CHAPTER 6

6 DISCUSSION AND FUTURE DIRECTIONS

6.1 Summary

In the current thesis I explore the molecular mechanisms involved in the maintenance of the first cell lineage decision in the mouse embryo, using *in vitro* stem cell models. I employ an unbiased genome-wide loss of function strategy using a CRISPR/Cas9 genetic screen coupled to the differentiation of ESCs to TSCs, to identify genes that suppress this lineage conversion in ESCs.

6.1.1 Proof-of-concept studies

For this screen I developed an *Elf5::Venus* ESC reporter cell line, which carries Cas9 stably integrated in the *Rosa26* locus, and a *T2A::H2B::Venus* cassette knocked-in the *Elf5* locus, thereby taking advantage of the endogenous control elements regulating *Elf5*. Bisulphite sequencing of the DMR regions in the *Elf5* promoter revealed the engineered line remained hypermethylated, reflecting an intact lineage barrier upon gene editing. Proof-of-concept experiments using knockout of *Oct4* proved that the *Elf5::Venus* reporter line was efficient for CRISPR/Cas9-mediated gene editing, and also allowed faithful monitoring of *Elf5* activation and thus, TSC differentiation. The reliability of this reporter line was assessed by RT-qPCR profiling of Venus positive cells, showing they activate a TSC transcriptional programme, and by immunostaining proving Venus cells co-express Elf5, as well as Cdx2. These proof-of-principle experiments laid the foundation for the genome-wide screen, as they allowed optimisation of editing and differentiation conditions. They also highlighted two screen limitations. First, detection of

genes essential for ESC self-renewal, as the *Oct4* knockout indel population is quickly taken over by *in-frame* mutations due to their growth advantage. Second, genes like *Dnmt1*, that might either need a complete protein depletion or prolonged ESC culture to allow downstream effects to take place before achieving a state permissive to TSC differentiation.

6.1.2 Genome-wide CRISPR/Cas9 loss-of-function screen identifies new suppressors of ESC differentiation to trophoblast

Using lentiviral delivery of a CRISPR gRNA library covering the mouse genome, I then performed a genome-wide loss-of-function screen to identify genes whose mutation allow activation of *Elf5*, and thus cells expressing Venus. I differentiated the *Elf5::Venus* mutant ESC library for 14 days in TSC medium and sorted Venus positive cells at the end of this process for analysis of enriched gRNAs by NGS. With this strategy I could identify 42 genes that result in ESC conversion to TSC, including factors such as *Nf2*, *Lats1* or *Tet1* (Nishioka, Inoue et al. 2009, Koh, Yabuuchi et al. 2011, Cockburn, Biechele et al. 2013), which were already described in the literature, demonstrating the screen could capture some of the known biology underlying this lineage barrier. Out of the 42 screen hits, I took 34 for validation experiments, confirming 22 genes with varying differentiation efficiency.

Genes related to histone acetylation, such as *Trim24*, *Taf5l* and *Ccdc101*, exhibited a weak activation of *Elf5*. The same was seen for *Tet1*, which was already shown to allow TSC differentiation (Dawlaty, Ganz et al. 2011, Koh, Yabuuchi et al. 2011) with *in vivo* contribution to placenta (Koh, Yabuuchi et al. 2011), indicating that even genes with weak phenotype in this experimental setting might reflect important biological functions.

An intermediate phenotype was observed for genes related to mTOR signalling such as *Tsc1* and *Tsc2*. In addition, *Kctd5* was also identified, which has been described as a negative regulator of the Akt pathway. This way, *Kctd5* knockout results in over-activation of the Akt pathway (Brockmann, Blomen et al. 2017), which acts upstream of *Tsc1/2*, promoting mTOR activation and further reinforcing the possible interplay of these two signalling pathways in early embryo development. Finally, strong phenotypes were detected for genes that compose the ncPRC1.1 polycomb complex, *Bcor*, *Rnf2*, *Rybp* and *Pcgf1*, implicating this complex has an essential role suppressing the trophoblast lineage in ESCs.

Overall, for most of the genes identified through my genetic screen, their role in the function and maintenance of the first cell lineage decision in ESCs was not previously known. Together, my findings shed new light on the mechanisms active at this early developmental checkpoint.

6.1.3 *Bcor* and *Rnf2* in the lineage barrier preventing ESC differentiation to TSC

Based on my genome-wide screen, I have established ncPRC1.1 in the maintenance of the TSC lineage restriction in ESCs. I next studied in higher detail the specific role for *Bcor* and *Rnf2* genes in this process. I confirmed that upon differentiation, both *Bcor-KO* and *Rnf2-KO* activate a trophoblast transcriptional programme. Derivation of clonal *Bcor-KO* and *Rnf2-KO* ESCs confirmed these do not affect ESC self-renewal, and are convertible to the naïve state. Furthermore, this lineage restriction is preserved in the epiblast-equivalent state, allowing differentiation of naïve knockout ESCs to TSCs. This phenotype could be partially rescued upon *TY1-Bcor* overexpression in *Bcor-KO* ESCs, and fully rescued in the equivalent experiment for *Rnf2-KO*. Importantly, *Bcor-KO* ESCs displayed hypermethylated *Elf5* promoter DMR regions demonstrating *Bcor* does not have a direct role on this epigenetic layer. For *Rnf2*, the results were not as clear and further studies will be necessary to address the effect of *Rnf2-KO* on the *Elf5* promoter methylation status.

RNA sequencing analysis of the differentiation profiles for *Bcor-KO* and *Rnf2-KO* Venus positive cells by day 18 indicated these are partially differentiated, upregulating TSC core genes but to levels below those of bona-fide TSCs. Importantly, *Rnf2-KO* cells had lower capacity for upregulation of *Cdx2*, *Eomes* and *Elf5* compared to *Bcor-KO* cells. Consistently, I could not establish a TSC self-renewing line from *Rnf2-KO* cells, but I succeeded in establishing one TSC-like line from *Bcor-KO* ESCs. This limitation from *Rnf2-KO* ESCs can either be due to a lower capacity for TSC differentiation compared to *Bcor-KO* ESCs, or due to *Rnf2* playing a role in TSC self-renewal.

Transcriptomics analysis of *Bcor-KO* and *Rnf2-KO* ESCs revealed extensive changes compared to *wild-type* ESCs. Both genes affect different processes, but have hundreds of upregulated genes in common, which form an important basis for future mechanistic studies.

Importantly, preliminary *in vivo* chimera experiments, with injection of *Bcor-KO* and *Rnf2-KO* ESCs into 8-cell host embryos, showed these mutant ESCs display impairment

of the first cell lineage restriction and can be found within extraembryonic compartments in E6.5 chimeric embryos. These results will have to be further corroborated with independent experiments, as well as immunostainings to definitely prove that cells incorporated into the extraembryonic tissues truly express the expected cell markers and are not mis-positioned ESCs.

6.2 Future Perspectives

6.2.1 The potential of genome-wide CRISPR/Cas9 screens as tools for dissection of lineage specification and restriction

In vitro stem cells represent valuable models for embryo development as they can overcome many of the limitations associated with the reduced throughput of *in vivo* developmental studies. Mammalian embryo biology has typically been addressed with hypothesis-driven approaches. Hypothesis-free forward genetics have the potential to uncover unexpected regulators and widen our view of key biological processes. To date, there have been a few genome-wide studies in early mouse development, mainly focused on pluripotency and ESC self-renewal (Hu, Kim et al. 2009, Guo, Huang et al. 2011, Zheng and Hu 2014) or exit from this state (Betschinger, Nichols et al. 2013, Li, Yu et al. 2018).

Pluripotency networks and ESC self-renewal have been widely characterised. In contrast, the mechanisms underlying the first cell lineage decision and its maintenance in ESCs remain poorly understood. Herein, I presented the results of an unbiased genome-wide CRISPR/Cas9 screen to address this question.

Due to the simplicity of the DNA recognition mechanism of CRISPR/Cas9, this tool was quickly adopted for high-throughput loss-of-function screens. Thus far these screens have mainly explored growth-related phenotypes such as cell survival and proliferation (Shalem, Sanjana et al. 2014, Wang, Wei et al. 2014, Zhou, Zhu et al. 2014, Hart, Chandrashekhar et al. 2015, Wang, Birsoy et al. 2015, Evers, Jastrzebski et al. 2016, Tzelepis, Koike-Yusa et al. 2016, Han, Perez et al. 2018). Although there are some examples of screens based on the expression of fluorescent reporter markers or cell surface markers followed by FACS (Koike-Yusa, Li et al. 2014, Parnas, Jovanovic et al. 2015, Golden, Chen et al. 2017, Li, Yu et al. 2018), these are still less common. My screen readout was based on the expression of Venus fluorescent protein, upon

activation of endogenous *Elf5* as a selection for trophoblast differentiation, adding another example to the successful CRISPR/Cas9 screens based on cell sorting.

Genes which affect differentiation of ESCs to TSCs with activation of *Elf5* expression are rare as demonstrated by the low percentage of Venus positive cell at the end of all my screen replicates (about 0.2%). Nevertheless, I could still identify 42 genes involved in this lineage conversion and I validated 22 of them with different phenotype strengths. This clearly shows the power of genetic screens using CRISPR/Cas9. This work directly contributes to knowledge of these genes' role in the mechanisms orchestrating the ESC/TSC lineage barrier.

In a broader context, my work clearly highlights the value of using similar approaches. Each differentiation screen is biased depending on the timings and conditions used, which means that employing different time-courses and different culture media will likely identify other factors regulating this lineage conversion. Furthermore, in the future, differentiation studies should be carried out using multiple lineage markers in order to gain a wider understanding of the differentiation process. For example, in my case, *Elf5* is a stringent marker for ESC differentiation to TSC, that fails to be activated in most published protocols (Cambuli, Murray et al. 2014) and hence was chosen for this study. However, *Elf5* is not specific to the TSC self-renewal network and can also be expressed in differentiating TSC cells (Latos, Sienerth et al. 2015). Consistently, most of the genes detected result in mixed populations of self-renewing and differentiating TSCs.

If one is interested in identifying factors affecting a particular cell population, then the use of multiple reporter systems should be considered given that each individual gene can be expressed in distinct cell types throughout development. Considering a self-renewing TSC state, reporter genes such as *Cdx2*, *Eomes* or *Tfap2c* could prove beneficial as these are core transcription factors of the TSC self-renewal machinery. Therefore, crosstalk between modulators of all these could yield the recipe that truly converts ESCs to self-renewing TSCs. Importantly, design of knock-in cell lines to follow the endogenous expression of specific markers coupled to genome-wide CRISPR/Cas9 recessive screens can be a valuable tool to provide hypothesis-free insights into countless biological questions.

6.2.2 ncPRC1.1 as key player in the maintenance of the first cell lineage decision

Polycomb group proteins were initially discovered in *Drosophila* (Lewis 1947) where they are part of an epigenetic cellular memory system that maintains the expression patterns of HOX gene throughout development (Ingham 1985, Cavalli and Paro 1998, Paro, Strutt et al. 1998, Poux, Horard et al. 2002). In mammalian systems, there are two main Polycomb complexes, PRC1 and PRC2 which are also involved in gene repression through histone post-translational modifications. In ESCs, PRC1 and PRC2 are widely established as regulators of differentiation, with PRC2 depositing the repressive histone mark H3K27me3 (Cao and Zhang 2004, Kuzmichev, Jenuwein et al. 2004) and PRC1 complementing this repression through H2AK119ub1 (Wang, Wang et al. 2004, Cao, Tsukada et al. 2005). Both PRC1 and PRC2 are essential for differentiation of ESCs to the three germ layers as exemplified by the fact that co-deletion of *Rnf2* and *Eed* (both essential for PRC1and PRC2 activity, respectively) severely impaired the formation of embryoid bodies from ESCs (Leeb, Pasini et al. 2010). They however seem to be redundant as embryoid body formation is possible in individual knockouts (Leeb, Pasini et al. 2010). Other components of Polycomb group proteins also affect differentiation of ESCs to the embryonic lineages, as discussed in chapter 5.1.3. Nevertheless, their role in preimplantation development has not been reported yet, although co-deletion of both *Ring1a* and *Rnf2* results in developmental arrest at the 2-cell state (Posfai, Kunzmann et al. 2012).

My work revealed that ncPRC1.1 is involved in maintaining the first cell lineage restriction in ESCs preventing their conversion to trophoblast. As such, knockout of *Bcor*, *Rnf2*, *Rybp* or *Pcgf1* results in an ESC state permissive to TSC differentiation, demonstrating a disrupted lineage barrier. It has been shown that H2AK119ub1 is a highly dynamic process in preimplantation embryo development and ncPRC1 is active as early as the zygote stage (Eid and Torres-Padilla 2016). The results herein complement this observation and suggest an active role specifically for ncPRC1.1 complex in early embryo development.

The mechanism behind the specific role of *Bcor* and *Rnf2* in the first cell lineage restriction remains to be elucidated and will be addressed in the future, as discussed in chapter 5.3.4. One of the key epigenetic distinctions between ESCs and TSCs is the presence of a small subset of genes, defined as gatekeepers, which display hypermethylated CGIs in their promoters and are silenced in ESCs state, but are hypomethylated and expressed in TSCs (Ng, Dean et al. 2008, Cambuli, Murray et al.

2014). This is considered an epigenetic memory as most of the published protocols for conversion of ESCs to TSCs fail to activate *Elf5* and the other 9 gatekeeper genes, maintaining a methylation state characteristic of their embryonic origin (Cambuli, Murray et al. 2014). *Bcor-KO* and *Rnf2-KO* cells have different capacities for the upregulation of the core transcriptional network for TSC self-renewal. However, they both successfully upregulate this class of genes, clearly demonstrating their common role in the maintenance of this epigenetic layer.

Bcor and *Rnf2* are expressed in both ESCs and TSCs. Both genes are dispensable for ESC self-renewal, but necessary for proper differentiation towards embryonic lineages. In TSCs, *Bcor* is dispensable for TSC self-renewal (Zhu, Fei et al. 2015). The exact role of *Bcor* and *Rnf2* in TSC self-renewal and differentiation is not reported and is another important factor to be addressed in follow up experiments. Nevertheless, both should be essential for proper trophoblast differentiation as knockout embryos arrest in development with placental defects (Voncken, Roelen et al. 2003, Cox, Vollmer et al. 2010). Considering this, it seems that ncPRC1.1 is important for the differentiation of both ESCs and TSCs. An equivalent paradox has been reported for *Sox2* and *Esrrb* genes, both necessary for TSC and ESC identity (Avilion, Nicolis et al. 2003, Martello, Sugimoto et al. 2012, Adachi, Nikaido et al. 2013, Latos, Goncalves et al. 2015). These two transcription factors were shown to be context-dependent and to be able to recruit different machinery in ESCs and TSCs, resulting in distinct downstream regulation (Adachi, Nikaido et al. 2013, Latos, Goncalves et al. 2015). Similarly, *Bcor* and *Rnf2* might be regulating a different subset of genes in the embryonic and extraembryonic lineages.

ncPRC1.1 was recently established as essential for hESC self-renewal (Wang, Gearhart et al. 2018). Analysis of their published datasets for chromatin binding profiles of Bcor, Rnf2 and Pcgf1 demonstrated that in hESCs ncPRC1.1 directly binds to the promoter regions of core TSC genes *Cdx2*, *Eomes* and *Gata3*. Based on this data, I raise the hypothesis that in mouse ESCs, ncPRC1.1 might also directly bind and repress these genes, ensuring proper restriction of the trophoblast lineage in the embryonic compartment. Furthermore, this data suggests that this particular lineage restriction mechanism might be conserved between mouse and human embryos and can be an exciting area of investigation in the future. If true, then it would implicate Polycomb genes as a mechanism of cellular memory throughout development, in a function conserved from *Drosophila* to mouse and human, and acting as early as the very first cell lineage decision in mammals.