesoderm formation.

For the first time, results from ChIP-seq data suggest a direct molecular mechanism whereby GATA6 directly controls the gene expression of endoderm markers such as SOX17 and CXCR4, placing it centrally in the regulation of endoderm specification. The direct binding of GATA6 to GATA4 at both the DE and pancreatic progenitor stages also indicates that GATA6 and GATA4 are interacting partners, a finding that has also been previously reported in the developing and postnatal myocardium (Charron et al., 1999). Thus, the down-regulation of *GATA4* expression in the *GATA6* mutants suggests that *GATA6* is directly responsible for this observation. Interestingly, results from my ChIP-seq data also shows PDX1 as a binding partner of GATA6 (data not shown), a finding that was previously not shown in an earlier study (Cebola et al., 2015). This suggests that *GATA6* also plays a direct role in pancreatic specification.

Performing ChIP and ChIP-seq on Patients A and B could further identify important direct binding partners of GATA6. Since *GATA6* mutations in Patients A and B did not lead to nonsense-mediated decay of the GATA6 protein (Figure 52), it would be possible to perform ChIP on these samples. Thus, subsequent bioinformatics analyses comparing Patient A and B ChIP-seq datasets to their respective isogenic corrected wild-type control cell lines could further elucidate the molecular mechanisms of GATA6. However, this was not done due to time constraints.

Overlapping the GATA6, EOMES and SMAD2/3 ChIP-seq datasets at the DE level suggests a fundamental role of GATA6 at the DE stage, and that EOMES is required in the interaction of GATA6-SMAD2/3. Thus, as EOMES limits the expression of mesodermal markers, it can be speculated that GATA6 mutations may impede DNA binding of EOMES-SMAD2/3 linked to endoderm formation (Figure 127).

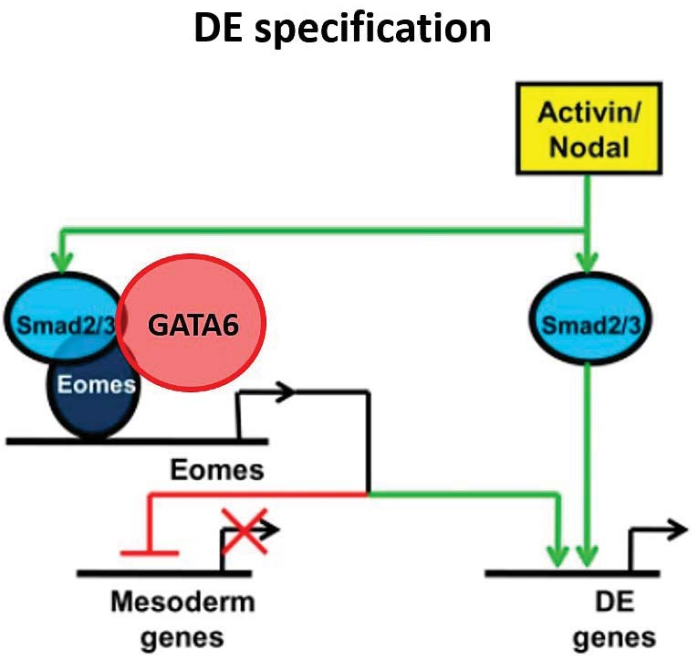

Figure 127. Model depicting the molecular mechanism of action for GATA6 in the formation of the DE. EOMES, SMAD2/3 and GATA6 interacts to initiate DE differentiation while repressing mesoderm genes. Illustration adapted and modified from Teo et al., 2011**.**