CHAPTER 6 : GENERAL DISCUSSION

Findings from this thesis can be divided into three components. First, a transposon-mediated genetic screen was performed to identify candidate genes which display the capacity to replace the ectopic requirement of Oct4 during the acquisition of pluripotency. Second, two genetic factors, BNC2 and RARβ, were identified from the genetic screen and exhibited competence in acting as exogenous Oct4 substitutes. Third, molecular dissection of the role played by RARβ to negotiate the need for ectopic Oct4 during the attainment of pluripotent derivatives inadvertently shed light on the ability of RARs to transcriptionally activate the Oct4 distal enhancer and in certain conditions, RAREoct. This activation triggers rapid and efficient reprogramming, providing insight to the epigenetic changes that accompany the acquisition of pluripotency. In this chapter, the three areas will be discussed further, culminating in the future prospects of this thesis.

6.1 Transposon assisted mutagenesis serves as a dependable tool to identify genetic candidates

The origin of iPSC technology was attributed to a candidate gene-based screen comprising of 24 ESC specific transcription factors. Although this screen has led to the inception of a landmark discovery in nuclear reprogramming, succeeding studies aimed at expanding the list of factors involved in the maintenance of pluripotency have employed alternative approaches. Diverse methods include chemical screens, siRNA library screens and transcriptional profiling of the reprogramming process (Ichida et al., 2009; Chia et al., 2010; Samavarchi-Tehrani et al., 2010). Transposon mediated mutagenesis is an emerging gene discovery tool (Guo et al., 2010). Initially portrayed as a tool for cancer gene discovery, its potential has diversified to cell-based assays in a spectrum of phenotypic assays (Guo et al., 2010; Guo et al., 2011; Wang et al., 2011b) transcending across species (Li et al., 2011a; You et al., 2011b).

Chapter 3 reiterates the ability of transposon mutagenesis to isolate genes capable of substituting exogenous Oct4 during the reprogramming process. Unbiased assessment of

genes is exemplified in the observation that most genes incurring transposon integrations are poorly characterised. To embody the reliability of the screen, a significant proportion of genes which received insertion sites represented pathways critical during development. As an evaluation of the dependability of the screen, candidate genes identified were validated through independent experiments. The abilities of BNC2 and RARβ in generating iPSC colonies in the absence of ectopic Oct4 reaffirmed the validity of the genetic screen. On the other hand, of a total of 177 genes identified to harbour integration sites only three genes were validated and two exhibited potency as exogenous Oct4 substitutes. 98.3% of candidate genes were ignored as they did not fulfil the selection criteria of possessing more than one independent insertion site. Keeping in mind that the genetic screen attained genomic coverage of an average of one integration per 10kb of genomic material, a similar genetic screen can be conducted in a larger scale to increase the prevalence of integrations and common insertion sites, ultimately improving the probability of uncovering causal mutations. Alternatively, introduction of a variety of transposons with distinct properties may facilitate saturation of the host genome (Izsvak and Ivics, 2005). However, the detection of integration sites in a transposon-mediated screen relies on several layers of experimental procedures. These include the amplification of transposon arms, insertion of the amplicons into an expression vector and capillary sequencing of the inserted fragments. Technical drawbacks at each stage results in a diminished breadth of integration sites. To circumvent this, next generation sequencing techniques can be employed (reviewed by Metzker, 2010). The low cost and large sequence data volume allows sequencing of genomic DNA and identification of every transposon integration site. This will result in an exhaustive list of candidate genes and permit improved analysis to identify prominent candidate genes.

Analysis of integration sites to ascertain candidate genes which drive the phenotype-of-interest requires intensive statistical computation (Brett et al., 2011) and persists as a bottleneck in the employment of transposon mediated mutagenesis in genetic screens. In contrast, directed genetic screens examine a pre-determined set of candidate genes (Takahashi and Yamanaka, 2006; Chia et al., 2010), dispensing the requirement of rigorous analysis. To circumvent this caveat, recent advances in piggyBac mediated

transposition have led to the generation of a library containing 14,000 individual gene-trap ESC clones (Wang et al., 2009). In co-operation with Bloom deficiency, homozygosity at mutated loci can be achieved to facilitate the performance of loss of function genetic screens. As each ESC clone carries a single genomic integration, laborious analysis of insertion site will be avoided. Alternatively, the use haploid cells which encompass a single set of chromosomes overcomes the need for bi-allelic mutants to observe a phenotype-of-interest and eliminates the requirement for extensive computational analysis (Leeb and Wutz, 2011; Yang et al., 2012; Carette et al., 2009).

Overall, the immense potential of transposon-mediated mutagenesis can be clearly illustrated from the discovery of novel genes involved in the pathogenesis in an array of cancer types (Rad et al., 2010; Dupuy et al., 2005). In addition to its inherent ability to integrate into the host genome in a random fashion, there has been heightened interest in its ability to deliver genes efficiently temporally with the option of excision (Balciunas et al., 2006; VandenDriessche et al., 2009; Kang et al., 2009). The pleiotropic properties of transposons highlight its prime candidacy to serve as gene delivery and discovery tools in a clinical setting. As a non-viral delivery method, transposon mediated gene delivery has been appealing in its ability to perform stable, ectopic and reversible manipulations an array of human cell types (Chen et al., 2010; Belay et al., 2011). Uses of transposon mediated gene delivery include *ex vivo* gene therapy (reviewed by Di Matteo et al., 2012), circumventing the occurrence of detrimental effects from the use of retroviruses (Blaese et al., 1995). Although there have been some success in performing gene therapy using retroviruses (Aiuti et al., 2009), a non-integrative and amenable approach would be an attractive alternative avenue.

6.2 Replacements of ectopic Oct4 during the initiation of reprogramming events

A number of studies have investigated the possibility to substitute exogenous requirements of Oct4 with alternative factors (Heng et al., 2010; Redmer et al., 2011). Findings from this thesis reiterate the observations that ectopic Oct4 can be substituted and two effective genetic replacements were identified, BNC2 and RARβ.

The identification of RARβ as a potent substitute for the ectopic requirement of Oct4 during the generation of iPSCs is worth deliberating over. Although a number of studies have attributed RA signalling to the loss of pluripotency, there are several observations which suggest otherwise (Ben-Shushan et al., 1995; Wang et al., 2011c). RA has been described to possess anti-proliferative and differentiation inducing properties, representing an effective deterrent during tumour formation (Schuldiner et al., 2001; Jiang et al., 2007; Yamashita et al., 2005; Metallo et al., 2008; Strickland and Mahdavi, 1978; Flynn et al., 1983). However, opposing cellular responses have been witnessed where RA signalling results in proliferation and the initation of tumorigenesis (Ledda-Columbano et al., 2004; Liu et al., 2011). Furthermore, exposure to RA in human and mouse subjects reveals confounding results and promotes tumorigenesis (Mikkelsen et al., 1998; Albright et al., 2004; Mollersen et al., 2004; Ommen et al., 1996). These pieces of evidence highlight our poor understanding of the RA signalling pathway. The discovery that RARs are potent instigators of the reprogramming process may provide a distinct dimension to their participation in developmental pathways.

Unravelling of factors which play co-operative roles in the activation of endogenous Oct4 during the acquisition of pluripotency offers insight to the roles performed by Oct4 during the reprogramming process. Epitomising the first transcription factor that participated in the maintenance of pluripotency, Oct4 has received voluminous amounts of attention in deciphering the molecular pathways that accompany it. Large scale genomic and proteomic studies have elucidated its role as the core transcriptional machinery in co-operation with Nanog and Sox2, and an exhaustive list of binding regions have been delineated (Boyer et al., 2005; Loh et al., 2006; Wang et al., 2006). However, post translational modifications of Oct4 have been side-lined. A number of studies have determined that phosphorylation, SUMO-ylation, ubiquitination, and glycosylation of Oct4 affect its transcriptional activity and stability (Brehm et al., 1997; Saxe et al., 2009; Brumbaugh et al., 2012; Wei et al., 2007; Zhang et al., 2007; Xu et al., 2009; Jang et al., 2012). To this end, it will be intriguing to determine if factors in the reprogramming milieu

exhibit affinities toward sites which are susceptible to post translational modifications. This knowledge will further enhance our knowledge on the mechanisms behind Oct4 and reprogramming.

Interestingly, porcine fibroblasts have been described to defy the notion that exogenous Oct4 or a suitable replacement is quintessential during the reprogramming process (Montserrat et al., 2011). c-Myc, Klf4 and Sox2 were described to be sufficient in generating porcine iPSCs. Reprogrammed cells exhibited Oct4 expression from its endogenous loci and were able to contribute to the formation of three germ layers when injected into immune-compromised mice. Conversely, the addition of ectopic Oct4 to the reprogramming mix deteriorated the reprogramming process and increased the propensity of reprogrammed cells to differentiate. However, Oct4 has been portrayed to be expressed at low levels in porcine fibroblasts and trophoectoderm, indicating a separate role of Oct4 in distinct vertebrate developmental processes. Elucidation of the roles played by Oct4 during reprogramming leaves much to be addressed and experiments performed in this thesis offer some insight to unravel the unknowns. Overall, this study reinforces the notion that ectopic Oct4 can be replaced or dispensed, but endogenous Oct4 remains crucial for the maintenance of pluripotency and the generation of iPSCs.

6.3 Epigenetic Regulation of Oct4

The regulatory elements of Oct4 are conserved across species (Nordhoff et al., 2001; van Eijk et al., 1999), reflecting their importance from an evolutionary perspective. Classified as Conserved Regions 1-4 (CR1-4), murine, human and bovine genomes exhibit these features within major histocompatibility complexes (MHC) on syntenic chromosomes (Abdelrahman et al., 1995; Takeda et al., 1992; van Eijk et al., 1999). Methylation profiles at these conserved regions have been described to indicate the transcriptional activity of Oct4. Differentiation of human pluripotent cells has been depicted to accompany methylation at the Oct4 distal enhancer and diminuition in Oct4 expression levels (Deb-Rinker et al., 2005; Yeo et al., 2007; Cheong et al., 2010), whereas DNA demethylation is synonymous with the activation of the Oct4 locus (Mikkelsen et al., 2008).

Hypermethylation can be observed as early as the first cell fate decision where the genomic locus of Oct4 is inflicted by increased methylation marks, resulting in the repression of Oct4 expression levels and differentiation into the trophoblastic lineage (Hattori et al., 2004). On the contrary, the presence of DNA demethylating agents or ablation of Dnmt1 resurrects Oct4 expression, highlighting the importance of epigenetic remodelling on the expression of Oct4.

This thesis proposes that RAR family members facilitate the activation of Oct4 cis-regulatory elements by imposing H3K4me3 activating histone marks. Extrication of histone modifications has been employed to study alternative cellular processes such as the silencing of p16(INK4a) during tumorigenesis and Jarid1a/b mediated H3K4 demethylation during senescence (Yao et al., 2010; Chicas et al., 2012). As specific epigenetic alterations associated to the acquisition of pluripotency remain concealed, an unbiased large scale analysis of promoter and enhancer sequences can be performed to elucidate novel molecular changes that complement the conversion of somatic cells to pluripotent derivatives (Patwardhan et al., 2012; Melnikov et al., 2012; Sharon et al., 2012).

Another layer of expression regulation is through alternative splicing to produce transcript variants and four protein isoforms. The human Oct4 locus is reported to generate three transcript variants and four protein isoforms. Clear delineation of the functions exhibited by each variant remains to be uncovered but it has been suggested that they exhibit distinct functions in the maintenance of pluripotency and cellular stress response (Cauffman et al., 2006; Lee et al., 2006; Papamichos et al., 2009). As an additional layer of complexity, Oct4 has been described to possess multiple transcription start sites and pseudogenes (van Eijk et al., 1999; Lin et al., 2007; Panagopoulos et al., 2008; Singh et al., 2012). Having elucidated the importance of the cis-regulatory elements of Oct4, it will be intriguing to determine their influence on the expression of the transcript variants and pseudogenes. Moreover, the participation and effect of these variants during the reprogramming process may also contribute to our understanding of Oct4 and its role in pluripotency.

6.4 Kinetics of Reprogramming

Since the inception of nuclear reprogramming, molecular kinetics that dictate the revival of transcriptional machinery underpinning the conversion of cellular identity have been of immense interest. This thesis attempts to dissect molecular events that influence the kinetics behind the dedifferentiation process. By retrieving epigenetic snapshots of the Oct4 regulatory elements over the first 72 hours of reprogramming, it was illustrated that activation of the distal enhancer of Oct4 and RAREoct is a key impetus to stimulate the dedifferentiation process. Alternative approaches to determine molecular changes during the generation of iPSCs include the examination of transcriptomes over the course of nuclear reprogramming (Foshay et al., 2012; Samavarchi-Tehrani et al., 2010; Mikkelsen et al., 2008). Despite the attainment and analysis of transcription profiles that accompany reprogramming intermediates, epigenetic changes reflect subtle molecular changes that precede fluctuations in transcript levels and possess better predictive value when delineating molecular events during reprogramming. To address this, methylomes have been illustrated in somatic cells and pluripotent cells (Meissner et al., 2008; Hawkins et al., 2010; Marks et al., 2012). However, an equivalent study to explore genome-wide histone modifications during the conversion of somatic cells to iPSCs has not been performed. To this end, this thesis presents as an initial effort to explore the tip of the iceberg, and inspects microscopic changes in histone conformation at precise regulatory elements of Oct4 over the course of reprogramming.

Exploration of kinetics that influence nuclear reprogramming events provides a perspective on cascades of molecular events that take place. This is exemplified through the examination of molecular changes during cell fusion which suggests biphasic reprogramming. The initial phase composes of trans-acting events that are independent of DNA replication and involve the cellular environment, whereas the subsequent phase embodies DNA replication dependent cis-acting chromatin modifiers that reawaken silent genes (Foshay et al., 2012). In contrast, the generation of iPSCs has been proposed to be divided into three phases: initiation, maturation and stabilisation (Samavarchi-Tehrani et

al., 2010). The initial phase involves mesenchymal-to-epithelial transition, whereas the maturation and stablisation phase are delineated by the up-regulation of a subset of pluripotency associated genes.

Findings from this study strengthen our understanding of acquired pluripotency, hence extrapolation of these findings and examination of histone modifications on a genome-wide scale will yield copious amounts of information about the clockwork behind the generation of iPSCs.

6.5 Future Direction

The objective of this thesis was to identify genetic replacements of exogenous Oct4 during the initiation of reprogramming events, thereby illuminating the role of Oct4 during the acquisition of pluripotency. To this end, two factors were identified and verified. Through the elucidation of the mechanism behind RARβ, it was delineated that activation of the distal enhancer and proximal promoter of Oct4 is integral during the reprogramming process. The presence of RARγ and LRH1 facilitates this process and accelerates iPSC colony formation and improves reprogramming efficiencies.

iPSC technology holds immense potential in a clinical setting. Derivation of patient specific iPSCs enables the execution of disease modelling, drug screening and cell therapy. As such, it will be interesting to determine if RARβ and BNC2 are able to replace the exogenous requirement of Oct4 in reprogramming human somatic cells. Interestingly, RARβ has been described to possess tumour suppressor properties (Widschwendter et al., 1997; Xu et al., 1997), whereas the conventional set of reprogramming factors have been established to promote tumorigenesis (Hochedlinger et al., 2005; Okita et al., 2007). Hence, application of RARβ in place of Oct4 may be desirable. However, reprogramming of mouse and human fibroblasts are distinct, where introduction of an identical reprogramming cocktail consisting of Oct4, c-Myc, Klf4 and Sox2 leads to iPSCs that

exhibit distinct properties and morphologies (reviewed by De Los Angeles et al., 2012). Attempts at replacing reprogramming factors have only been performed in limited cell types or with the use of chemicals (Giorgetti et al., 2009; Kim et al., 2009b; Kim et al., 2009c; Huangfu et al., 2008b), and have to be pursued with caution.

Although conventional reprogramming methods convert human somatic cells to iPSCs that resemble human ESCs, addition of RARγ and LRH1 to the traditional reprogramming mix has been described to generate mouse ESC-like derivatives (SH-iPSCs) which are disparate from human ESCs (Wang et al., 2011c). Unlike human ESCs, SH-iPSCs are easily amenable and able to survive in a single cell suspension. As such, it will be fascinating to determine if RARγ and LRH1 elicit a similar effect on the distal enhancer and proximal promoter of human Oct4, explaining the peculiar cellular phenotype of SH-iPSCs. Retinoic acid differentiation of human ESCs incites hypermethylation of the Oct4 distal enhancer (You et al., 2011a; Yeo et al., 2007). As such, it is tempting to speculate that a similar phenomenon takes place during the acquisition of pluripotency in human cells and activation of both the distal enhancer and proximal promoter of human Oct4 is imperative. It has been demonstrated that cell fusion results in epigenetic modifications of the distal enhancer and proximal promoter of Oct4 (Freberg et al., 2007). Keeping this in mind, it will be interesting to learn if RARγ and LRH1 also have the capacity to enhance reprogramming efficiencies through cell fusion.

Findings from this thesis set a foundation for deeper comprehension of the reprogramming process in human and murine cells. Elucidation of the activation of endogenous Oct4 during the early stages reprogramming may allow us to improve the reprogramming protocol, where the endogenous genes are targeted, dispensing the need for ectopic introduction of reprogramming factors. To this end, transcription activator-like effectors (TALEs) have been designed to recognise the Oct4 proximal promoter. Introduction of TALEs into ESCs stimulates Oct4 transcriptional activity, whereas epigenetic modifiers are required for TALEs to elicit an effect on neural stem cells (Bultmann et al., 2012). This surmises the possibility that activation of the Oct4 distal enhancer could overcome the

epigenetic barrier faced in neural stem cells. Overall, this thesis proposes that epigenetic changes at the endogenous Oct4 locus act as reprogramming checkpoints and can be surmountable by the presence of RARγ and LRH1. Accordingly, unravelling of microscopic alterations during the generation of iPSCs may resolve uncertainties in the molecular processes that occur during nuclear reprogramming.