Chapter 6

Conclusions and Future Perspectives

This thesis describes the use of CRISPR-Cas9-based genome-wide knockout library to study the transition from pluripotency to differentiation in mouse embryonic stem cells. It encompasses three major sections of work. The fist part focused on the preparation work and proof-of-principle studies, including generation of constitutive Cas9 expressing Rex1:GFPd2 cell line and analysis of the differentiation profiles of Tcf7l1 and Apc knockout mESCs. These studies established screening foundations and provided valuable insights into screening design optimisation. The second part describes a detailed analysis of the screening result. Recovery of several well-studied pathways and factors confirmed that the screen was effective in identifying drivers of commitment. The discovery of previously unknown pathways and pluripotency regulators suggested the powerfulness and sensitivity of the screen. The final section focused on the mechanistic studies of mTOR-related pathway in regulation of pluripotency and differentiation. Through detailed dissection of the knockout phenotypes of *Nprl2* and *Tsc2*, I demonstrated that GATOR1 and TSC2 complexes regulate pluripotency via distinct routes, albeit both function as mTORC1 inhibitors.

6.1 The immeasurable potential of CRISPR-Cas9-mediated genetic screen in stem cell biology

Since the establishment of ESC cultures more than three decades ago, studies have revealed a number of genetic and epigenetic machineries regulating pluripotency and differentiation. However, the mechanism behind the transition from pluripotency to differentiation has not been fully understood. This is partly due to the lack of scalable and high-throughput genetic methodology that allows comprehensive investigation of genes causing a specific phenotype. So far stem cell research has been typically carried out with resource-intensive hypothesis-driven approaches. The hypothesis-free forward genetics approach has been applied in yeast, Drosophila and *Caenorhabditis elegans* and provided deep insights into diverse biological processes [119] [178] [387]. A few large-scale screens were performed in stem cells (summarised in Table:4.1), but being limited by the technology available at the time, they were shown to be inefficient and lack of power.

The advent of CRISPR-Cas9 technology provided unprecedented opportunities to perform genetic screens and investigate genetic pathways underlying cellular processes in mammalian cells. The results described in this thesis demonstrated immense power and remarkable sensitivity in identifying genes involved in the regulation of differentiation

initiation. The CRISPR-Cas9-mediated screen is a step forward from RNAi-mediated screens and insertion-mediated screens in regard to knockout efficiency, off-target effect and simplicity in identifying causative mutations. It is also advantageous in offering the opportunity for a truly unbiased genome-wide comprehensive analysis of a particular biological process.

Most recently published CRISPR-Cas9 screens in mammalian cell culture used cell growth as a model [422] [153] [435]. Unlike those, the screen describe in this thesis used FACS-based reporter gene expression as a readout for positive selection. This mode of genetic screening is relatively less explored, possibly due to the difficulties in maintaining library complexity during cell sorting. To overcome such technical challenges, a series of optimisation experiments were performed for the best outcome of large-scale cell sorting. Under optimised conditions (described in Chapter 2), I could collect at least $5x10^6$ cells per replicate for downstream genomic DNA extraction and gRNA amplification. This achieves 50X coverage per gRNA, which is approximately half of the coverage in negative-selected cancer cell growth studies.

Furthermore, as shown in Chapter 4 and 5, by carefully choosing screening parameters, one can possibly uncover both enriched and depleted gRNAs from a screen and identify both positive and negative regulators of the process. Additionally, as the phenotype is assessed based on the gRNA abundance in control and treated samples, statistical analysis can be performed to quantitatively evaluate the phenotype caused by the knockout of a particular gene. The statistical significant value obtained from the screen is to some extent reflecting the level of association of a specific gene to the biological process, which provides a useful guidance for investigators to narrow down the candidate list and set up priorities for further investigation. The screening result described in this thesis demonstrated high sensitivity, which is well exemplified by uncovering genes whose mutation caused loss of Rex1-GFP expression. Although the screening conditions were optimised for positive selection to capture genes whose loss lead to sustained Rex1-GFP expression, 12 genes could be identified from negative enrichment with a more relaxed cut-off of FDR=25%. Even so, *Rictor*, which was ranked at 37 in the negative selection with an FDR of 51%, could be validated by a separate experiment in knockout mESCs (Figure 5.6). This indicates that, due to higher noise in negative selection, some genes failed to pass the statistical cutoff can still be meaningful and worth further investigation. Another gene *Sesn2*, encoding Sestrin2, is known to function as a leucine sensor whose loss results in continuous mTORC1 upregulation. *Sesn2* was ranked 46 in the negative selection with

an FDR of 0.61. Although it did not pass the statistical cutoff, the phenotype of *Sesn2* might be true, subject to further validation experiments. Re-designing the screen with optimised conditions for negative enrichment could be worthwhile to uncover the biology of mechanisms facilitates the initiation of differentiation.

The success of CRISPR-Cas9 genome-wide screen in unraveling genes involved in the exit of pluripotency using mouse ESCs demonstrated its potential in stem cell biology and opened up several lines of research work that could be performed in the near future. One of the apparent projects is to apply the same screen in human ESCs. Although mouse ESCs remain as invaluable materials in developmental biology, it is an imperfect model for studying human embryonic development [121]. Recent studies have developed several strategies to establish and maintain human ESCs in the naive state by over-expression of certain transcription factors or chemically manipulate the culture system [58] [128] [403] [437] [412]. CRISPR-Cas9-based genetic screening could facilitate the identification of genes that are required to safeguard human naive pluripotency state, as well as genes that drive the exit from pluripotency.

The discovery of iPSCs has provided a revolutionary platform for human disease modelling, stem cell-based therapy and drug discovery. Great progress has been made since its birth a decade ago. However, several important issues remain to be addressed: the current reprogramming process occurs stochastically with low efficiency and the resulted reprogrammed cells often exhibit variations in differentiation potential [371]. CRISPR-Cas9-based genetic screening could open a new avenue for understanding the reprogramming mechanism and basic biology of iPSCs to develop methods that facilitate the efficient generation of standardised human iPSCs. Another challenge for the research and therapeutic applications of iPSCs and ESCs is to differentiate them into specific cell types. Recent studies have been successful in discovering the key events during cell fate decisions; however, the currently available differentiation methods often result in heterogeneous population with low efficiency in derivation of desired cell types [427] [231] [31]. CRISPR-Cas9-based genetic screening holds great potential in understanding the genetic and molecular control that dictates the specification to distinct cellular identities, which facilitates the development of better strategies for directed differentiation of pluripotent cell to a specific cell type.

6.2 The complex regulatory network that drives dissolution of pluripotency

The screening result described in this thesis demonstrated a panoramic picture of cellular pathways that facilitate progression out of naive pluripotency. These include known signal transduction pathways such as Wnt, FGF/MAPK and PI3K pathways. The screenings hits also include processes that are indirectly involved in signal transductions such as the heparan sulphate biosynthesis pathway and endocytosis. As a positive regulator of the binding of Fgf and Fgfr, deficiency in heparan sulphate biosynthesis results in reduced FGF/MAPK signalling therefore negatively regulating the onset of differentiation. The endosomal trafficking pathway controls the intensity, duration and specificity of the signalling pathways. Of various vesicle trafficking and endocytosis complexes were identified, all 6 genes that compose the HOPS complex demonstrated noticeable differentiation defects. Other published pluripotency regulating pathways that were enriched in the screen include mRNA degradation and miRNA biogenesis pathways. To be specific, genes involved in mRNA decapping, deadenylation, nonsense-mediated decay and m^6A mRNA methylation appeared to be significant in the screen, which is consistent with published studies [224] [132] [14]. miRNAs are also known to regulate differentiation [187] [378]. The screen has identified genes involved in miRNA biogenesis such as *Dicer1* and *Drosha*, as well as the miRNA nuclear exporter *Xpo5*. The effect of transcriptional regulatory proteins and chromatin modifiers on regulation of pluripotency and differentiation has been extensively studied. A large number of them has been captured in the screens, such as Otx2, Zfp281, polycomb repressive complexes and HDAC1/2 containing complexes. Interestingly, the screen has identified *Pou5f1* (protein: Oct4) as a gene required for differentiation. As a well-established core pluripotency gene, this finding seems somewhat contradicting. However, it has has been shown recently that the effect of Oct4 is sensitive to its protein level and ESCs with heterozygous Oc4 expression exhibited enhanced self-renewal and delayed differentiation kinetics [328] [188]. As CRISPR-Cas9 is a DSB-based technology which generates various alleles, there are probably a subpopulation of heterozygously edited cells in the mutant library which demonstrated delayed differentiation phenotype, while homozygous *Pou5f1* knockout ESCs were depleted from the population.

The screen has not only identified the majority of the pluripotency and differentiation regulators, but also revealed a large number of previously unknown factors, which broadened the landscape of factors regulating ESC identify and differentiation. This dataset provides

an important resource for the scientific community and raised several interesting biological questions for further investigation, such as the role of SAGA complex and NuA4 complex, which are highly conserved chromatin regulatory complexes whose function in the stem cell context has not been well-studied. Also, compare to genetic and epigenetic regulations, the role of metabolism in pluripotent stem cells is ill defined. Nearly half of the 563 screening hits were nucleus-encoded mitochondrial genes, indicating the the functional importance of mitochondria and oxidative phosphorylation in the transition from naive to primed pluripotency. However, it was widely accepted that primed pluripotent cells mainly rely on glycolysis as the source of energy, and mitochondria respiration is reduced compare to naive pluripotent cells [489] [402] [184]. It is therefore a mystery whether ATP production is directly taking part in regulating the exit of naive pluripotency, and if so, what is the underlying molecular mechanism. Another possible explanation is that metabolism pathways integrate with genetic and epigenetic programs to regulate stem cell function and fate. Certain byproduct or intermediate product of metabolism such as acetyl-CoA can be transported into the nucleus for epigenetic modification. The screen revealed a crucial role of the under-appreciated metabolic pathways in pluripotency and differentiation, whose exact function and mechanism require further investigation. Another interesting module that worth further study is the ESCRT machinery and HOPS complex.Although likely to be involved in sustaining external signal transmission such as FGF-MAPK pathway, in the last few years, studies have revealed that late endosome and lysosome function not only in protein sorting and degradation, but also act as a platform for mTORC1 signalling pathway[355] [393]. It was also shown that the level of amino acids inside the lysosome lumen directly modulates mTORC1 activity [492]. It might be worth investigating any additional roles of the ESCRT and HOPS complexes.

Collectively, these pathways demonstrated strong interconnectivity and formed a functionally interdependent network. The dissolution of pluripotency is dependent on differentiation cues, which transmit into the cell via signalling cascades. One of the downstream effects of differentiation input is the repression of naive pluripotency network, which is achieved cooperatively by transcription repressors and chromatin modifiers. To ensure a rapid response to differentiation signals, the mRNA modification and degradation pathways are activated, which promotes the down-regulation of pluripotent transcripts. As the naive pluripotency network is dismantled, the cells initiate a new transcription programme which resembles the early post implantation epiblast. This transition process is supported by metabolic pathways, possibly through providing energy and intermediate metabolites

required for the function of genetic and epigenetic pathways.

6.3 The prominent role of mTOR and its related proteins in pluripotency and differentiation regulation

This thesis highlighted the power of CRISPR-Cas9 mediated genetic screen in identifying novel pathways underlying biological processes in stem cells. Among the novel pathways I focused on showing the pivotal role of the mTOR-centred pathway in controlling the exit of pluripotency. mTOR regulators were overrepresented in the candidate list: there were 19 mTORC1 regulators and 2 components of mTORC2. This was especially true among the negative selection gene hits, where 5 of the top 10 significant genes are mTOR related. Surprisingly, deficiency in TSC complex and GATOR1 complex resulted in opposite phenotypes during ESC differentiation despite both function to inhibit mTORC1 activities. Through knocking out two GATOR1 components (Nprl2 and Depdc5), I demonstrated that mTORC1 activation by loss of mTORC1 negative regulators increases Gsk3 activity and destabilise naive pluripotency. I have also demonstrated through Rictor knockout mESCs that mTORC2 deficiency causes reduction of Akt activity and consequently increases Gsk3 activity, thereby destabilising naive pluripotency.

Conversely, mTORC1 upregulation by TSC complex deficiency resulted in the opposite phenotype, that is, delayed differentiation. Data in this thesis showed that Tsc2 loss in mESCs led a unique phosphorylation pattern: Akt in Tsc2 knockout exhibited higher activity evidenced by the hyper-phosphorylated downstream targets Pras40 and the result of Akt kinase assay. This is contradicting the established knowledge in cancer cells that activated mTORC1 induces negative feedback loop and down-regulates Akt[252]. Knockout of Rictor in Tsc2 knockout cells demonstrated that high Akt activity and this was due to ectopically activated mTORC2. However, the activated Akt in Tsc2 KO mESCs does not play a major role in Gsk3 phosphorylation nor phenotypic outcome. It is still unclear why activated Akt does not phosphorylate its conventional target, Gsk3, and what the molecular basis of this phosphorylation rewiring is. S6K activation due to Nprl2 deficiency does not result in phosphorylation on Gsk3. As Tsc2 physically interact with the β-catenin degradation complex [246], TSC complex might serve as a molecular scaffold and recruit Akt to phosphorylate Gsk3. Upon loss of Tsc2, the β-catenin degradation complex is no longer attached to this platform and Gsk3 can now be accessed and phosphorylated by S6K. It would be interesting to further investigate these hypotheses.

Several mTORC1 regulators including Folliculin and Tsc2 have been identified as drivers of commitment in a siRNA-mediated screen, and a model has been proposed in which Folliculin acts downstream of mTOR and drives differentiation by restricting nuclear localisation [19]. This model was limited by the scarcity of knowledge of the biological function of Folliculin at the time. It was found later that the Folliculin-Fnip2 complex acts as a GAP for RagC/D that stimulates mTORC1 activity in the presence of amino acids [420] [321]. This agrees with GATOR1 phenotype which inhibits mTORC1 via acting as a GAP for RagA/B. Although there was a link between Folliculin-Fnip2 complex to the subcellular location of Tfe3, it is not clear how Folliculin-Fnip2 caused the nuclear exclusion of Tfe3 and how does mTORC1 involved this process. Therefore more work is to be done to unravel the detailed molecular mechanisms of mTOR-related pathways. Several questions were raised in this thesis, one of which is whether mTORC1 is directly involved in exit of pluripotency. Although many mTORC1 regulators were identified in the screen, knockout of mTORC1 itself does not seem to alter the differentiation profile given that none of its essential components exhibited any phenotype in the screen. Furthermore, it was observed that treating wild type cells with rapamycin did not affects their differentiation progress [19]. However, rapamycin treatment could rescue the aberrant differentiation phenotype of *Tsc2* knockout and *Nprl2* knockout. For further investigation, it may be worth adopting other perturbation approaches other than rapamycin such as conditional knockout or knockdown. It has also been reported two point mutations on mTOR confer constitutive activation of mTOR signalling even under starvation conditions [354]. These activated mutants could be helpful to reveal the role of mTOR by direct comparison with *Nprl2* knockout and *Tsc2* knockout. Another interesting finding that would be interesting to follow up is how *Tsc2* knockout resulted in mTORC2 activation and unconventional Akt regulation in mouse ESCs. It is hypothesised that TSC complex plays additional roles other than mTORC1 inhibitor. A promising approach is to perform CRISPR-Cas9-mediated genome-wide screen with Tsc2 knockout ESCs.