C/EBPα instructs trophectoderm and pluripotency through the *Il6* pathway

Marcos Plana Carmona

TESI DOCTORAL UNIVERSITAT POMPEU FABRA / 2020

Department of Experimental and Health Sciences

Thesis supervisor:

Dr. Thomas Graf

Gene Regulation, Stem Cells and Cancer

Centre for Genomic Regulation

A todas aquellas personas que cruzaron mi camino.

ABSTRACT

THESIS ABSTRACT

Studying how transcription factor overexpression induces cell fate changes in specialized cells has provided invaluable insights into how normal cells differentiate. Our lab has developed two highly efficient cell conversion systems: continuous C/EBPα expression converts B cells into macrophages while a pulse of the factor enhances their reprogramming into iPS cells induced by the Yamanaka factors. I found that C/EBPα directly regulates *Il6* pathway genes in both B cell transdifferentiation and reprogramming, activating *Il6* and its receptor in different cell subsets. While IL-6 signaling turned out to be dispensable for switching B cells into macrophages, it is strictly required for the stable activation of pluripotency and trophectodermal genes during reprogramming. I therefore examined the possibility that these findings reflect a role of C/EBPα and the *Il6* pathway in early embryo development. Indeed I detected C/EBPα expression in 4- to 8-cell mouse embryos and in the trophectoderm (TE) layer of blastocysts. I also found that *Il6* is predominantly expressed in the trophectoderm while *Il6ra* is enriched in the pluripotent inner cell mass (ICM). Furthermore, treatment of developing embryos with an antibody blocking IL-6 delayed the morula to blastocyst transition, supporting a functional role of the pathway in early embryogenesis. I also studied the effect of C/EBPα overexpression in ESCs and discovered that it rapidly

activates both trophectoderm and pluripotency genes within distinct subpopulations. In conclusion, my data suggest that the observed C/EBPα-enhancement of iPS cell reprogramming reflects an instructive role for the ICM/TE segregation, mediated in part through regulation of the IL-6 signaling pathway.

RESUMEN DE TESIS

El estudio de los cambios en la identidad celular provocados por la sobreexpresión de factores de transcripción ha proporcionado un conocimiento valioso sobre cómo las células se diferencian de manera natural. Nuestro laboratorio ha desarrollado dos sistemas de conversión altamente eficientes: la transdiferenciación de células B en macrófagos tras la expresión continuada de C/EBPα así como la reprogramación de estas mismas células en iPSCs tras un breve pulso de C/EBPα seguido de la activación de los factores de Yamanaka. En este proyecto he descubierto que C/EBPα regula directamente los genes de la vía del *Il6* durante la transdiferenciación y reprogramación de las células B, activando *Il6* y su receptor en diferentes grupos celulares. Mientras que la señalización por IL-6 resultó prescindible para el cambio a macrófagos, es por el contrario necesaria para la estabilización de genes de pluripotencia y trofectodermo durante la reprogramación a iPSCs. En ese aspecto también examiné la posibilidad de que estos descubrimientos reflejasen un papel de C/EBPα y la vía del *Il6* durante el desarrollo embrionario temprano. De hecho logré detectar expresión de C/EBPα en embriones de 4 a 8 células y en el trofectodermo de blastocistos. Más específicamente, también encontré que *Il6* se encuentra predominantemente expresado en la capa del trofectodermo mientras que su receptor *Il6ra* queda concentrado en la masa

interior de células pluripotentes. La neutralización de IL-6 durante el desarrollo de los embriones por mediación de anticuerpos retrasa la transición de mórulas a blastocistos, señalando un papel funcional de esta vía en embriogénesis. Finalmente estudié el efecto de la sobreexpresión de C/EBPα en ESCs y descubrí que el factor activa rápidamente genes de trofectodermo y pluripotencia en distintas poblaciones de estas células. En conclusión, mis datos sugieren que la estimulación de C/EBPα en la reprogramación hacia iPSCs refleja un papel instructivo para la segregación de la masa interior pluripotente y el trofectodermo en el embrión, mediado en parte por la regulación de la vía de señalización de IL-6.

TABLE OF CONTENTS

INTRODUCTION

1. Reprogramming cell fate.

1.1 Transcription factor-mediated transdifferentiation.

The first factor-induced lineage conversion was achieved by transdifferentiating mouse embryonic fibroblasts (MEFs) into myoblasts[1]. This was based on the earlier observation that treatment of MEFs with 5-azacytidine led to the formation of colonies containing myotubes[2]. Transfection of MEFs with cDNAs obtained from such 5-azacytidine treated cells also led to the formation of myoblast colonies. Screening cDNA libraries in this assay led to the identification of a single clone that was sufficient to drive the conversion. This cDNA was found to encode the helix-loop-helix transcription factor (TF) designated as *MyoD*[1].

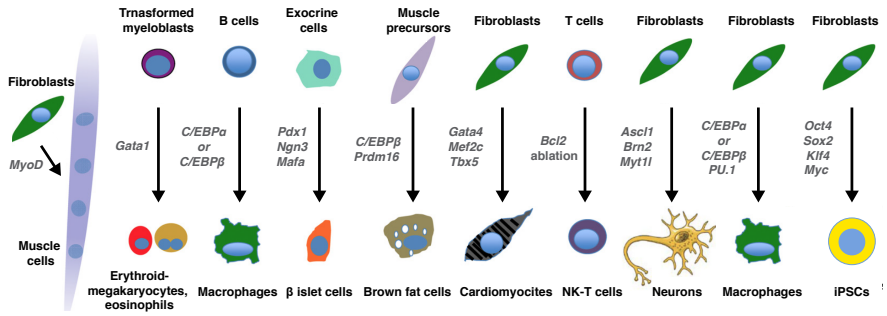


Fig. I1| Examples of cell fate conversions induced by transcription factors. Figure modified from Graf (2011).

Multiple groups subsequently succeeded in switching the identity of differentiated cells after the ectopic expression of defined transcription factors[3,4] (Fig. I1). Studies on the hematopoietic system have vastly contributed in the understanding of the molecular processes underlying cell differentiation. GATA-1 overexpression in myeloblasts was shown to induce the formation of erythroblasts, platelet precursors and eosinophils[5]. Subsequently, committed B lymphocyte precursors and mature B cells (together they will be designated as B cells from now on) could also be reprogrammed into macrophages by the myeloid restricted factors C/EBPα or C/EBPβ[6]. Furthermore, a combination of these factors with the pan-hematopoietic transcription factor PU.1 could convert fibroblasts into macrophages[7]. In parallel, it could be demonstrated that cells can be transdifferentiated by ablating a transcription factor required for their maintenance. Thus, deletion of *Pax5* in B cells led to the formation of myeloid cells and T cells[8] and deletion of *Bcl11b* in T cells led to their conversion into natural killer-like cells[9]. Among non-hematopoietic cell conversions, simultaneous PRDM16 and C/EBPβ overexpression could initiate a muscle precursor to brown fat switch[10] whereas combinations of three TFs (*Pdx1*, *Ngn3* and *Mafa*) could reprogram pancreatic exocrine cells into β-cells[11] or MEFs into functional cardiomyocytes upon *Gata4*, *Mef2c* and *Tbx5* induction[12]. However, a major achievement in the field was the conversion of fibroblasts into neurons with the combination of

Ascl1, *Brn2* and *Myt1l* factors, showing that cell identities could be switched among distinct germ layers[13].

1.2 Reprogramming of somatic cells into iPSCs by a transcription factor cocktail.

The landmark observation by John Gurdon in the 1950s that transplantation of nuclei of intestinal cells from *Xenopus laevis* into frog oocytes can generate tadpoles and even adult animals[14] (for a review see Graf 2011) showed that during development cells do not lose their genetic information. This also implied that the diversity of cell types is generated by the selective readout of their genome. Gurdon's findings inspired Yamanaka and Takahashi to perform their breakthrough experiment in 2006 in which they reprogrammed for the first time somatic cells into pluripotent stem (iPS) cells[15]. They identified a cocktail of four lineage instructive transcription factors consisting of OCT4, SOX2, KLF4 and MYC (OSKM) able to reprogram fibroblasts into induced pluripotent stem cells (iPSCs). Upon blastocyst injection, iPSCs form chimeras in which they participate into the development of all embryonic layers including germ cells, but they are excluded from extra-embryonic layer formation in the trophectoderm. Importantly, fertile animals can be entirely derived from iPS cells.

iPSC reprogramming was also achieved by replacing OSKM components by other transcription factors such as *Glis1*

substituting *Myc*[16] or *Esrrb* instead of *Klf4*[17]. Some of these substitutions, such as *Tbx3* replacing *Myc*, generated iPSCs with an improved contribution to the germ-line[18]. Other approaches improved reprogramming by repressing the p53/p21 pathway and/or by overexpressing *Lin28*[19].

A systematic survey of the reprogramming capacity of different somatic cell types revealed that most of them can be converted into iPSCs but at very low efficiencies (0.01-0.1%[20]). B cells were found among the least susceptible[21] but the simultaneous overexpression of C/EBPα with OSKM raised the efficiency 10-fold to approximately 1%[22]. Subsequent experiments by our laboratory showed that a transient overexpression of C/EBPα before OSKM activation in B cells further increased the iPSC reprogramming efficiency about 100-fold, involving 95% of the population and decreasing the time required to obtain iPSC colonies from 12 to approximately 6 days[23].

1.3 Alternative ways to reprogram somatic cell nuclei: cell fusion and extract treatment.

Besides the initial somatic cell nuclear transfer experiments, the establishment of embryonic carcinoma cells (ECCs) derived from germinal tumors[24] facilitated new insights on somatic nuclear reprogramming. Thus, the fusion of ECCs with thymocytes resulted in hybrids in which pluripotency

features were dominant[25]. When it became possible to generate embryonic stem cells (ESCs) from normal blastocysts[26,27] it was found that they likewise could reprogram T cells in fusion experiments[28]. Importantly, fusing different types of specialized cells also demonstrated dominance of one gene expression program over another. For example, fusion of human amniocytes with differentiated mouse muscle cells showed expression of the muscle program in a scenario where the hybrid cells retained the full complement of both sets of chromosomes[29].

Culturing permeabilized epithelial 293T cells with protein extracts from ECCs and ESCs also proved to reprogram somatic cells[30]. Analyses of those extracts identified components of the ATP-dependent BAF chromatin remodeling complex, which showed to significantly facilitate binding of OSKM factors to target promoters and therefore increase reprogramming efficiency of MEFs[31].

1.4 Epigenetics of reprogramming.

The initially observed low reprogramming efficiency of cultured MEFs led to postulate two alternative models that explaind this process: an elite model where only a fraction of cells are plastic enough to reprogram (‘elite cells’); alternatively, a stochastic model by which all somatic cells in a culture are equally susceptible and stochastically overcome

epigenetic roadblocks that prevent the establishment of pluripotency[32]. What are these roadblocks? One well-characterized barrier is the histone methylation mark H3K9me3 that decorates heterochromatin[33]. Erasure of DNA methylation in gene regulatory regions has also been reported to be required for the transition into the pluripotency program[34]. However, multiple layers including histone modifying enzymes, nucleosome remodelers, architectural proteins, non-coding RNAs and RNA binding proteins influence the chromatin state and the efficiency of reprogramming[35].

The highly efficient C/EBPα-enhanced B cell system showed that, under the right conditions, somatic cell reprogramming is a deterministic process[113]. This system also allowed to study molecular events in a time-resolved fashion, revealing that C/EBPα can recruit histone writers, readers and erasers[36,37] as well as reshape chromatin architecture[38]. C/EBPα-enhanced reprogramming to iPSCs is thus a powerful instrument to study mechanisms of pluripotency establishment and it might be suitable to model events that occur during early embryo development. A more detailed discussion of C/EBPα interactions with the epigenetic machinery will be provided below.

1.5 Single-cell sequencing defines reprogramming roadmaps.

The advent of technologies to analyze changes in gene expression at the single cell level has brought a refined understanding of cell transdifferentiation and reprogramming. Because of the low reprogramming efficiency of MEFs, it became useful to discover specific cell surface markers that could be used to separate the starting cells from cells destined to become pluripotent by FACS. Thus, THY1 resulted a convenient fibroblast marker that, together with pluripotent SSEA-1, allowed to capture intermediate populations during the process[34]. Another fruitful approach was the use of a NANOG-eGFP reporter to better isolate final time points of reprogramming[39].

Implementing single-cell sequencing during iPSC reprogramming using these tools identified two phases of pluripotency acquisition: a stochastic one where the expression of *Esrrb*, *Utf1*, *Lin28* and *Dppa2* predict the cells that will progress into iPSCs; and a second deterministic phase triggered by *Sox2* expression that finally establishes the pluripotency TF network[40].

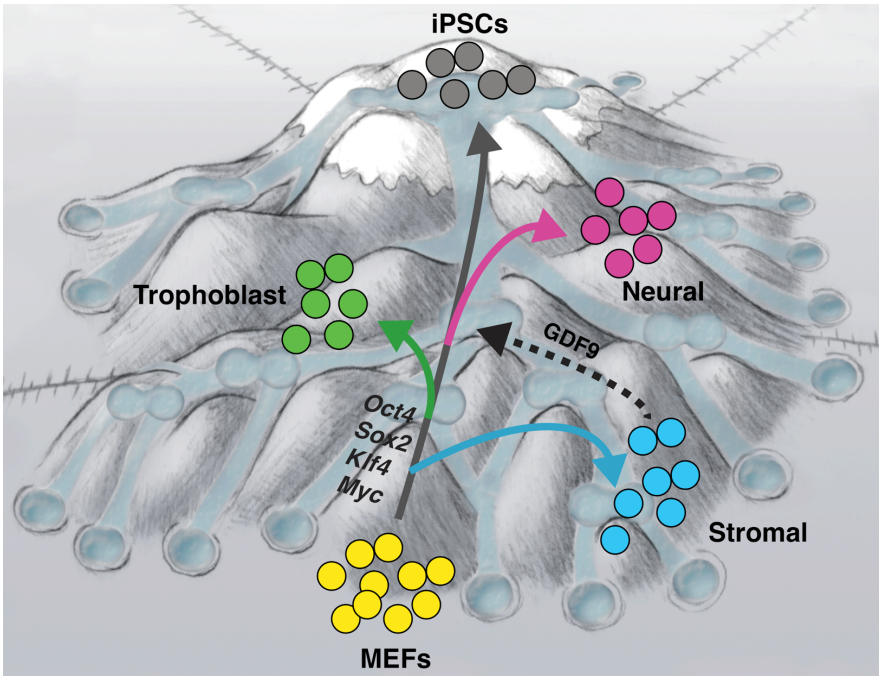


Fig. I2| Types of populations that emerge from MEF to iPSC reprogramming. Single-cell sequencing and pseudo-temporal trajectories of OSKM induced MEFs identify a lineage towards iPSCs and three additional gene signatures (stromal, trophoblast and neural). Such peripheral populations can influence the efficiency of reprogramming as detected for GDF9 secreted from stromal cells. Waddington landscape has been adapted by Thomas Graf and bifurcations from single cell MEF reprogramming come from the study by Schiebinger (2019).

Processing single-cell transcriptomics in the course of reprogramming led to the development of algorithms to reconstruct temporal cell trajectories. Several studies identified a single bifurcation that sorted cells into successfully reprogramming or into a “dead end” state where pluripotency failed[41,42]. Optimized algorithms allowed the detection of additional bifurcations that revealed stromal, neural and even trophoblast identities escaping the path

towards iPSC reprogramming from fibroblasts[43] (Fig. I2). Interestingly, some of these alternative differentiation pathways contributed to pluripotency acquisition through paracrine signaling, as shown for GDF9 expression from a stromal cell subset. Using a similar approach, trophoblast stem cells have been recently derived from iPSC reprogramming of human dermal fibroblasts by separately culturing cells exhibiting trophectodermal signatures[44].

Single-cell sequencing has also been applied in B cell to iPSC reprogramming and macrophage transdifferentiation where differences in *Myc* related genes helped to identify two initial sub-populations of B cells that could further be distinguished by size[45]. Small B cells with lower *Myc* activity performed better in macrophage transdifferentiation while large, *Myc*-high B cells yielded higher reprogramming efficiencies.

2. Molecular regulation of cell fate transitions: C/EBPα as a case study.

2.1 Chromatin binding and transcriptional control.

C/EBPα has been shown to be able to access closed chromatin, and thus acts as a pioneer TF. It is involved in central biological processes including cell fate decisions, metabolism or cell cycle, which have been widely described

during normal and malignant hematopoiesis[46,47,48]. Genome-wide data generated in recent years from *in vitro* cell-fate-conversion systems has opened the possibility to explore new layers of regulation exerted by C/EBPα as well as to understand its role in other physiological contexts during development.

C/EBPα was first purified from rat liver as a nuclear protein capable of binding to the 5'-CCAAT-3' sequences harbored in the promoters of the herpesvirus thymidine kinase gene and the Moloney murine sarcoma virus LTR[49,50]. Its ectopic expression in hepatoma cells further proved its ability to transcriptionally activate genes such as the serum albumin[51]. Protein-DNA crystallization solved the structure of the factor, consisting of a leucine-zipper basic region (bLZ) on the C-terminus and a transcription activation domain on the N terminus. The bLZ region interacts with DNA[52] and allows dimerization with other bLZ from TFs[53]. The transactivation domain recruits co-activators and repressors to modulate transcription[54].

C/EBPα is almost exclusively expressed in myeloid cells and its knock-out results in mice that lack the formation of granulocyte/macrophage progenitors (GMPs)[55]. It closely collaborates with the ETS family transcription factor PU.1 that is expressed both in lymphoid and myeloid cells and whose ablation leads to mice lacking both lineages[56]. A large

proportion of gene regulatory regions (GREs) in myeloid cells are bound by both PU.1 and C/EBPα. Moreover, a large proportion of PU.1 binding regions overlap with those bound by C/EBPα and, for the activation of myeloid genes in B cells, C/EBPα requires the expression of PU.1[57].

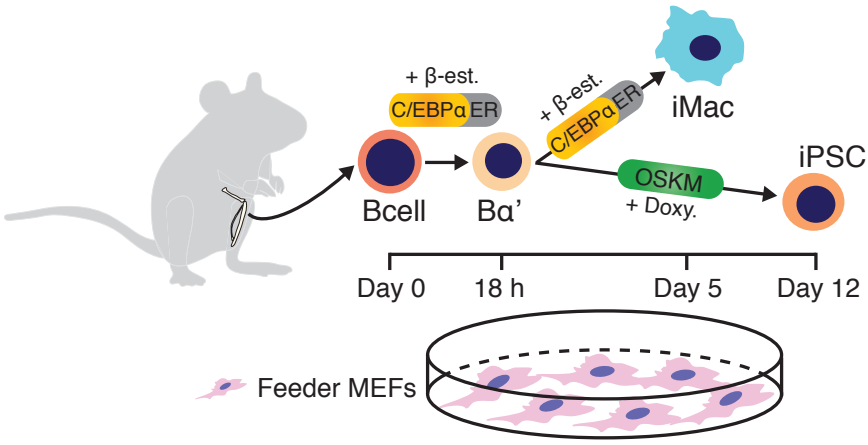


Fig. I3| Highly efficient B cell conversion systems induced by C/EBPα. Two protocols were used: a) B cells derived from mouse bone marrow were exposed to C/EBPα for 5 days to convert them into induced macrophages (iMacs). b) To reprogram B cells into iPSCs they are pulsed by C/EBPα for 18 hours followed by OSKM activation. In both protocols C/EBPα-induced B cells are cultured on a feeder layer of mouse embryonic fibroblasts (MEFs).

Imbalanced C/EBPα expression or truncations in the N-terminus region of the protein impair hematopoietic development[58] and lead to acute myeloid leukemias[59]. Of note, C/EBPα-induced transdifferentiation of human B cell leukemias into macrophages (inspired by the efficient system developed in mice, Fig. I3) reduces tumorigenicity after

xenotransplantation into immunodeficient mice[60]. Likewise, the spontaneous conversion of human BCR-ABL lymphomas into myeloid cells *in vivo* is accompanied with a loss of tumor progression that can be enhanced by C/EBPα-induced transdifferentiation[61].

As mentioned, during the establishment of pluripotency in somatic cell reprogramming, C/EBPα has been shown to dramatically enhance the efficiency at which B cells convert into iPSCs[22]. An optimized approach based on a transient 18 hours pulse of C/EBPα in B cells[23] preceding the Yamanaka cocktail[15] set up a highly efficient and uniform system to molecularly study the link between C/EBPα and pluripotency (Fig. I3). Using this system, it was found that C/EBPα binds to the enhancer regions of pluripotency genes in B cells (i.e.: *Klf4*, *Lefty2* or *Rarg*) already after the 18 hours pulse. This binding is followed by an increase of active histone mark H3K27ac, chromatin opening (revealed by ATAC-sequencing) and subsequent RNA upregulation[36]. Other pluripotency genes such as *Tcf7* also show *de novo* C/EBPα peaks in their enhancers after the pulse but their chromatin does not get activated or opened until later stages when the genes become expressed[37]. Endogenous C/EBPα has been also described to bind to pluripotency genes such as *Gdf3* during reprogramming of MEFs into iPSCs[62].

The fact that C/EBPα binding to chromatin precedes pluripotent gene activation in both B cell and MEF reprogramming systems raises the question as to whether it regulates pluripotency acquisition in pre-implantation development. Indeed, the combined ablation of *Cebpa* and *Cebpb* has been reported to display defective placental development and trophoblast implantation[63]. Likewise, the ectopic expression of early trophectodermal *Cdx2* in ESCs has shown to trigger the upregulation of endogenous *Cebpa* [64], which indicates that early embryonic TFs can also regulate it.

C/EBPα together with PPARγ coordinate adipogenesis with the reported ability of C/EBPα to directly bind and regulate key metabolic genes (i.e.: *Nmt*, *Pdk4*, *Ereg* or *Acsl1*)[65,66]. Adipocytes and liver hepatocytes fail to accumulate lipids in mouse neonates if C/EBPα is missing[67]. These latter cells also require fine-tuned binding of C/EBPα and C/EBPβ to control homeostatic and cell-cycle gene batteries during liver regeneration[68] and adipogenesis[69].

Interestingly, members of the C/EBP family can remain attached to mitotic chromatin during cell division[70] with C/EBPβ and C/EBPδ allocating to the centromeric regions of compact chromosomes in adipocyte differentiation[71]. Later studies depicted this property also for C/EBPα in hepatocytes showing how pioneering factors bind more strongly than other

factors to mitotic chromatin[72]. A recent project focused on the binding dynamics of HNF4α and C/EBPα during liver development attributes bookmarking properties to such mitotic retention[73]. Hence, they concentrate in a group of hepatic genes that become early activated by C/EBPα and still retain the factor bound to the regulatory chromatin after silencing their expression. Curiously, active H3K27ac mark decreases in this subset but the genes do not acquire repressive H3K27me3 marks typical of heterochromatin. Therefore, they speculate that some sort of epigenetic memory mediated by C/EBPα might prevent heterochromatin spreading in genomic regions so that genes stay ready for faster expression in future events.

2.2 C/EBPα partners with chromatin modifiers to shape the epigenome.

Methylation of specific lysines or arginines in histones is associated with either transcriptional activation or repression. There are approximately 50 histone methyltransferases that act in unique or redundant fashions rendering a wide range of epigenetic signatures and combinatorial modes of control[74]. For instance, WDR5 of the MLL trithorax complex mono- or di-methylates H3K4 histone lysines to activate enhancers[75] while members of the repressive polycomb complex deposit H3K27me3 mark to silence genes[76]. On the other hand, PRMT enzymes can also mono- or di-methylate arginine

residues, with mono-methylation occurring symmetric or asymmetrically, again raising a wide readout of silencing and activation[77].

LSD1 is the best described lysine demethylase[78] while the existence of specific arginine demethylases is still under debate[79,80]. TFs can also interact with reader proteins that recognize specific chromatin modifications, such as BRD4 whose bromodomain binds to acetylated histones[81,82] thus stimulating RNA polymerase II activity[83]. By bringing together chromatin modifiers and readers, TF mediated complexes fine tune chromatin regulation.

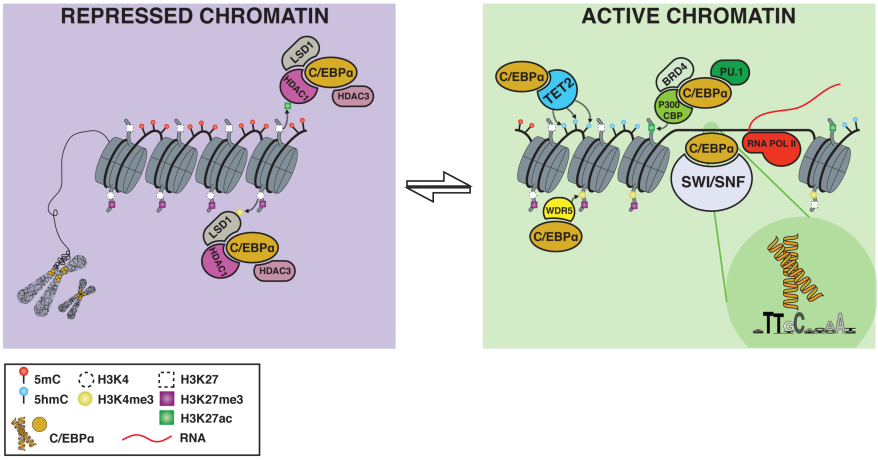


Fig. I4| C/EBPα may act as a pioneer factor that partners with chromatin-modifying enzymes to activate gene expression. (Left) C/EBPα complexes remove histone active marks to reduce chromatic accessibility and repress gene expression programs. C/EBPα binding to close heterochromatin remains in compact chromosomes during mitotic segregation. (Right) C/EBPα is capable of binding to closed chromatin as a pioneer factor and regulate it throughout multiple layers: it forms complexes with co-factors such as PU.1, it recruits chromatin modifying

enzymes (WDR5, P300, CBP, BRD4) to deposit active histone marks, it recruits chromatin remodeling enzymes (SWI/SNF) to displace nucleosomes, it recruits TET2 to induce DNA demethylation and it facilitates RNA pol II loading for subsequent activation of gene expression.

Applied to C/EBPα it has been shown that it tethers several chromatin modifiers that silence and activate gene programs during cell identity transitions (Fig. I4). For example, C/EBPα cooperates with CBP/p300[36] and MYST to acetylate histones[84] and with BRD4[85] to activate GREs embedded in chromatin. Similarly, C/EBPα interacts with the histone deacetylases HDAC1[86] and HDAC3[87] to repress gene expression. During C/EBPα-enhanced B cell to iPSC reprogramming, it forms a complex with HDAC1 and LSD1 to repress B cell genes while cooperating with BRD4 at H3K27ac marked pluripotency enhancers[36]. Interestingly, ectopic C/EBPα activates *Hdac1* and *Lsd1* expression in B cells and also increases BRD4 protein levels[36]. Furthermore, C/EBPα can interact with WDR5 in B cells during reprogramming[36] as it had been described for C/EBPα p30 isoform in myeloid leukemias[88]. Besides histone lysine methyltransferases, C/EBPα and C/EBPβ have been recently described to functionally interact with the arginine methyltransferases PRMT1[89] and CARM1 (PRMT4)[90] in both cancer and normal cells. Finally, C/EBPα also cooperates with members of the SWI/SNF chromatin remodeling complex and this interaction was shown to be essential for adipocyte differentaition[91] and myeloid cell proliferation[92].

Methylation of the fifth position of cytosines (5-mC) is a well studied mark that is tightly linked with mammalian development[93]. A set of DNA methyltransferases (DNMT1[93], DNMT3A and DNMT3B[94]) catalyze the deposition of these methyl groups after being passively diluted during cell division[95] to establish and maintain essential genome methylation signatures. DNA methylation of GREs is generally associated with transcriptional repression and is regulated by both passive and active mechanisms. Thus, methylation can be passively lost during cell division or removed actively by TET enzymes through hydroxymethylation of 5mCs (5-hmC)[96].

C/EBPα overexpression in B cells binds to and activates *Tet2* GREs during induced macrophage transdifferentiation[97]. This leads to the de-repression of myeloid genes and *Tet2* knock-down impairs macrophage identity acquisition, which can be counteracted by *Dnmt1* downregulation. Similarly, *Tet2* upregulation is also required for the establishment of pluripotency and promoter regions of the key pluripotency factors *Oct4* and *Nanog* become demethylated already after the pulse of C/EBPα[23].

Because TET2 lacks a DNA binding domain[98] it must be recruited to GREs with 5-mC residues[99]. Recent studies have shown that TET2 becomes recruited by specific TFs, such as

PU.1 in myeloid cells[100] or NANOG in ESCs[101]. Studies in our laboratory have shown that during B cell reprogramming C/EBPα directly interacts with TET2 and recruits it to GREs of myeloid and pluripotency genes (i.e.: *Klf4*, *Tet2* itself or *Cdh7*) where it erases 5-mC after hydroxymethylation[37]. In addition, the pluripotency TFs KLF4 or TFCP2L1 can also recruit TET2 to chromatin, a mechanism also shown during MEF reprogramming[37].

Given the implications of DNA methylation during different steps in embryo development (parental imprinting, X-chromosome inactivation, trophectodermal differentiation or primordial germ cell specification[93]) and the aforementioned control of some of this processes by C/EBPα, it would be interesting to study the role of this factor in the context of early lineage fate specification.

3. The IL-6 signaling pathway.

3.1 IL-6 signaling components and their regulation.

IL-6 cytokine was first identified as a secreted factor capable of stimulating B cells to proliferate and differentiate into antibody-forming cells[102,103]. In that line, *Il6* depletion in mice fails to activate T cell-dependent antibody formation and acute-phase responses upon infection[104]. As a pleiotropic

factor, IL-6 is also involved in inflammatory processes during tumor progression[105]. Interestingly, these functions have been recently extended to IL-6 synthesized by brown adipocytes showing that it can stimulate liver gluconeogenesis, fueling "fight or flight" responses to stress[106].

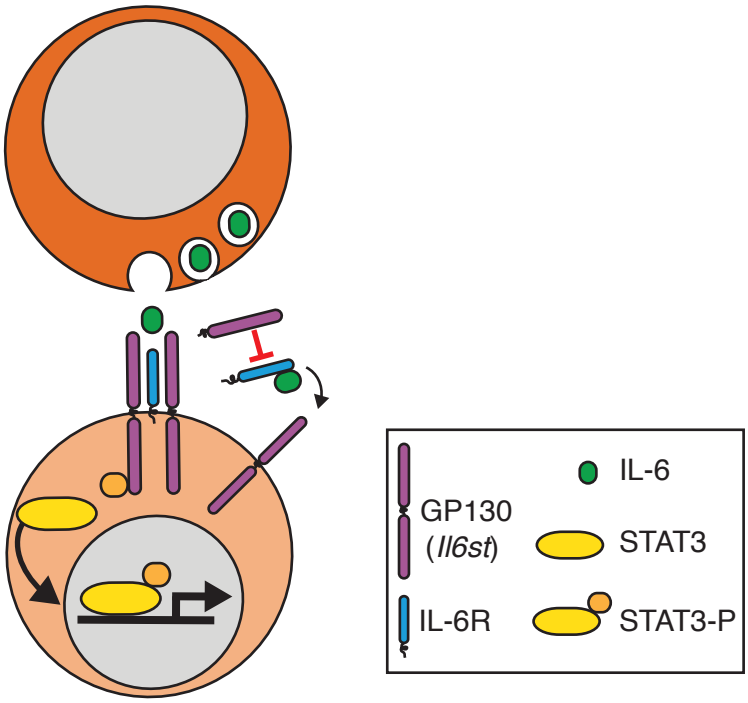


Fig. I5| IL-6 signaling pathway. Classical IL-6 signaling is initiated by the binding of the ligand to a specific receptor (IL-6R), which leads to heterodimerization with a common signal-transducing receptor (GP130, encoded by *Il6st*). In a more indirect transactivation mechanism, a soluble form of the receptor (encoded by a spliced form of *Il6ra*) is released by proteolysis and binds to IL-6. This complex allows the cytokine to act on cells that express GP130 but lack IL-6R. Antagonistically, a soluble form of GP130 can also block the enhanced signaling triggered by soluble IL-6R. Both modes of IL-6 receptor signaling end up with STAT3 phosphorylation, facilitating its translocation into the nucleus. After binding to DNA STAT3-P activates gene expression.

Molecularly, IL-6 binds to its specific receptor IL-6R and the shared signal transducer GP130 (also IL6ST)[107] forming dimeric structures[108]. IL-6R is mainly expressed and active in hepatocytes, myeloid cells and megakaryocytes[103,109]. On the other hand, the GP130 receptor is the effective signaling molecule shared with IL-11, leukemia inhibitory factor (LIF), oncostatin M, ciliary neurotrophic factor and cardiotrophin 1[110]. Not surprisingly, *Gp130* knock-outs result in embryonic death while *Il6*$^{-/-}$ and *Il6ra*$^{-/-}$ null mice are viable.

During IL-6 signaling, GP130 dimers activate JAK kinases that phosphorylate STAT3, STAT1 and STAT5[111]. Phosphorylated STAT proteins then shuttle into the cell nuclei and contribute to the transcriptional activation of specific gene programs such as the pluripotency signature regulated by STAT3[112,113,114]. IL-6 signaling is tightly controlled at different levels including inhibitors of activated STATs as well as cytokine and receptor supressors[111]. Additionally, soluble isoforms of the membrane receptors have been shown to play a role in signaling regulation[115]. Alternative splicing[116] and protein shedding mediated by ADAM members[117] can give rise to a soluble version of IL-6R that enhances IL-6 signaling[118]. Conversely, a soluble isoform of GP130 can further inhibit soluble IL-6R in the medium to prevent trans-signaling amplification of the IL-6 effect[119] (Fig. I5).

3.2 Connections between C/EBP family members and the *Il6* pathway.

Multiple studies have described a role of C/EBP members in inflammation and cytokine regulation[120]. Recent analyses have also detected evolutionarily conserved *Il6*-*Il6ra*-C/EBP gene modules that are important for myeloid differentiation in humans[121]. C/EBPβ (originally named NF-IL6) was first identified as a nuclear factor capable of binding to the IL-6 promoter and activate its expression during inflammatory responses[122,123]. Parallel work in our group identified the chicken equivalent of C/EBPβ (named nuclear factor myeloid or NF-M[124]) to also bind to the promoter of cMGF, a gene encoding an IL-6-like cytokine secreted by myeloid cells[125]. However, *Cebpb*$^{-/-}$ null mice still harbored high IL-6 levels[126], compatible with the fact that other transcription factors, such as NFkB[123], NOTCH[127] or JUN[128] can also control *Il6* expression.

In addition, C/EBPα was shown to be required for the proper expression of IL-6R in the liver and *Cebpa*$^{-/-}$ null hepatocytes failed to phosphorylate STAT3 upon LPS or IL-6 stimulation[129]. Curiously, IL-6 itself triggers C/EBPδ upreglation in hepatocytes by activating STAT3 and SP1 that will bind to C/EBPδ promoter[130]. However, more precise research into C/EBP members controlling *Il6* pathway genes will shed light on its functions during different processes.

3.3 IL-6 signaling in cell identity and pluripotency.

IL-6 was shown to participate in cell fate differentiation during monocyte maturation. Thus, monocytes can develop into dendritic cells or macrophages during hematopoietic differentiation and IL-6 secretion from feeder MEFs was shown to skew this decision towards macrophages by triggering the upregulation of the M-CSF receptor (*Csf1r*) in monocyte cultures *in vitro*[131].

In the context of pluripotency, the addition of a combination of IL-6 and soluble IL-6R to the culture medium of ESCs could replace LIF for their maintenance[132]. IL-6 is also rapidly upregulated in early stages during heterokaryon reprogramming of human fibroblasts by mouse ESCs[133]. In this study, they proved that IL-6 addition could enhance MEF to iPSC reprogramming and they further showed that *Il6* gene could functionally substitute *Myc* from the Yamanaka cocktail. IL-6 was also found to be strongly upregulated in the pancreas during *in vivo* reprogramming by OSKM[134]. This effect was attributed to the induction of senescence as IL-6 is one of the main cytokines secreted by injured or aged tissues[135]. Blocking IL-6 *in vivo* resulted in a reduced number of somatic cells expressing *bona fide* pluripotency markers, suggesting a paracrine effect of IL-6 released by surrounding senescent cells during *in vivo* iPSC formation[134]. A similar paracrine effect from senescent tissues was also described

for the reprogramming of skeletal muscle cells[136].

Il6 mRNA has been also detected in mouse late blastocysts[137], a developmental stage during embryo pre-implantation when the first pluripotent cells arise in the inner cell mass (ICM) layer. Additionally, blocking IL-6 during the *in vitro* development of mouse embryos induced a decrease in STAT3 phosphorylation, suggesting that the IL-6 pathway is active in this developmental window[138]. More precisely in humans, *Il6* mRNA expression has been detected in the extra embryonic trophectoderm (TE) while *Il6ra* expression was enriched in the ICM at the late blastocyst stage[139].

Future studies on the functionality and regulation of the *Il6* pathway will be required to better understand its role in pluripotency acquisition during somatic cell reprogramming and ICM/TE segregation in pre-implantation development.

4. Pre-implantation mouse development.

Understanding the development of a fertilized egg into a multicellular organism with functionally different cell types is a fascinating challenge in biology. In the mammalian embryo, the first cleavages of the zygote before its implantation in the uterus contain key information to specify the future lineages of the cells in the adult tissues[140,141]. These pre-implantation events include the loss of totipotency around the 8-cell

stage[142] and the first fate decision in development when blastomeres segregate into two cell populations: TE and ICM. Before implantation, the ICM further differentiates into the primitive endoderm (PrE) and the epiblast (EPI)[143]. During post-implantation development, the trophectoderm gives rise to the placenta, the primitive endoderm develops into the yolk sac and the epiblast forms the embryo[144]. This progression from totipotency to lineage specification involves multiple layers of regulation that gives this process a high degree of control and robustness (Fig. I6).

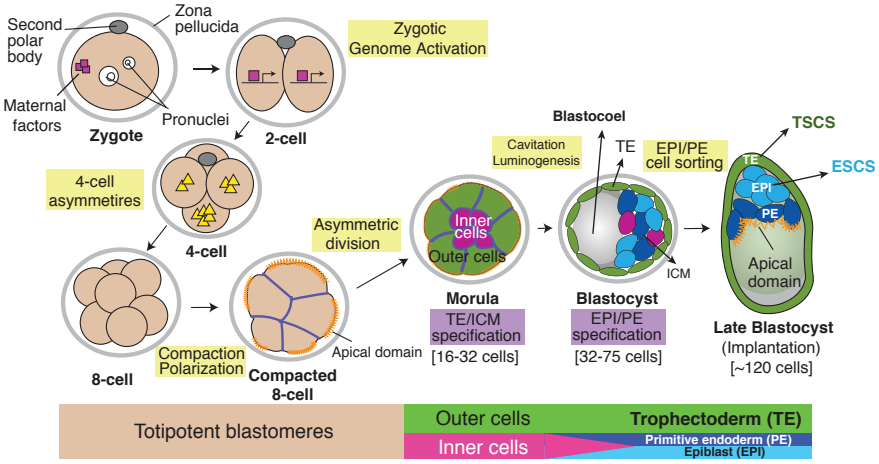


Fig. I6| Pre-implantation mouse development. Schematic of mouse pre-implantation embryo development. The coloured bars on the bottom indicate the progression from totipotent blastomeres to embryos with the first three lineages: trophectoderm (green), epiblast (light blue) and primitive endoderm (dark blue). The latter two derive from the inner cell mass in pink. Important morphogenetic events are highlighted in yellow boxes. Lineage specification events are highlighted in purple boxes. Figure modified from Rossant and Tam (2009); Chazaud and Yamanaka (2016).

4.1 Maternal regulation of the mouse embryo.

The mouse embryo is almost entirely dependent on the inherited components of the egg for survival prior to activation of the genome in later stages[145,146]. Removal of polyadenylated RNA debris through endogenous siRNAs and RISC-DICER[147] components is essential for the zygote to undergo subsequent cleavages. Intact ubiquitin-proteasome pathways for selective maternal protein degradation have been shown to be important in early developmental stages[148]. Methylating enzymes such as DNMT1[149], DNMT3[150] or STELLA[151] are required to prevent passive demethylation and misregulated gene expression during the first embryonic divisions. In addition to a potential role of maternal OCT4 enhancing zygotic genome activation[152], recent studies have identified members of the DUX transcription factor family (also known to be involved in muscular distrophy[153]) as the main regulators of this process[154,155,156]. Nonetheless, maternal depletion of *Oct4*[157] and *Dux*[158] have proved mouse development to still be possible, attesting to the robustness of early embryonic development. In addition, NFYA protein (a CCAAT-binding member of the NF-Y complex involved in the maintenance of embryonic stem cell identity *in vitro*[159]) has been shown to maternally contribute in the rearrangement of chromatin accessibility so that additional transcription factors may start the zygotic genome activation[160]. It is likely that new

factors will be identified that complement the function of OCT4, DUX and NFYA.

4.2 Zygotic genome activation at the 2-cell stage

Early studies demonstrated that individual blastomeres from split 2-cell embryos can give rise to fertile organisms[161]. At this stage, transcripts from the newly assembled diploid genome in the mouse start to be detected[146]. Among them, murine endogenous retroviral (MuERV-L) transcripts get highly expressed and then vanish during the transition towards 4- and 8-cell embryos[162]. Using an MuERV-L-GFP reporter it was discovered that around 1% of ESCs in culture can transiently express MuERV-L transcripts[163]. These 2-cell-like cells can contribute both to embryonic and extra embryonic tissues (TE) when injected into morula embryos and transcriptionally upregulate the embryonic 2-cell program including well-characterized genes such as *Zscan4*[164]. Additionally in 2-cell-like cells, pluripotent transcription factors become degraded at the protein level, chromatin de-condensates (reflected by the loss of chromocenters) and there is a switch in histone mobility and histone marks[165]. However, a limitation of this system is that 2-cell-like cells cannot be stably maintained in culture. Recent studies showed that DUX overexpression in ESCs can increase the proportion of these 2-cell-like cells to about 75%[156], suggesting that DUX is one of the main regulators of the 2-

cell embryo, providing an *in vitro* system to explore the molecular features of this earliest stage of development.

The question whether 2-cell stage blastomeres already develop asymmetries that foreshadow their developmental potential has been controversial in the field. Some studies reported an unequal distribution of mitochondrial, ribosomal[166] and long non-coding RNAs[167] between 2-cell blastomeres that already influence the allocation of subsequently derived cells into the blastocyst's ICM or TE. Complementary work has also detected that sister blastomeres present extensive transcriptomic differences and have distinct capacity to build a whole organism upon separation[168]. However, as it will be discussed in the following section, the consensus is now that symmetry breaking at the 4-cell stage defines a clearer picture to predict cell fate.

4.3 The 4-cell stage: breaking symmetries and cell fate commitment.

Single 4-cell blastomeres can still contribute to both embryonic and extra embryonic layers and they give rise to a whole organism when cultured with carrier cells[142,169]. The prediction that differences between sister blastomeres observed at the 2-cell stage would be amplified by zygotic genome activation in the 4-cell embryo[170] raised the question of whether blastomeres at this stage would exhibit stronger asymmetries that predict their developmental fate.

Elegant studies by the laboratory of Magdalena Zernicka-Goetz tracking the division planes of 2-cell embryos (combining labeled beads attached to the cell membrane with injected cytoplasmic reporters) firstly characterized those asymmetries at the 4-cell stage[171]. These studies revealed that blastomeres located in the vegetal pole of the 4-cell embryo, which is the farthest position from the polar body (a haploid DNA structure remaining from meiotic divisions during oogenesis), preferentially give rise to the TE layer in the blastocyst[171].

The first study that molecularly addressed 4-cell asymmetries identified distinct histone arginine methylation levels correlating with ICM and TE segregation[172]. Blastomeres with high levels of H3R26me preferentially participated in the ICM and expressed higher levels of pluripotency factors. The specific methyltransferase responsible for the observed H3R26 pattern was found to be the arginine specific methyltransferase CARM11 (PRMT4). Indeed, both the enzyme and the histone mark were lowest in the vegetal blastomere. Moreover, *Carm1* overexpression in a 2-cell blastomere resulted in the contribution of that cell predominantly to the ICM. These results suggested a role of *Carm1* in creating a functional asymmetry among 4-cell embryos that predict subsequent cell fate specification.

Initially, the observed histone methylation asymmetry was thought to introduce differences in chromatin accessibility among the four blastomeres. If so, uniformly expressed cell fate instructive transcription factors might segregate cells into ICM or TE depending on the cell's chromatin accessibility at specific genome regulatory regions. This idea was tested by injecting photoactivable forms of pluripotency-related transcription factors, allowing the measurement of multiple kinetics (residence binding time in chromatin, diffusion coefficients, etc.) that correlated with lineage patterning in the early embryo. In one study, OCT4 binding was shown to be heterogeneous, being bound more stably to the chromatin of blastomeres fated to become ICM[173]. Using similar approaches another study reported that the pluripotency factor SOX2 results less diffusible among ICM-committed blastomeres in a CARM1 dependent manner[174].

Single cell RNA sequencing also identified asymmetrically expressed lineage instructive transcription factors with *Sox21* ranking highest in the list[175]. As for SOX2 diffusion, asymmetric distribution of SOX21 protein levels in 4-cell blastomeres was found to depend on CARM1 activity. In addition, they found that *Sox21* downregulation led to a strong increase of CDX2 detection. Since previous work had demonstrated that the CDX2 TF is required for TE differentiation[176] and that SOX21 binds to *Cdx2* promoter repressing its expression in ESCs[177], they concluded that

Sox21 helped to maintained pluripotency in the embryo. Surprisingly, although *Carm1* downregulation by siRNAs showed no effects in the total number of cells contributing to the TE, it altered the proportion of cells contributing to the EPI and PrE layers within the ICM. Correspondingly, *Carm1* overexpression in the zygote triggered the upregulation of ICM related genes (*Sox21*, *Sox2* and *Nanog*) at the 8-cell stage without affecting *Cdx2* mRNA levels[175].

A recent study by the laboratory of Nicholas Plachta reported asymmetries in the expression of keratins 8 and 18 (keratins are intermediate filaments of the cytoskeleton), showing that their expression at the 8-cell stage predicts trophectodermal fate in mouse and human embryos[178]. KRT8 and KRT18 were first detected in 8-cell blastomeres derived from the vegetal cell in 4-cell embryos where the SWI/SNF chromatin remodeling component BAF155 is most highly expressed. *Baf155* knockdown reduced keratin expression, establishing BAF155 as a regulator of KRT8 and 18 required for TE formation[178]. This observation connected these studies to previous work showing that BAF155 is another direct target of CARM1[179]. Therefore, they further analyzed methylated BAF155 and found that it exhibits lowest levels in the vegetal 4-cell blastomere[178]. In that line, *Carm1* overexpression was also found to impair keratin detection, similarly to *Baf155* knockdown. For reasons that are unclear, in this study *Carm1* overexpression led to reduced CDX2 protein levels in

potential contradiction to the described lack of effect on *Cdx2* mRNA expression[175].

New investigations will likely identify additional histone modifiers and transcription factors that link with structural proteins to fill the missing pieces dictating early cell fate decisions.

4.4 Compaction and polarization at the 8-cell stage leading to the first cell fate decision.

In the progression towards the blastocyst, 8-cell blastomeres eliminate the intercellular space and flatten against each other in a process known as compaction. E-cadherin (*Cdh1*) is the main protein mediating cellular adhesion and its depletion leads to *Cdx2* upregulation and an increase in TE fated cells[180]. Recent studies have shown that microtubule organizing centers (MTOCs) create cytoplasmic bridges between blastomeres through which E-cadherin molecules are transported[180]. CDH1-dependent filopodia also protrude into the membrane of adjacent cells and create pulling forces to bring blastomeres together[181]. Other proteins such as β-catenin further participate in this compaction process[182].

8-cell blastomeres also undergo polarization. The apical domain of the cells accumulates more ezrin, microvilli, F-actin, PAR complex members and aPKC (among other

molecules) than the basal domain[141]. Polarized cells then either divide symmetrically, each inheriting half of the apical and basal cytoplasm, or asymmetrically by exclusively retaining either the apical or basal domains. Subsequently, the cells distribute into the outer or inner regions of the embryo. Although this cell positioning has been coupled with fate specification, with inner cells turning into ICM and outer cells turning into TE[183], inheriting the apical domain (regardless of cell position[184]) has been reported to be sufficient for the first lineage segregation. Thus, cells that inherit the apical domain commit to the TE lineage[185].

The Hippo/YAP pathway has been described as the molecular sensor that transforms these mechanical cues into the induction of TE cell identity[186], mainly through *Cdx2* upregulation[187]. Sequestering angiomotin (AMOT) in outer cells by apical PAR proteins, aPKC and F-actin filaments prevents AMOT phosphorylation. Consequently, the Hippo pathway co-activators YAP/TAZ shuttle into the nucleus where they couple with TEAD4 TF to upregulate expression of the TE effector *Cdx2*. Furthermore, *Cdx2* transcripts also become enriched in the apical domain of the cells and preventing such restricted localization spreads TE fate among pluripotent committed cells where ectopic CDX2 is detected[188]. In addition to *Cdx2*, *Gata3* and *Eomes* have also been reported to regulate TE development downstream of *Tead4*[189,190], with *Gata3* acting in parallel to *Cdx2*.

In ICM committed cells, phosphorylated AMOT is free in the cytoplasm and together with other factors (such as NF2) triggers Hippo pathway activation through the interaction with LATS1/2 kinases[191]. These enzymes phosphorylate and sequester YAP/TAZ co-activators in the cytoplasm thereby preventing *Cdx2* expression. Cytoplasmic YAP/TAZ localization is sufficient to induce ICM-restricted *Sox2* expression and to initiate the pluripotency program[192]. However, the molecular intermediates that control *Sox2* activation remain unclear.

Besides the Hippo pathway there are other molecular players involved in the first lineage decision. For instance, NOTCH signaling cooperates with TEAD4 to activate *Cdx2*[193] and, at least under low O_2 conditions, *Tead4* null embryos can still segregate into TE and ICM[194]. Furthermore, mechanical signals coming from a rigid extracellular matrix and cell shape are able to activate YAP/TAZ independently of the Hippo pathway[195], suggesting that other factors may also participate in this early fate decision.

Studies on the role of F-actin rings that form at the apical domain of blastomeres before polarization have shown that these structures disassemble after each cell division and that they are formed *de novo* instead of being directly inherited[196]. This observation challenged the proposed model by which

actomyosin flows led the initial concentration of proteins in the inherited apical domains during TE/ICM segregation. However, recent studies have managed to address this point by further characterizing that it is the expression timing of *Tead4* together with *Tfap2c* and activated RhoA what drives polarization in the embryo. These experiments indicate that TEAD4-TFAP2C complexes, rather then actomyosin flows, mediate the accumulation of proteins in the apical domain that will determine cell polarization and lineage fate[197].

4.5 Luminogenesis and emergence of the primitive endoderm: the second cell fate decision.

ICM-cells that appear after compaction around the 32-cell stage are not uniform but show a “salt-and-pepper” pattern for the expression of *Nanog* and *Gata6*[198]. These genes, which cross-antagonize each other, respectively specify the EPI and PrE layers at the ICM. Thus, *Nanog* is exclusively expressed in the EPI layer while *Gata6* is expressed in the PrE layer of the late the blastocyst. The mutual repression is mediated through FGF signaling[199,200,201]. Binding of FGF4 to FGFR2, which is homogeneously expressed in ICM-cells, activates ERK kinases, leading to an increase of *Gata6* expression and the silencing of *Nanog*. On the other hand, *Nanog* expressing cells amplify FGF4 secretion, thus repressing FGFR2 and silencing *Gata6* to stabilize pluripotency. Blocking FGF signaling during embryo development has been reported to induce ICM cells lacking the PrE layer[202]. Conversely, excess

FGF4 is sufficient to induce the differentiation of ICM cells into PrE[203]. Importantly, manipulating FGF signaling becomes ineffective at later stages when cells are already committed[204].

Additionally, *Fgf4*$^{-/-}$ null embryos fail to form PrE in the late blastocyst although they still express *Gata6* during compaction[205], indicating that other factors redundantly control PrE development. *Gata6*$^{-/-}$ embryos develop *Nanog* positive ICM cells in which FGF4 can no longer rescue PrE formation[206], showing that *Gata6* acts downstream of FGF4. Interestingly, *Nanog*$^{-/-}$ embryos show reduced FGF4 levels and still express *Gata6* in the ICM[207], indicating that the initiation of the PrE program is independent of both *Nanog* and FGF4 signaling. Stochastic local gradients of *Nanog*-expressing cells and of FGF4 in the ICM have been proposed to trigger EPI/PrE differentiation process in the ICM[208]. Further investigations will disclose yet unknown factors that might be required for the selective increase of PrE specific genes in a subset of ICM cells.

After *Gata6* restriction, PrE cells mature and express specific markers including *Sox17*, *Gata4*, *Dab2* and *Pdgfr*α[141]. The maintenance and progression of mature PrE has been shown to also depend on *Sox2* and *Oct4* pluripotency factors expressed in the EPI layer. Thus, *Sox2*$^{-/-}$ null embryos exhibit reduced *Sox17* and *Gata4*[192] expression levels while *Oct4*

overexpression has been shown to promote ESCs differentiation into PrE[209]. Fine tuning of the FGF4 signaling has been reported to be the mechanism by which *Oct4* and *Sox2* regulate PrE maturation[210]. Other signaling pathways like BMP[211] and apoptosis[212] have also been described to influence this second cell fate decision in the embryo.

Moreover, the formation of PrE involves cell polarization and migration given that these cells form of a basal epithelium beneath the EPI layer. aPKC[213] and structural proteins including laminins, collagen or F-actin have been described to participate in this process[213].

Concomitantly with PrE/EPI segregation, the embryo also forms a cavity (blastocoel) between the mural TE and the ICM in a process known as luminogenesis or cavitation. Na/K-ATPase pumps that transport Na^+ ions[214], aquaporins that form water channels[215] and structural proteins such as zona occludens-1[216] enlarge the intercellular space to form the cavity. Recent studies have also detected widespread secretion of cytoplasmic vesicles from outer and inner cells into the intercellular space driving early fluid accumulation[217]. In these experiments, FGF4 has been detected in early luminal structures, leading to the proposal that ICM lineage specification and spatial segregation depend on expansion of the cavity[217]. Other groups have shown that TE cells also

secrete FGF2 into the blastocoel thereby activating the FGF pathway in other TE cells that express FGFR2[198].

Integrating the effect of mechanical forces with the various signaling pathways that are active during luminogenesis will be key to fully understand early embryo organization and PrE specification[218].

4.6 Stem cell derivation: reconstructing the mouse embryo.

The first ESCs that resembled the ICM were independently derived by Martin Evans and Gail Martin from mouse blastocysts in 1981[26,27]. Years later, the group of Janet Rossant also succeeded in generating trophoblast stem cell lines (TSCs) from the TE layer[219]. Her group also identified that the TF CDX2 can convert ESCs into TSCs by repressing OCT4[220]. More recently, cell lines bearing features of PrE (XEN lines)[221] or the EPI layer (EpiSCs)[222] have also been successfully derived from post-implantational stages of the embryo. These cell lines have become instrumental to approach embryonic development and molecularly disentangle what governs cell lineage specification.

By aggregation of different stem cell lines and culturing them under specific conditions, several recent studies have reported the recreation of complex structures resembling distinct developmental stages of the embryo[223]. For example,

Zernicka-Goetz and colleagues were able to partially recapitulate post-implantational events such as mesoderm specification and primordial germ cell's emergence in embryo-like structures after co-culturing ESCs and TSCs in 3D scaffolds[224]. Some other groups have also managed to create pre-implantation blastocyst-like structures ("blastoids") either by combining ESCs and TSCs [225] or just from ESCs[226] . However, none of these attempts have yet succeeded in recapitulating all developmental stages from the zygote or give rise to structures that can functionally replace an embryo. Such embryo reconstruction approaches promise to identify new layers of regulatory events that control early developmental stages.

AIMS

The C/EBPα-enhanced Bcell reprogramming system that our laboratory developed[23] raised the question of whether C/EBPα itself would also play a physiological role in the establishment of pluripotency during early embryonic development.

Additional studies on the field also pinpointed the relevance of IL-6 signaling both during *in vitro*[133] and *in vivo*[134,136] somatic cell reprogramming. These new observations, together with former studies of certain C/EBP members regulating the *Il6* pathway[123,129], led us to further explore the molecular connection between C/EBPα and the IL-6 signaling in the context of pluripotency acquisition and cell fate changes.

Thus, this thesis project aimed to develop two main points:

- Investigating the activity of the *Il6* signaling pathway in cell fate transitions and its molecular connection to C/EBPα using our B cell to iPSC reprogramming and B cell to macrophage transdifferentiation systems.

- Interrogating the potential role of C/EBPα and the *Il6* pathway during the first cell fate decisions in early embryonic development.

RESULTS

CHAPTER 1: Role of C/EBPα and the IL-6 pathway during somatic cell reprogramming.

Previous studies had shown the relevance of the *Il6* pathway for the establishment of pluripotency during somatic cell reprogramming both *in vitro*[133] and *in vivo*[134,136]. Furthermore, IL-6 signaling has been reported to promote monocyte differentiation into macrophages in culture[131]. Parallel studies have also proved that transcription factors from the C/EBP family regulate the *Il6* pathway[123,129].

Based on this work, we decided to take advantage of our C/EBPα-enhanced reprogramming system in B cells to further explore the molecular regulation of the *Il6* pathway in the acquisition of pluripotency. Moreover, we also interrogated the potential role of IL-6 signaling in alternative cell fate transitions using our B cell to macrophage transdifferentiation system.

1. C/EBPα upregulates *Il6* pathway genes in transdifferentiation and in enhanced reprogramming of B cells.

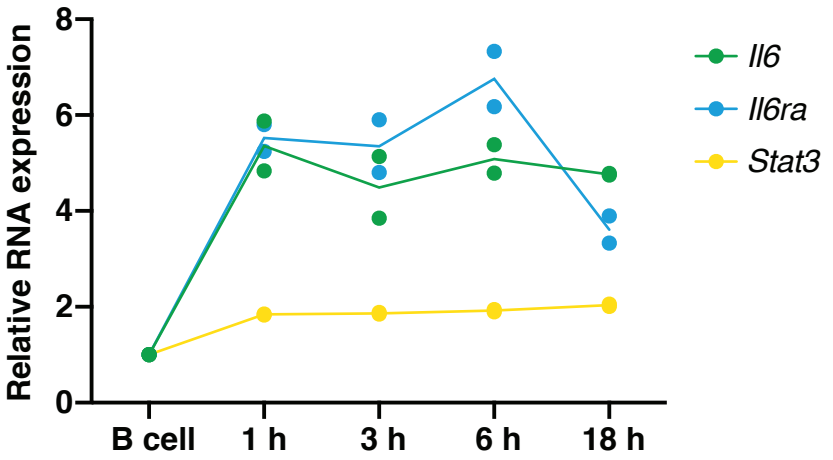


Fig. R1| Gene expression changes of *Il6* pathway genes induced by C/EBPα overexpression in B cells. RNA expression levels relative to uninduced B cells at different times after C/EBPα induction, derived from RNA-seq data.

The first step in our highly efficiency reprogramming protocol consists in the activation for 18 hours of C/EBPα in B cells expressing C/EBPαER by the addition of β-estradiol, generating Bα' cells. These cells correspond to an early myeloid cell state with maximal susceptibility for reprogramming into iPSCs by subsequent OSKM activation[23,36] (Fig. I4). To study a possible connection to the *Il6* pathway, we first determined the effect of C/EBPα expression by RNA-seq in B cells induced for 0, 1, 3, 6 and 18 hours by C/EBPα. Compared to uninduced cells, this

revealed 5 to 6-fold increases of *Il6* and *Il6ra* and about 2-fold increase of *Stat3* gene expression (Fig. R1).

2. C/EBPα binds to and activates regulatory regions of *Il6* pathway genes.

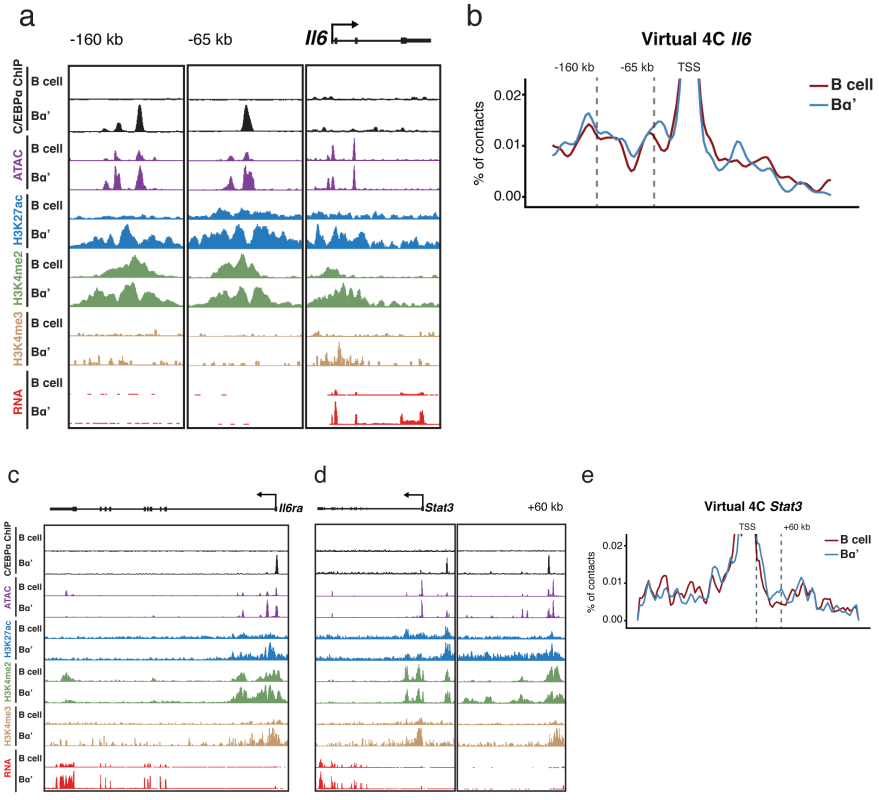


Fig. R2| Genomic changes at C/EBPα bound sites on the *Il6* gene locus in B and Bα' cells. a, Browser shots of regions upstream of *Il6* bound by C/EBPα (ChIP-seq) after 18h of induction (Bα' cells), showing changes in chromatin accessibility (ATAC-seq), histone activation marks (H3K27ac, H3K4me2 and H3K4me3) and RNA expression **b,** chromatin contacts at the *Il6* locus in B and Bα' cells. Virtual 4C data using the transcription start site (TSS) as a viewpoint. Dashed lines indicate C/EBPα binding sites. **c-e**, Similar browser shots of analyses as in described **a** and **b** for the *Il6ra* and *Stat3* loci.

To determine whether C/EBPα directly regulates *Il6* pathway genes we combined various datasets of ChIP-seq[36], ATAC-seq[38] and RNA-seq[38] generated in our laboratory, analyzing the *Il6*, *Il6ra*, and *Stat3* genes. For the ChIP-seq data this included C/EBPα as well as chromatin decoration by the activating histone marks H3K27Ac and H3K4me3, the latter of which is characteristic for activated promoters[74].

The following results were obtained for the *Il6* gene: C/EBPα was found to be bound to two upstream regions at -65 and at -160 kb (Fig. R2a). These regions are already accessible to the Tn5 transposase (used in the ATAC assay) in B cells, and the size of the corresponding peaks slightly increased in Bα' cells, suggesting that it is due to the addition of C/EBPα to already bound complexes in B cells. Together with the promoter, these regions also showed an increased histone H3K27 acetylation in Bα' cells, indicating chromatin activation. Additionally, the promoter exhibited an increase of H3K4me3 decoration in Bα' cells, correlating with the observed elevated mRNA reads of the gene's 4 exons. Finally, the -65 kb upstream region showed an increase in the number of contacts with the TSS upon C/EBPα binding in Bα' cells, suggesting an enhancing effect of the active chromatin to foster transcription (Fig. R2b).

Very similar chromatin changes were observed at the *Il6ra* gene where C/EBPα was found to be bound to the gene

promoter (Fig. R2c). Here, as for the *Il6* enhancers, at the C/EBPα bound site there was an increase in the ATAC peaks as well as an increase in the histone activation marks, again correlating closely with increased expression of the gene's RNA.

For the *Stat3* gene, C/EBPα was found to be bound to a promoter-proximal region (+2 kb) as well as to a +60 kb upstream region. These regions increased decoration with activating histone marks correlating with RNA upregulation upon C/EBPα binding (Fig. R2d). Here, C/EBPα-bound upstream region also increased the number of DNA contacts with the TSS of *Stat3* in Bα' cells (Fig. R2e).

Together, these data suggest that overexpression of C/EBPα in B cells directly activates the *Il6*, *Il6ra* and *Stat3* genes. They also suggest that the genes become activated by direct binding of C/EBPα to a promoter and/or enhancer proximal regions.

3. C/EBPα induces IL-6-dependent STAT3 phosphorylation.

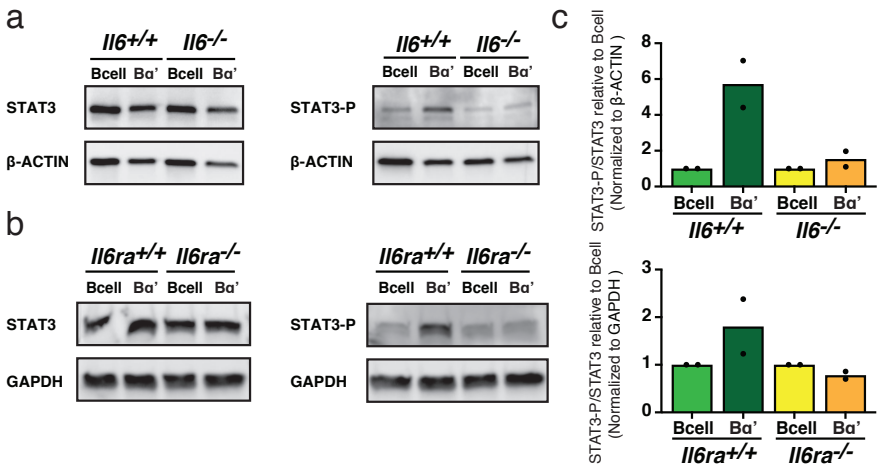


Fig. R3| Detection of phosphorylated STAT3 upon C/EBPα induction and effect of *Il6/Il6ra* ablation. a, Western blot showing bands corresponding to STAT3 and STAT3-P in wild type and *Il6*$^{-/-}$ B and Bα' cells. β-ACTIN was used as a loading control. **b**, Same as in **a** except here cells from wild type were compared with cells from *Il6ra*$^{-/-}$ mice. GAPDH was used as a loading control. **c**, Quantification of STAT3-P/STAT3 protein in B and Bα' cells with the indicated genotypes.

After observing that C/EBPα can directly induce the expression of *Il6* pathway genes we tested whether this results in the activation of the downstream signaling effector. For this purpose we examined the expression of STAT3 protein and its tyrosine 705 phosphorylated (activated) version before and after the pulse of C/EBPα. To determine whether STAT3 phosphorylation requires activation of the *Il6* pathway conditions, we infected B cells from wild type, *Il6*$^{-/-}$ or *Il6ra*$^{-/-}$ mice with a C/EBPαER retrovirus, sorted the infected cells and plated them onto mitomycin-C inactivated fibroblasts

(MEFs). After C/EBPα induction of the B cells by the treatment with β-estradiol for 18 hours they were removed from the feeders and tested by Western blot for the expression levels of STAT3 and STAT3-P proteins under the different conditions.

We observed enhanced STAT3-P bands specifically in C/EBPα-induced cells with a wild type genotype as the intensity of the bands corresponding to the STAT3-P protein did not increase in either *Il6*$^{-/-}$ *or Il6ra*$^{-/-}$ Bα' cells (Fig. R3). In addition, we did not detect substantial changes in STAT3 levels under any of the conditions tested. We therefore conclude that C/EBPα functionally activates the *Il6* pathway in B cells.

4. IL-6 signaling is not required for B cell to macrophage transdifferentiation.

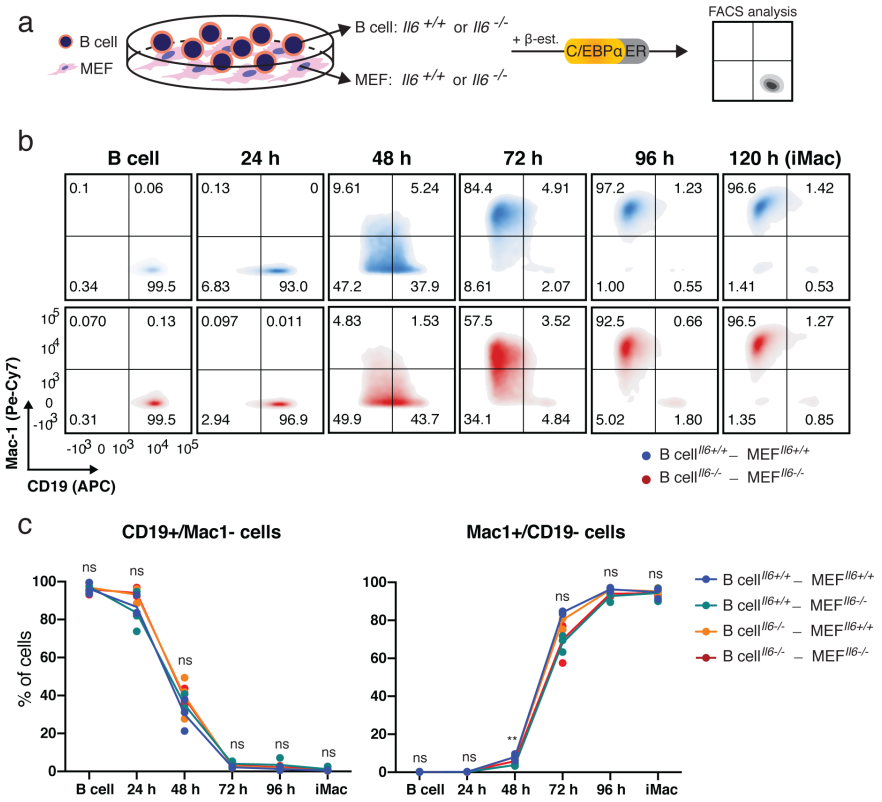


Fig. R4.1| Effect of the *Il6* genotype on C/EBPα-induced B cell transdifferentiation. a, Experimental protocol showing wild type or *Il6* knock-out B cells infected with C/EBPα-ER. Cells grow onto wild type or knock-out MEFs and C/EBPα is shuttled into the nuclei upon the addition of β-estradiol. Transdifferentiation dynamics are tracked by FACS analysis at different time points using labeled antibodies against the B cell marker CD19 and the macrophage marker Mac-1. **b**, FACS profiles of B cell$^{Il6+/+}$ and MEF$^{Il6+/+}$; B cell$^{Il6-/-}$ and MEF$^{Il6-/-}$ during transdifferentiation into induced macrophages (iMac). **c**, Transdifferentiation kinetics of the four possible combinations of *Il6* genotypes between B cells and MEFs, as indicated. Statistical significance between the various genotypes (n=3 biological replicates) was determined using two-way ANOVA and Tukey's multiple comparison test. Mac-1 upregulation was significantly higher (p=0.0028) for the B cells$^{Il6+/+}$ grown on MEF$^{Il6-/-}$ feeders at 48 hours post induction compared to B cells$^{Il6-/-}$ grown on MEF$^{Il6+/+}$ feeders.

To test whether the *Il6* signaling pathway has an impact on the cell identity switch during B cell to macrophage transdifferentiation, we derived primary B cells as well as mouse embryonic fibroblasts (MEFs) used as feeders from wild type and *Il6*$^{-/-}$ mice. In short, we isolated CD19 positive B lineage cells from bone marrow and infected them with a C/EBPαER-hCD4 containing retrovirus. Infected B cells were sorted on a magnetic column based on the presence of hCD4 cell surface marker before seeding them onto a feeder layer of MEFs inactivated by mitomycin C-treatment (Fig. R4.1a). C/EBPα was shuttled into the nuclei after the addition of β-estradiol during the whole transdifferentiation.

To test B cells and MEFs as possible sources of IL-6 (mitomycin C-treated MEFs become senescent and secrete IL-6 into the culture medium[134,135,227]), we performed the transdifferentiation experiments using different combinations of wild type and *Il6*$^{-/-}$ B cells and MEFs. We monitored the transdifferentiation dynamics by FACS through immunostaining of the B cell surface marker CD19 and the macrophage marker Mac-1. We essentially observed indistinguishable transdifferentiation kinetics with the four pairwise combinations of *Il6* genotypes tested (B cell$^{Il6+/+}$ and MEF$^{Il6+/+}$; B cell$^{Il6-/-}$ and MEF$^{Il6-/-}$; B cell$^{Il6+/+}$ and MEF$^{Il6-/-}$; B$^{Il6-/-}$ and MEF$^{Il6+/+}$) except for a slight delay in Mac-1 upregulation at the initial time points (Fig. R4.1b, c).

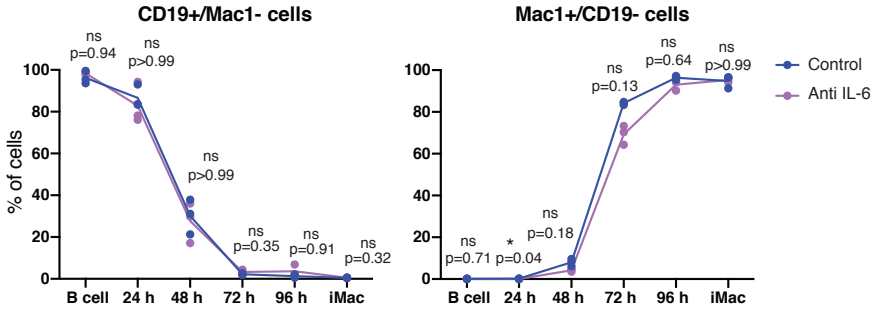


Fig. R4.2| Effect of IL-6 neutralization on transdifferentiation. Transdifferentiation kinetics of wild type B cells treated with 0.1 mg/mL antibody (BE0046) that blocks IL-6. The antibody was added at daily intervals to the culture medium. Experiments were performed with n=3 biologically replicates and statistical significance was determined applying one-tailed Student's t-test. No statistical differences were observed except for a slight impairment of Mac-1 upregulation by the antibody at 24 h post induction.

In another approach to test for a potential role of IL-6 signaling during transdifferentiation we added a blocking antibody against IL-6 into the culture medium during the whole transdifferentiation. This essentially did not change the kinetics of CD19 downregulation although it slightly impaired Mac-1 upregulation (Fig. R4.2).

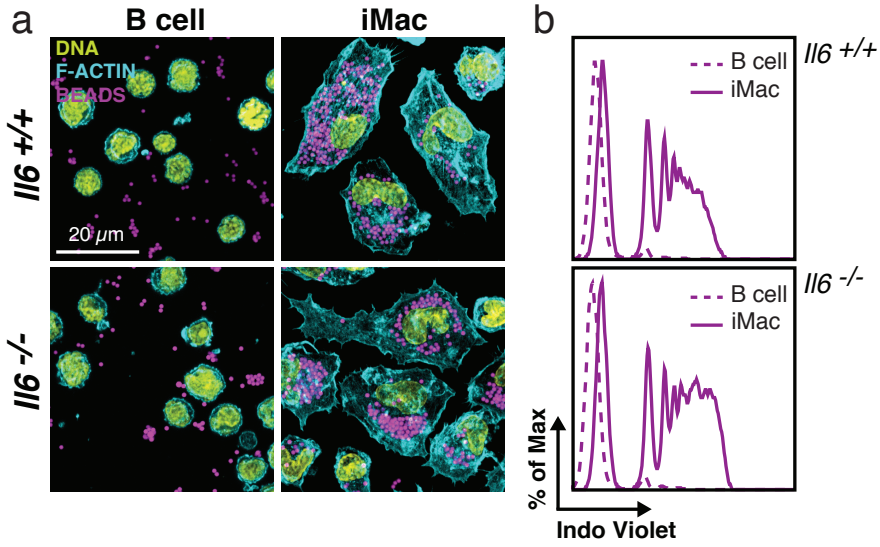


Fig. R4.3| Phagocytosis of induced wild type and *Il6*$^{-/-}$ B cells. a, Immunofluorescent images of B cells and 120 hours induced macrophages (iMacs) incubated overnight with purple fluorescent carboxylated beads. DNA was stained with picogreen (P7589) and F-ACTIN with phalloidin Alexa Fluor 568, shown in blue. **b**, FACS histograms depict fluorescent intensity of the internalized beads detected with indo violet laser in wild type and *Il6*$^{-/-}$ cells. Uninduced B cells are represented with a dashed line; iMacs with a continuous line.

We next tested whether the induced *Il6*$^{-/-}$ macrophages maintain functional properties of macrophages, by assessing their phagocytic capacity. For this purpose, wild type and *Il6*$^{-/-}$ iMacs (respectively grown on wild type and *Il6*$^{-/-}$ MEFs) were incubated with carboxylated microspheres and examined by fluorescence microscopy and FACS 24 hours later. As shown in Fig. R4.3a, both wild type and *Il6*$^{-/-}$ iMacs ingested the beads, while uninduced B cells did not exhibit phagocytic properties, as expected. FACS analyses confirmed these observations, showing that ~70% of both with wild type and *Il6*$^{-/-}$ iMacs ingested the beads, while uninduced B cells were negative and only showed a fluorescence negative peak (Fig. R4.3b). These results show that the C/EBPα-induced transdifferentiation into macrophages is essentially independent of IL-6 signaling, although it involves the upregulation of key genes within the *Il6* pathway.

Altogether, these experiments prove that the transdifferentiation of B cells into functional macrophages can be induced independently of *Il6*. Moreover, the efficiency and kinetics of this process was essentially unaffected by the impairment of IL-6 signaling.

5. IL-6 is strictly required for B cell to iPSC reprogramming and acts in a cell autonomous manner.

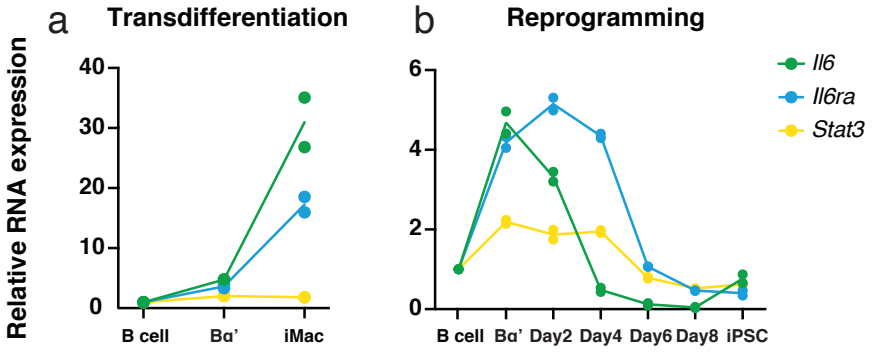


Fig. R5.1| RNA expression of *Il6*, *Il6ra* and *Stat3* during C/EBPα-induced transdifferentiation and C/EBPα-enhanced reprogramming of B cells. a, RNA expression during transdifferentiation at 18 hours (Bα' cells) and 120 hours post induction (iMac) relative to uninduced B cells. **b**, Relative RNA expression of Bα' cells and cells at different times after activation of the Yamanaka factors, including iPSCs as a control.

To examine the role of the *Il6* pathway in the C/EBPα-enhanced B cell to iPSC reprogramming, we first determined whether the *Il6* pathway genes are expressed during the process. As shown before, *Il6* and *Il6ra* become upregulated at 18 hours post induction with C/EBPα (Fig. R1), and further during transdifferentiation into iMacs, with *Stat3* remaining largely silent (Fig. R5.1). In contrast, during reprogramming *Il6* becomes downregulated in Bα' cells after OSKM induction while the expression of *Il6ra* and *Stat3* was maintained for another 2 days, dropping only by day 6.

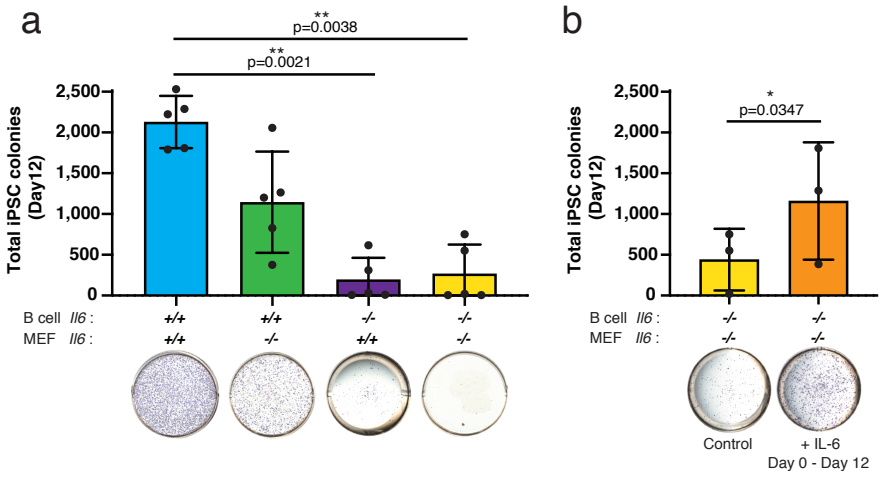


Fig. R5.2| *Il6* ablation impairs iPSC colony formation and can be partially rescued by addition of soluble IL-6. a, Effect of *Il6* genotype on reprogramming efficiency. All four possible combinations between wild type and *Il6*$^{-/-}$ B cells and MEFs were tested. Results show number of alkaline phosphatase (AP) positive iPSC colonies at day 12 of reprogramming obtained with n=5 biological replicates. Error bars indicate mean ± s.d. Statistical significance was determined using one-way ANOVA test with paired values and Dunnett's test for multiple comparisons relative to the control condition. Photographs of representative plates are shown on the bottom. **b**, Rescue experiment with IL-6 protein. Recombinant IL-6 was added at 20 ng/ml under daily intervals to the culture medium, control plates received none. Data show averages of n=3 biological replicates with bars representing mean ± s.d. Statistical significance was determined using one-tailed Student's t-test with paired values.

Next we determined whether IL-6 is required for reprogramming of B cells and if so whether the critical source of IL-6 is provided by the feeder activity of the MEFs. As before, we compared the reprogramming efficiency of cells tested in all possible combinations of wild type and *Il6*$^{-/-}$ B cells and MEFs. As shown in Fig. R5.2a, when *Il6* is absent in both B cells and MEFs, the number of alkaline phosphatase (AP+) colonies at day 12 decreased >10-fold. A similar

decrease was also observed in *Il6*$^{-/-}$ B cells grown on wild type MEFs. This difference was only 2-fold when the B cells were wild type and the MEF feeders were *Il6*$^{-/-}$. Together, these results show that IL-6 signaling is strictly required for C/EBPα-enhanced reprogramming as it has previously been reported for MEFs *in vitro*[133] and *in vivo*[134,136]. Our findings also suggested that IL-6 secreted by the B cells is much more important than the contribution of IL-6 from senescent MEFs.

We next attempted to rescue the reprogramming capacity of *Il6*$^{-/-}$ B cells grown on *Il6*$^{-/-}$ MEFs with recombinant IL-6. For this purpose we supplied 20 ng/mL of the factor on a daily basis starting from B cells 24 hours prior to the C/EBPα pulse and up to day 12 during C/EBPα enhanced reprogramming. The number of iPSC colonies increased significantly about 3-fold (Fig. R5.2b), not quite reaching the number of colonies obtained with wild type B cells grown on wild type MEFs (Fig. R5.2a), suggesting a partial rescue.

6. Early neutralization of secreted IL-6 reduces iPSC colony formation.

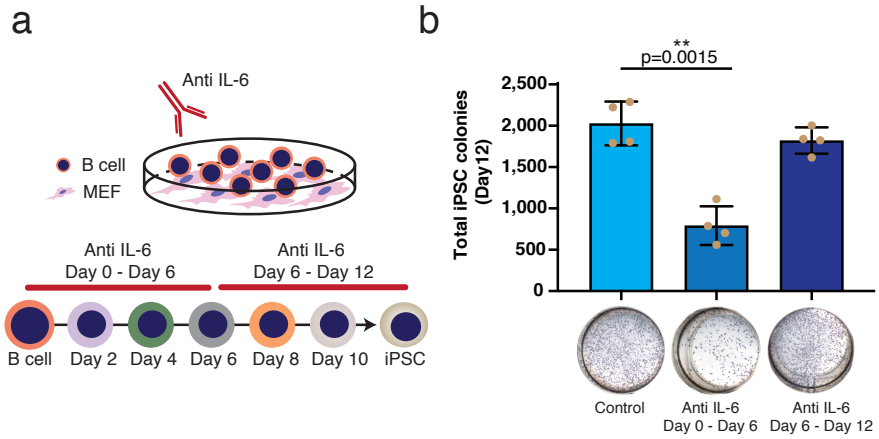


Fig. R6| Effect of blocking IL-6 on C/EBPα-enhanced reprogramming. **a**, Schematics of the experiment. Wild type B cells grown on a feeder layer of MEFs were induced with an 18 hours pulse of C/EBPα followed by OSKM activation. They were treated daily with 0.1 mg/mL of anti-IL-6 antibody (BE0046), either from day 0 to day 6 or from day 6 to day 12. **b**, Number of AP+ iPSC colonies scored at day 12. Representative images of the plates are shown on the bottom. The bars indicate mean ± s.d. of n=4 biological replicates. Statistical significance was determined using one-way ANOVA test with paired values and Dunnett's test for multiple comparisons relative to the control condition.

The observations described raised the question as to whether IL-6 is required in the early or late phases of C/EBPα-enhanced reprogramming, and if it acts non-redundantly from LIF (LIF is added to the medium after C/EBPα pulse onwards), which signals through the shared GP130 co-receptor encoded by *Il6st*. To test the IL-6 requirement during the early phase, we added IL-6 neutralizing antibody to the cultures from 24 hours before C/EBPα pulse to day 6 of

reprogramming. To test a requirement during the late period, blocking antibody was applied between day 6 and 12 (after the transient upregulation of *Il6* pathway genes, Fig. R5.1b), when the iPSC colonies were scored. As a control, the reprogramming experiment was performed without antibody addition. We observed that blocking of IL-6 significantly reduced the number of iPSC colonies (about 2.5-fold) when cells were treated early (Fig. R6b), while treatment during the late phase only caused an insignificant reduction in colony numbers.

Together, these results support the idea that IL-6 signaling results in functions required for B cell to iPSC reprogramming that are non-redundant with LIF. Additionally, they suggest that IL-6 secreted by C/EBPα-induced B cells acts in a cell autonomous manner not depending on feeder MEFs.

7. Reprogramming of *Il6*$^{-/-}$ B cells alters expression of myeloid, pluripotency, trophectoderm, proliferative and senescence programs.

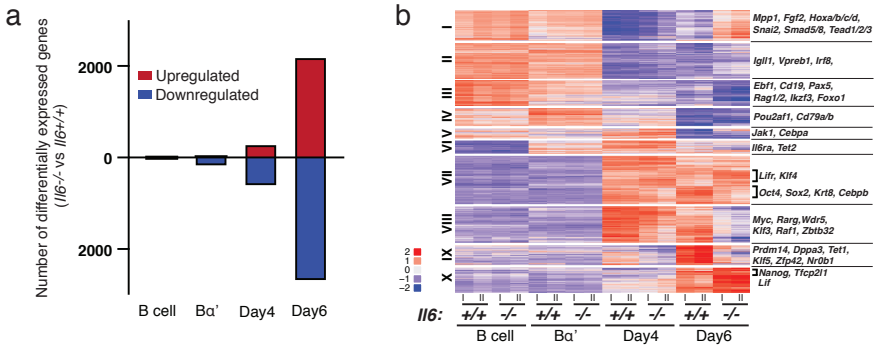


Fig. R7.1| Global gene expression changes between wild type and *Il6*$^{-/-}$ B cells during iPSC reprogramming. a, Total number of differentially expressed genes (p<0.01) at each time point in *Il6*$^{-/-}$ B cells compared to wild type B cells during reprogramming. **b**, Unsupervised K-means clustering of differentially expressed genes (p<0.01) between wild type and *Il6*$^{-/-}$ cells at different time points during reprogramming.

Given that the main impact of the IL-6 signaling occurred during early time points in our two-step B cell reprogramming system (Fig. R6b), we next decided to analyze the impact of *Il6* on gene expression changes upon C/EBPα pulse and OSKM activation. For this aim, RNA was collected from wild type and *Il6*$^{-/-}$ B cells grown on wild type MEF feeders and sequentially treated with β-estradiol and doxycycline. We observed that the lack of *Il6* in B cells led to the downregulation of ~600 genes at day 4 and ~2,500 at day 6 and somewhat upregulated genes at these two time points (Fig. R7.1a).

To further analyze the gene expression data, we performed an unsupervised K-means clustering analysis and divided genes that significantly changed ($p<0.01$) into 10 groups (Fig. R7.1b). In general, we observed that gene expression dynamics were similar between wild type and *Il6*$^{-/-}$ reprogramming up to day 4. However, some clusters showed major differences in gene expression at day 6. This included the B cell identity genes *Ebf1* and *Pax5*, as well as *Cd19*, which became downregulated more strongly in *Il6*$^{-/-}$ cells (cluster III). Some transiently upregulated genes, such as *Il6ra* and *Tet2* showed comparable kinetics under the two conditions (cluster VI). Clusters VIII and IX included pluripotency associated factors that became less upregulated in *Il6*$^{-/-}$ cells than in wild type cells, such as *Myc*, *Prdm14*, *Dppa3*, *Tet1*, *Klf5*, *Zfp42* and *Nr0b1*. The key pluripotency factors *Oct4* and *Sox2* as well as the myeloid factor *Cebpb* and the trophectoderm (TE) marker *Krt8* were more strongly upregulated in the wild type background (cluster VII) as for pluripotency genes *Nanog* and *Tfcp2l1* in cluster X. Curiously, other pluripotency factors including *Lifr*, *Klf4* (cluster VII) and *Lif* (cluster X) became more expressed in the *Il6*$^{-/-}$ mutant background than in wild type B cells. Finally, cluster I included a group of genes that became downregulated by day 4 but were re-expressed by day 6 only in *Il6*$^{-/-}$ cells. This cluster contained genes encoding for regulators of diverse signaling pathways such as *Fgf2* (FGF signaling); *Snai2* (Wnt

pathway); *Tead1*, *Tead2* and *Tead3* (Hippo pathway); *Smad5* (TGF-β signaling) as well as homeobox genes.

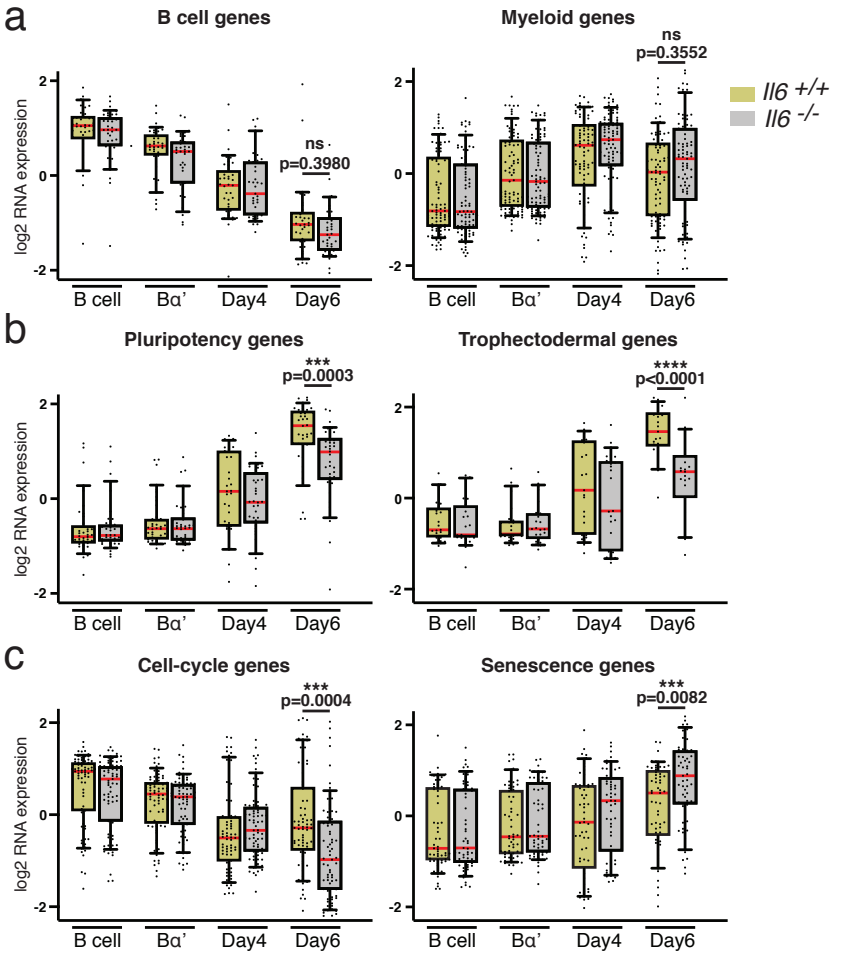


Fig. R7.2| Expression dynamics of specific gene signatures in reprogramming of wild type and *Il6*$^{-/-}$ B cells. a, RNA expression levels of selected B cell (n=37) and myeloid cell-specific genes (n=89). **b**, Expression of selected pluripotency (n=37) and trophectodermal (n=23) specific genes. **c**, Expression of selected senescence (n=59) and cell-cycle specific genes (n=73). Dots represent the average value of n=2 RNA-seq biological replicates for each gene. Boxplots and whiskers depict 10-90 percentiles with the median value represented by a red line. Statistical significance was determined using two-way ANOVA and Sidak's multiple comparison tests. See Table MM9 for gene lists.

To analyze our gene expression data for changes associated with cell identity, cell growth and senescence we interrogated the expression of signature genes (at least 23) for each functional category. We observed no significant differences in the erasure of B cell genes nor in the transient activation of myeloid genes between wild type and *Il6*$^{-/-}$ conditions (Fig. R7.2a). In contrast, the expression of pluripotency factor genes was reduced in *Il6*$^{-/-}$ cells by day 6 (Fig. R7.2b), as shown for the pluripotency genes highlighted in clusters VII and IX in Fig. R7.1b. Interestingly, *Il6*$^{-/-}$ cells also failed to upregulate trophectoderm restricted genes at day 6 of reprogramming (Fig. R7.2b). A possible explanation is that IL-6 is required for bipotent intermediates towards pluripotency or trophectodermal fates during reprogramming. In line with the known cell cycle inhibitory activity of C/EBPα and the pro-proliferative effect of the Yamanaka factors, the expression of cell cycle genes decreased after the pulse of C/EBPα and resumed by day 6 of reprogramming (Fig. R7.2c). Notably, the increased expression of cell cycle genes at day 6 was significantly reduced in *Il6*$^{-/-}$, suggesting that IL-6 is not only required for the formation of the postulated bipotent intermediate but also for its proliferation. Complementarily, the opposite effect was observed for genes associated with senescence: the increase observed in day 4 and day 6 among wild type cells was further exacerbated in *Il6*$^{-/-}$ cells (Fig. R7.2c).

In conclusion, our data showed that *Il6* dependent and independent changes in gene expression occur during C/EBPα-mediated B cell to iPSC reprogramming. Unexpectedly, the response of pluripotency genes was heterogenous, with most genes requiring *Il6* for their full upregulation, while for others the factor seems repressive. Additionally, intact *Il6* has an impact on the upregulation of trophectodermal genes during reprogramming. Finally, our data also suggest that the activation of the *Il6* pathway during reprogramming involves the upregulation of a cell proliferation-associated program with a concomitant decrease of cellular senescence.

8. The *Lifr* gene becomes upregulated late during reprogramming and it is partially repressed by the *Il6* pathway.

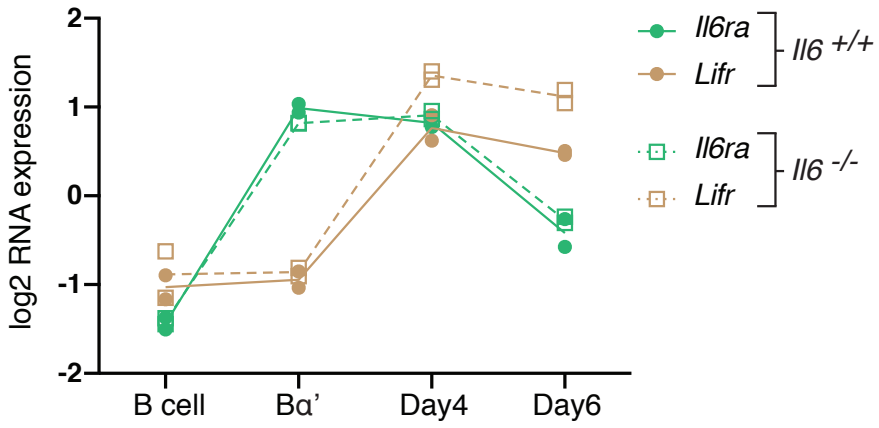


Fig. R8| *Il6ra* and *Lifr* expression during reprogramming of wild type and *Il6*$^{-/-}$ B cells. Data show RNA expression levels (RNA-seq) of *Il6ra* and *Lifr* from biological duplicates for each condition and time point.

As mentioned earlier, binding of LIF to its specific membrane receptor (LIFR) can also trigger the activation of the *Il6* pathway through the phosphorylation of STAT3 and signaling by the co-receptor GP130 (encoded by *Il6st*)[110,111]. The fact that in our experiments we add recombinant LIF in the medium raises again a possible redundancy or competition between the IL-6 and LIF signaling pathways. To probe into the relationship between the two pathways, we compared the expression dynamics of the *Lifr* and *Il6ra* genes. Whereas *Il6ra* becomes expressed as early at the Bα' stage and it is maintained up to the day 6 cell stage, *Lifr* only becomes upregulated at day 4. Interestingly, while the kinetics of *Il6ra* regulation were indistinguishable between wild type and *Il6*$^{-/-}$ cells, the upregulation of *Lifr* was higher in cells lacking *Il6*. This shows that *Il6ra* becomes upregulated before *Lifr* during reprogramming and that IL-6 signaling partially impairs *Lifr* gene expression, thus acting non-redundantly.

9. C/EBPα overexpression in B cells splits them into subsets predominantly expressing either *Il6* or *Il6ra*.

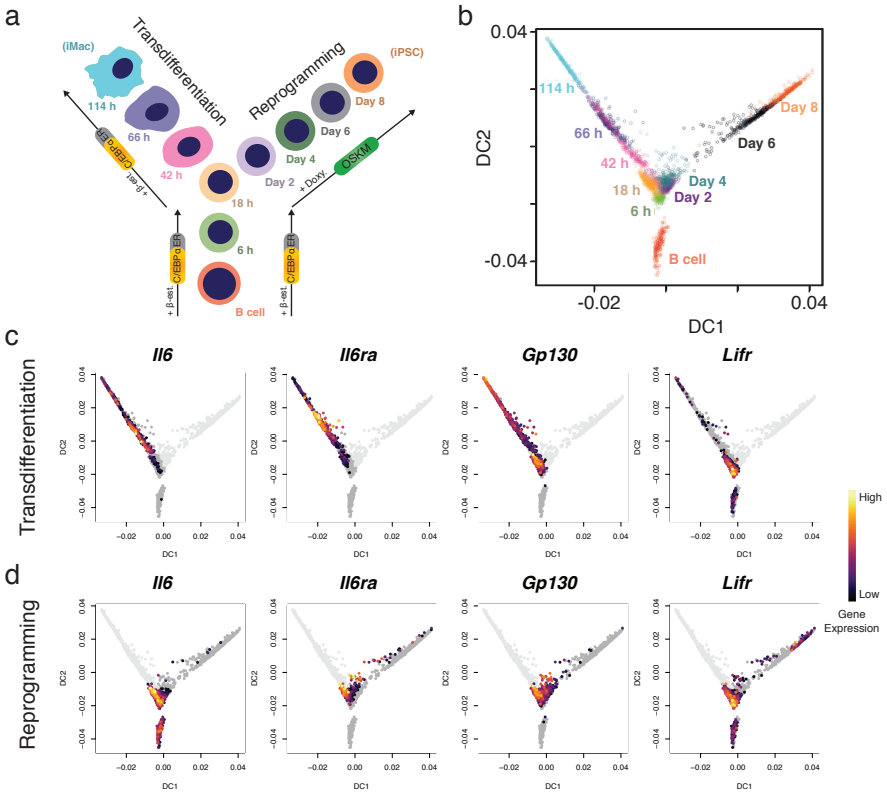


Fig. R9.1| Expression of *Il6* pathway genes at the single cell level during transdifferentiation and reprogramming. a, Overview of the experimental procedure showing the different time-points analyzed. **b**, Single cell projections onto the first two diffusion components (DC1 and DC2). **c**, As in **b** with top 50% of cells expressing selected *Il6* pathway markers during transdifferentiation and reprogramming. Data from Francesconi et al. (2019).

The findings that the main effect of the IL-6 signaling in reprogramming is cell autonomous (Fig. R5.2a) and that the emerging cell populations are heterogeneous[45] raised the question of how the *Il6* pathway signaling is coordinated

among different subpopulations. In particular, we wondered whether the *Il6* pathway operates in a paracrine fashion with some induced B cells expressing *Il6* and others the receptor (*Il6ra*), or in an autocrine mode with cells expressing both members of the pathway. To further investigate whether IL-6 or LIF drive the dominant signaling, we also interrogated the relationship between *Il6ra* and *Lifr* to the common *Gp130* transducer to see which of the specific receptors was preferentially co-expressed with the common chain.

To study this we re-analyzed a single cell expression dataset during B cell transdifferentiation and reprogramming obtained in the laboratory[45]. This consisted in RNA-seq data obtained from B cells either continuously treated with β-estradiol to activate C/EBPα for 0, 6, 18, 42, 66 and 120 hours resulting in induced macrophages (iMacs); or B cells treated for 18 hours followed by doxycycline addition to induce OSKM for 2, 4, 6 and 8 days, resulting in iPSCs (see scheme in Fig. R9.1a, b). During transdifferentiation into macrophages both *Il6* and *Il6ra* became strongly upregulated, reaching highest levels at 42 to 66 hours, while *Il6st* and *Lifr* were most highly expressed at the 18 hours stage (Fig. R9.1c). This confirmed, as expected, that B cells express low levels of *Il6* and essentially no *Il6ra*. During cell reprogramming there was also an early and transient upregulation of *Il6* coinciding with the pulse of C/EBPα (Fig. 9.1d) reflecting what we had observed in the bulk analysis (Fig. R5.1b).

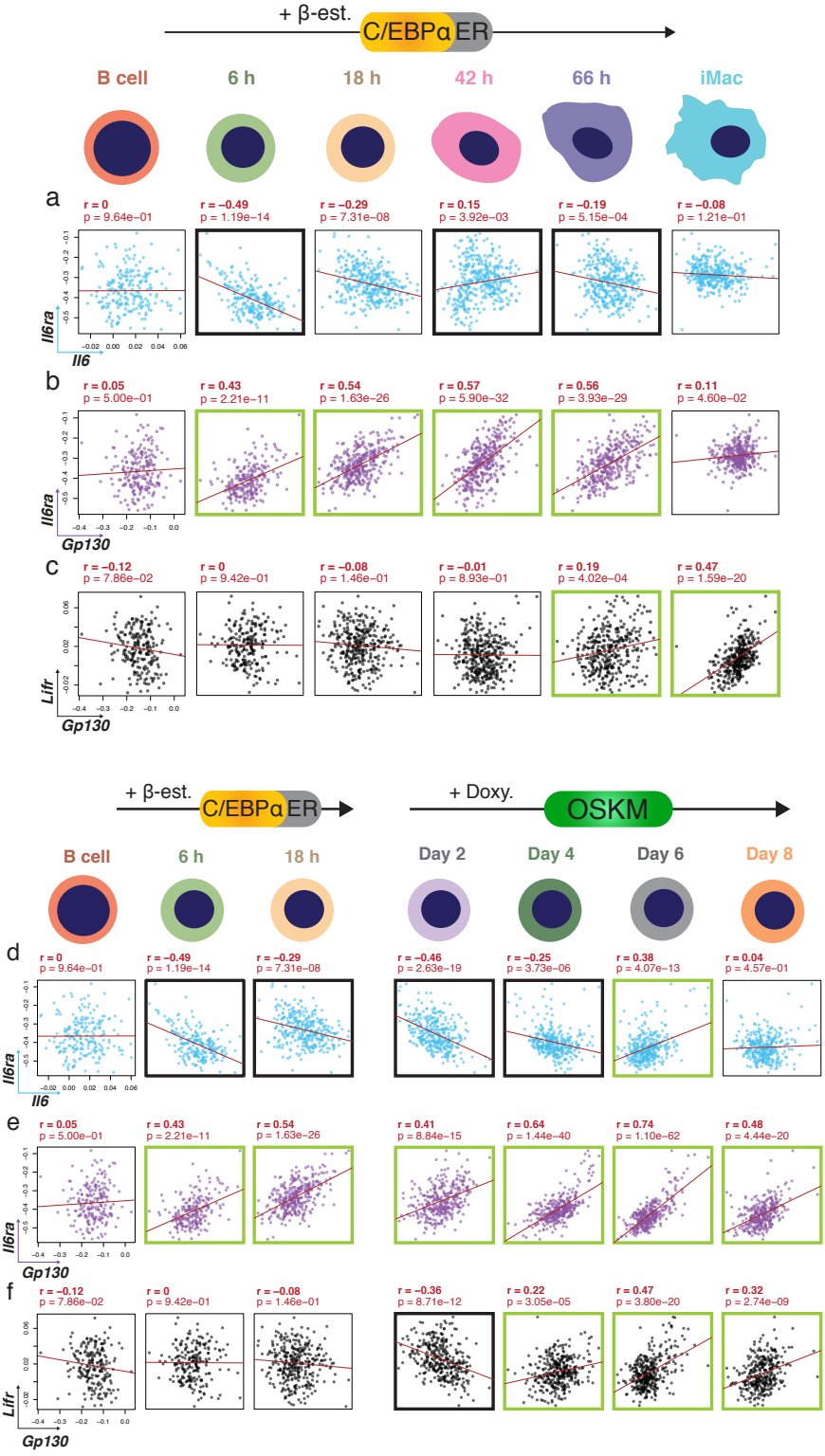

+ β-est.
C/EBPα ER
B cell
6 h
18 h
42 h
66 h
iMac
a
r = 0
p = 9.64e−01
r = −0.49
p = 1.19e−14
r = −0.29
p = 7.31e−08
r = 0.15
p = 3.92e−03
r = −0.19
p = 5.15e−04
r = −0.08
p = 1.21e−01
Il6ra
Il6
b
r = 0.05
p = 5.00e−01
r = 0.43
p = 2.21e−11
r = 0.54
p = 1.63e−26
r = 0.57
p = 5.90e−32
r = 0.56
p = 3.93e−29
r = 0.11
p = 4.60e−02
Il6ra
Gp130
c
r = −0.12
p = 7.86e−02
r = 0
p = 9.42e−01
r = −0.08
p = 1.46e−01
r = −0.01
p = 8.93e−01
r = 0.19
p = 4.02e−04
r = 0.47
p = 1.59e−20
Lifr
Gp130
+ β-est.
C/EBPα ER
+ Doxy.
OSKM
B cell
6 h
18 h
Day 2
Day 4
Day 6
Day 8
d
r = 0
p = 9.64e−01
r = −0.49
p = 1.19e−14
r = −0.29
p = 7.31e−08
r = −0.46
p = 2.63e−19
r = −0.25
p = 3.73e−06
r = 0.38
p = 4.07e−13
r = 0.04
p = 4.57e−01
Il6ra
Il6
e
r = 0.05
p = 5.00e−01
r = 0.43
p = 2.21e−11
r = 0.54
p = 1.63e−26
r = 0.41
p = 8.84e−15
r = 0.64
p = 1.44e−40
r = 0.74
p = 1.10e−62
r = 0.48
p = 4.44e−20
Il6ra
Gp130
f
r = −0.12
p = 7.86e−02
r = 0
p = 9.42e−01
r = −0.08
p = 1.46e−01
r = −0.36
p = 8.71e−12
r = 0.22
p = 3.05e−05
r = 0.47
p = 3.80e−20
r = 0.32
p = 2.74e−09
Lifr
Gp130

Fig. R9.2| Single cell expression correlations of *Il6* pathway genes and *Lifr* during reprogramming and transdifferentiation. a-c, Correlation plots between *Il6*-*Il6ra*, *Il6ra*-*Gp130* and *Lifr*-*Gp130* gene expression pairs in single cells at various time points during transdifferentiation. **d-f**, Same analyses as in **a-c** for expression pairs of cells during reprogramming. Pearson's correlation coefficients (r), p values and tendency lines are depicted in red. Thickened squares around the plots show significant p values ($<5x10^{-3}$) with black and green lines corresponding to negative and positive correlations, respectively. Data from Francesconi et al. (2019).

Using this data set we plotted the expression of cells in pairwise combinations at different time points, comparing *Il6* with *Il6ra*, *Gp130* with *Il6ra* and *Gp130* with *Lifr*. The plots permitted to calculate the Pearson's correlation coefficient, with r = 0 showing no correlation, r > 0 value showing a positive and r < 0 value a negative correlation. Correlations with a $p < 10^{-3}$ value were considered to be significant and were highlighted by a green box when positive, and when it is negative with a black box.

We observed that after 6 and 18 hours of C/EBPα induction, B cells showed a negative correlation between *Il6* and *Il6ra* expression, which leveled off at later stages of transdifferentiation (Fig. R9.2a). In contrast, *Gp130* correlated positively with *Il6ra* from 6 to 66 hours after induction (Fig. R9.2b), and a positive correlation was also found between *Gp130* and *Lifr*, but only at the latest two time points (Fig. R9.2c). These findings indicate that during transdifferentiation initial *Il6* and *Il6ra* anti-correlation is lost in later time-points. Both *Il6ra* and *Lifr* correlated positively with the expression of

the shared co-receptor *Gp130*, with the latter *Lifr-Gp130* combination showing a delay.

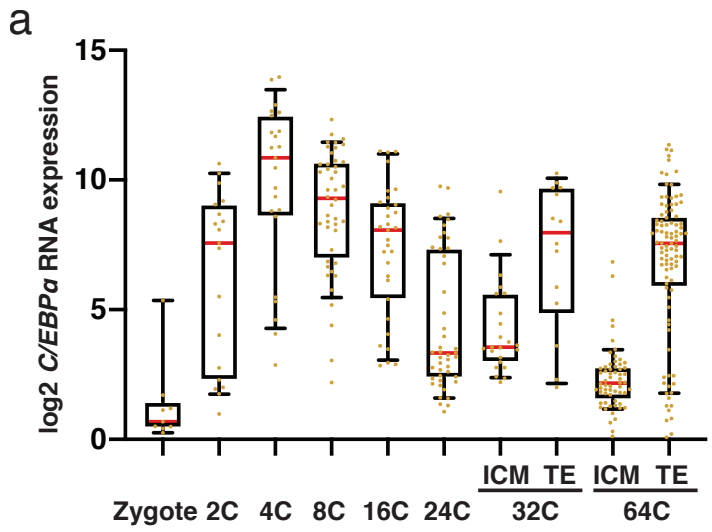
During induced reprogramming, day 2 and day 4 cells maintained the negative correlation between *Il6* and *Il6ra*, but then switched to a positive value by day 6 that was resolved at day 8 (Fig. R9.2d). Both *Il6ra* and *Lifr* correlated positively with *Gp130* between day 2 and day 8, with the *Lifr-Gp130* combination showing a delay and only starting to positively co-relate at day 4 (Fig. R9.2e, f).

Together, these findings suggest that early during reprogramming two types of cells are formed: one subset that produces IL-6 and another that expresses the receptor (*Il6ra-Gp130* positive cells) and becomes activated by paracrine IL-6. At later stages when *Lifr* and *Gp130* become co-expressed, LIF mediated signaling takes over.

CHAPTER 2: Role of C/EBPα and IL-6 during trophectoderm specification.

10. C/EBPα is highly expressed in 4- and 8-cell embryos and in the trophectoderm.

The observed interplay between C/EBPα and the *Il6* pathway described in chapter 1 raised the possibility that it reflects a mechanism operating in early embryo development. Thus, in our experimental system i) a pulse of C/EBPα poises B cells for OSKM-induced reprogramming into iPSCs; ii) C/EBPα-enhanced reprogramming strictly requires the *Il6* pathway; iii) the cells functionally secrete IL-6 during this process and iv) the *Il6* and *Il6ra* genes become upregulated in distinct cell sub-populations. Therefore we wondered whether C/EBPα is expressed in pre-implantation embryos, previous to the activation of key pluripotency factors.



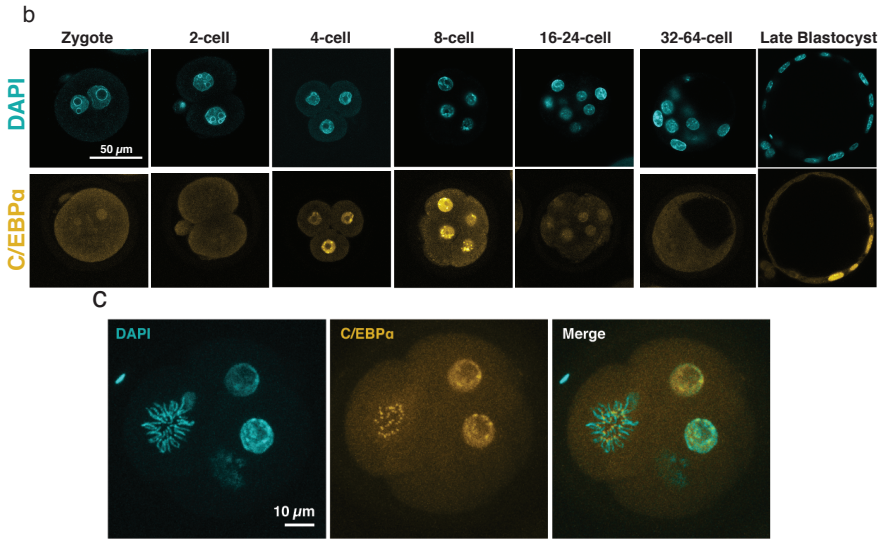


Fig. R10| C/EBPα expression during pre-implantation mouse development. a, Expression of C/EBPα assessed by RT-qPCR in single blastomeres normalized to *Gapdh* (data from Guo et al. (2010)). At the 32-cell stage blastomeres are grouped into TE or ICM committed fates according to their location in the outer or inner space of the embryo. At the 64-cell stage blastomeres are grouped into TE or ICM according to their characteristic gene expression signatures. Boxplots and whiskers depict 10-90 percentiles with the red line indicating median values. **b**, Immunofluorescence images (single z-planes) of C/EBPα in embryos at different developmental stages. DNA was stained with DAPI (upper panel) and C/EBPα with a primary rabbit monoclonal antibody (8178, Cell Signaling) followed by secondary Alexa Fluor 488 (A-11070, ThermoFisher) (lower panel). **c**, Same as above but representing 3D projection of C/EBPα immunostaining of a 4-cell embryo with a dividing blastomere exhibiting a metaphase.

We first searched the literature for gene expression data in single cells from early embryos and found a high quality dataset suitable for our purpose[200]. This revealed two waves of RNA upregulation (Fig. R10a). The first activation of *Cebpa* expression was detected at the 2-cell stage, peaking at the 4- to 8-cell stages. *Cebpa* RNA levels then decreased as the embryos divided to enter the morula stage. From this point

onwards, a second wave of *Cebpa* expression was exclusively detected in a subset of cells within the trophectodermal layer (the future placenta) in the blastocyst. To determine whether the expression of *Cebpa* in embryos can also be visualized at the protein level we set up an *in vitro* system to culture embryos up to the late blastocyst stage, starting with F1 zygotes retrieved from B6CBAF1/Crl crosses using super-ovulated mice. Zygotes were then harvested, placed in culture and embryos sampled at different stages were fixed and stained with a polyclonal antibody against C/EBPα. Confocal microscopy of these samples closely paralleled the RNA expression data: C/EBPα was first detected at the late 2-cell stage, further increased in some cells at the 4-cell stage (Fig. R10b) and became highly expressed in all blastomeres at the 8-cell stage. C/EBPα levels then dropped during morula development and in late blastocysts it was re-expressed in a subset of trophectodermal cells but not in the ICM (Fig. R10b). Immunofluorescence of C/EBPα also revealed its presence in the pericentromeric regions of mitotic chromosomes (Fig. R10c).

11. IL-6 is expressed in the trophectoderm and its blockage delays the morula to blastocyst transition.

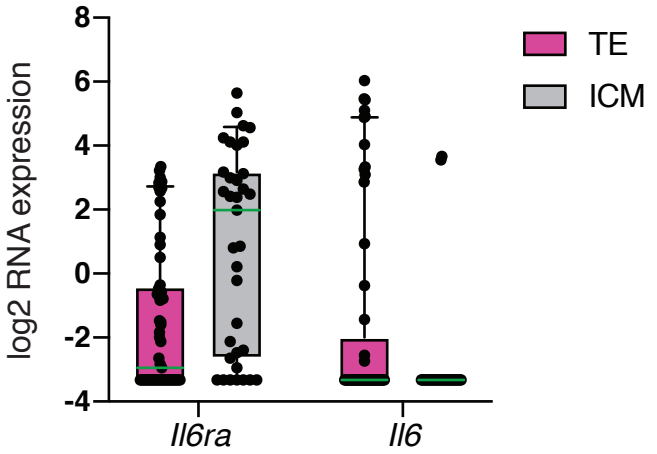


Fig. R11.1| *Il6* and *Il6ra* RNA expression in the TE and the ICM of late mouse blastocysts. Individual dots represent the expression of single cells determined by RNA-seq derived from late blastocysts with n=35 cells for the ICM and n=57 cells for the TE. All boxplots and whiskers depict 10-90 percentiles with the green line indicating median values. Figure represents an analysis of data from Wyatt et al. (unpublished, see materials and methods).

The selective expression of C/EBPα in the trophectoderm (or a subset of cells in this layer) raised the question as to whether it directly regulates TE genes. To begin testing this possibility we analyzed single cell data in late blastocysts[228] where cells had been clustered into distinct TE and ICM groups according to their specific gene expression signatures[200]. As shown in Fig. R11.1, ICM cells express high levels of *Il6ra* and essentially no *Il6,* whereas single TE cells variably express *Il6* and almost no *Il6ra*. The expression of *Il6* pathway genes in blastocysts is in line with experiments

reporting that antibody-mediated blocking of IL-6 in embryos leads to a decrease in the number of cells exhibiting nuclear STAT3-P^{138}.

Based on these observations we postulated that C/EBPα activates the expression of *Il6* pathway genes in blastocysts with TE cells secreting IL-6, which in turn activates the downstream effectors of the pathway in ICM cells expressing *Il6ra*. In other words, extraembryonic cells might act as a niche for the formation of pluripotent cells in the ICM through the secretion of IL-6.

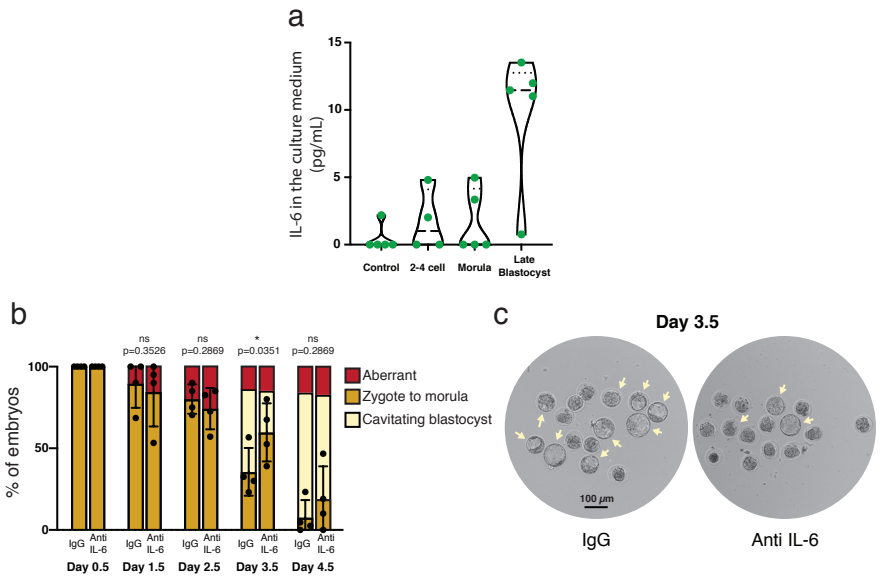


Fig. R11.2| IL-6 secretion by cultured mouse embryos and effect of IL-6 neutralization on their development. a, Detection of IL-6 protein by an ELISA assay in the supernatant of 30 µL cultures each containing 25 embryos. Values represent results obtained from 4-5 experiments. Control represents culture medium. **b**, Effect of IL-6 blocking antibody on development of embryos grown up to 4.5 days. A total of n=245 embryos were treated with 0.1 mg/mL of anti-IL-6 (BE0046) antibody and n=208

embryos with anti-horseradish peroxidase antibody (BE0088, IgG) as a control in 4 independent experiments. The transition from morula to blastocyst was scored by the formation of a blastocoel cavity. Statistical significance was determined using multiple paired Student's t-test applying Holm-Šídák's correction method. **c**, Representative phase contrast images of embryos cultured for 3.5 days with either control (IgG) or anti-IL-6 antibody. Yellow arrowheads pinpoint cavitated blastocysts.

To explore this idea, we first analyzed whether different stages of pre-implantation mouse secrete IL-6 into the medium. The presence of the cytokine was assayed by ELISA in the supernatant of embryo cultures. As shown in Fig. R11.2a, secretion of IL-6 was detected in embryos at day 4.5 of culture, corresponding to the late blastocyst stage, while morulas and 2- to 4-cell embryos were essentially negative. To determine whether IL-6 signaling has an effect on early embryo development we added to the medium blocking antibody against IL-6 from the zygote to the late blastocyst stage. We could indeed detect a significant delay of cavitation during the transition from morula to blastocysts at day 3.5 in the embryos treated with anti-IL-6 compared to the controls (Fig. R11.2b, c). Before morula formation, no differences were observed in the development of the various embryo stages in both anti-IL-6 and control conditions. Despite the marked delay in cavitation of embryos treated with anti-IL-6 antibodies at day 3.5, similar numbers of embryos reached the late blastocyst stage under both conditions.

In conclusion, our data have shown that C/EBPα is expressed in pre-implantation embryos both early (4- to 8-cell stages) and late (trophectoderm), that both *Il6* and *Il6ra* are expressed in blastocysts, that blastocysts secrete IL-6 and that the cytokine is required for the efficient transition from morulas to blastocysts.

12. The trophectoderm-fated vegetal blastomere of 4- cell embryos expresses the highest levels of C/EBPα.

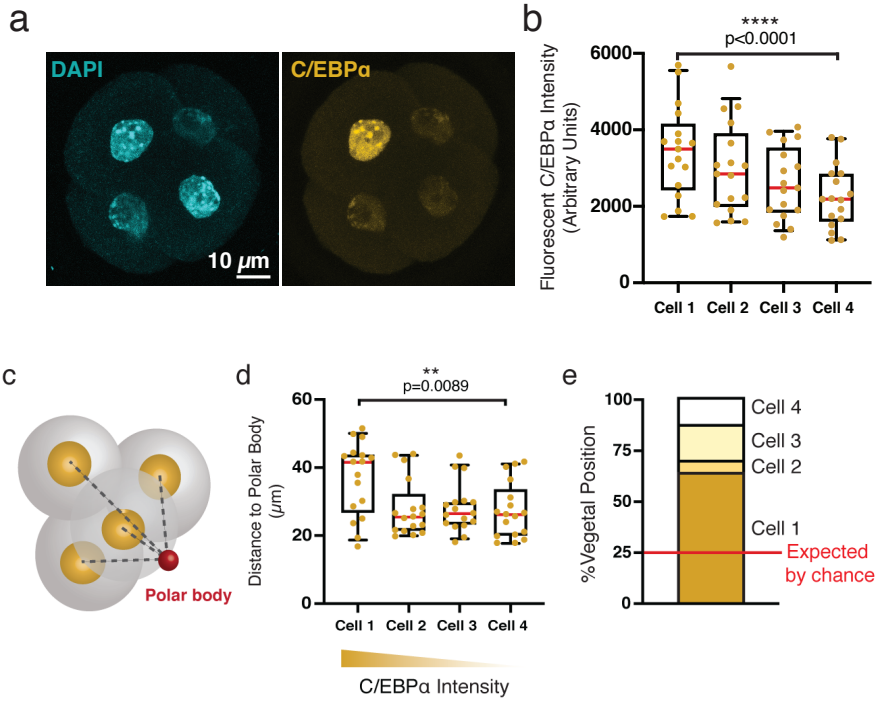


Fig. R12| Enrichment of C/EBPα protein in the vegetal blastomere at the 4-cell stage. a, Representative 3D projection of C/EBPα expression at the 4-cell stage as detected by confocal immunofluorescence (same antibodies as in Fig. R10). **b**, Quantification of nuclear C/EBPα intensity normalized by the cytoplasmic background for each single blastomere (cells 1 to 4) in n=17 4-cell embryos. **c**, Schematic 3D model of a 4-cell

embryo showing distances from nuclei centers to the polar body. The vegetal blastomere is the furthest from the polar body. **d**, Measurements of such distances in the same embryos as shown in **b**. Blastomeres showing the same relative C/EBPα intensity within each embryo are grouped into 4 categories (cells 1 to 4). All boxplots and whiskers depict 10-90 percentiles with a red band inside the box showing the median value. Statistical significance was determined using one-way ANOVA test with paired values. **e**, Percentage of blastomeres that hold the vegetal position at the 4-cell stage according to their relative C/EBPα intensity with 25% being expected by chance for each category.

The question whether localized determinants in the earliest stages of development (4- to 8-cells) establish a pre-pattern for the later segregation of the embryo into the ICM and TE, or whether cells are patterned at the morula/blastocyst stage has long been controversial[229]. However, recent experiments have compellingly shown that asymmetries in the expression of transcription factors and enzyme modifiers at the 4-cell stage predict commitment towards either the trophectoderm or pluripotent cell lineage in the mouse blastocyst[171]. More specifically, it has been unambiguously shown that the vegetal blastomere in the 4-cell embryo (the one which is located furthest from the polar body) preferentially gives rise to the trophectoderm[178]. Since we observed that the nuclear expression of C/EBPα is asymmetrically distributed in 4-cell embryos (Fig. R12a, b), we wondered whether its expression level predicts commitment towards one of the early blastocyst lineages. We therefore measured the distances from the center of the nuclei to the polar body in each blastomere of 17 4-cell embryos tested (Fig. R12c). We also quantified the cell's C/EBPα expression levels, revealing a significant correlation between C/EBPα expression levels and their

distance to the polar body (Fig. R12d). Moreover, in 11 out of the 17 embryos the blastomere with the highest C/EBPα expression occupied the vegetal position (Fig. R12e).

In conclusion, we observed an asymmetric expression of C/EBPα in 4-cell embryos, showing highest levels in blastomeres corresponding to the vegetal pole. As this has been shown to predict commitment to the TE lineage, our finding raises the possibility that C/EBPα has a TE lineage-instructive capacity in pre-implantation embryos.

13. *Dux* is a potential regulator of the initial wave of *Cebpa* expression in early embryos.

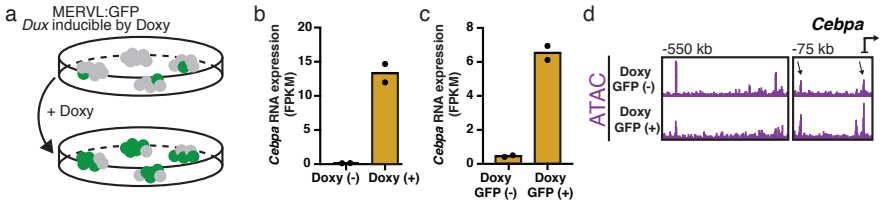


Fig. R13| *Cebpa* expression upon *Dux* induction in ESCs. a, Schematics of a reporter system where doxycycline (Doxy) mediated activation of an exogenous *Dux* gene in ESCs induces expression of an MERVL-GFP reporter and a 2-cell-like state. **b**, Changes in RNA expression of *Cebpa* after Doxy induction of *Dux* for 24 hours. **c**, Differences in RNA expression of *Cebpa* between sorted GFP (+) and GFP (-) cells from *Dux*-induced ESCs. **d**, Browser snapshots of ATAC-sequencing peaks in sorted GFP (-) and GFP (+) ESCs 24h after *Dux* induction, showing an increase in *Cebpa* promoter and at -75 kb after *Dux* activation. Analyses were performed using available datasets from Hendrickson et al. (2017).

A highly investigated topic within the early embryo development field is mechanisms by which the zygotic genes

become activated. Recently it has been reported that maternal DUX protein is involved in the first wave of zygotic genome activation (ZGA) at the 2-cell stage[154–156]. Cultured ESCs contain <1% of cells transcriptionally similar to 2-cell embryos (2-cell-like cells), including the expression of an endogenous retrovirus related gene designated MERV-L[163]. Using MERV-L-GFP as a reporter for the 2-cell state has permitted to identify and easily sort these cells from ESC cultures. A driver of 2-cell-like cell formation is the transcription factor *Dux*, whose overexpression in ESCs transiently increases the proportion of 2-cell-like cells up to almost 80%[156]. Together with our finding that *Cebpa* is first upregulated in 2-cell embryos (Fig. R10a) we wondered whether there is a connection with *Dux* and *Cebpa* activation.

We therefore analyzed published data where *Dux* was induced by doxycycline, followed by an RNA-seq analysis[156] (Fig. R13a). We found that *Cebpa* levels dramatically increased in clonal ESCs 24 hours after *Dux* induction (Fig. R13b). In addition, *Cebpa* levels were much higher in sorted MERV-L-GFP positive 2-cell-like cells than in GFP (-) cells (Fig. R13c). We next studied published data on chromatin accessibility changes after *Dux* induction. This revealed that the promoter of *Cebpa*, as well as an upstream region at -75 kb, became more nuclease accessible in the *Dux*-induced ESC population (Fig. R13d). The figure also shows a -550 kb

site as a negative control where chromatin becomes less accessible upon *Dux* expression.

In short, our preliminary experiments suggest that the observed upregulation of *Cebpa* in late 2-cell embryos is mediated by activation of the DUX protein, which might be directly sitting on *Cebpa* regulatory sites.

14. Analysis of cultured *Cebpa*$^{-/-}$ embryos identifies candidate genes regulated by the factor.

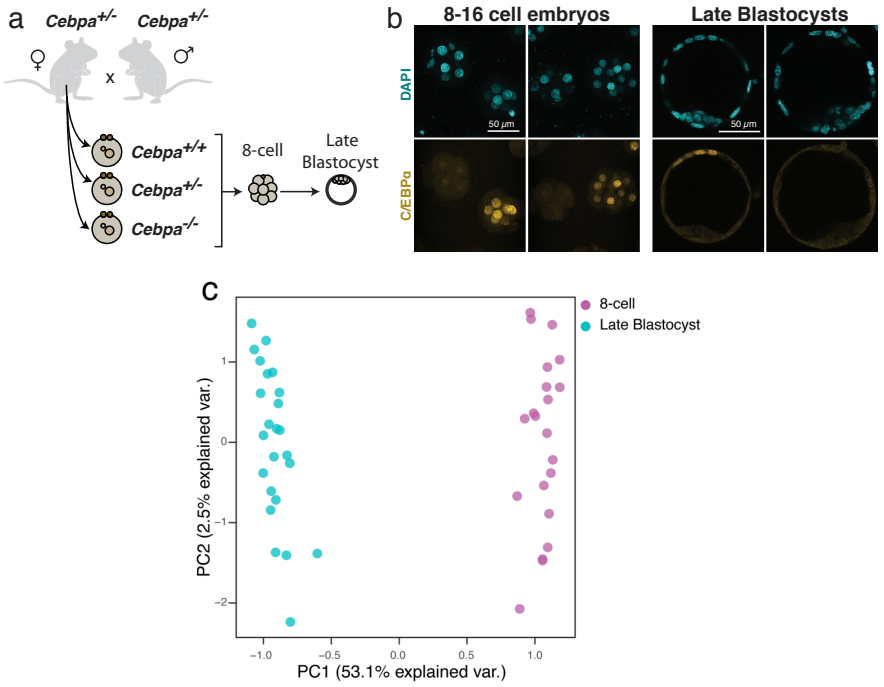


Fig. R14.1| Analysis of embryos from heterozygous *Cebpa* crosses by RNA-sequencing and immunostaining. a, Experimental strategy to study the impact of *Cebpa* ablation on early development. **b**, C/EBPα immunostaining of cultured 8-cell embryos and late blastocysts showing one C/EBPα positive and on one negative embryo at each stage. **c**, PCA

of RNA expression profiles of individual 8-cell embryos and late blastocysts obtained from F1 crosses of heterozygous C/EBPα mice.

In an attempt to identify genes regulated by C/EBPα in early embryos, we established in the laboratory a heterozygous *Cebpa* knock-out mouse line[67]. *Cebpa*$^{-/-}$ mice develop normally through gestation, exhibiting the expected mendelian ratios after an F1 cross of heterozygous animals. After birth, they die perinatally due to defects in the liver and the hematopoietic system.

To determine whether under the more stringent *in vitro* culture conditions the knock-out embryos exhibit a phenotype, we crossed heterozygous *Cebpa* mice to collect *Cebpa*$^{-/-}$, *Cebpa*$^{+/-}$ and *Cebpa*$^{+/+}$ zygotes that were placed in culture (Fig. R14.1a). We harvested a total of 25 8-cell embryos and 25 late blastocysts for RNA-seq analysis. These two stages of development were chosen as they represent the two waves of C/EBPα expression (Fig. R10). Immunostaining experiments performed in parallel proved that some 8-cell and late blastocyst embryos derived from these crosses were C/EBPα negative, also confirming that embryos without C/EBPα can reach the late blastocyst stage *in vitro* (Fig. R14.1b). After filtering out embryos with low RNA quality, we used 21 8-cell embryos and 24 late blastocysts for analysis of gene expression. As shown by the PCA in Fig. R14.1c, no clear subgroups within the two distinct developmental stages could

be identified based on transcriptional differences among the individual embryos.

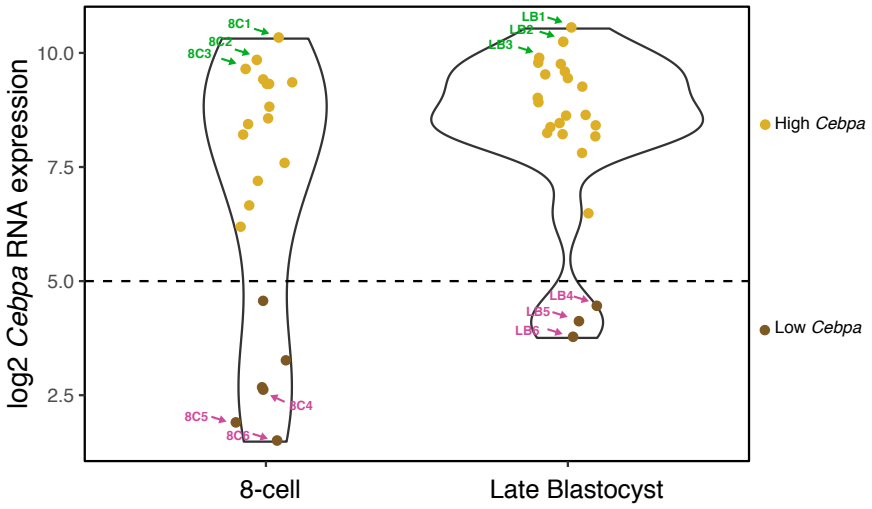


Fig. R14.2| *Cebpa* expression distribution in 8-cell and late blastocyst embryos derived from heterozygous *Cebpa* crosses. The dashed line distinguishes embryos with high or low *Cebpa* RNA levels. Three embryos of each group chosen for further studies are highlighted by green and pink arrows.

We first checked the RNA expression levels of *Cebpa* to determine whether the embryos that developed to the 8-cell and blastocyst stages correspond to the expected mendelian proportions. Although the RNA sequencing data permitted us to identify *Cebpa* low embryos most likely corresponding to the *Cebpa*$^{-/-}$ genotype, wild type and heterozygous embryos could not be distinguished and were considered as a single group (Fig. R14.2). According to mendelian ratios, we expected to observe 5.25 homozygous knock-out embryos for the 8-cell stage and indeed we found 6. In contrast, we only

observed 3 knock-outs instead of the expected 6 for the late blastocyst group. Although the numbers of embryos was probably too low to assess statistical significance, this observation suggests that C/EBPα is required for at least some embryos to reach the blastocyst stage while it is dispensable for the development of embryos to the 8-cell stage.

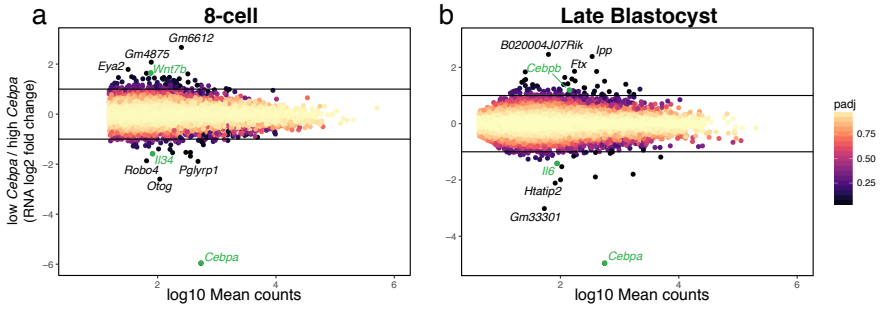


Fig. R14.3| Attempts to identify C/EBPα targets by gene expression analysis. a, Volcano plots of genes expressed in high and low *Cebpa* embryos at the 8-cell stage. **b**, Similar comparisons of genes expressed at the late blastocyst stage. Gene expression values (shown as dots) represent averages of the three selected embryos at each stage and condition indicated in Fig. R14.2. The upper and lower lines represent 2-fold change cut-off values. The genes highlighted in green are discussed in the text.

To investigate differences in the gene expression profile of knock-out and wild type embryos (containing one or two wild type alleles), we compared the 3 highest with the 3 lowest *Cebpa* expressing embryos from both groups (Fig. R14.2). Although there were no major differences in global gene expression between low and high expressing *Cebpa* embryos, we detected a small number of misregulated genes

for each developmental stage. Among these, at the 8-cell stage we found *Wnt7b* (a well-characterized gene involved in placental development[230]) to be upregulated in *Cebpa* low embryos, while *Il34* (encoding a myeloid interleukin and a potential C/EBPα target) was downregulated (Fig. R14.3a).

At the late blastocyst stage we detected an upregulation of *Cebpb* in the *Cebpa* low group. Since *Cebpb* encodes a factor that is structurally and functionally closely related to C/EBPα, this suggests a potential compensatory upregulation of *Cebpb*. In the downregulated group we also found *Il6*. Our data pinpointed a reduced number of candidate target genes of C/EBPα in early development whose function in embryogenesis remains to be further tested. However, because of the likely compensatory upregulation of *Cebpb* or other genes such as *Wnt7b*, it is possible that the majority of C/EBPα's target genes and its own function were obscured.

15. Activation of C/EBPα in ESCs upregulates *Il6ra* as well as trophectodermal genes.

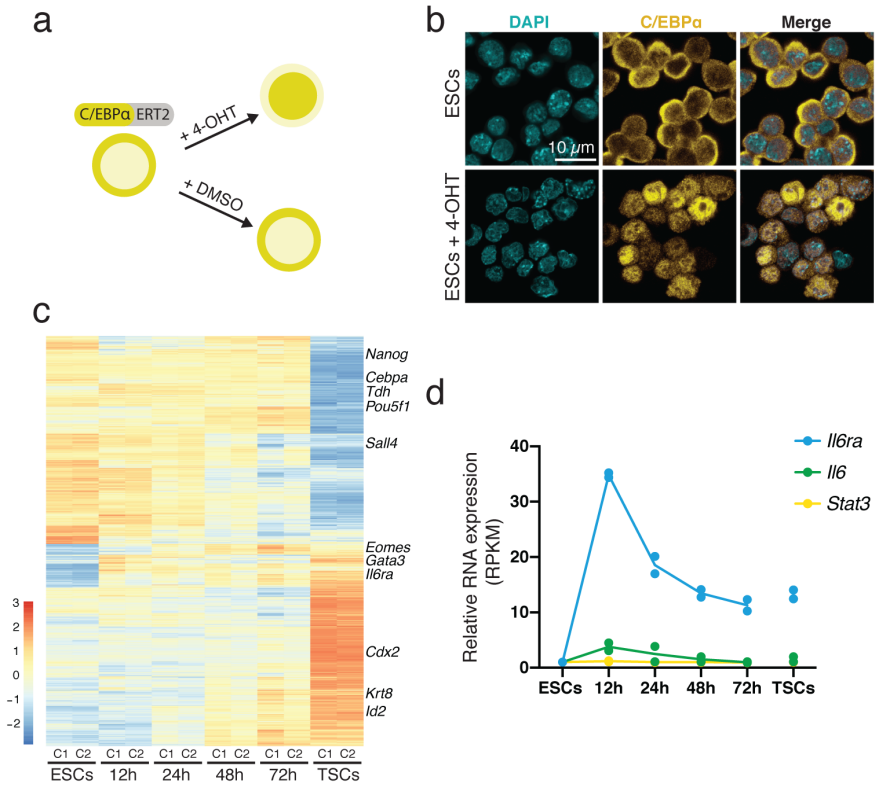


Fig. R15.1| Effects of inducing C/EBPα overexpression in ESCs. a, Schematics of C/EBPα induction in the ESC line clone C1 by addition of 4-hydroxytamoxifen (4-OHT) or DMSO as a control. **b**, Immunofluorescence images (confocal microscopy z-sections) from C/EBPα-inducible ESCs before and after 24 hours of 4-OHT treatment. **c**, Heatmap of gene expression (RNA-seq) from two inducible clones C1 and C2 at different time points after 4-OHT treatment. **d**, Expression of *Il6, Il6ra* and *Stat3* genes at different times after C/EBPα induction in clones C1 and C2 as well as in TSCs.

An alternative approach to studying the potential target genes of C/EBPα in early embryo cells consists in overexpression

experiments, also circumventing the problem of the observed redundancy with C/EBPβ and other potential compensations. For this purpose we constructed a plasmid containing C/EBPα fused to GFP and transcribed it *in vitro* with T7 polymerase into RNA. This RNA was injected into zygotes that were subsequently placed in culture. Although we were able to detect 2- to 8-cell embryos with GFP positive nuclei, we observed associated aberrant mitoses and induced cell death (also likely to the lack of experimental optimization), making this approach obsolete (data not shown).

Since C/EBPα must be co-expressed with pluripotency and trophectodermal factors at the 8-cell stage (Fig. R10), we therefore decided instead to test the effects of C/EBPα overexpression in ESCs. To avoid possible toxicity of C/EBPα overexpression, we constructed a lentiviral vector with an inducible form of C/EBPα fused to the estrogen receptor ERT2 and controlled by the EF1α promoter (Fig. R15.1a). Adding 4-hydroxytamoxifen (4-OHT) 1 μM to the cells expressing this construct shuttles C/EBPα from the cytoplasm to the nucleus (Fig. R15.1b) where it acts as a transcription factor. To obtain homogeneously responding cells, we sorted bulk infected ESCs with the C/EBPαERT2 construct and isolated two clones (C1 and C2) that were used for further studies. Both clones were treated in parallel with 4-OHT for 0, 12, 24, 48 and 72 hours, RNA extracted and analyzed by

RNA-seq in comparison to cultured trophoblast stem cells (TSCs).

Considering C/EBPα expression in pre-implantation and its predicted role in trophectodermal commitment, we compared global expression changes in ESCs upon C/EBPα induction with TSC profiles and observed an upregulation of trophoblast stem cell genes with the concomitant downregulation of genes expressed in ESCs (Fig. R15.1c).

Given their regulation by C/EBPα in B cells and our findings in the mouse embryo, we particularly checked for *Il6* pathway genes in ESCs where C/EBPα induction led to a 35-fold transient increase of *Il6ra* after 12 hours, a 3-fold increase of *Il6* and no considerable changes of *Stat3* (Fig. R15.1d).

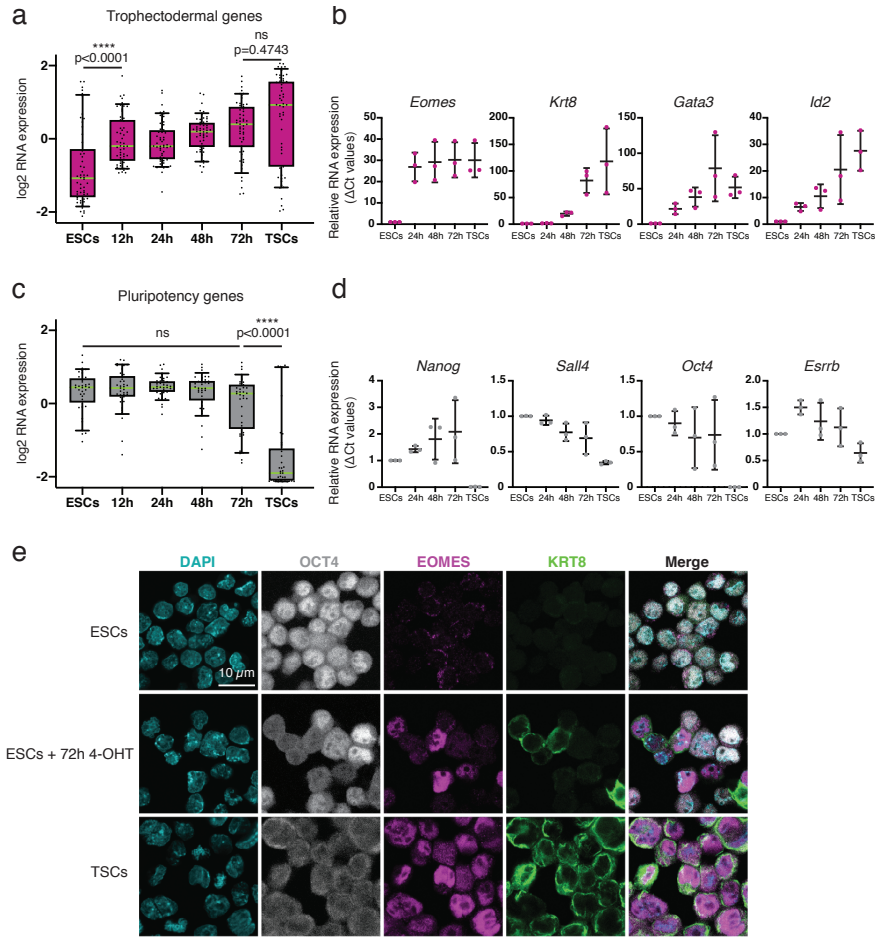


Fig. R15.2| Changes of trophectodermal and pluripotency gene expression induced by C/EBPα in ESCs. a, RNA expression kinetics of signature trophectodermal genes (n=54). RNA-seq data from clones C1 and C2 with dots representing average expression values for each gene. Boxplots and whiskers depict 10-90 percentiles with green line indicating the median value. Statistical significance was determined using two-way ANOVA and Tukey's multiple comparison tests. **b**, Selected TE genes validated by qRT-PCR in n=3 independent experiments. **c**, Same as in **a** for signature pluripotency genes (n=37). **d**, Validation of selected pluripotency genes as in **e**. Immunofluorescence images (confocal z-sections) showing expression of the pluripotency marker OCT4 and the TE markers EOMES and KRT8 in ESCs, C/EBPα-induced ESCs for 72 hours and in trophoblast stem cells (TSCs). See table MM9 for gene signatures.

To further examine the detected changes in gene expression, we concentrated on trophectodermal and pluripotency signatures (see gene lists in Material and Methods) assuming that they would be the main affected ones by C/EBPα induction in ESCs and we validated some of them by qRT-PCR (Fig. R15.2b, d).

Overall, we detected a fast and significant upregulation of trophectodermal genes right after 12 hours post-induction (Fig. R15.2a) with little decrease in pluripotency associated genes even after 72 hours (Fig. R15.2c). Most notably, the TE restricted transcription factors *Gata3*, *Eomes* and *Hand1* became upregulated 18-fold, 28-fold and 300-fold respectively, while *Cdx2* increased 3.3-fold. In addition, we observed a gradual upregulation of the TE specific keratin genes *Krt7*, *Krt8*, *Krt18* and *Krt19*, with expression of all genes increasing at least 50-fold at 72 hours post-induction. The response in the expression of pluripotency associated genes was heterogeneous. Some became downregulated (*Oct4*, *Sall4* and *Zfp42*), others became first up and then downregulated (*Esrrb* and *Gdf3*) while still others became upregulated (*Nanog* and *Dppa3*) (Fig. R21c). The functional relevance of these changes, however, still needs to be evaluated. Additionally, other gene categories should be analyzed in an unsupervised manner.

We also analyzed the effects of C/EBPα induction in ESCs by confocal microscopy of immunostained cells. As shown in Fig. R15.2e, at 72 hours post-induction some cells expressed the trophectodermal markers EOMES or KRT8 while losing expression of OCT4. Of note, these markers were expressed in different cell subsets either because they represent different stages of TE cell differentiation or reflect a heterogeneity in the TE-like population induced by C/EBPα.

16. C/EBPα activation in ESCs generates two distinct sub-populations enriched either in trophectoderm and pluripotency gene expression.

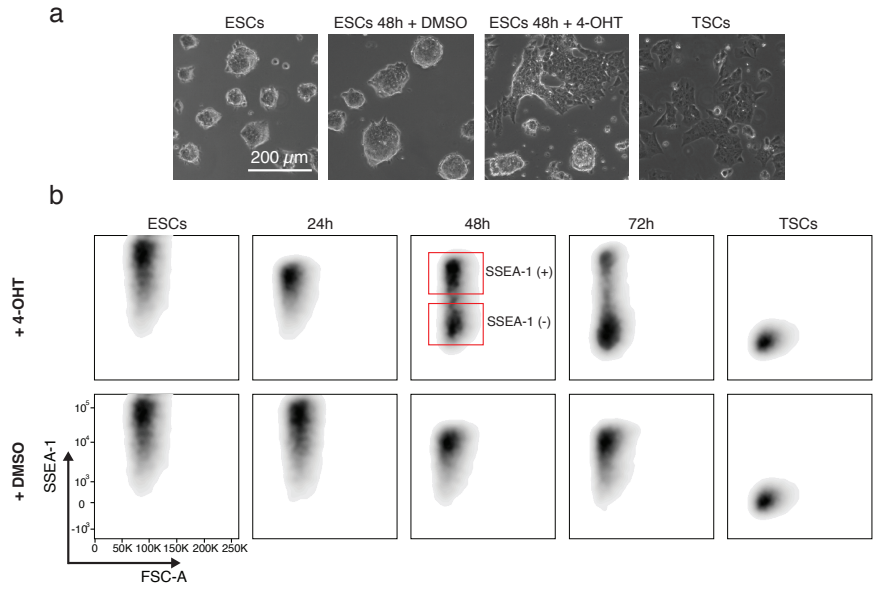


Fig. R16.1| C/EBPα activation in ESCs induces two distinct cell sub-populations. a, Phase contrast images of inducible ESC clone C1, untreated, treated with DMSO or with 4-OHT for 48 hours. TSCs are shown as a comparison. **b**, FACS plots of 4-OHT treated and DMSO treated clone C1 ESCs. Cells were analyzed at different time points

together for the expression of the pluripotent cell surface marker SSEA-1 relative to forward scatter (FSC-A). TSCs were used as controls.

The changes in gene expression of ESCs induced by C/EBPα (after addition of 4-OHT) were also reflected by changes in cell morphology. Thus, at 48 hours post-induction some ESC colonies lost their dome-shaped spherical structure, became flat-adherent and cells acquired a morphology resembling that of cultured TSCs (Fig. R16.1a). Other C/EBPα-induced colonies, however, retained their spherical shape. Control ESC cultures treated with the solvent DMSO increased in cell density and appeared to differentiate, probably as a result of over-crowding. In an attempt to separate the TE-like from the pluripotent stem cell population, we tested the expression of the cell surface antigen SSEA-1, previously described as an early indicator of pluripotency during the OSKM-induced reprogramming of MEFs into iPSCs[230].

We therefore analyzed the two clones at 24, 48 and 72 hours after C/EBPα induction by FACS for the expression of SSEA-1, in parallel with uninduced control cultures and TSCs (Fig. R16.1b). In the induced clones we first observed a decrease in SSEA-1 expression at 24 hours followed by the appearance at 48 and 72 hours of two sub populations: SSEA-1 negative and positive cells. In the uninduced controls we observed a slight global decrease of SSEA-1 at 48 and 72 hours possibly reflecting spontaneous differentiation.

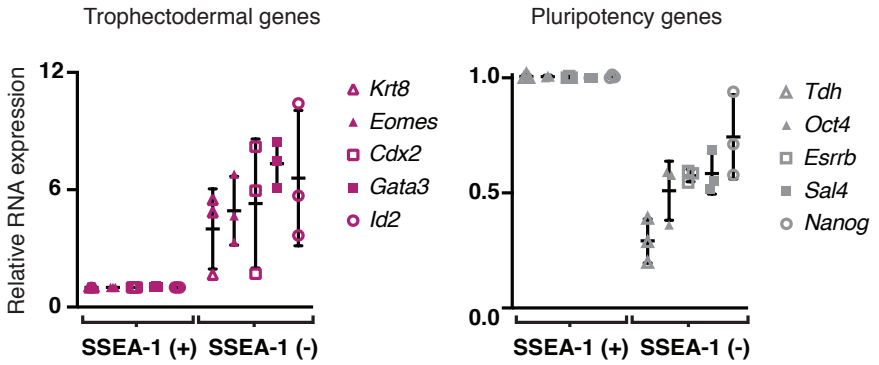


Fig. R16.2| Expression of selected genes in sorted sub-populations of C/EBPα induced ESCs. SSEA-1 (+) and SSEA-1 (-) populations from the experiment in Fig. R16.1 were FACS-sorted 48 hours after C/EBPα induction and expression levels of key trophectodermal and pluripotency genes were measured by qRT-PCR relative to *Actb*. Values represent n=3 biological replicates.

We then FACS-sorted the two subpopulations at 48 hours post-induction based on their SSEA-1 expression and monitored the levels of key trophectodermal and pluripotency genes by qRT-PCR, comparing the expression levels between the two cell fractions. This showed that TE markers are highly expressed in the SSEA-1 (-) population while pluripotency gene expression is slightly enriched in the SSEA-1 (+) cells (Fig. R16.2).

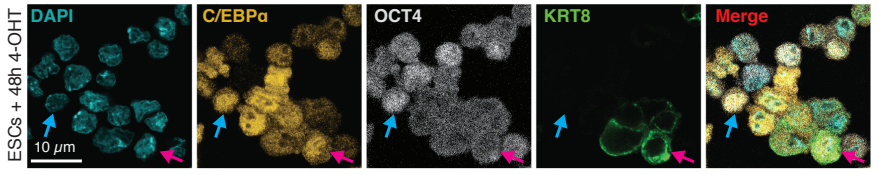


Fig. R16.3| Imaging of C/EBPα, pluripotency and trophectodermal marker expression in C/EBPα-induced ESCs. Immunofluorescence images (confocal z-sections) showing expression of C/EBPα, the

pluripotency marker OCT4 and the TE marker KRT8 in clone C1 ESCs induced for 48 hours. Blue arrow shows C/EBPα positive cell that co-expresses OCT4 but not KRT8. The pink arrow highlights a C/EBPα positive cell that co-expresses KRT8 but not OCT4.

To investigate C/EBPα expression relative to pluripotency or trophectodermal markers we co-stained 4-OHT treated cultures with antibodies against OCT4 and KRT8. Despite performing these experiments with a clonal cell line, the levels of nuclear C/EBPα were heterogeneous (Fig. R16.3). Strikingly, we found that cells containing comparable levels of nuclear C/EBPα can either co-express the pluripotency marker OCT4 or the TE marker KRT8. It will be interesting to extend these observations by single cell gene expression analyses.

In conclusion, C/EBPα overexpression in ESCs generated the formation of cells with a trophoblast-like morphology and TE gene expression, suggesting that C/EBPα is capable of inducing the formation of trophectodermal-like cells. In addition, it appears that the pluripotent-like subpopulation in the induced clones differs slightly from the starting ESCs, raising the possibility that C/EBPα exerts a second function in ESCs.

DISCUSSION

CHAPTER 1: Role of C/EBPα and the IL-6 pathway during somatic cell reprogramming.

In the first part of my thesis project I have found that overexpression of C/EBPα in B cells can upregulate the *Il6*, *Il6ra* and *Stat3* genes by directly binding to their regulatory regions and subsequently induce STAT3 phosphorylation. I found that activation of the IL-6 signaling pathway triggered by C/EBPα is dispensable for B cell to macrophage switch. On the contrary, upregulation of the *Il6* pathway genes at early time-points is required for the successful C/EBPα-enhanced B cell to iPSC reprogramming. The key effector of this requirement is not IL-6 secreted by feeder MEFs, but by the B cells themselves. Interestingly, distinct subsets of B cells mediate a paracrine IL-6 signaling with some B cells expressing *Il6* ligand and others the *Il6ra* receptor. Importantly, IL-6 depletion during reprogramming impacts not only the expression of pluripotency related genes, but also that of trophectodermal, senescence and cell cycle programs. Finally, we found a cross-talk between IL-6 signaling and LIF signaling, which is required during late stages of reprogramming.

1. C/EBPα as a direct regulator of the *Il6* pathway.

C/EBPβ has been reported to activate IL-6 by binding to the *Il6* promoter in the context of inflammation and senescence[123]. However, it is not the only factor at play since *Il6* expression was maintained even upon *Cebpb* depletion[126]. In my studies, I have observed that also C/EBPα can bind to and activate putative upstream enhancers of *Il6* (Fig. R2a). These C/EBPα-bound sites increase their contacts with the *Il6* promoter as shown by virtual 4C (Fig. R2b), suggesting a direct control by C/EBPα. Functional approaches such as destroying these C/EBPα binding sites and test their ability to drive the activation of a reporter gene would be needed to further strengthen our conclusion.

I considered the possibility that C/EBPα would directly control other members from the *Il6* pathway after a study where IL-6R protein was lost in *Cebpa* deficient hepatocytes. Additionally in this work, soluble IL-6 and LPS failed to phosphorylate STAT3 in the context of mutant *Cebpa*$^{-/-}$[129]. Thus, in addition to the direct *Il6* activation, I have observed that C/EBPα binds to and activates both the *Il6ra* and *Stat3* promoters (Fig. R2c-e). The finding that this is followed by STAT3 phosphorylation (Fig. R3) shows that a single transcription factor, C/EBPα, is capable of functionally activating several genes within the *Il6* pathway in B cells.

2. The *Il6* pathway during C/EBPα-induced B cell to macrophage transdifferentiation and the inflammatory response.

Recent studies in human cells have shown that blocking the *Il6* pathway prevents hematopoietic stem cells ($CD34^+$) from differentiating into monocytes upon infection with *Mycobacterium tuberculosis in vitro*[121]. Additionally, it has been reported that IL-6 production from feeder MEFs skews the differentiation of monocytes into macrophages rather than towards dendritic cells[131]. In contrast, B cells could differentiate into monocytes/macrophages under both control and IL-6 neutralizing conditions in our system where I observed no significant changes in the transdifferentiation kinetics of *Il6* deficient compared to wild type B cells (Fig. R4.1, R4.2), with the induced macrophages exhibiting full phagocytic capacity (Fig. R4.3). This is in line with previous studies that demonstrated that macrophages from $Il6^{-/-}$ mice can also develop in the absence of the cytokine.

Two types of macrophages have been described: M1 type macrophages are involved in tumor suppression and M2 type macrophages in inflammatory responses[231,232]. Moreover, studies with $Il6^{-/-}$ mice showed a bias towards M1 type macrophages in the context of induced inflammation, impairing the progression of dermatitis[233] and colorectal cancer[234]. Despite not affecting transdifferentiation kinetics or

phagocytosis, we speculate that IL-6 could still influence the type or proportions of induced macrophages (M1/M2 categories) that we obtain at the end of the B cell transdifferentiation process.

In a recent study our laboratory tested the requirement of the topological factor CTCF both for the formation of macrophages and their response to bacterial stimuli. Surprisingly, while the lack of CTCF attenuated the inflammatory response after LPS stimulation, transdifferentiation was not impaired and even slightly accelerated[235]. These observations suggest that two fundamentally different processes regulate *Il6* pathway genes during macrophage development: i) upregulation to high levels during the differentiation of hematopoietic progenitors into monocytes/macrophages, creating a poised 'state' and ii) further upregulation in response towards bacterial stimulus. In that regard, the transdifferentiation induced by C/EBPα in B cells recapitulates the first process. Therefore, despite not observing differences in the transdifferentiation kinetics, we predict that *Il6*$^{-/-}$ iMacs could still exhibit attenuated inflammatory responses as it happened for CTCF deficient iMacs challenged with LPS.

3. IL-6 signaling in C/EBPα-enhanced B cell to iPSC reprogramming indicates a paracrine mechanism involving two B cell-derived subpopulations.

Our observation that the genetic ablation of *Il6* or depletion of the cytokine by antibody neutralization leads to significantly reduced numbers of iPSC colonies during B cell reprogramming (Fig. R5, R6) is in line with previous studies using MEFs *in vitro*[133] or *in vivo*[134,136]. Thus, regardless of the somatic cell type used for reprogramming, IL-6 signaling is required for the establishment of pluripotency. We also found that in our system the early phase is critical since treatment with antibodies that block IL-6 during the C/EBPα pulse and the first 6 days of OSKM activation impaired reprogramming while they were ineffective at late time-points (Fig. R6). This early role of IL-6 signaling coincides in time with the observed transient expression of *Il6* and *Il6ra* in C/EBPα-pulsed B cells and in cells activated for 2 days with OSKM (Fig. R5.1b).

The fact that in our experiments B cells were grown on a feeder layer of IL-6-secreting MEFs, raised the possibility that the observed IL-6 dependence is due to a paracrine mechanism involving MEFs. Perhaps surprisingly, however, testing different genetic combinations of deficient *Il6* B cells and MEFs revealed that signaling among B cells is sufficient for iPSC reprogramming and that IL-6 secreted from the mitomycin-C inactivated feeders only had a minor effect (Fig. R5.2a).

Further analyses of single cell data during reprogramming revealed that the C/EBPα pulse splits the B cell population into two different subsets expressing either the ligand *Il6* or its receptor *Il6ra* (Fig. R9.2). These findings suggest that such cell autonomous IL-6 signaling described for B cells is further paracrine among them and it must be orchestrated by different populations.

4. Heterogeneity of B cell reprogramming to pluripotency: branching towards TE-like and senescence related cells.

Our RNA expression analyses of C/EBPα-enhanced reprogramming in *Il6*$^{-/-}$ B cells showed similarities and differences to wild type B cells. On the one hand, the dynamics of B cell silencing and transient upregulation of myeloid genes was indistinguishable among cell types (Fig. R7.2a). On the other hand, I observed an overall impairment of pluripotency and TE gene upregulation in the mutant cells and, additionally, a faster downregulation of cell cycle genes with the concomitant upregulation of a senescence program (Fig. R7.2b, c). As summarized in the diagram of Fig. D1 at the end of this chapter, our findings indicate highly specific functions of IL-6 signaling during reprogramming towards iPS cells.

The finding that the *Il6* pathway is required for the full activation of both pluripotency and trophectodermal fates

raises the possibility that it acts on the growth of a bipotent precursor. Further analysis of our single cell data will disclose whether there is a link between the formation of TE-like cells, senescence and IL-6.

We hypothesize that the two distinct cell subsets expressing either *Il6* or *Il6ra* reflect the observed branching heterogeneity during B cell to iPSC reprogramming. More precisely, we speculate that *Il6* expression is characteristic for the TE-like, and that these cells potentially activate a senescence program. In contrast, we propose that the *Il6ra* expressing subpopulation corresponds to an iPSC progenitor whose proliferation is induced by IL-6 uptake. In short, the overall idea is that TE-like cells act as a niche for the formation of pluripotent cells, as we also hypothesize happens during early embryo development (see below). Further analyses of single cell reprogramming trajectories of wild type and *Il6*$^{-/-}$ B cells will be required to test these ideas.

Recent single cell studies from other laboratories centered on reprogramming trajectories of MEFs have also identified several alternative branches from the main path towards iPSCs[43]. These bifurcations from the pluripotent fate include cells with neural, stromal and TE signatures, the latter of which gave rise to stable trophoblast stem cell lines in human dermal fibroblast reprogramming[44]. The stromal line was shown to secrete signaling molecules in a paracrine fashion

such as GDF19 that contributed to the establishment of pluripotency. These observations suggest that besides the described IL-6 signaling required for B cells, fibroblasts and *in vivo* reprogramming of different tissues[134], other signaling pathways also act in a paracrine fashion during the establishment of pluripotency. The best-studied among these, the LIF pathway, will be discussed below.

5. Coordinated IL-6 and LIF signaling. Redundancy or competition?

The IL-6 signaling pathway that leads to STAT3 phosphorylation is shared with LIF. Thus, binding of either IL-6 and LIF to their specific receptors induces a heterodimerization with the common GP130 signaling receptor encoded by *Il6st* gene[103]. This raises the question about whether IL-6 and LIF act redundantly.

Our evidence points to both redundant and non-redundant functions. The initial upregulation of *Il6* pathway genes preceding *Lifr* expression during B cell reprogramming suggests that the *Il6* pathway becomes active before LIF signaling (Fig. R5.1b, R8). In addition, pairwise correlations of single cell gene expression showed that *Gp130* is co-expressed with *Il6ra* during initial time points of reprogramming while *Lifr* expression only coincides with *Gp130* at later stages, when *Il6* pathway components are already downregulated (Fig. R9.2d-f). In our reprogramming

experiments LIF is added to the medium from the 18 hours of C/EBPα pulse onwards until day 12 of OSKM activation. The observed sequential upregulation of *Lifr* following expression of *Il6ra* suggests that LIF signaling takes over from IL-6 to phosphorylate STAT3 at mid to late reprogramming phases, eventually being required to maintain the proliferation of pluripotent cells.

The observed increase in *Lifr* expression in *Il6*$^{-/-}$ B cells at intermediate stages of reprogramming suggests that the IL-6 signaling represses the LIF pathway, although the physiological significance of this observation is presently unclear. In sum, our data support the notion that the action of IL-6 and LIF, although likely redundant at intermediate stages of reprogramming, is separated in time and that there is a cross-talk between the two pathways.

One way to analyze this cross-talk further would be to extend *Il6* expression during reprogramming and determine whether this delays *Lifr* upregulation. Likewise, overexpressing *Lifr* during the early time points of *Il6*$^{-/-}$ reprogramming could conclude whether LIF signaling is sufficient to rescue the lack of IL-6, indicating redundancy in this context.

Unexpectedly, the observed sequential regulation of the IL-6 and LIF pathways during mouse B cell reprogramming is reversed during pre-implantation where early LIF signaling

precedes later IL-6 effects[103]. In this study we have further characterize the IL-6 function in pre-implantation development (Fig. R11.2) and showed that TE and ICM layers in mouse blastocysts are respectively enriched for *Il6* and *Il6ra* expression (Fig. R11.1 discussed in chapter 2), which reproduces a pattern that is also conserved in human embryos[139].

6. Our current model: limitations and future perspectives.

I observed that IL-6 neutralization during reprogramming only partially abolishes iPSC colony formation (Fig. R6). Based on the fact that B cells grow in tight aggregates, this may provide an obstacle for the IL-6 blocking capacity of the anti IL-6 antibody.

Another point to consider is that *Il6* ablation in either B cells, feeder MEFs or both does not completely abrogate iPSC reprogramming. Here, late activation of LIF could be rescuing pluripotency, leading to the formation of 'escapee' iPSC colonies. Thus, depletion of both *Il6* and *Lif* (or their receptors) would be necessary to assess whether the acquisition of pluripotency is exclusively dependent on GP130 and STAT3 signaling. In addition, more detailed studies on differential gene expression would help to understand the regulation of pluripotency networks that allow iPSC formation in the context of *Il6* ablation.

It needs to be mentioned that addition of recombinant IL-6 during *Il6*$^{-/-}$ cell reprogramming only partially rescued the number of iPSC colonies observed with wild type cells (Fig. R5.2b). Here, a possible role of soluble IL-6R and GP130 receptors, described to respectively enhance or repress IL-6 signaling in BAF/3 B cells[119], may be consider to understand the whole phenotype. Visualizing IL-6 and LIF signaling components with fluorescent reporters will be key to better understand the interplay between these pathways during cell reprogramming.

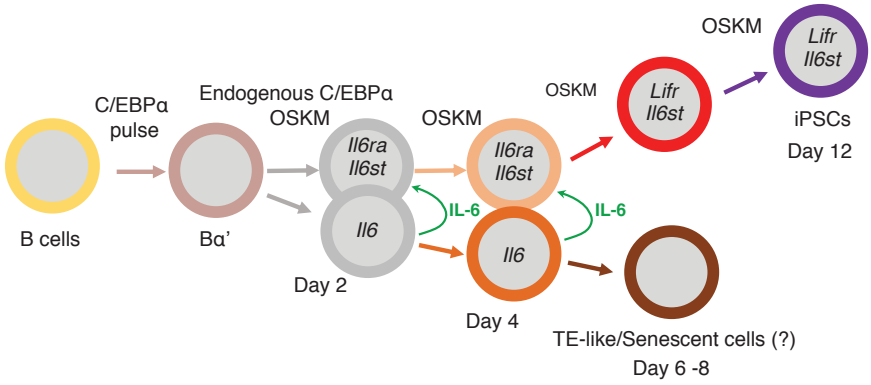


Fig. D1| Model of how IL-6 signaling is necessary for C/EBPα-enhanced B cell to iPSC reprogramming. The diagram shows the various stages of C/EBPα enhanced reprogramming to iPSCs as well as to TE-like cells and their connection to the expression of *Il6* and *Lif* signaling pathway genes. We postulate that C/EBPα transiently activates *Il6ra* in the subset of B cells fated towards pluripotency. At a later stage these cells express *Lifr*. *Gp130* (*Il6st*) expression overlaps with both of these receptor genes at early and late stages of reprogramming. The branch fated towards TE and senescent cells transiently expresses C/EBPα-activated *Il6*. This branch might act as a niche necessary for the formation of pluripotent cells through secretion of IL-6 in a paracrine fashion.

CHAPTER 2: Role of C/EBPα and IL-6 during trophectoderm specification.

For the second part of my study I will discuss the role of C/EBPα during pre-implantation mouse development. I observed expression of C/EBPα in pre-implantation embryos at both the RNA and protein levels. *Cebpa* RNA is first detectable at the 2-cell stage, then its protein becomes asymmetrically expressed in the blastomeres of 4-cell embryos and reaches its highest expression levels in 8-cell embryos. After this stage, C/EBPα expression drops in the morula and then becomes re-expressed specifically in TE cells of the late blastocyst (Fig. R10). I have also found that blastocysts produce IL-6 and that IL-6 neutralization in cultured embryos delays the morula to blastocyst transition (Fig. R11.2), indicating an as yet undescribed role of the IL-6 signaling pathway in pre-implantation embryo development. Finally, I found that C/EBPα overexpression in ESCs induces dramatic gene expression changes resulting from the generation of TE-like cells and alterations in pluripotency related gene signatures (Fig. R 15.1, 2).

7. C/EBPα as a lineage instructive transcription factor for trophectoderm cell fate.

Previous studies reported that C/EBPα is highly expressed in the trophectoderm of mouse blastocysts[200]. In addition, embryos from double knock-outs of *Cebpa* and its family

related member *Cebpb* displayed defective placentas[63]. I have likewise detected the expression of C/EBPα at the protein level in the trophectoderm of late blastocysts, although in only a subset of the cells (Fig. R10 and R14.1). Most importantly however, I found that C/EBPα overexpression in ESCs dramatically upregulates several TE regulators such as *Gata3* and *Eomes* (and to a lesser extent *Cdx2*) within only 12 hours of induction and more gradually the trophectodermal markers *Krt8* and *Krt18* in later time points (Fig. R15.2). These results suggest that C/EBPα has the potential to instruct TE lineage fate in the early mouse embryo.

The finding that *Cdx2* only becomes upregulated modestly by C/EBPα could be explained by the lack of molecular partners in ESCs that may cooperate with C/EBPα in the totipotent embryo to downstream activate *Cdx2* in subsequent stages. Another explanation could simply be that C/EBPα does not directly regulate *Cdx2*, being C/EBPα expression in a TE subset of cells disconnected from *Cdx2* expression. Indeed, trophectoderm can be further divided into mural and polar TE from which different types of cells (mostly primary and secondary giant cells) and tissues arise[236,237]. Embryo implantation in the uterus and placental development rely on the proper differentiation of these cell types and on their phagocytic capacity[238]. Therefore, characterizing and tracking

the specific cells that express C/EBPα in the TE may help to understand the diversification of cell types in the placenta.

8. Is C/EBPα a symmetry-breaking factor in 4-cell blastomeres that forecasts TE specification?

Using immunofluorescence and confocal microscopy I discovered that C/EBPα is expressed in 4-cell embryos with varying levels among the blastomeres. More precisely, C/EBPα was most highly expressed in the vegetal blastomere, the farthest cell from the polar body (Fig. R12). Since the vegetal blastomere has been described to preferentially give rise to the trophectoderm[171,178] this suggests that C/EBPα expression at the 4-cell stage is an early determinant of trophectodermal lineage specification.

To further support this conclusion it would be desirable to study expression of C/EBPα together with other symmetry breaking factors such as SOX21, whose high level expression in 4-cell blastomeres biases the cells towards an ICM fate[175]. Our results would predict that cells with the highest levels of C/EBPα exhibit the lowest levels of SOX21 and the other way around. Another experiment that would more directly address the issue of whether C/EBPα acts as a symmetry breaker is to inject *Cebpa* RNA together with a tracer in one blastomere of 2-cell embryos and asses whether C/EBPα overexpressing cells preferentially contribute to the TE-lineage, as predicted.

The ablation of both *Cebpa* and *Cebpb* in one of the two blastomeres, such as by CRISPR, would be another interesting experiment. This should result in a decreased proportion of TE cells in blastocysts, unless there is yet another layer of redundancy.

9. Does a CARM1-induced arginine methylation of C/EBPα modulate the ICM/TE lineage bifurcation?

As reviewed in the Introduction, CARM1 is critically required for early embryo development. Moreover, the asymmetric expression in 4-cell embryos of this arginine methylating enzyme has been shown to determine ICM or TE specification, as proved by gain and loss of function experiments, with CARM1 high cells being fated towards the ICM[172]. Candidate targets so far are a specific histone modification (H3R26me) and the BAF155 component of the SWI/SNF complex[179].

However, it is also possible that C/EBPα is directly modified by CARM1 since the closely related TF C/EBPβ has been identified as a CARM1 target, with methylation of specific arginines causing distinct changes in the factor's biological properties[90]. Furthermore, ongoing studies in our laboratory have shown that CARM1 asymmetrically methylates arginine in position 35 of C/EBPα and that an alanine replacement of this residue abrogates iPSC reprogramming while

dramatically increasing the speed by which C/EBPα converts B cells into macrophages (unpublished). Based on these observations, we speculate that methylated form C/EBPα is required for pluripotency specification and, by default, that the unmethylated form biases a hypothetical bipotent progenitor to differentiate towards the TE fate. Nonetheless, it is also possible that high C/EBPα levels alone are sufficient to specify TE cells, or that a combination of high C/EBPα levels and CARM1-induced arginine methylation is critical for the earliest embryonic cell fate decision.

10. Functional approaches to study the role of C/EBPα during pre-implantation development.

Mice with a knock-out of *Cebpa* undergo normal embryo development and die perinatally[67]. However, culturing fertilized zygotes from an F1 cross of *Cebpa*$^{+/-}$ animals revealed a reduction in the number of *Cebpa*$^{-/-}$ embryos reaching the blastocyst stage, since fewer embryos as expected from Mendelian ratios were detected (Fig. R14.2). Although the significance of this observation requires larger numbers of mice, it suggests that not all *Cebpa*$^{-/-}$ null embryos survive under the more stringent conditions of *in vitro* culturing compared to the *in vivo* environment.

In the *Cebpa*^{low} embryos that survived, we found only very few differences in gene expression compared to *Cebpa*^{high}

embryos (Fig. R14.1). However, as discussed in the Introduction, early embryo development is highly robust and there are few examples where, as a consequence of a single gene ablation, embryos get arrested at the 4- or 8-cell stage. For example, knock-outs of the key lineage instructive factors *Oct4*[157] or *Dux*[157] do not alter cell fate specification during embryo development. A likely reason is that compensatory upregulation of redundant factors obliterate the observation of possible phenotypic alterations caused by the knock-outs.

This could also be the case for our experiments since we found that *Cebpa*^{low} blastocysts show high *Cebpb* expression, suggesting a compensatory upregulation of the latter (Fig. R14.3). Supporting this interpretation, it has been reported by Achim Leutz's laboratory that while double knock-outs impair placental labyrinth formation and embryo implantation, single knock-out embryos of *Cebpa* and *Cebpb* show no clear deficiencies[63].

Among the handful of genes whose expression differed between *Cebpa*^{low} and *Cebpa*^{high} cells, the most interesting is *Wnt7b*. This gene, which becomes upregulated in *Cebpa*^{low} 8-cell embryos is a well-characterized trophectodermal factor[230] and it might also contribute to the compensation of *Cebpa* loss. It would therefore be interesting to explore the effects on early embryo development of neutralizing WNT7B protein. It would also be interesting to study whether the simultaneous

depletion of *Cebpa* and *Cebpb* leads to a greater impact on *Wnt7b* expression and more overt developmental deficiencies. Another approach to bypass redundancy with other TFs would be to overexpress C/EBPα through injection of *Cebpa* mRNA into zygotes or single blastomeres in 2-cell embryos.

11. Does C/EBPα have a role in the ICM/TE lineage segregation?

Ectopic C/EBPα overexpression in our inducible ESC system turned out to be a powerful tool to study its role in early lineage commitment. Thus, I found that tamoxifen mediated induction of the factor triggers an almost immediate upregulation of trophectodermal markers, generating a subpopulation of cells that downreglate the pluripotent SSEA-1 cell surface marker while another subset of cells maintained or even increased SSEA-1 detection (Fig. R16.1). In addition, exogenous C/EBPα increased the bulk upregulation of *Nanog* (Fig. R15.2) while maintaining higher expression of several pluripotency markers in the SSEA-1 (+) fraction (Fig. R16.2). Therefore, we speculate that C/EBPα overexpression not only segregates TE-like cells but potentially ICM-like cells as well, which recapitulates, at least in part, the first cell fate decision in the embryo.

In this model, C/EBPα would create bipotent precursors that co-express trophectodermal and pluripotency restricted TFs before splitting into TE-like cells (SSEA-1 (-)) and ICM-like cells (SSEA-1 (+)). Such segregation could also reflect the proposed bifurcation during B cell reprogramming into *Il6ra* positive iPSCs and *Il6* positive TE/senescent cells. Single cell analyses at early time points after C/EBPα induction in ESCs will be required to validate this model. It will also be interesting to determine whether blocking or adding IL-6 during C/EBPα induction alters the balance between TE- and ICM-like populations in ESCs.

It is remarkable that the length of C/EBPα expression required for enhancement of reprogramming[23] and the expression period in early embryos preceding the establishment of pluripotency is roughly similar (12-20 hours). Fine tuning of C/EBPα expression in ESCs for different lengths of time might further improve the resemblance of the cells obtained to those that form under physiological conditions. Additionally, injecting into morulas C/EBPα-induced ESCs before and after their segregation into SSEA-1 (+) or SSEA-1 (-) subsets and assessing their preferential contribution to either TE or ICM in the blastocyst might reveal the transient formation of bipotent TE/ICM-like precursors driven by C/EBPα. Finally, transcriptomics, chromatin accessibility, C/EBPα chromatin binding, histone methylation, protein levels and interaction partners of C/EBPα in SSEA-1

(-) and (+) populations could reveal the role of different C/EBPα isoforms in specifying ICM and TE lineages.

12. Does C/EBPα target the *Il6* pathway in the embryo?

Earlier work of the laboratory has described a remarkable specificity in the ability of C/EBPα to enhance reprogramming of B cells, as neither the erythroid TF GATA1, the muscle regulator MyoD nor the neuronal cell fate inducer ASCL1 showed such a capacity[23]. This might be explained by the observation that C/EBPα, but none of the other TFs tested, activates myeloid genes, including three members of the *Il6* pathway (Fig. R2), and that the *Il6* pathway is strictly required for B cell to iPSC reprogramming (Fig. R5, R6).

The role of IL-6 during reprogramming in turn appears to reflect what is happening during pre-implantation development since *Il6* is expressed in the trophectoderm and *Il6ra* in the ICM of the blastocyst (Fig. R11.1). Moreover, I detected the production of IL-6 by cultured blastocysts and discovered that neutralization of IL-6 in the culture medium delays the morula to blastocyst transition (Fig. R11.2). The latter findings extend the previously reported *Il6* global expression in blastocysts[137] and that blocking IL-6 during embryo development reduces STAT3 phosphorylation[138].

However it is still not entirely clear whether C/EBPα directly controls the *Il6* pathway during pre-implantation or whether IL-6 signaling itself influences TE/ICM lineage segregation. Our finding that *Cebpa*^{low} blastocysts show a slight decrease in *Il6* expression supports this hypothesis (Fig. R14.3), but more embryos need to be analyzed to determine the significance of the observed difference.

As discussed earlier, we hypothesize that different post-translational modifications of C/EBPα activate the *Il6* and *Il6ra* genes in TE and ICM respectively. Therefore, IL-6 signaling from TE cells could contribute to the establishment of pluripotency, raising the possibility that the trophectoderm acts as a niche for the ICM (see model in Fig. D2 at the end of the chapter). Future experiments such as genetic ablations of *Il6* and *Il6ra* in *Lifr* null backgrounds, visualization of IL-6 and IL-6R in the embryo and additional molecular approaches will be required to test our model.

The observed upregulation of *Il6ra* mRNA 12 hours after C/EBPα induction in ESCs (Fig. R15.1d) raises the question of whether this occurs in the subset of cells that remain fated to pluripotency state. Studying the levels of IL-6R and of phosphorylated STAT3 between the SSEA-1 (+) and (-) populations together with single cell analyses at early time points after C/EBPα induction might clarify this question. The fact that in this system we could not detect *Il6* upregulation

could be due to the presence of LIF that represses *Il6* or because our cultures do not exactly reproduce the physiological conditions required for *Il6* expression.

Of note, recent studies have identified *Gata3* as an evolutionarily conserved target of the Hippo pathway that specifies trophectodermal fate in mouse, human and cow embryos[239]. Interestingly, *Gata3* mRNA was also found to be highly upregulated after C/EBPα overexpression in ESCs (Fig. R15.1, 2). This observation raises the possibility that *Gata3* lies in a pathway shared between C/EBPα and the Hippo signaling or that different pathways can redundantly activate the gene.

13. What regulates C/EBPα expression in the embryo?

What causes the upregulation of C/EBPα in the first place? We observed that *Cebpa* mRNA gets upregulated as early as at the 2-cell stage (Fig. R10) coinciding with the zygotic genome activation by maternal factors in the mouse. In addition, by data mining of the published literature, we found that *Dux* overexpression in ESCs dramatically upregulates *Cebpa* during the cells' transition into a 2-cell-like state. In addition, such upregulation is preceded by chromatin opening of the *Cebpa* promoter suggesting a direct effect on the regulation of the gene (Fig. R13). These findings suggest that

maternal DUX initiates the expression of *Cebpa* during pre-implantation embryogenesis. In this context it would be interesting to explore whether a double knock-out of *Cebpa* and *Cebpb* (the latter of which is not normally expressed at this stage but could potentially compensate) leads to a block of embryo development at the 2-cell stage. What causes the observed downregulation of *Cebpa* after the 8-cell stage and its later re-expression in a subset of TE cells is unclear (Fig. R10) but it might be caused by newly expressed factor(s) that respectively act as repressors and activators of *Cebpa*.

14. Non-canonical roles of C/EBPα.

One of the unexpected observations in my study is that C/EBPα protein remains bound to mitotic chromosomes and that it is also enriched in heterochromatin dense regions among interphase nuclei (Fig. R10). Since the pericentromeric regions to which C/EBPα binds are highly enriched in satellite DNA, this suggests that it is sequestered by repetitive DNA elements. This might prevent ectopic transcriptional activation as it has been suggested for C/EBPα in 3T3-L1 pre-adipocyte cells[239].

However, other explanations are equally possible. For example, chromosome-bound C/EBPα could also act as a "bookmarking" for transcription sites that quickly resume their expression after cell division, as it has been shown for

ESRRB expressed in the embryo[239]. Mitotic C/EBPα could also represent an anchoring point for the microtubule spindle to stabilize chromosome segregation during mitosis, as shown for other TFs such as OCT1 or WT1[240,241]. Finally, the observed pattern of C/EBPα localization in mitotic chromosomes might also be a consequence of local liquid-liquid phase separation and condensate formation[242]. *In vivo* tracking of endogenous C/EBPα labeled with a reporter or degrading the protein in specific phases of the cell cycle might help to clarify the functionality of this observation.

Summing up, by studying forced cell fate transitions of hematopoietic and embryonic stem cells in rather artificial model systems, my work has led me to obtain novel insights into the earliest bifurcation during mammalian development. This revealed an unexpected role of a classical lineage regulator and its link to a cytokine pathway that is typically associated with an inflammatory response of immune cells.

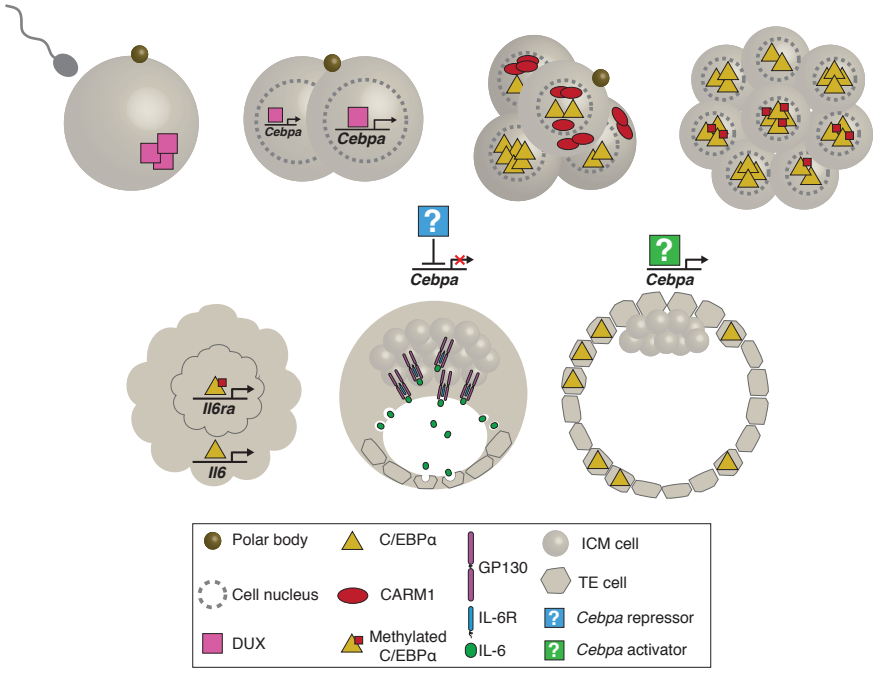


Fig. D2| Model of the role of C/EBPα and IL-6 signaling in pre-implantation mouse development. *Cebpa* is first expressed at the 2-cell stage, possibly by the direct binding and activation by the maternal TF DUX. At the 4-cell stage C/EBPα is asymmetrically expressed, with the highest levels located in the blastomere at the vegetal pole. This blastomere is fated to turn into TE. C/EBPα levels and/or its arginine methylation state mediated by CARM1 determine TE and ICM commitment in the transition between morula and blastocyst, with *Il6* being expressed in the outer layer destined to become TE and *Il6ra* in the inner layer, fated to become ICM.

CONCLUSIONS

From the results obtained in this thesis project we draw the following conclusions:

1. C/EBPα directly upregulates the *Il6, Il6ra and Stat3 genes* in B cells and activates its downstream signaling effectors.
2. IL-6 signaling is dispensable for B cell to macrophage transdifferentiation but strictly required for C/EBPα-enhanced reprogramming to iPS cells.
3. During iPSC reprogramming C/EBPα transiently triggers paracrine IL-6 signaling among different B cell subsets, preceding LIF signaling.
4. *Il6* depletion alters the expression of genes associated with pluripotency, trophectoderm (TE) and cell senescence in B cell reprogramming.
5. There are two waves of *Cebpa* expression in pre-implantation mouse embryos: the first starts at the 2-cell stage and reaches maximum levels in 8-cell embryos; and the second restricts C/EBPα detection to the TE in late blastocysts.
6. DUX is a likely early transcriptional regulator of *Cebpa* expression during embryogenesis.
7. A high C/EBPα protein level asymmetrically detected among 4-cell blastomeres predicts TE cell fate specification.

8. Depletion of *Cebpa* partially impairs pre-implantation development *in vitro* and reveals few potential target genes not compensated by *Cebpb* upregulation.
9. Blocking IL-6 protein secreted by cultured blastocysts delays the morula to blastocyst transition, suggesting a non-redundant role with LIF.
10. The enrichment of *Il6* expression in TE and of *Il6ra* in the inner cell mass (ICM) suggests that the TE is a niche for ICM formation.
11. C/EBPα overexpression in embryonic stem cells induces the formation of TE-like and ICM-like subpopulations, supporting a role of the factor as a critical player in embryonic TE/ICM specification.

MATERIALS AND METHODS

Mice.

To obtain B cells for the transdifferentiation experiments we used 8-12 week old C57BL/6J mice, isolated their bone marrow and used a biotin-conjugated anti-CD19 antibody (553784, BD Biosciences) together with streptavidin microbeads (130-048-101, Miltenyi Biotec) to isolate CD19 positive cells through magnetic columns (130-042-401, Miltenyi Biotec). These cells, which were used for further studies and designated as “B cells”, represent committed B-lymphocyte lineage cells consisting of pro-B and pre-B cells. As a source of the primary B cells used in our reprogramming experiments, we crossed ‘‘OSKM-reprogrammable mice’’ containing a doxycycline-inducible OSKM cassette and the tetracycline transactivator[243] with an OCT4-GFP reporter strain[244], as previously described[23,36]. We further crossed these reprogrammable reporter strain with *Il6*$^{-/-}$ mice[104] kindly provided by Dr. Manuel Serrano’s laboratory (IRB). Mouse embryo fibroblasts (MEFs) used as feeders were derived either from E13.5 embryos of C57BL/6J and *Il6*$^{-/-}$ mice. To study pre-implantation development *in vitro*, B6CBAF1/Crl females (purchased from Charles River laboratories) were super-ovulated by injecting pregnant mare’s serum gonadotropin (100 µL of 50UI/mL PMSG, Foligon) followed by human chorionic gonadotropin (100 µL of 5UI/mL hCG, Veterin Corion) after 48 hours. Females were then mated with B6CBAF1/Crl males and zygotes harvested from swollen

ampullas 20 hours after hCG injection. A line of heterozygous *Cebpa*$^{+/-}$ mice[67] was obtained from Dr. Achim Leutz (MDC, Berlin) and processed in the same way. Mice were housed in standard cages under 12 hours light–dark cycles and fed ad libitum with a standard chow diet. All experiments were approved by the Ethics Committee of the Barcelona Biomedical Research Park (PRBB) and performed according to Spanish and European legislation.

Cell culture.

B cell transdifferentiation into macrophages.

Transdifferentiation of primary B cells isolated from the bone marrow of wild type and *Il6*$^{-/-}$ mice was performed as previously described[6]. Briefly, B cells were infected with C/EBPα-ER-hCD4 retroviruses and hCD4 positive cells were plated at 500 cells/cm^2 in gelatinized 24-well plates onto mitomycin-C (M0503, Sigma) treated MEFs (10µg/mL mitomycin-C for 3 hours to inactivate MEFs). Cells were grown in RPMI culture medium (12633012, GIBCO) containing 20%-FBS (10270-106, GIBCO), which was further supplemented with 10 ng/mL each of IL-7 (217-17, Preprotech), IL-3 (213-13, Preprotech), FLT-3 (250-31, Preprotech), mCSF-1 (315-03B, Preprotech) and 100 nM β-estradiol (3301, Merck Millipore) to shuttle C/EBPα into the cell nucleus. The final culture medium also contained 100

U/mL Penicillin- 100 ng/mL Streptomycin (15140122, GIBCO), 2 mM L-Glutamine (25030081, GIBCO) and 0.1 mM 2-Mercaptoethanol (31350010, Invitrogen).

Culture medium was renewed at day 3 with the same composition but without IL-7. Cells were harvested at the indicated time points by trypsinization to be processed for FACS analysis or phagocytosis assays.

B cell reprogramming to iPSCs.

Reprogramming of primary B cells from wild type or *Il6*$^{-/-}$ reprogrammable/OCT4-GFP mice was performed as previously described[23,36]. Briefly, freshly isolated B cells were infected with C/EBPα-ER-hCD4 retrovirus and hCD4 positive cells were plated at 500 cells/cm^2 in gelatinized 12-well plates onto wild type or *Il6*$^{-/-}$ inactivated MEF feeders in 20%-FBS RPMI medium with 10 ng/mL IL-7. Transcription factor was activated by the addition of 100 nM β-estradiol for 18 hours. After β-estradiol washout, the cultures were switched to N2B27 medium (50% DMEM-F12 (12634010, GIBCO), 50% Neurobasal (21103049, GIBCO), 100X N2 supplement (17502048, GIBCO) and 50X B27 supplement (17504044, GIBCO)) in addition to 10 ng/ml IL-4 (217-14, Preprotech, 10ng/mL IL-7 and 2 ng/mL IL-15 (210-15, Preprotech). The final culture medium also contained 100X MEM Non-Essential Amino Acid solution (11140068, GIBCO), 1 mM sodium

pyruvate (11360070, GIBCO), 100 U/mL Penicillin- 100 ng/mL Streptomycin, 2 mM L-Glutamine and 0.1 mM 2-Mercaptoethanol.

OSKM cassette was activated by the addition of 2 µg/mL of doxycycline (D9891, Sigma). Samples needed for western blots or RNA-sequencing were harvested by trypsinization followed by removal of feeder cells through differential adherence to tissue culture dishes for 40 minutes. iPSC colony formation was assessed at the end of the process by alkaline phosphatase staining as detailed below.

- IL-6 neutralization and rescue experiments.

Anti-IL-6 blocking treatments were performed by adding 0.1 mg/ml of the antibody BE0046 (BioXCell) to the cultures. For rescue experiments, an additional 20 ng/ml of mouse recombinant IL-6 protein (406-ML, R&D Systems) was added. Both neutralization and rescue experiments consisted in adding the compounds daily from the pulse of C/EBPα until the end of the processes.

Pre-implantation mouse embryo cultures.

After collecting fertilized zygotes from the ampulla, cumulus cells were removed by incubation with 300 µg/mL hyaluronidase (H4272, Sigma) in M2 medium (M7167,

Sigma). After washing the embryos in a few drops of KSOM medium (MR-106-D, Millipore), they were cultured in KSOM microdrops under mineral oil (NO-400K, Nidacon) in an incubator with 5% CO_2 at 37°C. Embryos were handled with a mouth aspirator (A5177-5EA, Sigma) coupled to fire-polished glass Pasteur pipettes and collected at different stages of development from the *in vitro* cultures either for RNA-sequencing or protein immunostaining as detailed in the sections below. To track their development in control or IL-6 blocking conditions, phase contrast pictures of the developing embryos were taken with a Leica inverted microscope (DMI6000B) and embryos were counted using Fiji software (source).

- IL-6 secretion assessment and neutralization.

The presence of IL-6 secreted into the microdrops was assessed using an ELISA kit (M6000B, R&D Systems) according to manufacturers' instructions. Emissions of the samples at 450 nm wavelength were measured in a plate reader (SPECTROstar Nano, BMG Labtech). Anti-IL-6 blocking treatments were performed by adding 0.1 mg/mL of antibody (BE0046, BioXCell) or anti-horseradish peroxidase antibody (BE0088, BioXCell) as a control in culture microdrops from the zygote stage up to late blastocysts.

Embryonic stem cells and infection with C/EBPα.

Embryonic stem cells (E14TG2) were cultured on gelatinized plates in 15% ES-FBS (16141079, Life Technologies) Knockout-medium (10829018, GIBCO) containing 1000 U/ml leukemia inhibitory factor (LIF) (ESG1106, Merck Millipore). For C/EBPα expression, cells were infected with inducible C/EBPα-ERT2-dTomato lentiviral vector and single cell clones were expanded after FACS-sorting. For C/EBPα induction, cells were treated with 1 μM 4-hydroxytamoxifen (4-OHT) (H7904, Sigma) that shuttle the factor into the cell nucleus. The final culture medium also contained 100 U/mL Penicillin- 100 ng/mL Streptomycin, 2 mM L-Glutamine, 0.1 mM 2-Mercaptoethanol, 100X MEM Non-Essential Amino Acid solution and 1 mM sodium pyruvate.

Trophoblast stem cells

Trophoblast stem cell lines (TSCs) were originally established in Janet Rossant's laboratory and kindly provided by Dr. Manuel Irimia (CRG). They were cultured in feeder-free conditions as previously described[245]. Briefly, TSCs were grown in 20%-FBS RPMI medium mixed with conditioned medium in a 3:7 proportion. To obtain feeder-conditioned medium, mitomycin-C inactivated MEFs were cultured in the presence of 20%-FBS RPMI, the supernatants collected from

day 3 to 9 of culture and cleared through 0.45 µm filters (SLHV033RB, Millex). The final culture medium also contained 100 U/mL Penicillin- 100 ng/mL Streptomycin, 2 mM L-Glutamine, 0.1 mM 2-Mercaptoethanol, 100X MEM Non-Essential Amino Acid solution and 1 mM sodium pyruvate.

Both ESCs and TSCs were further used for immunostaining, RNA isolation and FACS-sorting as detailed in the following sections.

Virus production and cell infection.

Lentiviruses used for transduction of ESCs.

Lentiviruses were produced by transfecting HEK-293T cells with 6 µg of pCMV-VSV-G, 15 µg of pCMVDR-8.91, and 20 µg of the phage Ef1α-C/EBPα-ERT2-IRES-dTomato (sequence available upon request) plasmids using the calcium phosphate transfection method. Briefly, calcium phosphate-DNA precipitates were prepared by pooling the upper amounts of the three plasmids in a 2.5M $CaCl_2$ aqueous solution. While vortexing, one volume of the calcium phosphate-DNA solution was added dropwise to an equal volume of HBS 2X (HEPES-buffered saline solution pH = 7.05, 280 mM NaCl, 0.05 M HEPES and 1.5 mM Na_2HPO_4).

The mixture was incubated for 15 minutes at room temperature and added dropwise to HEK-293T cells grown in 10%-FBS DMEM (12491015, GIBCO) onto 150 mm dishes. After 16 hours of incubation at 37°C, the transfection medium was replaced with fresh 20%-FBS DMEM and the supernatant collected after 24 hours. The medium was replaced again and a second round of supernatant was collected after another 24 hours and mixed with the previous batch. The combined supernatants were centrifuged for 5 min at 300 rcf and filtered through 0.45 µm strainers to remove cell debris. Lentiviral particles were then concentrated by centrifugation for 2 hours at 20,000 rcf (Optima L-100K, Beckman Coulter) in round bottom polypropylene tubes (326823, Beckman Coulter). After discarding the supernatants, the lentiviral pellets obtained from one 150 mm dish were thoroughly re-suspended in 80 µL of PBS. 10^6 freshly trypsinized ESCs were then collected in 100 µL of 15%-ES-FBS Knockout-DMEM and 5 µL of lentiviral suspension were added. Subsequently, the virus-cell mixture was centrifuged at 1,000 rcf for 2 hours at 32°C (Allegra X-30R, Beckman Coulter). Infected ESCs were then cultured as described above and subsequently FACS-sorted for the establishment of clonal cell lines. DMEM medium for HEK-293T cells was further supplied with 100 U/mL Penicillin- 100 ng/mL Streptomycin and 2 mM L-Glutamine.

Retroviruses used for transduction of hematopoietic cells.

The C/EBPα-ER-hCD4 retroviral vector was generated as described before[246]. Viral production with platinum E cells (RV-101, Cell Biolabs) and infection of B cells also were performed as previously described[23,36].

Protein extraction and Western Blotting.

Whole B cell extracts were re-suspended in lysis buffer and quantified using Bradford solution (Bio-Rad) for absorbance measurements (UV/VIS photo-spectrometer) at 595 nm wavelength. 40 µg of sample were mixed with 4X Laemmli buffer (1610747, Bio-Rad) and boiled for 7 min at 96°C. Samples were loaded in 10% Mini-PROTEAN TGX gels (Bio-Rad) and resolved by electrophoresis in running buffer. Protein samples were transferred to a methanol pre-activated PVDF membrane (1620177, Bio-Rad) by running them in transfer buffer for 1 hour at 300 mA and 4°C. Membranes were rinsed in milliQ water and protein transfer was checked by Ponceau staining (Sigma). Transferred membranes were washed once with transfer buffer and three times in TBS-Tween (TBST) for 10 minutes followed by a block of unspecific antibody staining in 5% milk in TBST for 45 minutes. Membranes were then incubated with primary antibodies in 5% milk TBST rotating overnight at 4°C. Next

morning, membranes were washed three times for 10 minutes with TBST followed by incubation with the secondary antibodies conjugated to horseradish peroxidase in 5% milk TBST for 1 hour rotating at room temperature. After three TBST washes for 10 minutes, proteins were detected using enhanced chemiluminescence reagents (Amersham ECL Prime Western Blotting detection) in an Amersham Imager 600 analyzer. Quantification of band intensity from scanned blots was performed with Fiji software. See tables MM1 and MM2 for buffers and table MM3 for antibodies used.

Lysis buffer
50mM Tris-HCl pH=7.4
1% Triton X-100
0.1% SDS
Cocktail protease inhibitors

Running buffer	Transfer buffer	TBST
25mM Tris-base	25mM Tris-HCl H=8.3	10mM Tris-HCl pH=7.5
200mM glycine	200nM glycine	100mM NaCl
0.1% SDS	20% methanol	0.1% Tween 20

Antibody	Company	Catalogue	Species	Dilution
β-Actin	Sigma	A1978	Mouse	1:2000
Gapdh	Abcam	Ab8245	Mouse	1:5000
Stat3	Cell Signaling	9139	Mouse	1:1000
Stat3-P	Abcam	Ab76315	Rabbit	1:5000
Anti-rabbit	Bethyl	A120-201P	Goat	1:20000
Anti-mouse	Bethyl	A90-116P	Goat	1:10000

Tables MM1, MM2 and MM3. Chemical reagents used to prepare buffers and washing solutions (MM1-MM2). Antibodies used for protein detection in Wester Blotting experiments (MM3).

Immunostaining.

B cells and intermediate stages during transdifferentiation towards iMacs.

Cell cultures were trypsinized and centrifuged at 300 rcf for 5 minutes. Cells were re-suspended in 100 µL PBS containing 1 µg/mL of mouse Fc block (purified rat anti-mouse CD16/CD32 553141, BD Pharmingen) for 10 minutes. Conjugated primary antibodies were added to the blocking solution and cells were further incubated at 4°C in the dark for 20 minutes. Cells were washed with additional 500 µL of PBS and centrifuged at 300 rcf for 5 minutes. The supernatant was discarded and cells were re-suspended in 300 µL of PBS containing 5 µg/mL of DAPI. Samples were processed in a FACS analyzer (LSR II, BD) with DiVa software and data analyzed using FlowJo software.

Pre-implantation mouse embryos.

Embryos were fixed in 4% PFA for 10 minutes at room temperature. They were then washed twice in PBS for 5 minutes before permeabilization with 0.5% Triton X-100 PBS (0.5% PBST). Zygotes to morulas were permeabilized for 10 minutes while blastocysts were permeabilized for 15 minutes. Embryos were washed twice in 0.1% PBST for 5 minutes

before incubation in 0.1% PBST containing 3% bovine serum albumin (BSA) (9048-46-8, Sigma) for 45 minutes at room temperature to block unspecific immunostaining. Embryos were then treated with primary antibodies diluted in 0.1% PBST containing 1% BSA overnight at 4°C inside a moistened chamber. Next morning, embryos were sequentially washed in 0.1% PBST for 5, 15, 20 and 30 minutes at room temperature. A second blocking was performed in 0.1% PBST containing 3% BSA for 45 minutes at room temperature. For the secondary staining, embryos were placed in 0.1% PBST containing 1% BSA with the corresponding antibodies and 5 µg/mL of DAPI (D1306, Invitrogen). Embryos were left in secondary staining solution for 90 minutes at room temperature inside a moistened chamber in the dark. Three washes in 0.1% PBST were performed before mounting the embryos in 10 µL drops of PBS on 35 mm coverglass plates (P35G-1.0-14-C, MatTek) covered in light oil (M5310, Sigma). Embryos were imaged in a Leica TCS SPE or SP5 inverted confocal microscopes and further processed in Fiji and Imaris software for intensity quantifications and distance measurements. Of note, all the incubation steps were performed on shaking platforms.

Mouse embryonic and trophoblast stem cells.

Cell cultures were trypsinized and fixed in 4 % PFA for 10 minutes on a spinning wheel. Tubes were centrifuged at 300

rcf, pellets re-suspended in 75% ethanol and spun onto charged glass slides (SuperFrost plus 10149870, Thermo Scientific) by centrifugation at 800 rpm for 5 minutes (Cystospin 4 Shandon, Thermo Fisher). Areas with cells were circled with a hydrophobic marker (Mini PAP Pen 008877, Thermo Fisher). Cells were permeabilized with 0.1% PBST for 10 minutes at room temperature, washed twice in PBS for 5 minutes and incubated in 0.1% PBST and 3% BSA at room temperature for 45 minutes inside a moistened chamber. Cells were then incubated with primary antibodies in staining solution with 0.1% PBST and 1% BSA overnight at 4°C in a moistened chamber. The next morning, cells were washed twice in PBS and incubated with secondary antibodies and 5 µg/mL of DAPI in staining solution with 0.1% PBST and 1% BSA for 60 minutes at room temperature in the dark in a moistened chamber. Cells were subsequently washed twice in PBS for 5 minutes. Excess PBS was removed with a piece of blotting paper and after adding 14 µL of mounting medium (7µL Dako (S3023, Agilent) + 7µL 0.1% PBST) a glass coverslip (0101030, Marienfeld Superior) was placed on top and the structure was sealed with nail polish. Imaging of the cells was done in a Leica TCS SPE inverted confocal microscope and further processed using Fiji software. All washing and staining steps were performed using a hand pipette to add or remove 20 µL of the indicated solutions to the cells within the hydrophobic circle. See attached tables MM4 and MM5 for used antibodies.

FACS				
Antibody	**Company**	**Catalogue**	**Species**	**Dilution**
CD19	BD Pharmingen	550992	Rat	1:200
Mac-1	BD Pharmingen	552850	Rat	1:200
hCD4	BD Pharmingen	555347	Mouse	1:20

Immunofluorescence				
Antibody	**Company**	**Catalogue**	**Species**	**Dilution**
C/EBPα	Cell Signaling	8178	Rabbit	1:100
OCT4	Santa Cruz	Sc-5279	Mouse	1:100
KRT8	DSHB	N/A	Rat	1:20
EOMES	Abcam	Ab23345	Rabbit	1:200
Anti-rabbit	ThermoFisher	A-11070	Goat	1:1000
Anti-mouse	ThermoFisher	A-21236	Goat	1:1000
Anti-rat	ThermoFisher	A-11081	Goat	1:1000

Tables MM4 and MM5. Antibodies used against cell surface markers detected by FACS (MM4). Antibodies used against intracellular proteins detected by confocal microscopy (MM5).

Phagocytosis assay.

Around ~200,000 B cells per well in 12-well plates and the induced macrophages (iMacs) after 5 days of transdifferentiation were cultured overnight in the presence of 1:1000 diluted carboxylate microspheres (Fluoresbrite 17458-10) added to the culture medium.

Fluorescence imaging and FACS analysis.

Following previous removal of feeder inactivated MEFs through differential adherence to tissue culture dishes for 40 minutes, B cells and trypsinized iMacs were transferred to

other wells in 12-well plates containing 0.01% poly-L-lysine (Sigma) treated coverslips. Cells were centrifuged at 300 rcf for 5 minutes and, in the case of iMacs, left in fresh medium for additional 2 hours in the well so that induced macrophages would better attach to the coverslips. The supernatant was removed and the cells were washed once with PBS. For fixation, 4% PFA was added to the wells for 20 minutes, cells were washed twice with PBS and cell membranes permeabilized with 0.1% Triton X-100 PBS (0.1% PBST) for 15 minutes at room temperature. Actin filaments were subsequently stained with 1:100 diluted red phalloidin (Alexa Fluor 568, Thermo Fischer Scientific, A12380) while DNA was stained with a 1:500 diluted yellow probe (Quant-iT PicoGreen dsDNA Assay Kit, Thermo Fischer Scientific, P7589). Cells were incubated with the two dyes in 0.1% PBST containing 1% BSA at room temperature for 1 hour in the dark and washed twice with PBS afterwards. Coverslips carrying the the attached cells in the well were then recovered with tweezers and mounted upside-down onto a charged glass slide containing a 14 µL drop of mounting medium (7µL Dako + 7µL 0.1% PBST). Coverslips were sealed with nail polish and imaged in a Leica TCS SPE inverted confocal microscope. Image analysis and processing was done using Fiji software.

Cells were also used for FACS analysis by staining with CD19-APC and Mac-1-Pe-Cy7 antibodies as described in the

transdifferentiation protocol but without adding DAPI. Cells were analyzed using an indo violet laser to excite and detect bright blue emission from the carboxylated beads. In the case of B cells $CD19^+$/$Mac\text{-}1^-$ cells were selected while $Mac\text{-}1^+$/$CD19^-$ cells were gated to analyze iMacs with the indo violet laser. Samples were processed in a LSR II equipment with DiVa software and further analyzed using FlowJo software.

FACS-sorting of ESCs and TSCs.

For the establishment of cell clones expressing C/EBPα-ERT2-IRES-dTomato, ESCs were infected with a lentiviral vector and grown in culture for 3 days before sorting. Bulk ESCs were trypsinized and re-suspended in medium with 5 µg/mL of DAPI at a concentration of 10^7 cells/mL. Then they were sorted based on red fluorescence at one cell per well in 96-well plates using a FACSAria II SORP instrument. Most wells developed into colonies that were further expanded, generating the C/EBPα-inducible cell lines “clone 1” and “clone 2” referred in the results.

To sort cells based on SSEA-1 expression, TSCs, uninduced ESCs and C/EBPα-induced ESCs (48 hours treatment with 1µM 4-OHT) were trypsinized. 10^7 cells were re-suspended in 400 µL PBS and incubated with 1 µg/mL of Fc block, 12 µg/mL of APC-conjugated SSEA-1 antibody (50-8813-42,

Thermo Fisher) and 5 µg/mL of DAPI for 20 minutes at 4°C in the dark. To remove the staining solution cells were centrifuged for 5 minutes at 300 rcf, the pellets re-suspended in 1 mL of culture medium and separated with a FACSAria II SORP sorter into SSEA-1 high and low cell fractions. RNA from these fractions was further isolated and gene expression was assessed by qRT-PCR.

Scoring and counting of iPSC colonies.

Reprogrammable B cells cultured in 12-well plates containing feeders MEFs were analyzed at day 12 after OSKM induction. Cultures were washed twice with 0.05% PBS-Tween 20 (PBSTw), fixed for 2 minutes in 4% PFA and washed once again in 0.05% PBSTw. Cells were then incubated in freshly prepared alkaline phosphatase staining solution at room temperature in the dark for 10-20 minutes and washed twice in PBS. Alkaline phosphatase expression was detected by the purple color of the iPSC colonies. Plates were scanned (Perfection V850 Pro, Epson) and colonies were counted in Fiji software. See table MM6 for iPSC staining solution composition.

Alkaline phosphatase staining solution
100mM Tris-HCl pH=8.8
100mM NaCl
50mM $MgCl_2$
NBT/BCIP 1:100 (11681451001, Roche)

Table MM6. Chemical reagents used to freshly prepare alkaline phosphatase staining solution to detect iPSC colonies.

RNA isolation, quantitative RT-PCR and RNA-seq.

RNA was extracted with a miRNeasy mini kit (217004, Qiagen), quantified with a NanoDrop spectrophotometer and its quality examined in a fragment Bioanalyzer (Aligent 2100 Bioanalyzer DNA 7500 assay). cDNA was synthesized with a High Capacity RNA-to-cDNA kit (4387406, Applied Biosystems). qRT–PCR analyses were performed in technical triplicate reactions with SYBR Green qPCR Master Mix (4367659, Applied Biosystems) and run in a Viia7 Real-Time PCR instrument (Applied Biosystems). For RNA-sequencing, libraries were prepared with a TruSeq Stranded mRNA Library Preparation Kit (Illumina) followed by single-end sequencing (50 bp) on an HiSeq2500 instrument (Illumina), obtaining at least 40 million reads per sample. See table MM7 for qPCR primer list.

For the isolation of RNA from 8-cell and late blastocysts, RNA was extracted and retro-transcribed into cDNA using a SMART-Seq v4 Ultra Low Input RNA kit (634894, Takara). RNA concentration and quality was determined as above. Libraries for RNA-sequencing were prepared as previously described[247] and sequenced as above.

Gene	Direction	5' to 3' sequence
Eomes	Forward	GGCTGAGCCTGAGAGTCAAG
	Reverse	AGAGCCAGCCCTACAACAAA
Gata3	Forward	CTGGAGGAGGAACGCTAATGG
	Reverse	CATCTTCCGGTTTCGGGTCT
Cdx2	Forward	CTGCCACACTTGGGCTCT
	Reverse	CTTGGCTCTGCGGTTCTG
Krt8	Forward	TGGAAGGACTGACCGACGAGAT
	Reverse	GGCACGAACTTCAGCGATGATG
Id2	Forward	GGACCTGCAGATCGCCCTGG
	Reverse	TCAGATGCCTGGAAGGACAGGATG
Nanog	Forward	CAGTTTTTCATCCCGAGAAC
	Reverse	CTTTTGTTTGGGACTGGTAG
Sall4	Forward	AAGAACTTCTCGTCTGCC
	Reverse	AGTGTACCTTCAGGTTGC
Oct4	Forward	GTCCCTAGGTGAGCCGTCTTT
	Reverse	AGTCTGAAGCCAGGTGTCCAG
Esrrb	Forward	AAAGCCATTGACTAAGATCG
	Reverse	AATTCACAGAGAGTGGTCAG
Tdh	Forward	CAGACTGAAGATAAAAGGCAG
	Reverse	GCATCTGTTCTTCTGATACC

Table MM7. 5' to 3' sequences of forward and reverse primers used in qRT-PCR analyses to assess pluripotency and trophectodermal gene expression.

Computational analysis of RNA-seq data.

Reads were mapped using STAR[248] (standard options) and the Ensembl mouse genome annotation version mm10vM21. Gene expression was quantified using STAR (--quantMode GeneCounts). Sample scaling and statistical analysis were performed using the R package DESeq2[249] (R 3.3.2 and Bioconductor 3.0). Statistical power of gene expression variation at any given time point was identified using the nbinomLRT test. Log2-vsd (variance stabilized DESeq2)

counts were used for further analysis unless stated otherwise. Clustering was performed using the K-means method.

Data mining of public datasets.

Some of the data analyzed in this project were retrieved from available public data sets (see Table MM8). Single-cell expression trajectories and correlations in B cell transdifferentiation and reprogramming were processed by Romain Bulteau and Mirko Francesconi (LBMC, Lyon). *Il6* and *Il6ra* expression in the ICM and trophectodermal mouse cells was processed by Dr. Chris Wyatt at the CRG. More precisely, available single cell gene expression data[228] was re-analyzed and gene expression tables were taken from Wyatt et al. (unpublished). These cells were processed using *vast-tools* (align module), which provides a normalized count measure for each gene (cRPKM, corrected (for mappability) Reads Per Kilobase of transcript per Million mapped reads[250]). For blastocysts, individual cells were clustered into ICM or TE using specific transcription factor markers[200] with samples that did not cluster well being excluded from the analysis. RNA obtained in transdifferentiation and reprogramming experiments were processed and analyzed as described above.

Fig.	Content	Publication	ID
R1	Early RNA expression in B cells after C/EBPα overexpression	(Torcal, G. et al. In preparation)	N/A
R2	C/EBPα, H3K27ac, H3K4me3 ChIP-seq in B cells and Bα'	(Di Stefano, B. et al. 2016)	GEO: GSE71218
	Virtual 4C, H3K4me2 ChIP-seq, ATAC-seq and RNA-seq in B cells and Bα'	(Stadhouders, R. et al. 2018)	GEO: GSE96611
R5.1	RNA expression in B cell to macrophage transdifferentiation	(Torcal, G. et al. In preparation)	N/A
	RNA expression in B cell to iPSC reprogramming	(Stadhouders, R. et al. 2018)	GEO: GSE96611
R9.1,2	Single-cell RNA-seq in B cell transdifferentiation and reprogramming	(Francesconi, M. et al. 2019)	GEO: GSE112004
R10a	C/EBPα RNA expression in pre-implantation mouse development	(Guo, G. et al. 2010)	MGI, J:140465
R11.1	*Il6* and *Il6ra* RNA expression in single ICM and trophectodermal mouse cells	(Deng, Q. et al. 2014)	GEO: GSE45719
R13	C/EBPα RNA expression and ATAC-seq in Dux-induced mouse ESCs	(Hendrickson, P.G. et al, 2017)	GEO: GSE85632

Table MM8. Collection of public datasets analyzed in the project.

Statistics.

Statistical analyses were performed using Prism 9 software. To be able to calculate significance, samples from at least 3 biologically independent experiments were analyzed. Two biological replicates were used for RNA-sequencing experiments and statistics applied to the expression of a collection of genes (see Table MM9). For samples with $n \geq 3$, values shown in the figures represent median ± standard deviation with 10-90 percentile boxplots and whiskers. One-

way, two-way ANOVA (with the corresponding multiple comparison analyses) and Student's t-tests were applied accordingly. P values appear indicated in each figure.

B cell genes – Figure R7.2a left
Ebf1, Foxo1, Ikzf1, Spi1, Rag1, Rag2, Irf4, Irf8, Il7r, Bcl11a, Spib, Ikzf3, Pou2f2, Cd2, Cd19, Igll1, Vpreb1, Vpreb3, Vpreb2, Pou2af1, Blk, Cd79a, Cd79b, Lef1, Msh2, Igkc, Tlr9, Myb, Card11, Hdac9, Ms4a1, Ighm, Bcl2, Tcf3, Ankle1, Mybl2, Pax5.
Myeloid genes – Figure R7.2a right
Jun, Csf1, Tspan2, Notch2, Efna4, Fcer1g, Sell, Lbr, Psen2, Id1, Id2, Itga4, Casp8, Creb1, Tgfbr2, Ccr1, Mitf, Rbpj, Inpp4b, Tlr2, Pik3r1, Csf1r, Mapk14, Trem2, Vegfa, Cd109, Ostm1, L3mbtl3, Cited2, Fam20c, Snx10, Ikzf1, Klf10, Trib1, Hhex, Adam8, Cd81, Traf6, Spi1, Esrra, Batf2, Tcirg1, Gab2, Ltbr, Il23a, Tesc, Rb1, Pou4f1, Bmp4, Zfp36l1, Psen1, Fos, Batf2, Fes, Itgal, Itgam, Cbfa2t3, Pafah1b1, Dhrs7b, Ccl3, Rara, Pecam1, Prkca, Cd300lf, Sbno2, Zbtb7a, Tnfsf9, Icam1, Junb, Cebpa, Tyrobp, Relb, Spib, Rassf2, Src, Mafb, Cebpb, Runx1, Itgb2, Tfe3, Atp6ap1, Gab3, Epha2, Klf4_UTR, Chd7, Cd14, Ifitm6, Irf8, C1qc.
Pluripotency genes – Figure R7.2b left
Esrrb, Klf2, Tfcp2l1, Sall4, Klf5, Tbx3, Nr0b1, Zfp42, Lin28a, Gdf3, Tdh, Fbxo15, Nr5a2, Prdm14, Dppa5a, Dppa3, Fgf4, Lefty1, Lefty2, Dnmt3b, Dnmt3l, Nodal, Morc1, Pou5f1_UTR, Sox2_UTR, Klf4_UTR, Myc_UTR, Tex19.1, Eras,Nf2, Pcgf6, Slc2a3, Raf1, Dido1, Capn10, Tcl1, Nanog.
Trophectodermal genes – Figure R7.2b right
Elf5, Cdh1, Spry4, Zic3, Ccne1, Fzd5, Krt8, Cebpb, Tead4, Bmp1, Tspan8, Atp12a, Grhl1, Grhl2, Aqp3, Krt14, Krt17, Krt13, Cldn3, Dsc2, Cdcp1, Cd40, Tfap2c.
Cell cycle genes – Figure R7.2c left
Abl1, Anapc1, Anapc11, Anapc2, Anapc4, Anapc5, Anapc7, Bub1b, Bub3, Ccnb2, Ccnd2, Ccnd1, Ccnd3, Ccne1, Ccne2, Ccnh, Cdc14b, Cdc23, Cdc25a , Cdc25b, Cdc26, Cdc6, Cdk2, Cdk4, Cdk6, Cdk7, Cdkn1b, Cdkn1c, Cdkn2a, Cdkn2c, Cdkn2d, Chek2, E2f1, E2f3, E2f2, E2f4, E2f5, Fzr1, Gadd45a, Gadd45b Hdac1, Mad2l2, Mcm3, Mcm4,

Mcm6, Mcm7, Mdm2, Myc_UTR, Orc3, Orc4, Orc5, Pcna, Pttg1, Rad21, Rb1, Rbl1, Rbl2, Sfn, Skp2, Smad3, Smad4, Smc1a, Smc3, Stag1, Stag2, Tfdp1, Tgfb1, Tgfb3, Wee1, Ywhag, Ywhah, Ywhaz, Zbtb17.
Senescence genes – Figure R7.2c right
Axl, Btc, Ccl2, Ccl4, Ccl6, Ccl7, Csf2ra, Csf3, Cxcl1, Cxcl13, Cxcl2, Cxcl3, Cxcl5, Egfr, Fgf2, Fgf7, Hgf, Icam1, Igfbp3, Igfbp4, Igfbp6, Igfbp2, Il11, Il15, Il1a, Il1b, Il1r1, Il2ra, Il6, Mif, Mst1, Osm, Pdgfb, Serpine1, Tac1, Tnfrsf18, Tnfrsf11b, Tnfrsf1a, Tnfrsf1b, Tubgcp2, Vegfa, Agrn, Cebpb, Cst3, Cstb, Fn1, Mmp14, Nt5e, Oaf, Pxdn, Sfrp1, Tfrc, Tgfb1, Thbs1, Mmp3, Cxcl10, Csf1, Il1rn, Cdkn2a.
Trophectodermal genes (Extended) – Figure R15.2b
Tfap2c, Fgfr2, Bmp4, Elf5, Krt1, Cdh1, Bptf, Spry4, Zic3, Ccne1, Fzd5, Krt9, Id2, Krt8, Cebpb, Krt18, Msx2, Tead4, Bmp1, Tspan8, Atp12a, Mbnl3, Lcp1, Grhl1, Grhl2, Aqp3, Krt19, Krt7, Krt14, Krt17, Krt13, Cldn3, Dsc2, Cdcp1, Cd40, Cdx2, Eomes, Furin, Dlx3, Esx1, Ets2, Hand1, Gcm1, Peg10, Tmprss2, Bmp8a, Wnt6, Dppa1, Cdkn1a, Atpaf1, Atp1b1, Akt3, Rab13, Gata3.
Pluripotency genes – Figure R15.2b
Esrrb, Klf2, Tfcp2l1, Sall4, Klf5, Tbx3, Nr0b1, Zfp42, Lin28a, Gdf3, Tdh, Fbxo15, Nr5a2, Prdm14, Dppa5a, Dppa3, Fgf4, Lefty1, Lefty2, Dnmt3b, Dnmt3l, Nodal, Morc1, Pou5f1_UTR, Sox2_UTR, Klf4_UTR, Myc_UTR, Tex19.1, Eras,Nf2, Pcgf6, Slc2a3, Raf1, Dido1, Capn10, Tcl1, Nanog.

Table MM9. Gene signatures obtained from published collections. Trophectodermal genes were extended for C/EBPα induction in ESCs since the number of differentially expressed genes was grater than in *Il6*$^{-/-}$ reprogramming.

REFERENCES

1. Davis, R. L., Weintraub, H. & Lassar, A. B. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51**, 987–1000 (1987).
2. Lassar, A. B., Weintraub, H. & Paterson, B. M. Transfection of a DNA locus that mediate the conversion of 10T1/2 fibroblasts to myoblasts. *Cell* **47**, 649–656 (1986).
3. Graf, T. & Enver, T. Forcing cells to change lineages. *Nature* **462**, 587–594 (2009).
4. Graf, T. Historical origins of transdifferentiation and reprogramming. *Cell Stem Cell* **9**, 504–516 (2011).
5. Kulessa, H., Frampton, J., and Graf, T. GATA-1 reprograms avian myelomonocytic cell lines into eosinophils, thromboblasts, and erythroblasts. *Genes Dev.* **9**, 1250–1262 (1995).
6. Xie, H., Ye, M., Feng, R. & Graf, T. Stepwise reprogramming of B cells into macrophages. *Cell* **117**, 663–676 (2004).
7. Feng, R. *et al.* PU.1 and C/EBPα/β convert fibroblasts into macrophage-like cells. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 6057–6062 (2008).
8. Rolink, A. G., Nutt, S. L., Melchers, F. & Busslinger, M. Long-term in vivo reconstitution of T-cell development by Pax5-deficient B-cell progenitors. *Nature* **401**, 603–606 (1999).
9. Li, P. *et al.* Reprogramming of T cells to natural killer-like cells upon Bcl11b deletion. *Science (80-.).* **329**, 85–89 (2010).
10. Kajimura, S. *et al.* Initiation of myoblast to brown fat switch by a PRDM16-C/EBP-β transcriptional complex. *Nature* **460**, 1154–

1158 (2009).
11. Zhou, Q., Brown, J., Kanarek, A., Rajagopal, J. & Melton, D. A. In vivo reprogramming of adult pancreatic exocrine cells to β-cells. *Nature* **455**, 627–632 (2008).
12. Ieda, M. *et al.* Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. *Cell* **142**, 375–386 (2010).
13. Vierbuchen, T. *et al.* Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **463**, 1035–1041 (2010).
14. Gurdon, J. B., Elsdale, T. R. & Fischberg, M. Sexually mature individuals of Xenopus laevis from the transplantation of single somatic nuclei. *Nature* **182**, 64–65 (1958).
15. Takahashi, K. & Yamanaka, S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* **126**, 663–676 (2006).
16. Maekawa, M. *et al.* Direct reprogramming of somatic cells is promoted by maternal transcription factor Glis1. *Nature* **474**, 225–228 (2011).
17. Feng, B. *et al.* Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nat. Cell Biol.* **11**, 197–203 (2009).
18. Han, J. *et al.* Tbx3 improves the germ-line competency of induced pluripotent stem cells. *Nature* **463**, 1096–1100 (2010).
19. Hanna, J. *et al.* Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* **462**, 595–601 (2009).
20. Hochedlinger, K. & Plath, K. Epigenetic reprogramming and induced pluripotency.

Development **136**, 509–523 (2009).
21. Eminli, S. *et al.* Differentiation stage determines potential of hematopoietic cells for reprogramming into induced pluripotent stem cells. *Nat. Genet.* **41**, 968–976 (2009).
22. Hanna, J. *et al.* Direct Reprogramming of Terminally Differentiated Mature B Lymphocytes to Pluripotency. *Cell* **133**, 250–264 (2008).
23. Di Stefano, B. *et al.* C/EBPα poises B cells for rapid reprogramming into induced pluripotent stem cells. *Nature* **506**, 235–239 (2014).
24. Kleinsmith, L. J. and G. B. P. Multipotentiality of Single Embryonal Carcinoma Cells. *Cancer Res.* **24**, 1544–1551 (1964).
25. Miller, R. A. & Ruddle, F. H. Pluripotent Teratocarcinoma-Thymus Somatic Cell Hybrids. *Cell* **9**, 45–55 (1976).
26. Kaufman, M. H. & Evans, M. J. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154–156 (1981).
27. Martin, G. R. 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**, 7634–7638 (1981).
28. Tada, M., Takahama, Y., Abe, K., Nakatsuji, N. & Tada, T. Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. *Curr. Biol.* **11**, 1553–1558 (2001).
29. Blau, H. M., Chiu, C. P. & Webster, C. Cytoplasmic activation of human nuclear genes in stable heterocaryons. *Cell* **32**, 1171–1180 (1983).
30. Taranger, C.K., Noer, A., Sorensen, A.L.,

Hakelien, A.M., Boquest, A.C., A. & Collas, P. Induction of dedifferentiation, genomewide transcriptional programming, and epigenetic reprogramming by extracts of carcinoma and embryonic stem cells. *Mol. Biol. Cell* **16**, 5719–5735 (2005).

31. Singhal, N. *et al.* Chromatin-remodeling components of the baf complex facilitate reprogramming. *Cell* **141**, 943–955 (2010).
32. Yamanaka, S. Elite and stochastic models for induced pluripotent stem cell generation. *Nature* **460**, 49–52 (2009).
33. Soufi, A., Donahue, G. & Zaret, K. S. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* **151**, 994–1004 (2012).
34. Polo, J. M. *et al.* A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell* **151**, 1617–1632 (2012).
35. Apostolou, E. & Hochedlinger, K. Chromatin dynamics during cellular reprogramming. *Nature* **502**, 462–471 (2013).
36. Di Stefano, B. *et al.* C/EBPα creates elite cells for iPSC reprogramming by upregulating Klf4 and increasing the levels of Lsd1 and Brd4. *Nat. Cell Biol.* **18**, 371–381 (2016).
37. Sardina, J. L. *et al.* Transcription Factors Drive Tet2-Mediated Enhancer Demethylation to Reprogram Cell Fate. *Cell Stem Cell* **23**, 727-741.e9 (2018).
38. Stadhouders, R. *et al.* Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming. *Nat. Genet.* **50**, 238–249 (2018).

39. O'Malley, J. *et al.* High-resolution analysis with novel cell-surface markers identifies routes to iPS cells. *Nature* **499**, 88–91 (2013).
40. Buganim, Y. *et al.* Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell* **150**, 1209–1222 (2012).
41. Biddy, B. A. *et al.* Single-cell mapping of lineage and identity in direct reprogramming. *Nature* **564**, 219–224 (2018).
42. Guo, L. *et al.* Resolving Cell Fate Decisions during Somatic Cell Reprogramming by Single-Cell RNA-Seq. *Mol. Cell* **73**, 815-829.e7 (2019).
43. Schiebinger, G. *et al.* Optimal-Transport Analysis of Single-Cell Gene Expression Identifies Developmental Trajectories in Reprogramming. *Cell* **176**, 928-943.e22 (2019).
44. Liu, X. *et al.* Reprogramming roadmap reveals route to human induced trophoblast stem cells. *Nature* **586**, 101–107 (2020).
45. Francesconi, M. *et al.* Single cell RNA-seq identifies the origins of heterogeneity in efficient cell transdifferentiation and reprogramming. *Elife* **8**, 1–22 (2019).
46. Friedman, A. D. C/EBPα in normal and malignant myelopoiesis. *Int. J. Hematol.* **101**, 330–341 (2015).
47. Ohlsson, E., Schuster, M. B., Hasemann, M. & Porse, B. T. The multifaceted functions of C/EBPα in normal and malignant haematopoiesis. *Leukemia* **30**, 767–775 (2016).
48. Pulikkan, J. A., Tenen, D. G. & Behre, G. C/EBPα deregulation as a paradigm for leukemogenesis. *Leukemia* **31**, 2279–2285 (2017).

49. Graves, B. J., Johnson, P. F. & McKnight, S. L. Homologous recognition of a promoter domain common to the MSV LTR and the HSV tk gene. *Cell* **44**, 565–576 (1986).
50. Johnson, P. F., Landschulz, W. H., Graves, B. J. & McKnight, S. L. Identification of a rat liver nuclear protein that binds to the enhancer core element of three animal viruses. *Genes Dev.* **1**, 133–146 (1987).
51. Friedman, A. D., Landschulz, W. H. & McKnight, S. L. CCAAT/enhancer binding protein activates the promoter of the serum albumin gene in cultured hepatoma cells. *Genes Dev.* **3**, 1314–1322 (1989).
52. Landschulz, W. H., Johnson, P. F. & McKnight, S. L. The leucine zipper: A hypothetical structure common to a new class of DNA binding proteins. *Science (80-.).* **240**, 1759–1764 (1988).
53. Landschulz, W. H., Johnson, P. F. & McKnight, S. L. The DNA binding domain of the rat liver nuclear protein C/EBP is bipartite. *Science (80-.).* **243**, 1681–1688 (1989).
54. Nerlov, C. & Ziff, E. B. Three levels of functional interaction determine the activity of CCAAT/enhancer binding protein-α on the serum albumin promoter. *Genes Dev.* **8**, 350–362 (1994).
55. Zhang, P. *et al.* Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBPα. *Immunity* **21**, 853–863 (2004).
56. Iwasaki, H. *et al.* Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. *Blood* **106**,

1590–1600 (2005).
57. Heinz, S. *et al.* Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cell* **38**, 576–589 (2010).
58. Hasemann, M. S. *et al.* C/EBPα Is Required for Long-Term Self-Renewal and Lineage Priming of Hematopoietic Stem Cells and for the Maintenance of Epigenetic Configurations in Multipotent Progenitors. *PLoS Genet.* **10**, (2014).
59. Ohlsson, E. *et al.* Initiation of MLL-rearranged AML is dependent on C/EBPα. *J. Exp. Med.* **211**, 5–13 (2014).
60. Rapino, F. *et al.* C/EBPα Induces Highly Efficient Macrophage Transdifferentiation of B Lymphoma and Leukemia Cell Lines and Impairs Their Tumorigenicity. *Cell Rep.* **3**, 1153–1163 (2013).
61. McClellan, J. S., Dove, C., Gentles, A. J., Ryan, C. E. & Majeti, R. Reprogramming of primary human Philadelphia chromosome-positive B cell acute lymphoblastic leukemia cells into nonleukemic macrophages. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 4074–4079 (2015).
62. Chronis, C. *et al.* Cooperative Binding of Transcription Factors Orchestrates Reprogramming. *Cell* **168**, 442-459.e20 (2017).
63. Begay, V., Smink, J. & Leutz, A. Essential Requirement of CCAAT/Enhancer Binding Proteins in Embryogenesis. *Mol. Cell. Biol.* **24**, 9744–9751 (2004).
64. Huang, D. *et al.* The role of Cdx2 as a lineage specific transcriptional repressor for pluripotent

network during the first developmental cell lineage segregation. *Sci. Rep.* **7**, 1–10 (2017).
65. Lefterova, M. I. *et al.* PPARγ and C/EBP factors orchestrate adipocyte biology via adjacent binding on a genome-wide scale. *Genes Dev.* **22**, 2941–2952 (2008).
66. Madsen, M. S., Siersbaek, R., Boergesen, M., Nielsen, R. & Mandrup, S. Peroxisome Proliferator-Activated Receptor and C/EBP Synergistically Activate Key Metabolic Adipocyte Genes by Assisted Loading. *Mol. Cell. Biol.* **34**, 939–954 (2014).
67. Wang, N. D. *et al.* Impaired energy homeostasis in C/EBPα knockout mice. *Science (80-.).* **269**, 1108–1112 (1995).
68. Jakobsen, J. S. *et al.* Temporal mapping of CEBPA and CEBPB binding during liver regeneration reveals dynamic occupancy and specific regulatory codes for homeostatic and cell cycle gene batteries. *Genome Res.* **23**, 592–603 (2013).
69. Porse, B. T. *et al.* E2F repression by C/EBPα is required for adipogenesis and granulopoiesis in vivo. *Cell* **107**, 247–258 (2001).
70. Martínez-Balbás, M. A., Dey, A., Rabindran, S. K., Ozato, K. & Wu, C. Displacement of sequence-specific transcription factors from mitotic chromatin. *Cell* **83**, 29–38 (1995).
71. Tang, Q. Q. & Lane, M. D. Activation and centromeric localization of CCAAT/enhancer-binding proteins during the mitotic clonal expansion of adipocyte differentiation. *Genes Dev.* **13**, 2231–2241 (1999).
72. Caravaca, J. M. *et al.* Bookmarking by specific and nonspecific binding of FoxA1 pioneer factor

to mitotic chromosomes. *Genes Dev.* **27**, 251–260 (2013).
73. Karagianni, P., Moulos, P., Schmidt, D., Odom, D. T. & Talianidis, I. Bookmarking by Non-pioneer Transcription Factors during Liver Development Establishes Competence for Future Gene Activation. *Cell Rep.* **30**, 1319-1328.e6 (2020).
74. Hyun, K., Jeon, J., Park, K. & Kim, J. Writing, erasing and reading histone lysine methylations. *Exp. Mol. Med.* **49**, (2017).
75. Wysocka, J. *et al.* WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* **121**, 859–872 (2005).
76. Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B. & Cavalli, G. Genome Regulation by Polycomb and Trithorax Proteins. *Cell* vol. 128 735–745 (2007).
77. Guccione, E. & Richard, S. The regulation, functions and clinical relevance of arginine methylation. *Nat. Rev. Mol. Cell Biol.* **20**, 642–657 (2019).
78. Shi, Y. *et al.* Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**, 941–953 (2004).
79. Chang, B., Chen, Y., Zhao, Y. & Bruick, R. K. JMJD6 is a histone arginine demethylase. *Science (80-.).* **318**, 444–447 (2007).
80. Walport, L. J. *et al.* Arginine demethylation is catalysed by a subset of JmjC histone lysine demethylases. *Nat. Commun.* **7**, (2016).
81. Dey, A., Chitsaz, F., Abbasi, A., Misteli, T. & Ozato, K. The double bromodomain protein Brd4 binds to acetylated chromatin during

interphase and mitosis. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 8758–8763 (2003).
82. Filippakopoulos, P. *et al.* Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell* **149**, 214–231 (2012).
83. Moon, K. J. *et al.* The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Mol. Cell* **19**, 523–534 (2005).
84. Bararia, D. *et al.* Proteomic identification of the MYST domain histone acetyltransferase TIP60 (HTATIP) as a co-activator of the myeloid transcription factor C/EBPα. *Leukemia* **22**, 800–807 (2008).
85. Roe, J. S., Mercan, F., Rivera, K., Pappin, D. J. & Vakoc, C. R. BET Bromodomain Inhibition Suppresses the Function of Hematopoietic Transcription Factors in Acute Myeloid Leukemia. *Mol. Cell* **58**, 1028–1039 (2015).
86. Wang, G. L. *et al.* HDAC1 cooperates with C/EBPα in the inhibition of liver proliferation in old mice. *J. Biol. Chem.* **283**, 26169–26178 (2008).
87. Liu, G. F. *et al.* C/EBPα negatively regulates SIRT7 expression via recruiting HDAC3 to the upstream-promoter of hepatocellular carcinoma cells. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1859**, 348–354 (2016).
88. Grebien, F. *et al.* Pharmacological targeting of the Wdr5-MLL interaction in C/EBPα N-terminal leukemia. *Nat. Chem. Biol.* **11**, 571–578 (2015).
89. Liu, L. M. *et al.* Methylation of C/EBPa by PRMT1 inhibits its tumor-suppressive function

in breast cancer. *Cancer Res.* **79**, 2865–2877 (2019).
90. Kowenz-Leutz, E., Pless, O., Dittmar, G., Knoblich, M. & Leutz, A. Crosstalk between C/EBPB phosphorylation, arginine methylation, and SWI/SNF/Mediator implies an indexing transcription factor code. *EMBO J.* **29**, 1105–1115 (2010).
91. Pedersen, T. Å., Kowenz-Leutz, E., Leutz, A. & Nerlov, C. Cooperation between C/EBPα TBP/TFIIB and SWI/SNF recruiting domains is required for adipocyte differentiation. *Genes Dev.* **15**, 3208–3216 (2001).
92. Müller, C., Calkhoven, C. F., Sha, X. & Leutz, A. The CCAAT Enhancer-binding Protein α (C/EBPα) Requires a SWI/SNF Complex for Proliferation Arrest. *J. Biol. Chem.* **279**, 7353–7358 (2004).
93. Smith, Z. D. & Meissner, A. DNA methylation: Roles in mammalian development. *Nat. Rev. Genet.* **14**, 204–220 (2013).
94. Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247–257 (1999).
95. Leonhardt, H., Page, A. W., Weier, H. U. & Bestor, T. H. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* **71**, 865–873 (1992).
96. Yin, Y. *et al.* Impact of cytosine methylation on DNA binding specificities of human transcription factors. *Science (80-.).* **356**, (2017).
97. Kallin, E. M. *et al.* Tet2 facilitates the derepression of myeloid target genes during

CEBPα-Induced transdifferentiation of Pre-B cells. *Mol. Cell* **48**, 266–276 (2012).
98. Ko, M. *et al.* Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. *Nature* **497**, 122–126 (2013).
99. Hu, L. *et al.* Crystal Structure of TET2-DNA Complex: Insight into TET-Mediated 5mC Oxidation. *Cell* **155**, 1545–1555 (2013).
100. de la Rica, L. *et al.* PU.1 target genes undergo Tet2-coupled demethylation and DNMT3b-mediated methylation in monocyte-to-osteoclast differentiation. *Genome Biol.* **14**, R99 (2013).
101. Costa, Y. *et al.* NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature* **495**, 370–374 (2013).
102. Hirano, T. *et al.* Complementary DNA for a novel human interleukin (BSF-2) that induces B lymphocytes to produce immunoglobulin. *Nature* **324**, 73–76 (1986).
103. Hunter, C. A. & Jones, S. A. IL-6 as a keystone cytokine in health and disease. *Nat. Immunol.* **16**, 448–457 (2015).
104. Kopf, M. *et al.* Impaired immune and acute-phase responses in interleukin-6-deficient mice. *Nature* **368**, 339–342 (1994).
105. Rebouissou, S. *et al.* Frequent in-frame somatic deletions activate gp130 in inflammatory hepatocellular tumours. *Nature* **457**, 200–204 (2009).
106. Qing, H. *et al.* Origin and Function of Stress-Induced IL-6 in Murine Models. *Cell* **182**, 372-387.e14 (2020).
107. Taga, T. *et al.* Interleukin-6 triggers the association of its receptor with a possible signal

transducer, gp130. *Cell* **58**, 573–581 (1989).
108. Skiniotis, G., Boulanger, M. J., Garcia, K. C. & Walz, T. Signaling conformations of the tall cytokine receptor gp130 when in complex with IL-6 and IL-6 receptor. *Nat. Struct. Mol. Biol.* **12**, 545–551 (2005).
109. Jones, G. W. *et al.* Loss of CD4 + T Cell IL-6R Expression during Inflammation Underlines a Role for IL-6 Trans Signaling in the Local Maintenance of Th17 Cells . *J. Immunol.* **184**, 2130–2139 (2010).
110. Yoshida, K. *et al.* Targeted disruption of gp130, a common signal transducer for the interleukin 6 family of cytokines, leads to myocardial and hematological disorders. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 407–411 (1996).
111. Heinrich, P. C. *et al.* Principles of interleukin (IL)-6-type cytokine signalling and its regulation. *Biochem. J.* **374**, 1–20 (2003).
112. Boeuf, H., Hauss, C., De Graeve, F., Baran, N. & Kedinger, C. Leukemia inhibitory factor-dependent transcriptional activation in embryonic stem cells. *J. Cell Biol.* **138**, 1207–1217 (1997).
113. Niwa, H., Burdon, T., Chambers, I. & Smith, A. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev.* **12**, 2048–2060 (1998).
114. Raz, R., Lee, C. K., Cannizzaro, L. A., D'Eustachio, P. & Levy, D. E. Essential role of STAT3 for embryonic stem cell pluripotency. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 2846–2851 (1999).
115. Rose-John, S. & Heinrich, P. C. Soluble receptors for cytokines and growth factors:

Generation and biological function. *Biochem. J.* **300**, 281–290 (1994).
116. JONES, S. A., HORIUCHI, S., TOPLEY, N., YAMAMOTO, N. & FULLER, G. M. The soluble interleukin 6 receptor: mechanisms of production and implications in disease. *FASEB J.* **15**, 43–58 (2001).
117. Garbers, C. *et al.* Species specificity of ADAM10 and ADAM17 proteins in interleukin-6 (IL-6) trans-signaling and novel role of ADAM10 in inducible IL-6 receptor shedding. *J. Biol. Chem.* **286**, 14804–14811 (2011).
118. Make Peters, Stephan Jacobs, Marc Ehlers, PetraVollmer, Jiirgen Mfillberg, Eckhard Wolffl Gottfried Brem, K.-H. M. zum B. and S. R.-J. The Function of the Soluble Interleukin 6 (IL-6) Receptor In Vivo: Sensitization of Human Soluble IL-6 Receptor Transgenic Mice Towards IL-6 and Prolongation of the Plasma Half-life of IL-6. *J. Exp. Med.* **183**, 1399–1406 (1996).
119. Jostock, T. *et al.* Soluble gp130 is the natural inhibitor of soluble interleukin-6 receptor transsignaling responses. *Eur. J. Biochem.* **268**, 160–167 (2001).
120. Poli, V. The role of C/EBP isoforms in the control of inflammatory and native immunity functions. *J. Biol. Chem.* **273**, 29279–29282 (1998).
121. Delgobo, M. *et al.* An evolutionary recent IFN/IL-6/CEBP axis is linked to monocyte expansion and tuberculosis severity in humans. *Elife* **8**, 1–32 (2019).
122. Akira, S. *et al.* A nuclear factor for IL-6 expression (NF-IL6) is a member of a C/EBP

family. *EMBO J.* **9**, 1897–1906 (1990).
123. Matsusaka, T. *et al.* Transcription factors NF-IL6 and NF-κB synergistically activate transcription of the inflammatory cytokines, interleukin 6 and interleukin 8. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 10193–10197 (1993).
124. Müller, C., Kowenz-Leutz, E., Grieser-Ade, S., Graf, T. & Leutz, A. NF-M (chicken C/EBPβ) induces eosinophilic differentiation and apoptosis in a hematopoietic progenitor cell line. *EMBO J.* **14**, 6127–6135 (1995).
125. Leutz, A. *et al.* Molecular cloning of the chicken myelomonocytic growth factor (cMGF) reveals relationship to interleukin 6 and granulocyte colony stimulating factor. *EMBO J.* **8**, 175–181 (1989).
126. Screpanti, I. *et al.* Lymphoproliferative disorder and imbalanced T-helper response in C/EBP beta-deficient mice. *EMBO J.* **14**, 1932–1941 (1995).
127. Wongchana, W. & Palaga, T. Direct regulation of interleukin-6 expression by Notch signaling in macrophages. *Cell. Mol. Immunol.* **9**, 155–162 (2012).
128. Xiao, W. *et al.* NF-κB activates IL-6 expression through cooperation with c-Jun and IL6-AP1 site, but is independent of its IL6-NFκB regulatory site in autocrine human multiple myeloma cells. *Cancer Biol. Ther.* **3**, 1007–1017 (2004).
129. Mackey, S. L. & Darlington, G. J. CCAAT Enhancer-binding Protein α Is Required for Interleukin-6 Receptor α Signaling in Newborn Hepatocytes. *J. Biol. Chem.* **279**, 16206–16213 (2004).

130. Cantwell, C. A., Sterneck, E. & Johnson, P. F. Interleukin-6-Specific Activation of the C/EBPδ Gene in Hepatocytes Is Mediated by Stat3 and Sp1. *Mol. Cell. Biol.* **18**, 2108–2117 (1998).
131. Davoust, J., Palucka, A. K., Chomarat, P. & Banchereau, J. IL-6 switches the differentiation of monocytes from dendritic cells to macrophages. *Nat. Immunol.* **1**, 510–514 (2000).
132. Jennifer Nichols, I. C. and A. S. Derivatoin of Germline Competent Embryonic Stem Cells with a Combination of Interleukin-6 and Soluble Interleukin-6 Receptor. *Exp. Cell Res.* **215**, 237–239 (1994).
133. Brady, J. J. *et al.* Early role for IL-6 signalling during generation of induced pluripotent stem cells revealed by heterokaryon RNA-Seq. *Nat. Cell Biol.* **15**, 1244–1252 (2013).
134. Mosteiro, L. *et al.* Tissue damage and senescence provide critical signals for cellular reprogramming in vivo. *Science (80-.).* **354**, (2016).
135. Kuilman, T. & Peeper, D. S. Senescence-messaging secretome: SMS-ing cellular stress. *Nat. Rev. Cancer* **9**, 81–94 (2009).
136. Chiche, A. *et al.* Injury-Induced Senescence Enables In Vivo Reprogramming in Skeletal Muscle. *Cell Stem Cell* **20**, 407-414.e4 (2017).
137. Gerwin, N., Jia, G. Q., Kulbacki, R. & Gutierrez-Ramos, J. C. Interleukin Gene Expression in Mouse Preimplantation Development. *Dev. Immunol.* **4**, 169–179 (1995).
138. Do, D. V. *et al.* A genetic and developmental pathway from STAT3 to the OCT4-NANOG circuit is essential for maintenance of ICM

lineages in vivo. *Genes Dev.* **27**, 1378–1390 (2013).
139. Bourillot, P. Y., Santamaria, C., David, L. & Savatier, P. GP130 signaling and the control of naïve pluripotency in humans, monkeys, and pigs. *Exp. Cell Res.* **386**, 111712 (2020).
140. Rossant, J. & Tam, P. P. L. Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. *Development* **136**, 701–713 (2009).
141. Chazaud, C. & Yamanaka, Y. Lineage specification in the mouse preimplantation embryo. *Dev.* **143**, 1063–1074 (2016).
142. Kelly, S. J. Studies of the developmental potential of 4- and 8-cell stage mouse blastomeres. *J. Exp. Zool.* **200**, 365–376 (1977).
143. Leung, C. Y. & Zernicka-Goetz, M. Mapping the journey from totipotency to lineage specification in the mouse embryo. *Curr. Opin. Genet. Dev.* **34**, 71–76 (2015).
144. Takaoka, K. & Hamada, H. Cell fate decisions and axis determination in the early mouse embryo. *Development* **139**, 3–14 (2012).
145. Li, L., Zheng, P. & Dean, J. Maternal control of early mouse development. *Development* **137**, 859–870 (2010).
146. Schulz, K. N. & Harrison, M. M. Mechanisms regulating zygotic genome activation. *Nat. Rev. Genet.* **20**, 221–234 (2019).
147. Murchison, E. P. *et al.* Critical roles for Dicer in the female germline. *Genes Dev.* **21**, 682–693 (2007).
148. Suzumori, N., Burns, K. H., Yan, W. & Matzuk, M. M. RFPL4 interacts with oocyte proteins of

the ubiquitin-proteasome degradation pathway. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 550–555 (2003).
149. Hirasawa, R. *et al.* Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. *Genes Dev.* **22**, 1607–1616 (2008).
150. Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247–257 (1999).
151. Payer, B. *et al.* stella Is a Maternal Effect Gene Required for Normal Early Development in Mice. *Curr. Biol.* **13**, 2110–2117 (2003).
152. Foygel, K. *et al.* A novel and critical role for Oct4 as a regulator of the maternal-embryonic transition. *PLoS One* **3**, (2008).
153. Gabriëls, J. *et al.* Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. *Gene* **236**, 25–32 (1999).
154. De Iaco, A. *et al.* DUX-family transcription factors regulate zygotic genome activation in placental mammals. *Nat. Genet.* **49**, 941–945 (2017).
155. Whiddon, J. L., Langford, A. T., Wong, C. J., Zhong, J. W. & Tapscott, S. J. Conservation and innovation in the DUX4-family gene network. *Nat. Genet.* **49**, 935–940 (2017).
156. Hendrickson, P. G. *et al.* Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat. Genet.* **49**, 925–934

(2017).
157. Wu, G. *et al.* Establishment of totipotency does not depend on Oct4A. *Nat. Cell Biol.* **15**, 1089–1097 (2013).
158. Chen, Z. & Zhang, Y. Loss of DUX causes minor defects in zygotic genome activation and is compatible with mouse development. *Nat. Genet.* **51**, 947–951 (2019).
159. Oldfield, A. J. *et al.* Histone-Fold Domain Protein NF-Y Promotes Chromatin Accessibility for Cell Type-Specific Master Transcription Factors. *Mol. Cell* **55**, 708–722 (2014).
160. Lu, F. *et al.* Establishing chromatin regulatory landscape during mouse preimplantation development. *Cell* **165**, 1375–1388 (2016).
161. Tarkowski, A. K. Experiments on the development of isolated blastomeres of mouse eggs. *Nature* **184**, 1286–1287 (1959).
162. Ribet, D. *et al.* Murine Endogenous Retrovirus MuERV-L Is the Progenitor of the “Orphan” Epsilon Viruslike Particles of the Early Mouse Embryo. *J. Virol.* **82**, 1622–1625 (2008).
163. Macfarlan, T. S. *et al.* Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**, 57–63 (2012).
164. Rodriguez-Terrones, D. *et al.* A molecular roadmap for the emergence of early-embryonic-like cells in culture. *Nat. Genet.* **50**, 106–119 (2018).
165. Ishiuchi, T. *et al.* Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly. *Nat. Struct. Mol. Biol.* **22**, 662–671 (2015).
166. Zheng, Z., Li, H., Zhang, Q., Yang, L. & Qi, H. Unequal distribution of 16S mtrRNA at the 2-

cell stage regulates cell lineage allocations in mouse embryos. *Reproduction* **151**, 351–367 (2016).
167. Wang, J. *et al.* Asymmetric Expression of LincGET Biases Cell Fate in Two-Cell Mouse Embryos. *Cell* **175**, 1887-1901.e18 (2018).
168. Casser, E. *et al.* Totipotency segregates between the sister blastomeres of two-cell stage mouse embryos. *Sci. Rep.* **7**, 1–15 (2017).
169. Tarkowski, A. K. & Wróblewska, J. Development of blastomeres of mouse eggs isolated at the 4- and 8-cell stage. *J. Embryol. Exp. Morphol.* **18**, 155–180 (1967).
170. Chen, Q., Shi, J., Tao, Y. & Zernicka-Goetz, M. Tracing the origin of heterogeneity and symmetry breaking in the early mammalian embryo. *Nat. Commun.* **9**, 1–11 (2018).
171. Piotrowska-Nitsche, K., Perea-Gomez, A., Haraguchi, S. & Zernicka-Goetz, M. Four-cell stage mouse blastomeres have different developmental properties. *Development* **132**, 479–490 (2005).
172. Torres-Padilla, M. E., Parfitt, D. E., Kouzarides, T. & Zernicka-Goetz, M. Histone arginine methylation regulates pluripotency in the early mouse embryo. *Nature* **445**, 214–218 (2007).
173. Plachta, N., Bollenbach, T., Pease, S., Fraser, S. E. & Pantazis, P. Oct4 kinetics predict cell lineage patterning in the early mammalian embryo. *Nat. Cell Biol.* **13**, 117–123 (2011).
174. White, M. D. *et al.* Long-Lived Binding of Sox2 to DNA Predicts Cell Fate in the Four-Cell Mouse Embryo. *Cell* **165**, 75–87 (2016).
175. Goolam, M. *et al.* Heterogeneity in Oct4 and

Sox2 Targets Biases Cell Fate in 4-Cell Mouse Embryos. *Cell* **165**, 61–74 (2016).
176. Strumpf, D. *et al.* Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. *Development* **132**, 2093–2102 (2005).
177. Kuzmichev, A. N. *et al.* Sox2 acts through Sox21 to regulate transcription in pluripotent and differentiated cells. *Curr. Biol.* **22**, 1705–1710 (2012).
178. Lim, H. Y. G. *et al.* Keratins are asymmetrically inherited fate determinants in the mammalian embryo. *Nature* **585**, 404–409 (2020).
179. Panamarova, M. *et al.* The BAF chromatin remodelling complex is an epigenetic regulator of lineage specification in the early mouse embryo. *Development* **143**, 1271–1283 (2016).
180. Stephenson, R. O., Yamanaka, Y. & Rossant, J. Disorganized epithelial polarity and excess trophectoderm cell fate in preimplantation embryos lacking E-cadherin. *Development* **137**, 3383–3391 (2010).
181. Fierro-González, J. C., White, M. D., Silva, J. C. & Plachta, N. Cadherin-dependent filopodia control preimplantation embryo compaction. *Nat. Cell Biol.* **15**, 1424–1433 (2013).
182. de Vries, W. N. *et al.* Maternal β-catenin and E-cadherin in mouse development. *Development* **131**, 4435–4445 (2004).
183. Maître, J. L. *et al.* Asymmetric division of contractile domains couples cell positioning and fate specification. *Nature* **536**, 344–348 (2016).
184. Anani, S., Bhat, S., Honma-Yamanaka, N., Krawchuk, D. & Yamanaka, Y. Initiation of Hippo signaling is linked to polarity rather than

to cell position in the pre-implantation mouse embryo. *Dev.* **141**, 2813–2824 (2014).
185. Korotkevich, E. *et al.* The Apical Domain Is Required and Sufficient for the First Lineage Segregation in the Mouse Embryo. *Dev. Cell* **40**, 235-247.e7 (2017).
186. Royer, C. *et al.* Establishment of a relationship between blastomere geometry and YAP localisation during compaction. *Development* **147**, (2020).
187. Nishioka, N. *et al.* The Hippo Signaling Pathway Components Lats and Yap Pattern Tead4 Activity to Distinguish Mouse Trophectoderm from Inner Cell Mass. *Dev. Cell* **16**, 398–410 (2009).
188. Skamagki, M., Wicher, K. B., Jedrusik, A., Ganguly, S. & Zernicka-Goetz, M. Asymmetric Localization of Cdx2 mRNA during the First Cell-Fate Decision in Early Mouse Development. *Cell Rep.* **3**, 442–457 (2013).
189. Ralston, A. *et al.* Gata3 regulates trophoblast development downstream of Tead4 and in parallel to Cdx2. *Development* **137**, 395–403 (2010).
190. Yagi, R. *et al.* Transcription factor TEAD4 specifies the trophectoderm lineage at the beginning of mammalian development. *Development* **134**, 3827–3836 (2007).
191. Cockburn, K., Biechele, S., Garner, J. & Rossant, J. The hippo pathway member nf2 is required for inner cell mass specification. *Curr. Biol.* **23**, 1195–1201 (2013).
192. Wicklow, E. *et al.* HIPPO Pathway Members Restrict SOX2 to the Inner Cell Mass Where It Promotes ICM Fates in the Mouse Blastocyst.

PLoS Genet. **10**, (2014).
193. Rayon, T. *et al.* Notch and Hippo Converge on Cdx2 to Specify the Trophectoderm Lineage in the Mouse Blastocyst. *Dev. Cell* **30**, 410–422 (2014).
194. Kaneko, K. J. & De Pamphilis, M. L. TEAD4 establishes the energy homeostasis essential for blastocoel formation. *Dev.* **140**, 3680–3690 (2013).
195. Dupont, S. *et al.* Role of YAP/TAZ in mechanotransduction. *Nature* **474**, 179–184 (2011).
196. Zenker, J. *et al.* Expanding Actin Rings Zipper the Mouse Embryo for Blastocyst Formation. *Cell* **173**, 776-791.e17 (2018).
197. Zhu, M., Wang, P., Handford, C. E., Na, J. & Zernicka-Goetz, M. Transcriptional control of apical protein clustering drives de novo cell polarity establishment in the early mouse embryo. *bioRxiv* (2020).
198. Yang, J. *et al.* Binding of FGF2 to FGFR2 in an autocrine mode in trophectoderm cells is indispensable for mouse blastocyst formation through PKC-p38 pathway. *Cell Cycle* **14**, 3318–3330 (2015).
199. Kurimoto, K. *et al.* An improved single-cell cDNA amplification method for efficient high-density oligonucleotide microarray analysis. *Nucleic Acids Res.* **34**, (2006).
200. Guo, G. *et al.* Resolution of Cell Fate Decisions Revealed by Single-Cell Gene Expression Analysis from Zygote to Blastocyst. *Dev. Cell* **18**, 675–685 (2010).
201. Ohnishi, Y. *et al.* Cell-to-cell expression variability followed by signal reinforcement

progressively segregates early mouse lineages. *Nat. Cell Biol.* **16**, 27–37 (2014).
202. Chazaud, C., Yamanaka, Y., Pawson, T. & Rossant, J. Early Lineage Segregation between Epiblast and Primitive Endoderm in Mouse Blastocysts through the Grb2-MAPK Pathway. *Dev. Cell* **10**, 615–624 (2006).
203. Yamanaka, Y., Lanner, F. & Rossant, J. FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. *Development* **137**, 715–724 (2010).
204. Krawchuk, D., Honma-Yamanaka, N., Anani, S. & Yamanaka, Y. FGF4 is a limiting factor controlling the proportions of primitive endoderm and epiblast in the ICM of the mouse blastocyst. *Dev. Biol.* **384**, 65–71 (2013).
205. Kang, M., Piliszek, A., Artus, J. & Hadjantonakis, A. K. FGF4 is required for lineage restriction and salt-and-pepper distribution of primitive endoderm factors but not their initial expression in the mouse. *Dev.* **140**, 267–279 (2013).
206. Bessonnard, S. *et al.* Gata6, Nanog and Erk signaling control cell fate in the inner cell mass through a tristable regulatory network. *Dev.* **141**, 3637–3648 (2014).
207. Frankenberg, S. *et al.* Primitive Endoderm Differentiates via a Three-Step Mechanism Involving Nanog and RTK Signaling. *Dev. Cell* **21**, 1005–1013 (2011).
208. Dietrich, J. E. & Hiiragi, T. Stochastic patterning in the mouse pre-implantation embryo. *Development* **134**, 4219–4231 (2007).
209. Niwa, H., Miyazaki, J. I. & Smith, A. G. Quantitative expression of Oct-3/4 defines

differentiation, dedifferentiation or self-renewal of ES cells. *Nat. Genet.* **24**, 372–376 (2000).
210. Ambrosetti, D. C., Schöler, H. R., Dailey, L. & Basilico, C. Modulation of the activity of multiple transcriptional activation domains by the DNA binding domains mediates the synergistic action of Sox2 and Oct-3 on the Fibroblast growth factor-4 enhancer. *J. Biol. Chem.* **275**, 23387–23397 (2000).
211. Graham, S. J. L. *et al.* BMP signalling regulates the pre-implantation development of extra-embryonic cell lineages in the mouse embryo. *Nat. Commun.* **5**, (2014).
212. Plusa, B., Piliszek, A., Frankenberg, S., Artus, J. & Hadjantonakis, A. K. Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. *Development* **135**, 3081–3091 (2008).
213. Saiz, N., Grabarek, J. B., Sabherwal, N., Papalopulu, N. & Plusa, B. Atypical protein kinase C couples cell sorting with primitive endoderm maturation in the mouse blastocyst. *Dev.* **140**, 4311–4322 (2013).
214. Madan, P., Rose, K. & Watson, A. J. Na/K-ATPase β1 subunit expression is required for blastocyst formation and normal assembly of trophectoderm tight junction-associated proteins. *J. Biol. Chem.* **282**, 12127–12134 (2007).
215. Barcroft, L. C., Offenberg, H., Thomsen, P. & Watson, A. J. Aquaporin proteins in murine trophectoderm mediate transepithelial water movements during cavitation. *Dev. Biol.* **256**, 342–354 (2003).
216. Wang, H. *et al.* Zonula occludens-1 (ZO-1) is

involved in morula to blastocyst transformation in the mouse. *Dev. Biol.* **318**, 112–125 (2008).
217. Ryan, A. Q., Chan, C. J., Graner, F. & Hiiragi, T. Lumen Expansion Facilitates Epiblast-Primitive Endoderm Fate Specification during Mouse Blastocyst Formation. *Dev. Cell* **51**, 684-697.e4 (2019).
218. Chan, C. J. & Hiiragi, T. Integration of luminal pressure and signalling in tissue self-organization. *Dev.* **147**, 1–10 (2020).
219. Tanaka, S., Kunath, T., Hadjantonakis, A. K., Nagy, A. & Rossant, J. Promotion to trophoblast stem cell proliferation by FGF4. *Science (80-.).* **282**, 2072–2075 (1998).
220. Niwa, H. *et al.* Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell* **123**, 917–929 (2005).
221. Kunath, T. *et al.* Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts. *Development* **132**, 1649–1661 (2005).
222. Brons, I. G. M. *et al.* Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**, 191–195 (2007).
223. Shahbazi, M. N. & Zernicka-Goetz, M. Deconstructing and reconstructing the mouse and human early embryo. *Nat. Cell Biol.* **20**, 878–887 (2018).
224. Harrison, S. E., Sozen, B., Christodoulou, N., Kyprianou, C. & Zernicka-Goetz, M. Assembly of embryonic and extraembryonic stem cells to mimic embryogenesis in vitro. *Science (80-.).* **356**, (2017).
225. Rivron, N. C. *et al.* Blastocyst-like structures generated solely from stem cells. *Nature* **557**,

106–111 (2018).
226. Li, R. *et al.* Generation of Blastocyst-like Structures from Mouse Embryonic and Adult Cell Cultures. *Cell* **179**, 687-702.e18 (2019).
227. Mahmoudi, S. *et al. Heterogeneity in old fibroblasts is linked to variability in reprogramming and wound healing*. *Nature* vol. 574 (2019).
228. Qiaolin Deng, Daniel Ramsköld, B. R. and R. S. Single-Cell RNA-Seq Reveals Dynamic, Random Monoallelic Gene Expression in Mammalian Cells. *Science (80-.).* **343**, 193–196 (2014).
229. Hiiragi, T., Louvet-Vallée, S., Solter, D. & Maro, B. Does prepatterning occur in the mouse egg? *Nature* **442**, 2005–2006 (2006).
230. Parr, B. A., Cornish, V. A., Cybulsky, M. I. & McMahon, A. P. Wnt7b regulates placental development in mice. *Dev. Biol.* **237**, 324–332 (2001).
231. Geissmann, F., Jung, S. & Littman, D. R. Blood monocytes consist of two principal subsets with distinct migratory properties. *Immunity* **19**, 71–82 (2003).
232. Wynn, T. A., Chawla, A. & Pollard, J. W. Macrophage biology in development, homeostasis and disease. *Nature* **496**, 445–455 (2013).
233. Lerin Luckett-Chastain, Kaitlin Calhoun, T. S. and R. M. G. IL-6 influences the balance between M1 and M2 macrophages in a mouse model of irritant contact dermatitis. *J. Immunol.* **196**, (2016).
234. Chen, L. *et al.* IL-6 influences the polarization of macrophages and the formation and growth of

colorectal tumor. *Oncotarget* **9**, 17443–17454 (2018).
235. Stik, G. *et al.* CTCF is dispensable for immune cell transdifferentiation but facilitates an acute inflammatory response. *Nat. Genet.* **52**, 655–661 (2020).
236. Gardner, R. L., Papaioannou, V. E. & Barton, S. C. Origin of the ectoplacental cone and secondary giant cells in mouse blastocysts reconstituted from isolated trophoblast and inner cell mass. *J. Embryol. Exp. Morphol.* **30**, 561–572 (1973).
237. Rossant, J. & Cross, J. C. Placental development: Lessons from mouse mutants. *Nat. Rev. Genet.* **2**, 538–548 (2001).
238. Rassoulzadegan, M., Rosen, B. S., Gillot, I. & Cuzin, F. Phagocytosis reveals a reversible differentiated state early in the development of the mouse embryo. *EMBO J.* **19**, 3295–3303 (2000).
239. Gerri, C. *et al.* Initiation of a conserved trophectoderm program in human, cow and mouse embryos. *Nature* **587**, 443–447 (2020).
240. Yokoyama, H. Chromatin-Binding Proteins Moonlight as Mitotic Microtubule Regulators. *Trends Cell Biol.* **26**, 161–164 (2016).
241. Somma, M. P. *et al.* Moonlighting in Mitosis: Analysis of the Mitotic Functions of Transcription and Splicing Factors. *Cells* **9**, 1–27 (2020).
242. Boija, A. *et al.* Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell* **175**, 1842-1855.e16 (2018).
243. Carey, B. W., Markoulaki, S., Beard, C., Hanna,

J. & Jaenisch, R. Single-gene transgenic mouse strains for reprogramming adult somatic cells. *Nat. Methods* **7**, 56–59 (2010).
244. Boiani, M., Eckardt, S., Schöler, H. R. & John McLaughlin, K. Oct4 distribution and level in mouse clones: Consequences for pluripotency. *Genes Dev.* **16**, 1209–1219 (2002).
245. Kubaczka, C. *et al.* Derivation and maintenance of murine trophoblast stem cells under defined conditions. *Stem Cell Reports* **2**, 232–242 (2014).
246. Bussmann, L. H. *et al.* A Robust and Highly Efficient Immune Cell Reprogramming System. *Cell Stem Cell* **5**, 554–566 (2009).
247. Picelli, S. *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–181 (2014).
248. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
249. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 1–21 (2014).
250. Labbé, R. M. *et al.* A Comparative transcriptomic analysis reveals conserved features of stem cell pluripotency in planarians and mammals. *Stem Cells* **30**, 1734–1745 (2012).

ANNEX I: PUBLICATIONS

A manuscript from this thesis' results is under preparation.

Francesconi, M., Di Stefano, B., Berenguer, C., de Andrés-Aguayo, L., **Plana-Carmona, M.**, Mendez-Lago, M., Guillaumet-Adkins, A., Rodriguez-Esteban, G., Gut, M., Gut, I.G., Heyn, H., Lehner, B. and Graf, T. Single cell RNA-seq identifies the origins of heterogeneity in efficient cell transdifferentiation and reprogramming. *Elife* **8**, 1–22 (2019). https://doi.org/10.7554/eLife.41627.001

Ghose, R., Aranguren-Ibáñez, A., Arecco, N., Balboa, D., Bataller, M., Beltran, S., Benisty, H., Bénard, A., Bernardo, E., Carbonell Sala, S., Casals, E., Ciampi, L., Condemi, L., Corvó, A., Cosín-Tomás, M., Cuenca-Ardura, M., Duran Serrano, J. M., Espejo Díaz, M. I., Fernandez Callejo, M., Gañez-Zapater, A., Garcia-Castellanos, R., Garrido, R., Henkin, G., Hermoso Pulido, T., Hernandez-Alias, X., Herrero Vicente, J., Ingham, M., Lim, W. M., Llonch, S., Marmesat Bertoli, E., Miguel-Escalada, I., Montero-Blay, A., Navarrete Hernández, C., Neguembor, M. V., Ní Chárthaigh, R.A., Pardo-Lorente, N., Pascual-Reguant, L., Pérez-Lluch, S., Perza, R., Pesaresi, M., Picó Amador, D., Pifarré, P., Piscia, D., **Plana-Carmona, M.**, Ponomarenko, J., Radusky, L., Rivero, E., Rogalska, M., Torcal Garcia, G. and Wojnacki, J. From research to rapid response: mass COVID-19 testing by volunteers at the Centre for Genomic Regulation [version 1; peer review: awaiting peer review]. *F1000Research* **9**:1336 (2020) https://doi.org/10.12688/f1000research.27497.1

ANNEX II: ABBREVIATIONS

TFs	transcription factors
GREs	gene regulatory elements
C/EBP	CCAAT/enhancer-binding protein
B cell	B lymphocyte
Bα'	B cell pulsed for 18 hours with C/EBPα
iMacs	C/EBPα-induced macrophages
GMPs	granulocyte/macrophage progenitors
MEFs	mouse embryonic fibroblasts
OSKM	OCT4, SOX2, KLF4 and MYC
Yamanaka factors	OCT4, SOX2, KLF4 and MYC
iPSCs	induced pluripotent stem cells
ECCs	embryonic carcinoma cells
ESCs	embryonic stem cells
TSCs	trophoblast stem cells
ICM	inner cell mass
TE	trophectoderm
EPI	epiblast
PrE	primitive endoderm
EpiSCs	epiblast-derived stem cells
XEN cells	extra embryonic endoderm cells
mRNA	messenger RNA
AP	alkaline phosphatase
Doxy	doxycycline
β-estr	β-estradiol
DMSO	dimethyl sulfoxide
4-OHT	4-hydroxytamoxifen
ER	estrogen receptor

ERT2	tamoxifen-inducible estrogen receptor 2
IL-6	interleukin 6
LIF	leukemia inhibitory factor
ChIP	chromatin immunoprecipitation
ATAC	assay for transposase-accessible chromatin
-seq	sequencing
4C	chromosome conformation capture-on-chip
bLZ	basic leucine-zipper
HDAC	histone deacetylases
HMT	histone methyltransferase
5-mC	5-methylcytosine
5-hmC	5-hydroxymethylcytosine
FACS	fluorescence-activated cell sorting
IF	immunofluorescence
F1	first offspring generation

ACKNOWLEDGEMENTS

Quiero agradecer este trabajo a mi familia. En especial a mi hermana Eva, mi madre y mi padre. Por recorrer los espinguetes, apartar las barceras y encontrar un camino.

Al pueblo de Esplús por inspirarme a buscar fuera de los límites. A mis amigas y amigos de siempre. A su escuela rural, al campo y algunas viejas costumbres.

Me acuerdo también de las profesoras, profesores y compañeros de clase del instituto. De todas las personas con quienes coincidí en mis viajes al extranjero aprendiendo inglés. Gracias por transmitirme y compartir el gusto por saber, las ganas de aprender y disfrutar pensando cada detalle.

A mi etapa universitaria por encender la mecha. A mis primeros compañeros de laboratorio por hacerme vibrar con la genética del desarrollo y aventurarme a querer más. A mis amigas y amigos de Barcelona. A los inicios Penya, los delirios en Còrsega 112 y a “Amada Barna” por enseñarme que la vida, la amistad y los secretos para hacer girar el mundo se esconden detrás de unas cañas (tapa). A toda la secta voleibolera del PRBB por descubrirme este deporte y traerme personas increíbles.

Thank you Thomas for accepting me in your group and to guide me through this period. Thanks for inspiring me with

your endless passion for science and nourish my curiosity to break the secrets of nature. To all the people in the laboratory that set fire to my neurons coming up with the most creative strategies to address problems and challenging my brain to make the proper questions. Thanks for all your help, our shared laughs, stress and unique moments together.

Finally, thanks to the group of PhDs “Compactos” that started this trip with me four years ago. We made it guys. I feel lucky to have met you and that we lived this experience together: deep conversations by the sea, coffees, retreats, bars and trips. I hope our paths keep crossing wherever we are.