Abbreviations

4HT	4-hydroxytamoxifen
7-AAD	7-aminoactinomycin D
α-MEM	alpha - Minimum Essential Medium
AGM	Aorta-Gonad-Mesonephros Region
ALM	Anterior Lateral Plate Mesoderm
APC	Allophycocyanin
bHLH	Basic Helix-Loop-Helix Domain
BHK21	Glasgow Minimum Essential Medium
BMP	Bone Morophogenic Protein
BrdU	Bromodeoxyuridine
cDNA	copy DNA
ChIP	Chromatin Immuno-Precipitation
cTnT	Cardiac Troponin T
DMEM	Dulbecco's Modified Eagle Medium
Dox	Doxycycline
DMP	2,2-Dimethoxypropane
DN	Double Negative (Flk-1- PDGFRα–)
DNA	Deoxyribonucleic acid
DP	Double Positive (Flk-1+ PDGFRα+)
DSG	Di(N-succinimidyl) Glutrate
DTT	Dithiothreitol
E	Embryonic Day
EB	Embryoid Body
EF1α	Elongation Factor 1 alpha
EGS	Ethyleneglycol-bis-succinimidyl-succinate
EHT	Endothelium Haematopoietic Transition
Eomes	Eomesodermin
ER	Oestrogen Receptor
ES cell	Embryonic Stem Cell
FACS	Fluorescence Activated Cell Sorting

FCS	Foetal Calf Serum
FITC	Fluorescein Isothiocyanate
Flk-1	VEGF Receptor 2
GATA4	GATA Binding Protein 4
Gsc	Goosecoid
HATs	Histone Acetyltransferases
hCMV	Human Cytomegalovirus
HDAC	Histone Deacetylase
HRP	Horseradish Peroxidase
HSC	Haematopoietic Stem Cell
ICM	Inner Cell Mass
IGF	Insulin-like Growth Factor
Ihh	Indian Hedgehog
IP	Immuno-Precipitation
IMDM	Iscove's Modified Dulbecco's Medium
LDB	LIM Domain Binding Protein
LIF	Leukaemia Inhibitory Factor
LMO	LIM-only Domain Protein
LPA	Linear Acrylamide
MeC	Methylcellulose
MEL	Murine Erythroleukemia Cell Line
MTG	Monothioglycerol
Neo	Neomycin
PBS	Phosphate Buffered Saline
PBST	PBS containing 0.1% Tween-20
PCR	Polymerase Chain Reaction
PDGFRα	Platelet Derived Growth Factor Receptor alpha
PE	Phycoerythrin
PFA	Paraformaldehyde
PI	Propidium iodide
PKG	Phosphoglycerate Kinase 1
PLM	Posterior Lateral Plate Mesoderm

PRC	Polycomb Repressive Complex
P-Sp	Para-Aortic Splanchnopleura
qPCR	Quantitative PCR
Rb	Retinoblastoma (Protein)
RNA	Ribonucleic acid
SCL	Stem Cell Leukaemia (Protein)
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	SDS Polyacrylamide Gel Electrophoresis
SP-F	Single Positive Flk-1 (Flk-1^{+} PDGFRα^{-})
SP-P	Single Positive PDGFRα (Flk-1^{-} PDGFRα^{+})
T-ALL	T-cell Acute Lymphoblastic Leukaemia
Tet	Tetracycline
TGF	Transforming Growth Factor
TRE	Tet Responsive Element
TSS	Transcriptional Start Site
tTA	Tet Trans-activator
VE	Visceral Endoderm
VEGF	Vascular Endothelium Growth Factor
VSM	Vascular Smooth Muscle
WT	Wild Type
Zeo	Zeocin

1. Introduction

Embryonic development is a complex process through which tissue stem cells, progenitors and differentiated specialised cells are generated. This process is tightly controlled by environmental cues, such as cell-cell interactions and cytokines, which activate highly specific networks of signalling pathways and gene expression programmes that subsequently determines cell fate. Transcription factor proteins are important downstream targets of signalling pathways. They play key roles in the molecular processes underlying embryonic development, acting in a combinatorial manner to activate and repress the expression of specific genes to ensure correct cell differentiation. Transcription factors are modular proteins which are able to bind DNA, often as part of multi-protein complexes, and recruit transcriptional and chromatin modulators in order to regulate gene expression. Many have more than one described role in development, often acting through different molecular mechanisms depending on the cellular context and timing of expression.

Developmental haematopoiesis is a well-studied system which has provided a powerful model to study that transcriptional control underlying cell fate decisions. The transcription factor SCL (also known as TAL-1) has emerged as an essential regulator of embryonic blood development [1-6]. However, the first cells in which SCL is transcriptionally active have not been defined, and the mechanisms of action of SCL in the early stages of haematopoiesis are poorly understood. This is of particular interest from a developmental point of view and may also be relevant to clinical practice as these populations lie at the heart of how haematopoietic stem cells (HSCs) are formed. The main objective of this project was to investigate the role of

SCL in haematopoietic specification, from mesoderm patterning to production of the first blood cells.

1.1 Primitive Streak and Mesoderm Patterning

After the onset of embryogenesis, cells from the inner cell mass (ICM) of the blastocyst divide and proliferate to generate the egg cylinder. Through an extensive program of cell proliferation and differentiation, these cells give rise to the three germ layers (ectoderm, endoderm and mesoderm) and subsequently all the cells constituting [7-9]. In the mouse embryo, mesoderm development starts shortly after proximal-distal polarity and the anterior-proximal axis are established in the egg cylinder at around embryonic day 5.5 (E5.5) [7, 9]. Shortly after the onset of gastrulation, the primitive streak forms in the proximal-posterior region of the embryo. Cells migrating through the primitive streak give rise to embryonic mesoderm and endoderm lineages, while primitive ectoderm-derived lineages differentiate in anterior regions of the embryo [7, 10] (Figure 1.1A).

Molecular analysis and fate mapping studies have identified spatially organised regions of the primitive streak where different mesoderm lineages are generated in a time controlled manner. The first mobilised mesoderm cells migrate from the proximal-posterior region of the primitive streak into extra-embryonic regions, forming the yolk sac endothelium and blood islands [11-13]. Following this, lateral plate mesoderm cells are specified in the intra-embryonic, posterior region of the primitive streak. More anterior

regions of the primitive streak migrate laterally and distally and form cardiac and cranial mesoderm (Figure 1.1A) [7, 9]. Paraxial mesoderm lineages arise from the trunk region of the primitive streak, and the most anterior region of the primitive streak generates the definitive endoderm (Figure 1.1A) [7, 14, 15].

Lineage commitment in the primitive streak is controlled by a balance of Wnt, BMP and TGFβ/Activin/Nodal signalling [7, 14, 16, 17]. Exposure of cells to extra-cellular signals depends on spatial organisation, and causes expression of specific target genes, which in turn specify cell fate. For example, synergism between Wnt and BMP signalling in the posterior primitive streak directs cells towards a mesoderm fate. In contrast, high levels of Activin/ Nodal signalling in the anterior primitive streak specifies endoderm lineages [17].

1.1.1 Hallmarks of Early Mesoderm Populations

Molecular analysis and lineage mapping studies have characterised specific transcription factors and cell surface proteins that can define embryonic cell populations, depending on their level of commitment to any given lineage fate. Primitive streak cells can be identified by expression of the T-box transcription factor *Brachyury*, which is transiently expressed by all mesoderm cells that migrate from the primitive streak [12].

In the posterior region of primitive streak, co-expression of Brachyury with the vascular endothelium growth factor (VEGF) receptor Flk-1 (VEGFR2) marks cells that are destined for the extra-embryonic yolk sac [12, 18] and intra-embryonic blood, endothelial, vascular smooth muscle (VSM) and cardiac lineages [19, 20]. In the distal primitive streak, expression of the

T-box transcription factor *Eomesodermin* (*Eomes*), and/or platelet derived growth factor (PDGF) receptor alpha chain (PDGFRα), marks cells destined for cranial, cardiac and paraxial mesoderm lineages, such as bone, cartilage and axial skeleton [13, 21-23]. In the most anterior-distal region of the primitive streak, cells co-express *Brachyury* and the forkhead box transcription factor *Foxa2*. These cells form definitive endoderm lineages and can be identified by expression of the homeobox transcription factor Goosecoid (*Gsc*) and the cell-cell adhesion protein E-Cadherin [9].

1.2 Developmental Haematopoiesis

1.2.1 Haematopoiesis in the Mouse Embryo

Haematopoietic development occurs in distinct waves. The first primitive and definitive haematopoietic cells are found in the extra-embryonic yolk sac. This is followed by waves of intra-embryonic definitive haematopoiesis and the emergence of haematopoietic stem cells (HSCs) (Figure 1.1B).

1.2.1.1 Yolk Sac Haematopoiesis

In the mouse embryo, cells migrate from the most posterior region of the primitive streak into extra-embryonic regions of the egg cylinder in response to BMP signalling from the visceral endoderm around E6.5 [7, 13, 14] (Figure 1.1A). These migrating cells go on to form the extra-embryonic yolk sac mesoderm, consisting of VSM, endothelium and haematopoietic cells (Figures 1.1B and 1.2) [10-12, 24-27].

Ex vivo culture of Flk-1^+ cells from E7 and E7.5 extra-embryonic yolk sacs and intra-embryonic posterior regions has shown these cells to have the potential to generate haematopoietic and endothelial cells. Cells with such potential are often referred to as hemangioblasts or blast-colony forming cells (BL-CFCs) [12]. Despite this *in vitro* hemangioblast potential of yolk sac Flk-1^+ cells, commitment of cells to a haematopoietic or endothelial fate *in vivo* occurs in intra-embryonic regions, prior to cell migration out of the primitive streak [12, 18, 27]. For example, the first committed blood cells (identified by low levels of CD41 expression) are found in posterior regions of the primitive streak at E7.25. These cells later migrate into the extra-embryonic yolk sac [28].

It addition, it has been shown that yolk sac endothelial and haematopoietic cells arise independently during early gastrulation [11, 27]. While original studies suggested that endothelial and haematopoietic cells formed blood island structures in the yolk sac [24, 29, 30], later studies demonstrated (using whole mount *in situ*) that primitive erythrocytes form distinct band of blood cells around the circumference of the proximal yolk sac region at E7.75 [27]. This band in interspersed with endothelial cells, which eventually subdivided the band of primitive blood cells, thus forming the yolk sac capillary networks [27].

This first wave of extra-embryonic haematopoiesis is transient and generates primitive blood cells that are only found within a small temporal window starting at E7.5 (Figure 1.1B) [29]. The majority of these blood cells are primitive erythrocytes, which are larger than adult red blood cells and express a unique set of goblin genes. They exist as nucleated cells in the yolk sac, but do mature once they enter the embryo after the onset of

circulation (E8.25) and become enucleated in the foetal liver between embryonic days 12.5 and 16.5 [31-33].

A second wave of extra-embryonic yolk sac haematopoiesis generates definitive haematopoietic cells [29, 34] (Figure 1.1B). Small populations of definitive erythrocytes and megakaryocytes are present in the yolk sac at E8.25 to E9 [29, 34, 35], and myeloid lineages have been found prior to the onset of circulation in E8.25 yolk sacs [28, 29]. Whereas the first wave of extra-embryonic primitive erythrocytes arise from primitive streak haemangioblast cells [10, 12], the second wave of yolk sac haematopoietic cells are thought to emerge from a subset of specialised haemogenic endothelium [36, 37]. *Ex vivo* culture of yolk sac blood cells has shown they possess a very limited capacity to generate lymphoid cells [34, 35, 38]. This, plus the absence of multi-potent cells capable of self-renewal and long-term haematopoietic re-constitution in adult mice [38-40] indicates that definitive HSCs are not generated in the yolk sac.

There is some evidence that co-culture of E8.5 yolk sac cells with AGM derived stromal cells generates functional HSCs that can reconstitute adult mouse bone marrow (the functional hallmark of a HSC) [41]. This indicates that some HSC precursors may be specified in the yolk sac and migrate to the AGM to mature after the onset of circulation. In addition, cells which are which are capable of reconstituting neo-natal, but not adult, haematopoiesis have found in yolk sacs at E9 (termed neto-natal repopulating HSCs) [34]. However, since this is after the onset of circulation, it is possible these neto-natal repopulating HSCs arise from an intra-embryonic source. *In vivo* tracing of E7.5 *Runx1* expressing yolk sac cells tracked cells to the umbilical cord, AGM region and foetal liver, where they contributed to blood development up to 15 months post birth [42].

1.2.1.2 Intra-Embryonic Haematopoiesis

Definitive, intra-embryonic haematopoiesis can be detected from E8 onwards, coinciding with the decline in primitive blood cell production (Figure 1.1B) [29, 34]. Definitive haematopoiesis occurs in lateral plate mesoderm-derived structures, formed from migrating vascular endothelial cells of primitive streak origin [43, 44]. Cells with erythroid, lymphoid and myeloid potential are first specified in the para-aortic splanchnopleura region (P-Sp) at E8, before the start of circulation [38].

The first adult HSCs, capable of self-renewal and long-term haematopoietic reconstitution in an adult mouse, are detected in the AGM at E10.5 to E11.5 [45-47]. The first HSCs emerge from specialised endothelial cells in the lining of the ventral wall of the dorsal aorta (the haemogenic endothelium) that undergo an endothelium to haematopoietic transition (EHT) and bud-off from the vessel wall into the blood stream [46, 48-52]. This population of specialised endothelial cells distinctly differs from the haemogenic endothelium that gives rise to definitive yolk sac haematopoietic cells in terms of location within the embryo and cell marker expression, with yolk sac haemogenic endothelium unable to give rise to HSCs, and intra-embryonic haemogenic endothelium enable to generate yolk sac-like definitive erythroid and myeloid progenitors [51].

By E10.5, proper circulation is achieved and erythrocytes are dispersed at a steady density throughout the embryo, with definitive erythrocytes outnumbering primitive erythrocytes by E12.5 [29] (Figure 1.1.B). Around this time, there is an expansion of HSC generation in the foetal liver as it takes over as the main haematopoietic organ and HSC-derived definitive haematopoiesis takes over embryonic blood production.

Bone marrow haematopoiesis is initiated at E15 and becomes the main site of adult haematopoiesis shortly after birth (Figure 1.1B) [31, 32, 45].

1.2.2 Haematopoiesis in the Zebrafish Embryo

Our understanding of haematopoiesis has been greatly aided by the use of the zebrafish as a developmental model. The zebrafish has been a powerful model of development as the transparency of the embryo, and short gestation period, provide a highly accessible system. Like in mammalian models, zebrafish haematopoiesis occurs in distinctive waves. Shortly after gastrulation, cells in the ventral mesoderm migrate bilaterally to form anterior lateral plate mesoderm (ALM) and posterior lateral plate mesoderm (PLM) [53] (Figure 1.3).

Cells in the PLM migrate towards the trunk of the embryo, in response to BMP signalling from the visceral endoderm, and proliferate rapidly to generate the intermediate cell mass (ICM), which is thought to be the zebrafish equivalent of the mammalian yolk sac [53, 54]. In turn, the ICM gives rise to the posterior blood islands and the first primitive erythroblasts and endothelial cells [55, 56]. In addition, the PLM is the origin of the vascular tissue that generates the dorsal aorta [57]. In the embryonic head region, the rostral blood islands emerge from the ALM. These contain uni-potential cells that divide infrequently and differentiate into primitive macrophages [54]. The ALM also gives rise to an anterior haemangioblast population which subsequently differentiates into myeloid lineages and endothelial cells [58].

Like in the mouse, zebrafish HSCs emerge from a subset of specialised aortic endothelium cells in the ventral wall of the dorsal aorta. These HSCs bud-off in to the blood stream and differentiate into definitive blood cells and populate the sites of definitive zebrafish haematopoiesis, i.e. the thymus and kidney [53, 55, 59, 60] Later, the main site of adult haematopoiesis switches to the kidney marrow, which provides the niche for HSC maintenance throughout adult life [53].

1.2.3 Markers of Haematopoietic Development

Co-expression of *Brachyury* and Flk-1 in the early proximal-posterior primitive streak and in the extra-embryonic yolk sac marks cells capable of generating haematopoietic and endothelial cells when re-plated *ex vivo* [12]. Flk-1 expression is also associated with intra-embryonic definitive haematopoiesis, as P-Sp derived blood cells and AGM-derived HSCs [49, 61, 62] have been shown to arise from Flk-1^+ population. However, Flk-1 expression is only transient in blood cells and lost as immature cells differentiate further [20]. In contrast, endothelial cells retain Flk-1 expression throughout their life cycle [20]. Immuno-phenotypically, the expression of CD41, and later CD45, can be used to identify Flk-1^+ cells which have committed to a haematopoietic fate [28, 63].

1.2.4 Haematopoietic Transcription Factors

Early blood cell populations can also be identified by expression of specific transcription factors that are critical to haematopoietic differentiation. Many transcription factors have been associated with haematopoietic development, some of which have been shown to be critical for blood formation in the embryo. For example, the bHLH transcription factor SCL is essential for the formation of the first embryonic blood cells [2, 4, 5], while the Runt-related transcription factor RUNX1 is necessary for the emergence of HSCs [48, 56, 73, 74].

SCL is expressed in haematopoietic and endothelia primitive streak-derived populations from E7.5 onwards [64, 65]. Later, SCL is expressed in the yolk sac blood islands, HSCs and multi-potent blood cell progenitors, where it is essential for the terminal differentiation of erythrocytes and megakaryocytes [66-69]. In the zebrafish SCL is associated with blood formation from both the PML and AML [70, 71]. See Section 1.7 for further details.

Expression of RUNX1, another critical haematopoietic transcription factor, also marks haematopoietic cells in the yolk sac and in intra-embryonic regions. Expression of RUNX1 is not absolutely required for the formation of primitive erythrocytes, however *Runx1* $^{-/-}$ mice are embryonic lethal due to haemorrhaging in the central nervous system and server anemia in the foetal liver [72-75]. Additionally, loss of RUNX1 expression results in an absence of definitive blood cells and HSC emergence [50, 73, 75]. Importantly, RUNX1 is expressed in the haemogenic endothelium and in underlying mesenchyme tissue in the dorsal aorta, and is thought to be essential for the endothelial-haematopoietic transition required for HSC formation [52, 60, 76, 77].

Two other transcription factors which are important for haematopoietic development, and will be further discussed later, are GATA2 and LYL-1. The GATA-binding transcription factor GATA2 has also been associated with haematopoietic and endothelial differentiation. It is expressed in the yolk sac, P-Sp and AGM, and has been shown to be necessary for the expansion and proliferation of early blood progenitors [78]. *Gata2* $^{-/-}$ mice are embryonic lethal as they suffer from impaired primitive haematopoiesis and yolk sac anaemia. No HSCs or definitive haematopoietic cells are specified in *Gata2* $^{-/-}$ embryos. In addition, HSC production is severely impaired in the AGM regions, as is the proliferation of adult HSCs in the bone marrow [79]. In Xenopus, GATA2 is required for hemangioblast formation, but not in the zebrafish [80]. GATA2 is not critical for the maturation of definitive blood cells, and is down regulated in erythroid lineages, as cells switch from a more mature and which off expression of developmental transcription factors [81].

The bHLH transcription factor LYL-1 is thought to play a role in haematopoietic and endothelial differentiation as its embryonic expression pattern highly overlaps that of *Scl* in developing blood and endothelial lineages [82]. Although loss of LYL-1 does not inflict a drastic haematopoietic phenotype itself (mice exhibit decreased B-cell activity) [83], it is thought that LYL-1 may play a compensatory role when SCL expression is detected.

1.3 Cardiac Development

1.3.1 Cardiac Development in the Mouse Embryo

In the mouse, the heart develops from two spatially and temporally distinct regions which contribute to either the primary or secondary heart field [84, 85]. It is thought that these two heart fields develop from distinct progenitor pools and differentiate in response to different cytokine environments [86-88].

This process begins around E7.5, when a subset of primitive streak cells, located adjacent to the lateral plate mesoderm (Figure 1.1A), migrate antero-laterally in two bivalent regions in response to BMP signalling and Wnt inhibitors [89]. Cells in these migrating regions meet and fuse at the embryonic midline to form the cardiac crescent which gives rise to the primary heart field. Cells within the cardiac crescent differentiate into the endocardium and myocardium that forms the linear primary heart tube structure [86-88].

Looping of the primary heat tube is initiated around E8-8.5. This, along with the the onset of a cardiac action potential, initiates blood circulations throughout the embryo (Figure 1.2) [86]. Complex looping and re-modelling of the primary heart tube occurs between E8.5 and E10.5, generating the left two chambers of the heart [87, 90]. Initial cardiac contractions occur with the induction of an action potential around E8.5, resulting in the onset of blood circulation (Figure 1.2) [86, 87]. Cells recruited to the dorsal-anterior regions of the primary heart from E9.5 onwards, forming the secondary heart field, which form the two right heart chambers and out-flow track of the mature heart [87, 90].

1.3.2 Cardiac Development in the Zebrafish

Zebrafish cardiac development differs from mouse mainly due to the lack of a secondary heart field, which results in a two chambered heart rather than four [58]. Like in mammalian primary heart field formation, pre-cardiac mesoderm cells localise in two bilateral regions in the posterior region of the embryo. These regions fuse at the midline and form the primary heart tube and subsequent cardiac structures (Figure 1.3). In addition, zebrafish have an anterior hemangioblast population that is absent in mammalian embryo. These anterior hemangioblast cells arise from the ALM and are capable of differentiation into blood, endothelial and cardiac cells [58], thus indicating a strong developmental link between lateral mesoderm and cardiac development in this species.

There has however been some recent evidence that a secondary heart field exists in the zebrafish. Three independent studies have identified a discrete population of cells located at the arterial embryonic pole which express secondary heart field associated genes (*Isl1* and *Mef2c*) and contribute to myocardium and cardiac smooth muscle tissues after initial heart tube formation [91-93].

1.3.3 Markers of Cardiac Development

In addition to its expression in lateral mesoderm associated lineages, Flk-1 is expressed in regions of the primitive streak that give rise to endocardial and myocardial cells *in vivo* [20]. *Ex vivo* culture of intra-embryonic Flk-1^+ cells from E8.5 embryos generates endocardium and myocardium lineages, suggesting these cells are precursors to the heart tube

[19]. Like in the haematopoietic system, Flk-1 expression is also transient in cardiac differentiation, and is lost as cardiomyocytes switch on cardiac markers, such as Nkx2.5. However, Flk-1 expression is retained in the endocardium [94]. PDGFRα is expressed by pre-cardiac cells in the mesoderm and in the cardiac crescent in populations that co-express primary heart field genes [95]. Note that Brachyury, Flk-1 and PDGFRα are co-expressed in a subset of mesoderm cells at E7.5 [22], but the fate of these cells has not been traced *in vivo.*

Cardiac mesoderm populations can be identified by expression of the posterior mesoderm bHLH transcription factor MESP1. MESP1 is expressed in primitive streak cells destined for a cardiac lineage from E6.5, and in cardiac progenitors which contribute to both the primary and secondary heart fields [96-98]. Components and progenitors of the primary and secondary heart fields can be tracked and distinguished by the expression of key lineage specific markers. Cells of the primary heart field lineage are specified and regulated by interaction between three transcription factors which are expressed in the cardiac mesoderm, cardiac crescent and developing primary heart field; the homo-box protein NXK2.5, the zinc-finger protein GATA4 and the T-box protein TBX5 [87, 99, 100].

Cells that contribute to the secondary (anterior) heart field can be identified by expression of specific genes, such as *Isl-1*, *Tbx20*, *Hand2 and FoxH1* [58, 85, 101]. Terminal differentiation into a cardiomyocyte lineage can be identified by the expression of mature cardiac muscle associate proteins such as Mef2C and cardiac Troponin T (cTnT) [85, 102].

1.4 Paraxial Mesoderm Lineages

In the mouse embryo, the more anterior regions of the primitive streak give rise to paraxial mesoderm lineages [7]. As these cells differentiate, they migrate distally through the primitive streak and gather at the embryonic midline around E8-8.5. This paraxial mesoderm tissue separates into blocks of cells, known as somites [15, 103]. Somites are transient structures that, along with cells from the neural crest, undergo condensation and differentiation to generate precursor cells which are able to terminally differentiate into bone, cartilage, brown adipose tissue and muscle [15, 103, 104].

1.4.1 Markers of Paraxial Mesoderm Fate

Precursors of paraxial lineages can be identified as they are generated from a subset of mesoderm cells that express PDGFRα [13, 22]. Further lineage commitment of PDGFRα expressing cells can be determined by the expression of lineage-specific genes.

Cells expressing the transcription factor *Sox9* in the primitive streak at E8.5 have been shown to be bi-potent, and can give rise to chondrogenic and osteogenic precursors, but not adipose or skeletal muscle [104, 105]. Chondrocytes are specified by sustained high *Sox9* expression, while repression of *Sox9* activity results in high levels of *Runx2* expression and induces an osteogenic fate [104, 106]. Once terminally differentiated, chondrocytes and osteoblasts can be later identified by their expression of lineage specific proteins. For example, chondrocytes express specific Collagens and cell surface proteoglycans [107, 108], while differentiating

osteoblasts express alkaline phosphatase and later bone specific markers such as Osteocalcin [109]. Skeletal muscle precursors are generated through a separate pathway, and can be first identified by expression of the bHLH transcription factor Myogenin [110]. More terminally differentiated muscle cells can be detected by expression of the muscle specific bHLH transcription factor MyoD and structural muscle proteins such as skeletal actin [110, 111].

1.5 Embryonic Stem Cells as a Developmental Model

Mouse embryonic stem (ES) cells can be isolated from the inner cell mass (ICM) of the blastocyst 4–5 days post fertilization, and are capable of contributing to all embryonic tissues upon re-injection into a blastocyst. ES cells can be cultured for several passages *in vitro* in an undifferentiated state in the presence of LIF without losing their pluri-potent nature [112].

ES cells are used to study embryogenesis as they successfully recapitulate early developmental events that occur within the embryo. Upon the removal of LIF, ES cells lose their pluri-potency and differentiate to form 3D structures, known as Embryoid Bodies (EBs) consisting of cells from the three germ layers; ectoderm, endoderm and mesoderm [113-115]. Under the appropriate stimuli, these have been shown to differentiate *in vitro* into many of the cell lineages found in the embryo. These include haematopoietic cells [115-119], endothelial cells [119-121], vascular smooth muscle cells [120], cardiomyocytes [19, 94, 96, 122, 123], skeletal muscle [124], osteocytes and chondrocytes [106, 125], liver and pancreatic precursors [126, 127] and neuronal cells [114, 128]. Differentiation of haematopoietic cells from ES cells occurs with a comparable temporal progression as in the embryo [3, 19,

115, 122, 125, 129, 130]. Therefore, the ES cells / EB system provides us with an accessible system to investigate molecular processes involved in blood development.

1.5.1 Advantages of the ES/EB System

The small size and limited number of cells in the early embryo makes direct study of transcriptional control difficult. Using differentiating ES cells allows access to increased numbers of attainable cells that can be used to investigate developmental processes. For example, use of ES cells has allowed study into the molecular mechanism of sustained pluri-potency versus the onset of development and embryogenesis [131, 132] which would have been very difficult using *in vivo* techniques. Cell number is of particular importance for biochemical assays, such as co-immuno-precipitation (IP), chromatin-IP (ChIP) and mass spectrometry, which are difficult to carry out on primary material.

The ES/EB system offers a unique system to investigate early embryonic events without the constraints of embryonic lethality. For example, GATA4 and GATA6 null embryos are embryonic lethal due to defects in visceral endoderm formation, prior to cardiac defects in the embryo [133, 134]. The *in vitro* differentiation of *Gata4* $^{-/-}$ and *Gata6* $^{-/-}$ ES cells was able to show these genes are essential for cardiac lineages in the mouse [135, 136].

The differentiation of ES cells provides a system in which the effect of genetic modification can be determined with relative ease. EBs differentiate with a phenotypic variation which reflects natural heterogeneity, and therefore gives a much more reliable indication of *in vivo* developmental events than conventional cell lines. Furthermore, although ES cells

differentiating *in vitro* lack the surrounding environment of the embryo, our growing understanding of developmental cytokines and signalling pathways allows us to mimic instructional signalling pathways. This allows the induction of ES cell differentiation into chosen lineage of interest, thus permitting further study into that specific developmental pathway.

1.5.2 Weaknesses of the ES/EB System

As mentioned above, differentiating cell populations within EBs lack the polarised, highly ordered, regionalisation of the embryo. This is due to the absence of a proper niche environment and extra-cellular signalling pathways that cells would be exposed to *in vivo*, and therefore a lack of precise spatial-temporal signalling events.

Uncommitted multi-potent cells are very sensitive to environmental cues, and can easily switch phenotype and cell fate when exposed to certain *in vitro* conditions. Therefore, without proper spatial exposure to these signals, we can only determine the potential of cells to differentiate into a lineage fate under a given environment, which, although it does reflect their nascent ability, may not reflect what the cells would do in their natural environment. Therefore, although the ES/EB system does provide a platform to advance our understanding of complex developmental event, confirmation of results *in vivo* is essential.

1.5.3 Mesoderm Development

Mesoderm development, including haematopoietic and cardiac differentiation, occurs in developmental waves in EB structures. Like in the embryo, these populations can be fractionated and tracked using gene and cell surface markers (Figure 1.5).

Reflecting early gastrulation events, ES cells differentiate as they lose expression of pluri-potent genes, such as *Oct4*, *Sox2* and *Rex1*. A population of Brachyury expressing, primitive streak-like cells form around day 2 of differentiation [3]. A subset of these cells, which co-express E-Cadherin and PDGFRα, have a mesendoderm potential, and differentiate further to generate E-Cadherin^{+} Gsc^{+} Sox17^{+} definitive endoderm, and Brachyury^{+} PDGFRα^{+} Flk-1^{+} mesoderm [113]. PDGFRα^{+} Flk-1^{-} precursors share a similar cell fate potential as anterior primitive streak paraxial mesoderm cells, and are able to generate osteocytes, chondrocytes, adipocytes and myocytes *in vitro* [106, 125]. In contrast, Flk-1^{+} PDGFRα^{-} populations resemble the proximal-posterior primitive streak, and are enriched for haematopoietic, endothelial and VSM precursors [125], whereas cardiac precursors emerge from Flk-1^{+} PDGFRα^{+} populations [97, 137].

1.5.4 Blood Development

Haematopoietic development has been closely studied in the ES/EB system. Like *in vivo*, expression of Flk-1 marks the onset of blood cell specification, with Flk-1 $^{-/-}$ ES cells unable to generate any haematopoietic cells [138]. Flk-1 expression is initiated around day 2.5-3 of EB differentiation

[3]. Brachyury^{+} Flk-1^{+} cells from day 2.75 to 3.75 EB have the potential to generate hemangioblast-like colonies containing both endothelial and haematopoietic cells, known as blast colonies [3, 12, 19, 116, 117, 122]. Re-plating of Flk-1^{+} cells from day 3.25 EBs generates mainly primitive erythroid cells, with a smaller percentage of macrophages and very little lymphoid potential [129], suggesting that these cells represent a wave of primitive haematopoiesis, similar to that of the extra-embryonic yolk sac. This is a transient wave of blood development. Indeed, 24 hours later, Flk-1^{+} cells isolated from day 4.25 EBs (secondary Flk-1^{+} population) have greatly reduced haematopoietic potential [19, 122].

More recently, a secondary wave of EB haematopoiesis has been reported [129]. Between day 3.25 and 5.25 of EB differentiation, a separate population of Brachyury^{+} Flk-1^{+} cells emerge which express genes associated with AGM haematopoiesis (HoxB4, c-Kit, Sox17) and T-cell lineages (Gata3). Upon re-plating, these cells generate less primitive erythrocytes but do differentiate into definitive erythrocytes, macrophages and mixed colonies. Importantly, they also have significantly increased lymphoid potential [129]. Together this indicated that a secondary, more definitive wave of haematopoiesis occurs within differentiating EBs. This most likely reflects the wave of definitive haematopoiesis seen in the yolk sac at E9 [36, 37]. Additionally, like in the embryo, a haemogenic endothelium population has been derived *in vitro* from differentiating ES cells, Using time-lapse imagining, non-adherent CD41^{+} CD45^{+} nascent blood cells can be observed budding off endothelial cells [130, 139].

Definitive HSCs, that are capable of self-renewal and long-term re-population upon serial transplantation in mice, have not yet been isolated from differentiating ES/EB cells. Forced expression of HoxB4 (an AGM/HSC

associated gene) in differentiating EBs has however resulted in the formation of cells that were able to, capable of self-renewal and differentiation into multiple haematopoietic lineages [140]. These cells possessed limited lymphoid potential, suggesting they represent a multi-potent precursor population, rather than definitive HSCs. However, this data may still imply that it may be possible to derive HSCs from ES cells *in vitro*, but the culture conditions to allow this to occur without genetic alterations are not yet defined.

1.5.5 Cardiac Differentiation

During embryonic development, cardiac precursors arise from the primitive streak around E7-8, closely following the first wave of primitive haematopoiesis (Figure 1.4). Several *in vitro* studies have confirmed that this temporal progression of mesoderm differentiation is present in the ES/EB system, with the presence of Flk-1^+ population arising in-between waves of haematopoietic cells [19, 122]. Flk-1^+ cells arise from day 3.25 Brachyury$^+$ Flk-1^- cells after 24 hours re-culture. These day 4.25 Flk-1^+ cells express increased levels of cardiac genes (*Nxk2.5, Gata4, Tbx5, Mef2c*) and have *in vitro* potential to generate functional cardiomyocytes [19, 122]. These cardiac colonies are similar to those generated by re-plating E8.5 Brachyury$^+$ Flk-1^+ intra-embryonic cells *ex vivo* [19], suggesting a model of primary heart field development.

Recent studies have further pinpointed cardiac precursors as being the Flk-1^+ PDGFRα^+ cells in day 3-4 EBs [137] and more specifically, the Brachyury$^+$ Flk-1^+ PDGFRα^+ CXCR4$^+$ Mesp1^+ cells in day 4 EBs [97]. Flk-1^+ Isl-1^+ cardiac precursors have also been described in day 5 EBs [94], with

the expression of Isl-1 suggesting the secondary heart field is also modelled in EBs.

1.6 Cross-talk between Developing Mesoderm Lineages

The early embryo is made up of populations of relatively uncommitted cells which differentiate into specific lineages. Cell fate decisions at these early stages of development are governed by gene expression programmes and cell signalling pathways that often act to both specify cell fate and suppress differentiation of alternative lineages. This is important as even after commitment to a specific differentiation pathway, cells often retain high plasticity, and can change fate upon instruction [141].

This high plasticity and lineage flexibility is very apparent in the mesoderm, where cross-antagonising signalling pathways ensure correct cell differentiation. Furthermore, changes in mesoderm cell fate can be induced by changes in extra-cellular signalling, demonstrating the high degree of plasticity that exists in anterior and posterior epiblast cells. For example, the Activin/pMAPK pathways specify mesendoderm differentiation while actively repressing ectoderm formation in the primitive streak. Reversely, MAPK inhibitors induce ectoderm commitment at the expense of the endoderm [141]. Exposure to high levels of Activin can also re-direct posterior primitive streak cells into adopting an endodermal fate [16, 17, 142]. Similarly, culture of anterior ectoderm cells *ex vivo* in the presence of Indian hedgehog (Ihh) re-programmes cells to adopt a blood and endothelial fate [143], while enforced BMP signalling results in increased Flk-1 mesoderm formation at the expense of FOXA2 endoderm populations [16, 17].

Fate change can also be induced by changes in expression levels of important transcription factors. For example, expression of GATA1 and PU.1 controls the balance of erythroid versus myeloid differentiation from a multi-potent progenitor population, through a mechanism of cross-antagonism and auto-regulation [144]. GATA1 specifies erythroid differentiation while repressing PU.1 expression, and PU.1 specifies myeloid lineages while repressing GATA1 expression [145]. In a similar fashion, expression of Sox9 and Runx2 controls chondrogenic and osteogenic lineages decisions in differentiating bi-potent progenitors [104, 106]. The ability of even terminally differentiated cells to adopt a new lineage fate is highlighted by the generation of induced pluri-potent (iPS) cells. Mouse embryonic and adult fibroblasts can be re-programmed into ES-like cells by the introduction of *Oct4*, *Sox2*, *c-Myc* and *Klf4* expression [146]. Moreover, fibroblast cells can be directly re-programmed into alternative cell fates, for example, enforced *Gata4*, *Mef2c* and *Tbx5* expression results in cells adopting a cardiac fate, both *in vitro* and *in vivo* [147].

Cross-talk between blood and cardiac pathways has been documented, particularly during zebrafish development. In the zebrafish, FGF signalling is thought to control the boundary between cardiac and haematopoietic lineage commitment in the anterior hemangioblast [148]. Inhibition of FGF signalling results in the loss of cardiac tissue, coupled with an increase in blood and endothelium [148], whereas treatment with FGF antagonists causes increased expression of cardiac genes at the expense of haematopoietic and endothelial genes [149].

Other studies using the zebrafish model have shown that ectopic expression of central cardiac transcription factors, such as *Nkx2.5*, enhances the cardiac programme, while repressing hemangioblast development [148].

Reversely, forced expression of essential haematopoietic transcription factors, such as *Scl*, results in increased blood and vascular development, coupled with a decrease in cardiac gene expression in the rostral mesoderm and myocardium tissue formation, resulting in a smaller heart field [150]. Over-expression of haematopoietic genes was also able to convert somatic mesoderm cells into cells with the potential to generate blood and endothelium [71, 150, 151]. Whereas in the absence of blood and vessel specification (*cloche* mutant) there is ectopic cardiomyocyte production and the heart field is expanded [150]. In addition, induced Wnt3a expression in the Xenopus cardiac crescent has been shown to reprogram cells into adopting a haematopoietic fate [152].

Links have also been made between blood and cardiac pathways during mouse development. Normally, blast colonies, generated from either the yolk sac or from differentiating ES cells, have little or no cardiac potential [119]. However, forced induction of Notch4 results in alterations in Wnt and BMP signalling pathways can re-programme these hemangioblast-like cells into functional cardiomyocytes [153].

Together, these studies show the lineage flexibility exhibited not only by multi-potent progenitor cells, but also by cells which have undergone lineage commitment. It is unsurprising that early embryonic cells have the potential to form several different cell types, given the vast number of terminally differentiated lineages that are generated from relatively few cells in the egg cylinder. Cross-antagonism between developing lineages highlights the importance of signalling pathways and gene expression programmes in determining cell fate. The high levels of plasticity exhibited by more lineage committed cells may have arisen as a form of compensatory mechanism to ensure correct development.

1.7 The Transcription Factor SCL

SCL is a class II tissue-specific basic helix-loop-helix (bHLH) transcription factor and an essential transcriptional regulator involved in haematopoiesis. SCL was first discovered though its mis-expression in T-cell acute lymphoblastic leukaemia (T-ALL), which is usually as a result of a chromosomal translocation or a tumour-specific chromosome deletion [154].

The SCL protein contains a proline-rich N-terminal domain, a basic DNA-binding domain and a helix-loop-helix domain which allows protein-protein interactions (Figure 1.5A). This bHLH structure is highly conserved throughout evolution, and is shared with many other transcriptional regulators, such as MyoD [155]. The *Scl* gene encodes three major polypeptides, a 42kD, a 39 kD and a 22 kD protein, depending on whether the initiation of translation occurs at the first, second, or fifth ATG site [156]. All three proteins contain the highly conserved bHLH domain, with the 42kD and 39kD protein differing only by the loss of the first 25 N-terminal residues. The 22kD protein is a truncated version, lacking a larger region of the N-terminus [157]. Different isoforms of SCL are expressed depending on cellular context [158].

Like other class II bHLH proteins, SCL forms heterodimers with enhancer-binding proteins (E-proteins) such as E47 and E12 (both E2A gene products) [159]. These E-protein/bHLH protein heterodimers have been shown to specifically bind E-box DNA sequences (CANNTG), with a very high affinity, to activate target gene expression [155, 159, 160]. SCL/E-protein complexes have been shown to preferentially bind the E-box sequence AACAGATGGT [161], ACAGGTGG [162, 163] and ACAGCTGC [162-164].

SCL/E-protein heterodimers bind DNA as part of a multi-protein complex, with LIM-only domain protein LMO-2, the LIM domain binding protein LDB-1, and the haematopoietic transcription factor GATA1 [163]. This complex binds to conserved E-box and GATA motifs in the genome, typically separated by 9 to 12 nucleotides: GATA(n9-12)CANNTG [165, 166], to activate expression of target genes in erythrocytes [164] and megakaryocytes [167] (Figure 1.5C).

1.7.1 Expression of SCL in the Embryo

Using LacZ reporter constructs and immuno-histochemical staining, SCL expression has been tracked throughout murine embryonic development. SCL expression has been reported in primitive and definitive haematopoietic blood cells, in vascular endothelium (including the aortic wall), in the foetal liver and in the central nervous system [64, 65].

SCL expression is first apparent in the extra-embryonic mesoderm at E7-7.5 in both haematopoietic and endothelial cells. SCL is expressed in yolk sac-derived blood cells, and subsequently expressed in nucleated, circulating primitive erythrocytes between E9.5 and E12.5 [64, 65]. SCL is strongly expressed in the foetal liver from E11.5 until birth [64, 65], with high expression in erythroid and megakaryocyte lineages, and at low levels in myeloid cells [64]. By E15.5 SCL is expressed in the foetal spleen, coinciding with the appearance of haematopoietic cells in the organ [65].

SCL expression is also found in the developing vascular system between E10.5 and E13.5, in the dorsal aorta and the haematopoietic aortic clusters in the AGM. Furthermore, SCL expression has been reported in the developing endocardium between E8 and E11 [65]. SCL is expressed in the

developing neural tube and presumptive spinal cord from E9.5 onwards and in the developing mid-brain and hind-brain from E12.5 [64, 65].

Post-natal expression of SCL occurs mainly in the bone marrow and spleen [65]. At a cellular level, SCL expression is enriched in HSCs, their subsequent multi-potent progenitors and in erythrocyte and megakaryocyte lineages [69]. SCL expression has also been observed in the adult thymus, the vascular smooth muscle of the aorta and bladder, and a few rare cells in the brain (the thalamus) [65, 168]. In addition, SCL is expressed in adult osteoblasts which line the bone marrow surface, but not in mature osteoclasts or embryonic osteoblasts [169, 170].

1.7.2 SCL Expression in the ES/EB System

SCL expression in the ES/EB system has been studied using an ES cells line with human CD4 knocked into the *Scl* locus [171]. Flow cytometric analysis showed CD4 expressing was closely associated with Flk-1 expression, with expression first detected after 2.75 days of EB differentiation and sustained throughout EB differentiation until day 8 [171].

1.7.3 SCL in Haematopoiesis

Loss-of–function studies have highlighted the importance of SCL in haematopoietic development. *Scl* $^{-/-}$ mice die between E8.5 and E10.5 due to the lack of primitive blood cells [4] (Figure 1.5B). Yolk sac cells from *Scl* $^{-/-}$ E8.5 embryos are unable to generate any blood lineages *ex vivo* [4], and *Scl* $^{-/-}$ ES cells do not contribute to any haematopoietic lineages in mouse chimeras [2, 5]. *In vitro*, SCL is required for the specification of blood cells

from Flk-1^{+} precursor cells. Haematopoietic blast colonies cannot be generated from *Scl* $^{-/-}$ ES/EB cells, reflecting the absence of yolk sac haematopoiesis *in vivo* [1-3]. Flk-1^{+} *Scl* $^{-/-}$ EB cells do however form colonies consisting of endothelial and VSM cells (known as transitional colonies). However, further haematopoietic specification is blocked, suggesting that SCL is critical for mesoderm commitment to a haemogenic fate [3]. It has also been reported that SCL expression negatively influences smooth muscle differentiation of Flk-1^{+} ES cells, in favour of blood and endothelial development [172]. Re-introduction of SCL expression in *Scl* $^{-/-}$ cells rescues these haematopoietic defects completely [173].

Conditional knock out studies have shown that although SCL expression is enriched in HSCs and subsequent multi-potent progenitors, SCL is dispensable for HSC engraftment, self-renewal and differentiation into myeloid and lymphoid lineages [69, 174]. SCL has been implicated in the production of the dorsal aorta and haemogenic endothelium [175], but is not essential for the emergence of HSCs once the haemogenic endothelium is formed [176, 177]. Additionally, SCL has a role in controlling HSC quiescence, where it is thought to promoter HSC dormancy by negatively regulating the G_0 to G_1 transition [178].

SCL is required for definitive erythroid and megakaryocyte maturation, with the loss of SCL in the bone marrow impairing erythrocytic and megakaryocytic development [66-69]. Furthermore, enforced SCL expression in the mouse bone marrow enhances erythropoiesis and megakaryopoesis at the expense of myeloid lineages [179]. SCL has been shown to bind promoters of haematopoietic regulators in primary megakaryocyte cells [67, 180] and erythrocytes [164]. Conditional knock-out

studies have also shown mast cell development to be perturbed in the absence of SCL [181].

During zebrafish development, SCL is expressed in the ALM and the PML is implicated in hemangioblast and blood lineage formation [151, 175, 182]. Moreover, SCL expression is able to rescue the vascular and haematopoiesis defects exhibited by *cloche* mutants [183] suggesting SCL is critical to blood and vascular specification. SCL is also thought to specify haematopoietic mesoderm during Xenopus frog development [184].

There has been some evidence that SCL has a pre-haematopoietic function. Drug-induced pulses of ectopic SCL expression during pre-haematopoietic stages of EB differentiation has been shown to result in increased an Flk-1 expressing population, thus suggesting a role for SCL in mesoderm patterning [185]. Morpholino knock-down of *Scl* expression in the zebrafish results in the absence of primitive blood lineages and causes disruption in endothelial development and dorsal aorta formation [175]. Loss of SCL expression also results in increased expression of cardiac genes, such as *Hand2*, and a misshaped heart with an enlarged atrium due to an increase in cardiomyocytes [150]. Reversely, ectopic expression of SCL in the zebrafish induces vessel and blood marker expression at the expense of myocardium production [150, 151]. Furthermore, ectopic expression of SCL in zebrafish paraxial mesoderm tissues has also been shown to re-direct these cells into a lateral mesoderm fate, generating mature blood and endothelial cells [71].

1.7.4 Direct DNA-Binding can be Dispensable for SCL Function

During some stages of haematopoietic development, even though direct SCL-DNA binding is dispensable, SCL protein-protein interactions remain essential [173, 186]. *Scl* $^{-/-}$ ES cells expressing different versions of a DNA binding mutant form for SCL (containing a deleted basic DNA-binding region, or harbouring mutations which inhibit DNA-binding) exhibited rescue of the early stages of haematopoiesis [173, 187]. However, the introduction of mutations into the HLH domain of SCL which inhibit the ability of SCL to bind protein partners (e.g. E-proteins or LMO2), prevented haematopoietic rescue [173, 186].

In vivo, mouse embryos derived from ES cells expressing a DNA-binding mutant version of SCL survive past the stage of *Scl* $^{-/-}$ embryonic lethality at E9.5 [4], with most dying between E14.5 and E17.5 due to embryonic anaemia [187] (Figure 1.5B). These mutants express a version of SCL which harbours three amino acid substitutions with its basic DNA-binding domain (replacing RER with AAA; *Scl* $^{RER/RER}$ mutants) [173, 187]. *Scl* $^{RER/RER}$ embryos have functional yolk sac haematopoiesis, but present a block in maturation of erythroid cells in the yolk sac and foetal liver. This block occurs in the transition between $CD71^{+}$ $Ter119^{-}$ to $CD71^{+}$ $Ter119^{+}$ pro-erythroblasts [187]. This suggests that direct SCL-DNA binding is not required for the specification of both waves of haematopoiesis, but is required later on for erythroid maturation.

ChIP-sequencing has revealed that the majority of SCL gene targets in foetal liver primary erythrocytes require direct DNA-binding [164]. However, around 20% of target sequences were still pulled down from *Scl*

$^{RER/RER}$ cells, demonstrating that SCL can be recruited to DNA independently of its own DNA-binding domain. It has been suggested this may occur through association with the haematopoietic transcription factor GATA1 [164]. This provides evidence that SCL may act through different mechanisms depending on the cellular and temporal context.

1.7.5 SCL in Other Lineages

Many key developmental transcription factors are involved in the differentiation pathways of more than one lineage. In addition to its role in haematopoiesis, SCL expression has been linked to endothelial, endocardial, osteoclast and neuronal development.

Despite the lack of blast colony potential of *Scl* $^{-/-}$ yolk sac derived cells and EBs, endothelial cells can be derived from *Scl* $^{-/-}$ EBs *in vitro* [3, 116]. However, conditional knock-out models of *Scl* in endothelial lineages has shown that in the absence of Scl expression *in vivo*, the endothelia cells that are produced are not able to migrate properly within the embryo [188]. Endothelial cells in *Scl* $^{-/-}$ yolk sacs fail to undergo correct vascular remodelling, which results in defective angiogenesis and a disorganised, under developed extra-embryonic vascular networks [64, 188].

Intra-embryonic vasculogenesis was also defective, with *Scl* $^{-/-}$ embryos showing a dilated dorsal aorta and peri-neural veins. In addition, cardiac looping was not initiated, with the heart tube in E8-9 embryos remaining linear [64]. In the zebrafish, endocardial cells are formed, but they are not able to migrate properly and aggregate at the ventricular pole of the heart. This causes *Scl* mutants to lack arterial endocardium [182].

SCL has been linked to osteoclast formation. Osteoclasts exist in the bone marrow, where they act to maintain bone metabolism by resorbing bone [189]. Osteoclasts are thought to originate from the same haematopoietic precursors which give rise to monocytes [170]. These are not produced in the absence of SCL, therefore accounting for the defect in osteoclast formation [170].

SCL is important for neuronal differentiation, which occurs in the anterior region of the embryos [7]. SCL is co-expressed with GATA2 and GATA3 in the spinal cord from E10.5 onwards [64, 190] and with GATA1 in the developing brain from E12.5 [191]. Conditional deletion of *Scl* expression specifically in neuronal lineages gives rise to 60% post-natal lethality due to impaired neuronal growth, with surviving mice suffering from growth retardation [192]. Furthermore, SCL expression is required in the neural tube for the specification of astrocytes and V2b interneurons from multi-potent precursors. This occurs in a fashion similar to that described for the specification of blood and endothelial cells from the haemangioblast, where the level of SCL expression determines the lineage fate by specifying astrocyte and V2b interneuron differentiation over oligodendrocyte V2a interneuron differentiation pathways [190].

1.7.6 SCL as an Activator and a Repressor of Gene Expression

SCL mediates either activation or repression of gene expression depending on the target gene and cellular context. SCL often acts within a DNA-protein complex with E-proteins (e.g. E2A and E47), LMO2, LDB1 and

GATA1, which recruits co-activators, co-repressors and chromatin remodelers to control transcription of target genes (Figure 1.5C).

To date, the function and mechanisms of action of SCL have been best studied in erythroid cells, where SCL has been shown to bind, as part of a pentrameric complex, to haematopoietic target genes such as *Gata1* [164, 166], *Eklf* [164, 165], *Gpa* [164, 193], *c-kit* [194], *pb4.2* and globulin genes [164]. In addition, SCL has been shown to bind the Nfe2 promoter in megakaryocytic cell lines [67] and c-kit in erythroid-myeloid cell lines, where SCL forms a pentameric complex including GATA2 rather than GATA1 [195]. SCL has also been detected in a pentameric complex including GATA3 in lymphoid cells in T-ALL patient cell lines, where is drives Nkx3.1 expression [196, 197].

The SCL-protein complex is able to interact with co-activators and co-repressors to modulate gene expression. SCL is involved in the recruitment of co-activators, such as p300 and CBP [198] and in turn, chromatin remodeling factors, such as the histone acetyltransferase (HAT) pCAF [199]. This results in histone acetylation, relaxation of the chromatin/DNA structure, and enhances the ability of DNA polymerases to bind, hence driving target gene expression. [199-201].

In addition to its role as a direct transcriptional repressor, SCL can suppress differentiation by sequestering E-proteins from binding other bHLH protein regulators. For example, muscle differentiation is subject to transcriptional control by heterodimer complexes of bHLH protein regulators, e.g. MyoD, and E-proteins. Ectopic expression of SCL in myoblast cells caused inhibition of muscle differentiation as E-proteins preferentially dimerise with SCL over MyoD, resulting in blocked transcriptional activation

of muscle specific growth and differentiation genes [202, 203]. In a similar fashion, ectopic expression of SCL in T-ALL sequesters E47, thus prevents E47 binding to the bHLH protein HEB in T-cells [204]. This inhibits the expression of E47/HEB target genes that are critical for thymocyte differentiation, such as CD4, CD5, Rag 2 and pre T-cell receptors (Pre-TCRs), and leads to a block in T-cell maturation and induces leukaemia [204].

1.8 Project Aims

SCL is a multi-functional protein exhibiting multiple mechanism of action. The goal of this project was to focus on the role of SCL during the early stages of haematopoietic development, using the ES/EB system as a model of early mouse embryogenesis.

Specifically, the aims of this project were:

1. To understand when, and in which cells, SCL is first expressed (Results Chapter One).

2. To describe the function of SCL in early developmental populations through loss and gain of function studies (Results Chapter Two)

3. To create a cell line with inducible SCL expression, allowing further temporal characterisation of SCL function (Results Chapter Three).

4. To study SCL mechanisms of action and to begin dissecting the molecular pathways regulated by SCL during mesoderm lineage specification (Results Chapter Four).

2. Materials & Methods

2.1. Cell Culture Techniques

2.1.1 Cell Lines

J1 wild type ES cells were originally isolated from the 129S4/SvJae mouse strain [205]. *Scl* $^{-/-}$ ES cells are as previously described [2, 4]. *Scl* $^{RER/RER}$ ES cells were previously generated in the lab [187]. *Gata4* $^{-/-}$ ES cells were a kind gift from Dr Akira Murakami (Department of Viral Oncology, Kyoto, Japan).

SCL and GATA4 over-expressing ES cell lines were created on a J1 wild type background. *Scl* $^{-/-}$ ES cells were modified to re-express SCL (*Scl* $^{-/-}$ /SCL cell line) as described in Section 2.7. ES cell lines with inducible SCL expression were generated on a wild type and a *Scl* $^{-/-}$ background (iSCL/WT and iSCL/*Scl*$^{-/-}$ lines) as described in Section 2.8.

2.1.2 Maintenance of Mouse ES Cell Lines

Mouse ES cell lines were maintained on plastic plates coated with gelatin (0.1% in PBS). Cells were cultured in Dulbecco's Modified Eagle Medium (DMEM; Gibco) supplemented with 15% (v/v) foetal calf serum (FCS; PAA), 2% (v/v) Leukemia Inhibitory Factor (LIF) conditioned medium (obtained from a LIF-expressing CHO line), 50 µg/ml penicillin-streptomycin, 2mM L-glutamine and $1.5x10^{-4}$ M monothioglycerol (MTG; Sigma). Cultures were maintained in a humidified chamber at 37°C in a 5% CO_2/air mixture. For maintenance, cells were passaged every 2-3 days.

Cells were trypsinised with 0.25% (v/v) Trypsin/EDTA (Gibco) for 3 minutes. Correct trypsinisation of cells is essential avoids spontaneous differentiation of un-dissociated cell clusters and ensure a single cell suspension. Trypsiniated cells were re-plated at a 1:7 volume dilution in ES cell maintenance medium for continued culture.

2.1.3 *In vitro* Differentiation of ES Cells into EBs

Twenty-four hours prior to the onset of differentiation, ES cells were passaged onto gelatinised plates and cultured in Iscove's modified Dulbecco's Medium (IMDM; Gibco), containing the same supplements as the DMEM medium described in Section 2.1.2, to let the cells adapt to the differentiation medium base component.

2.1.3.1 Cell Suspension Method

In vitro differentiation into EBs was induced by the removal of LIF. ES cells were cultured in suspension in IMDM supplemented with 15% (v/v) FCS (batch-tested to promote haematopoietic differentiation; PAA), 2mM L-glutamine, 200 µg/ml human transferrin (for iron transportation; Roche), 0.5mM ascorbic acid (Sigma), 50 µg/ml penicillin-streptomycin and 4.5 $x10^{-4}$ MTG. 10^4 to 10^5 ES cells were plated per 10 ml of medium. These conditions have been previously described to sustain mesoderm / haematopoietic differentiation [115]. Unless otherwise stated, all EB differentiation assays were performed using this method.

2.1.3.2 Hanging Drop Method

ES cells were re-suspended in Glasgow Minimum Essential Medium (BHK21; Gibco) supplemented with (v/v) 20% FCS, 1 mM non-essential amino acids, 1mM sodium pyruvate, 1mM L-glutamine and $1x10^{-4}$ M mercaptoethanol. 20 µl hanging droplets, each containing 500 ES cells, were incubated for 48 hours at 37°C to induce EB formation. Day 2 EBs were re-suspended in 10 ml of medium for continued culture. These conditions have been previously described for the generation of EBs with cardiac potential [123].

2.1.4 Dis-aggregation of EBs

Prior to plating assay, flow cytomerty analysis and/or cell sorting, EBs were dis-aggregated using 0.25% (v/v) Trypsin/EDTA for 2 minutes at 37°C. Cells were passed through a 30 µM cell filter (Partec) to ensure a single cell population.

2.1.5 Re-aggregation of EBs

Following EB dis-aggregation and cell sorting, some assays required re-aggregation of single cell population in order to re-create cell-cell interactions that are often indispensable for cell differentiation. Single cell populations were re-aggregated for 20 hours in StemPro-34 SFM serum-free medium (Gibco) supplemented with 5% (v/v) Knock-Out serum replacement (Invitrogen), 2 mM L-glutamine, 200 µg/ml transferrin, 0.5 mM ascorbic acid and $4.5x10^{-4}$ M MTG, at a density of 4 to 6 $x10^5$ cells per ml in ultra-low-attachment 24-well plates (Corning) at 37°C (adapted from Kouskoff *et al*, 2005) [122].

2.1.6 Haematopoietic Differentiation

2.1.6.1 Blast colonies

To investigate haematopoietic/endothelial potential, blast colonies were generated by plating EB-derived, single cell suspensions in IMDM containing 2% (w/v) Methylcellulose, 10% (v/v) FCS, 20% (v/v) D4T conditioned medium (obtained from a Kit ligand expressing D4T embryonic endothelial cell line), 2mM L-glutamine, 300 µg/ml transferrin, 25 µg/ml ascorbic acid, $4.5x10^{-4}$ MTG and 5 ng/ml VEGF, at a density of $5x10^{5}$ cells per ml [1, 3]. The number of blast colonies generated was scored after 4 days incubation at 37°C.

2.1.6.2 Hemoglobin Production

EBs were generated as described in Section 2.1.3.1, with the addition of 5 % Protein-Free Hybridoma Medium (PFHM II; Gibco) as a supply of iron for hemoglobin production [115]. After 7 days of differentiation, EBs were visualised by light microscopy and the percentage of EBs containing red cells was calculated.

2.1.7 Cardiac Differentiation

2.1.7.1 Unsorted EBs

EBs, generated as described in Section 2.1.3, were re-plated onto gelatinised plates and maintained in BHK21 medium supplemented with 20% (v/v) FCS, 1 mM non-essential amino acids, 1mM sodium pyruvate, 1mM L-

glutamine and $1x10^{-4}$ M mercaptoethanol) at 37°C [123]. The number of contracting colonies was scored after a further 2-10 days of culture.

2.1.7.2 Sorted EB Populations

Single EB derived cell populations were re-aggregated as described in Section 2.1.5. Aggregates were transferred onto gelatinised plates and cultured at 37°C in IMDM with 10% Knock-Out serum replacement [122]. The proportion of beating cells was scored after 2-6 days of culture.

2.1.8 Osteogenic Differentiation

Whole EBs, or re-aggregated EB-like structures, were re-plated at a low density ($2.5x10^5$ cells per well) onto gelatinised 24-well plates. Cells were cultured in Knock-Out DMEM (Gibco) supplemented with 10% (v/v) FCS, 50 µM ascorbic-acid-2-phosphate (Sigma), 0.1 µM dexamethasone (Sigma), 10 mM ß-glycerophosphate (Sigma) and 10 ng/ml bone morophogenic protein 4 (BMP4; R&D Systems) at 37°C (adapted from Sakurai *et a*l, 2006) [125]. After 28 days, osteogenic activity was determined using Alizarin Red Staining (see Section 2.2.3).

2.1.9 Chondrogenic Differentiation

EB derived, single cell suspensions were re-plated at a density of $8x10^4$ cells/ml in Alpha Minimum Essential Medium (α-MEM, Gibco), supplemented with (v/v) 10% FCS, 50 µg/ml penicillin-streptomycin, 0.1 µM dexamethasone and 0.17 mM ascorbic-acid 2-phosphate. 10 ul droplets

were spotted onto 24-well culture plates and incubated for 30 minutes at 37°C to allow the cells to aggregate. Following the addition of 1 ml of the chondrogenic medium containing 10 ng/ml transforming growth factor beta 3 (TGF-ß3; R&D Systems), cells were incubated at 37°C for 7 days. Cells were then cultured for a further 7 days in the presence of 10 ng/ml BMP-2 (R&D Systems) (adapted from Sakurai *et a*l, 2006) [125]. Chondrogenic activity was assayed for by Alcian Blue staining (see Section 2.2.3).

2.1.10 Myogenic Differentiation

EB derived single cell populations were cultured on 24-well collagen type I coated dishes (R&D Systems) in KnockOut DMEM supplemented with 5% (v/v) horse serum (Gibco) and 2 ng/ml insulin-like growth factor-1 (IGF-1; R&D Systems). Cells were plated at a density of $2.5x10^5$ cells per well and cultured at 37°C for 14 days (adapted from Sakurai *et a*l, 2006) [125]. Expression of myocyte-specific markers was determined using immuno-histochemistry (see Section 2.2.4).

2.2 Cellular Staining Assays

2.2.1 Alkaline Phosphatase Staining of ES Colonies

Pluripotent ES cells can be characterised by their expression of Alkaline phosphatase; a hydrolase enzyme responsible for dephosphorylating molecules such as nucleotides, proteins, and alkaloids under alkaline conditions.

ES cells were plated onto gelatinised 6 well plates at a low density (100-500 cells per well) and cultured for 5 days to generate distinct colonies. Colonies were fixed using 4% (v/v) paraformaldehyde (PFA; Electron Microscopy Sciences; diluted in PBS) for 2 minutes at room temperature. Cells were washed twice with PBS containing 0.1% Tween-20 (PBST; Sigma) and stained with Fast Red Violet in the presence of Napthol-phosphate (Millipore Kit) for 15 minutes at room temperature, according to the manufacturer's protocol. Colonies were washed with PBST until all residual staining was removed, and visualised by light microscopy. The appearance of a red or purple stain indicated undifferentiated, pluripotent ES cells, while more differentiated cells were colourless. ES cells cultured in parallel in absence of LIF were used as a control.

2.2.2 Alizarin Red Staining for Osteogenic Differentiation

Osteogenic potential was measured by the formation of calcium deposits within cell clusters. These calcium deposits were detected using Alizarin Red S compound, which reacts with calcium ions to form a red salt.

Cells were washed twice with distilled water (to ensure a neutral pH) and fixed with ice cold 70% ethanol for 20 minutes at 4°C. After two more washes, cells were incubated with a 2% Alizarin Red S solution (in distilled water, pH 4.2; Sigma) for 30 minutes at room temperature. Cells were washed several times to remove excess dye and visualised by light microscopy.

2.2.3 Alcian Blue Staining for Chondrogenic Differentiation

As a measure of chondrogenesis, Alcian blue dye was used to screen for cartilage nodules as it selectively stains muco-molecules, including cartilage specific proteoglycans.

Prior to staining, cells were washed twice with PBS and fixed with 4% (v/v) PFA for 30 min at room temperature. Cells were stained by overnight at room temperature with a 0.05% (w/v) Alcian blue solution (in 3% (v/v) acetic acid, pH 1.5; Sigma). Cells were washed several times with PBS to remove excess dye and visualised by light microscopy.

2.2.4 Immuno-histochemistry for Myogenic Markers

Myogenic differentiation was detected by immuno-histochemical staining for Myogenin expression. Myogenin is a muscle-specific bHLH transcription factor involved in the coordination of skeletal muscle development and repair. Myogenin is first expressed during early muscle cell specification and throughout myocyte differentiation, and therefore suitable for use as a marker of myogenic potential.

Prior to antibody staining, cells were fixed for 30 minutes at room temperature with 4% (v/v) PFA. Endogenous peroxidase activity was sequestered by treatment with a 0.3 % (v/v) hydrogen peroxide (Sigma) solution for 30 minutes at 4°C. To permeablise the cells, so that the antibodies can penetrate the cell membrane and stain their nuclear targets, cells were incubated with 0.5% (v/v) Triton-X100 (in PBS; Invitrogen) for 10 minutes at room temperature. Cells were washed twice with PBST and incubated in a blocking solution (PBST containing 2% (w/v) skimmed milk

and 1% (v/v) goat serum (Invitrogen)) for 1 hour at room temperature. Cells were stained overnight at 4°C with the primary antibody, rat anti-Myogenin (Santa Cruz Biotechnologies), diluted 1:100 in blocking solution.

Cells were stained for two hours at room temperature with the secondary antibody, horseradish peroxidase (HRP) conjugated goat anti-rat IgG secondary (Santa Cruz Biotechnologies), diluted 1:250 in 2% (w/v) skimmed milk. HRP activity was assessed using a DAB-HRP detection kit (Vector Laboratories) according to the supplied protocol. The reaction was carried out for 8 minutes, and cells were fixed post-staining with 4% (v/v) PFA at 4°C overnight before visualization by light microscopy.

2.3. Immuno-phenotypic Analysis and Cell Sorting

2.3.1 Antibody Staining for Extracellular Proteins

Prior to flow cytomertic analysis and cell sorting, EBs were dis-aggregated into single cell suspensions as described in Section 2.1.4. Single cell suspensions were stained with FITC conjugated anti-mouse E-cadherin antibody (BD Biosciences), PE-conjugated anti-mouse Flk-1 antibody (eBiosciences), APC conjugated anti-mouse PDGFRα antibody (eBiosciences) and/or PE-Cy7 conjugated anti-mouse CD41 (BD Biosciences). Cells were stained in PBS containing 10% (v/v) FCS for 20 minutes at 4°C. After two PBS washes, cells were re-suspended in PBS containing 10% (v/v) FCS. Hoechst was added to prior to FACS analysis or cell sorting.

2.3.2 Antibody Staining for Intracellular Proteins

Following staining for extra-cellular proteins, cells were fixed with ice cold 2% (v/v) PFA for 10 minutes at 4°C. Cytoplasmic and nuclear cell membranes were permeablised with 0.1 % (v/v) Triton-X 100 detergent (Sigma) for 10 minutes at 4°C to enable antibodies to enter the cell nucleus.

For SCL intra-cellular flow cytomerty, cells were blocked in 5% (v/v) donkey serum (Invitrogen) in PBS for 10-20 minutes at 4°C. 10 ul of a mouse monoclonal anti-human SCL antibody (clone BTL73, kindly supplied by Dr Karen Pulford, Oxford University, UK; [206]) was added, and cells were incubated for a further 20 minutes at 4°C. Cells were washed with PBST and re-suspended in PBST containing 1:5000 dilution of a secondary Alexa Flour 488 conjugated (donkey, anti-mouse) antibody and incubated at 4°C for a further 20 minutes.

For GATA4 intra-cellular flow cytomerty, cells were re-suspended in PBS / (v/v) 10% FCS and stained with FITC-conjugated anti-mouse GATA4 antibody (BD Biosciences) for 20 minutes at 4°C. For Cardiac Troponin T (cTNT), a purified anti-mouse cTNT antibody (Thermo Scientific) was directly conjugated to APC using the Zenon® Labelling Technology Kit (Molecular Probes) according to the supplied protocol. Cells, re-suspended in PBS containing 10% FCS, were stained for 30 minutes at 4°C.

2.3.3 Flow Cytometry

All cells were washed twice with PBS and re-suspended in PBS containing (v/v) 10% FCS. Hoechst dead cell stain (Invitrogen) was added to un-fixed cells prior to analysis. Flow cytometric analysis and cell sorting was

performed using a CyAn™ flow cytometer and MoFlo™ cell sorter respectively.

2.3.4 Analysis of Cell Cycle and Proliferation

Cell cycle analysis was performed using the BrdU-APC flow kit (BD Biosciences). Cells isolated from EBs (as in Section 2.1.4) were incubated in 10 μM BrdU for 20 minutes, prior to staining with FITC E-cadherin and PE Flk-1 antibodies (as in Section 2.3.1) and a PE-Cy7 conjugated PDGFRα antibody (eBiosciences). Cells were fixed for 15 minutes at 4°C with BD Cytofix Buffer (PFA based; BD Biosciences) and permeablised with BD Cytoperm Buffer (Saponin based; BD Biosciences) for 10 minutes at 4°C. Cells were treated with 30 μM DNase to expose BrdU epitopes, and stained with an APC conjugated anti-BrdU antibody for 20 minutes at room temperature. After two PBS washes, cells were re-suspended in PBS containing 10% (v/v) FCS. 7-AAD was added to prior to flow cytometric analysis.

2.3.5 Analysis of Apoptosis

Cell death and apoptosis was measured using the Annexin V based Vybrant® Apoptosis Assay Kit (Molecular Probes). Prior to analysis, EBs were dis-aggregated as described in Section 2.1.4, and stained with FITC E-cadherin, PE Flk-1 and PE-Cy7 PDGFRα antibodies as described in Section 2.3.4. Cells were re-suspended in the supplied Annexin V Binding Buffer (Molecular Probes) and incubated with 5 μl of APC conjugated anti-Annexin V antibody for 15 minutes at room temperature. Cells were washed and re-suspending in Binding Buffer, and treated with 100 μg/ml of PI prior to flow cytometric analysis.

2.4. Protein Expression

2.4.1 Nuclear Extracts

Nuclear extracts were isolated from single cell suspensions. Cells were first re-suspended in 10x volume of ice cold Buffer A (10mM Hepes pH7.9, 1.5mM $MgCl_2$, 1mM DTT, 1% (w/v) mixed protease inhibitors (mini-complete, EDTA-free; Roche)) for 10 minutes. Once swelled, cells were vortexed for 10 seconds and nuclei were collected by a brief centrifugation (5 seconds at 1000 rpm). Pellets were re-suspended in ice cold Buffer C (420 mM NaCl, 20mM $MgCl_2$, 0.2M EDTA, 1mM DTT, 1% (w/v) mixed protease inhibitors and incubated at 4°C, with rotation, for 30-60 minutes, then centrifuged for 4 minutes at 4°C at 14,000 rpm. Supernatant was removed and stored at -80°C. Protein concentration was quantified using the Bradford Assay. Briefly, 2 ul of sample was added to 1ml Coomassie Blue G-2510 solution (Thermo Scientific) in a clear cuvette (Eppendorf), and absorbance was read at 595nm. Protein concentration was calculated using a BSA protein standard curve.

2.4.2 Whole Cell Lysates

Cells were incubated for 5 minutes at room temperature in PBS containing 4% SDS, 1 mM DTT and 1% (w/v) mixed protease inhibitors. 1µl (250 units) Benzonase® endonuclease (Sigma) was added to digest native nuclear acids. Lysates were then stored at -80°C.

2.4.3 Immuno-Precipitation

500 µg of nuclear extracts were diluted in PBS (containing 100mM DTT and 1% (w/v) mixed protease inhibitors) to reach a 150 mM NaCl concentration. Nuclear extracts were incubated overnight at 4°C with purified α-mouse SCL or α-mouse Gata4 monoclonal antibodies (Santa Cruz), cross-linked to protein G Dynabeads (Molecular Probes). Beads were washed twice with PBS (containing 100mM DTT and % (w/v) mixed protease inhibitors) and proteins were denatured and eluted from beads by incubation at 95°C for 10 minutes.

2.4.4 Western Blots

20-30 µg of nuclear extracts, or 10 µl of whole cell lysates, were denatured by heating to 95°C for 5 minutes. Proteins were subsequently resolved by SDS-PAGE onto a 4-12% gel (NuPage; Invitrogen) and blotted onto a nitrocellulose membrane (Hybond; Amersham Biosciences) using standard protocols.

Blocking was performed by incubating the membranes in PBST containing 5 % (w/v) milk, overnight at 4°C. After 3 washes in PSBT, membranes were probed with primary antibodies, diluted in blocking buffer, for 45 minutes at room temperature. Goat anti-mouse SCL and goat anti-mouse GATA4 antibodies (Santa Cruz) were used at a 1:200 dilution. The rabbit anti-mouse histone 3 (Abcam) was used at 1:10,000. After 3 PBST washes, membranes were incubated with HRP conjugated secondary antibodies (diluted in PBST) for 45 minutes at room temperature. Donkey anti-goat HRP (Santa Cruz) was used at 1:10,000 and conjugated protein G

(Invitrogen) at 1:20,000. Membranes were developed using the ECL Plus western blotting system (GE Healthcare), according to the supplied protocol, and exposed to Kodak film.

2.5. Gene Expression Analysis

Real-time PCR and Microarray analysis were used to assess the relative levels of mRNA transcripts in a variety of cell types at different stages of differentiation.

2.5.1 RNA Preparation

Total RNA was extracted with Trizol® (Invitrogen) according to the manufacturer's protocol, with starting cell numbers ranging between $5x10^5$ and $2x10^6$. Briefly, homogenised cells were lysed with 1 ml Trizol® for 5 minutes at room temperature and passed through a 25 gauge needle to shear genomic DNA and to ensure complete dissociation of nucleoprotein complexes. Samples were mixed and incubated with chloroform for 5 minutes at room temperature to allow phase separation. Following centrifugation, the aqueous phase was mixed with isopropanol and 5 µg linear acrylamide (LPA; Sigma) to precipitate the RNA. RNA pellets were recovered by centrifugation, washed with ice cold 70% ethanol, re-suspended in RNase free water (Gibco) and stored at -80°C.

2.5.2 cDNA Synthesis

Residual DNA was removed from RNA samples by treatment with DNaseI (Qiagen). RNA samples were diluted in the supplied enzymatic buffer before incubation with DNaseI for 30 minutes at 37°C. cDNA was generated using the Sensiscript RT Kit (Qiagen) according to the manufacturer's guidelines. Briefly, 50 ng of DNaseI treated RNA was reverse transcribed using the supplied reverse transcriptase enzyme, in the presence of random primers (Invitrogen) and dNTPs (Invitrogen), for 1 hour at 37°C. cDNA was then quantified and stored at -20°C.

2.5.3 Quantitative PCR

Quantitative real-time PCR was performed using either 400nM of forward and reverse primers diluted in SYBR Green PCR mix (Applied Bioscience) or using 500nM forward and reverse primers and 250nM of TaqMan® fluorescent probe, diluted in TaqMan® PCR mix (Applied Bioscience). qPCR were performed in a 25µl reaction, using the following cycling conditions: 95^0C for 10 minutes, followed by 40 cycles of amplification (95^0C denaturation for 15 seconds 60^0C annealing for 1 minute and 72^0C elongation for 2 minutes). Expression of *Gapdh* and/or *18s* housekeeping genes were analysed to provide a control for normalisation (see Table 1 for primer sequences).

2.5.4 Statistical Analysis

Unless other stated, a Student t test was used to determine statistical significance. Confidence intervals were set at 95%, 99% and 99.9%, where * indicates a P-value of < 0.05, * * indicates a P-value of < 0.01 and * * * indicates a P-value of < 0.001 in all figures.

2.5.5 Microarray and Bioinformatic Analysis

Expression profiling was conducted using RNA isolated from Flk-1 expressing EB cells, derived from wild type and *Scl* $^{-/-}$ ES cells, after 3.5 days of differentiation (cell populations were isolated by flow cytomery as described in Section 2.3). Microarray was performed using the Ilumina Sentrix MouseWG-6 v2.0 Expression BeadChip, covering 48,318 probes. Statistical analysis was used to compare expression profiles in the wild type and *Scl* $^{-/-}$ condition. RANK normalisation was conducted prior to data filtration, followed by statistical tests using Limma, to generate a list of genes differentially expressed in between wild type and *Scl* $^{-/-}$ cells.

2.6. Chromatin Immuno-Precipitation (ChIP)

2.6.1 α-SCL ChIP

DNA binding proteins were cross-linked to DNA by incubating cells with 1% (v/v) formaldehyde (Sigma) for 10 minutes at room temperature with rotation. The reaction was quenched by addition of 1 M glycine (Sigma) for 10 minutes at room temperature with rotation. To increase the strength and distance of fixed protein-protein interaction, thus allowing cross-linking of

protein complexes to DNA, cells were incubated in 2 mM EGS (ethyleneglycol-bis-succinimidyl-succinate; Sigma), 2mM DSG (Di(N-succinimidyl) glutarate; Sigma) or 20mM DMP (2,2-Dimethoxypropane; Sigma) for 45 minutes at room temperature prior to formaldehyde fixation. Cell pellets were washed twice with ice cold PBS (containing protease inhibitors) and lysed for 10 minutes at 4^{o}C in 2% (w/v) SDS solution (Invitrogen). Formaldehyde treated chromatin was sonicated for 10 minutes (30 seconds on/off using Biorupter). Chromatin treated with EGS or DSS plus formaldehyde was sonicated for 30 minutes (30 seconds on/off) to generate DNA fragments of approximately 200-500bp. 5 ul of sonicated lysates were run on a 2% (w/v) agroase gel to ensure DNA fragments were of a correct size.

Prior to immuno-precipitation, chromatin solutions were incubated with protein-A beads (either salmon-sperm; Millipore, or Dynabeads; Invitrogen) for 1 hour at 4^{o}C to pre-clear any endogenous proteins capable of binding the beads and therefore reducing background noise. Immuno-precipitation was carried out overnight at 4^{o}C using 15 ul of a rabbit α-SCL whole serum antibody [173], followed by a 2 hour incubation with protein-A beads at 4^{o}C to pull down the complexes. Beads were recovered and washed with buffers of decreasing salt concentrations, according to the supplied protocol (Upstate ChIP Kit; Millipore).

DNA/Protein complexes were eluted from the beads by incubation in PBS containing 1 % (w/v) SDS and 100 mM $NaHCO_3$ for 30 minutes at room temperature. To recover the DNA fragments bound by SCL, this elution was incubated with 5mM NaCl for 4 hours at 65^{o}C to reverse the DNA-Protein cross-linking. Cells were lysed with 2 µg/µl of Proteinase K (diluted in Tris-HCL pH 6.8 containing 5mM EDTA) for 1 hour at 65^{0}C. DNA fragments were

isolated by phenol-chloroform extraction (as described in Section 2.7.4). DNA was precipitated with 100% ethanol and 20 µg glycogen (overnight at -20^{0}C) and centrifuged at 4^{0}C for 20 minutes at 1400 rpm. Pellets were washed with 70% ethanol, air dried for 10 minutes and re-suspended in water.

2.6.2 Quantitative PCR Analysis of ChIP

Relative enrichment of specific DNA regions eluted from ChIP experiments was measured by real-time PCR performed using TaqMan® or SyberGreen® PCR, using the aforementioned cycling conditions (Section 2.5.3). Relative enrichment of DNA sequences in IPs was calculated as a fold change over relative levels found in input material. Primer sequences are shown in Table 2.

2.6.3 ChIP-Sequencing

To globally identify genomic regions bound by SCL in early EB mesoderm development, ChIP experiments were followed by next generation Solexa (Illumina) sequencing, performed by the Genomic Services Department, Wellcome Trust Centre for Human Genomics, Oxford, UK. Reads were mapped to the mouse genome (build m37) using Maq (Mapping and Assembly with Quality Programme). Peak calling was performed using MIG (Multi-Image Genome); an in-house programme developed by Steve Taylor, CBRG Department, Oxford, UK.

2.7. ES Cell Lines with Ectopic SCL or GATA4 Expression

2.7.1 Constructs

SCL or Gata4 cDNA were amplified from total RNA isolated from J1 wild type EBs after 4 days of differentiation (RNA extraction and reverse transcription reactions as described in Sections 2.5.1 and 2.5.2). Coding sequences were amplified using Phusion high fidelity DNA polymerase (Finnzymes) and primers show in Table 3. PCR products were digested with XbaI (New England Biolabs) and gel purified using QIAquick gel extraction kit (Qiagen).

DNA fragments were ligated into a XbaI linearised pEF1α-Neo vector (generated in the lab by Dr Isla Hamlett) using Rapid Ligation Kit (Roche). DNA constructs was transformed into DH5α competent bacteria cells (Invitrogen) according to the manufacturer's protocol. Cultures were plated onto agarose/LB plates containing 100 µl/ml Ampicillin for 48 hours at 37°C. Clones grown from antibiotic resistant cells were picked and expanded overnight in LB (10% (w/v) Bacto-tryptone, 10% (w/v) NaCl, 5% (w/v) yeast extract, pH7.5). Cells were lysed and DNA was isolated using a plasmid Mini Kit (Qiagen). Insertion of cDNA was confirmed by restriction enzyme digest using XbaI. Correct orientation was confirmed by sequencing (see Table 3 for primer sequences).

Since *Scl* $^{-/-}$ ES cell lines already express a Neo resistance gene, the Neo cassette in the pEF1α-SCL construct was replaced with a gene encoding for Zeocin resistance. The Neo cassette was removed by digestion with ClaI and Asp718 restriction enzymes (NEB). A Zeo cassette (provided by Dr Sarah Hoosdally) was digested using the same enzymes. Fragments were ligated and transformed into DH5α competent cells as described

previously. DNA was recovered from Ampicillin resistant cells using a Plasmid Mini Kit (Qiagen). Insertion of the Zeo fragment was confirmed by restriction enzyme digest. Bacterial clones expressing the final pEF1α-SCL-Zeo and pEF1α-GATA4-Neo targeting constructs were expanded overnight in LB and DNA was isolated using the Plasmid Maxi Kit (Qiagen).

2.7.2 Electroporation

150 µg of plasmid construct was precipitated with 100 % ethanol at -20°C for 30 minutes and re-suspended in 200 µl sterile PBS. Trypsinised ES cells were passed through a 30 µM filter (ensuring a single cell population is critical). Approximately $5x10^7$ cells were re-suspended on 0.6 ml PBS and transferred into a 0.4 cm gap cuvette (BIORAD) containing 100 µl of the linearised targeting construct. A proportion of the remaining cells were re-plated as a control. Cells were electroporated at 450 volts, 25 µF, with a time constant of 0.6 milliseconds. After 10 minutes at room temperature, cells were diluted in 9 ml of medium (see Section 2.1.2) and 1 ml of this mix was transferred onto 10 pre-gelatinised 10 cm plates containing 9 ml of medium.

2.7.3 Selection of ES clones by Antibiotic Resistance

Both constructs electroporated contained a resistance cassette, either for Neomycin (pEF1α-GATA4-Neo) or Zeocin (pEF1α-SCL-Zeo). After 48 hours incubation at 37°C, media was replaced on 9 out of 10 electroporated plates with selection medium containing 300µg/ml G148 (Invitrogen; for Neomycin resistance) or 400 µg/ml Zeocin (Invitrogen).

2.7.4 DNA Lysis, Extraction and Precipitation

Cells were lysed by overnight incubation at 37°C in lysis buffer (100 mM Tris pH8.5, 5 mM EDTA, 0.2% (w/v) SDS, 200 mM NaCl), containing 200 µg/ml Proteinase K (Invitrogen). Lysed samples were mixed with 20 µl of 3 mM NaAc, and 200 µl Phenol/Chloroform (Fluka) and centrifuged at room temperature for 5 minutes at 1300 rpm. To precipitate the DNA, 120 µl of ice cold isopropanol and 20 µg glycogen (Invitrogen) was added to each sample before incubation at -20°C for 30 minutes and centrifugation at 1300 rpm at 4°C for 20 minutes. Pellets were then washed with 70% ethanol, air dried for 10 minutes and re-suspended in water.

2.7.5 Selection of ES Clones by PCR and Western Blot

Antibiotic resistant clones were screened using PCR to confirm genomic integration of the construct (see Table 3 for primers). 50 µl PCR reactions were set up using 100 ng of genomic DNA (extracted from ES clones as described in Section 2.7.4), 300 ng of forward and reverse primers, 1 mM dNTPs and 2.5 Units of Pfu Ultra polymerase (Stratagene). The PCR cycling condition were as follows; denaturation for 2 minutes at 95°C, 30 cycles of denaturation for 30 seconds at 30°C, annealing for 30 seconds at 55.6°C and elongation for 1 minute at 72°C, followed by a final 10 minute elongation at 72°C. PCR products were visualised on a 1% agarose gel. To ensure that the selected clones correctly expressed SCL or GATA4 proteins, nuclear extracts or whole cell lysates were prepared from ES cell cultures, and western blots were performed using anti-SCL and anti-GATA4 antibodies (as described in Section 2.4.4).

2.8. ES Cell Lines with Doxycycline Inducible SCL Expression

To study the impact of SCL expression on EB differentiation in a time dependent manner, cell lines were created where the expression of a *Scl* cDNA could be induced, silenced and controlled depending on the addition and removal of doxycycline (Dox; Sigma). Cell lines were generated on wild type and *Scl* $^{-/-}$ backgrounds, creating the WT/i*Scl* and *Scl* $^{-/-}$/i*Scl* ES cells.

2.8.1 Constructs

Plasmid vectors referred to as pROSA26-Tch, pEX-Oct3/4 and pCre-recombinase were kindly donated by Dr Hitoshi Niwai (RIKEN Center for Developmental Biology, Kobe, Japan).

2.8.2 Generating ROSA26/Tetracycline Expressing ES Cells

150 µg of the pROSA26-TcH targeting construct (Figure 5.1B) was linearised using AscI restriction enzyme (New England Biosciences) for 2 hours at 37°C, precipitated with 100 % ethanol at -20°C for 30 minutes and re-suspended in 200 µl sterile PBS. Electroporation was performed as described in Section 2.7.2. Targeted ES cells were selected for by antibiotic resistance using medium containing 100 µg/µl of Hygromycin B (Invitrogen). Genomic DNA was isolated from Hygromycin resistant ES cell clones as described in section 2.7.4. PCR was used to confirm genomic incorporation of the targeting construct, as described in Section 2.7.5.

2.8.2.1 Southern Blot

Southern blot was used to confirm homologous recombination occurred correctly in the ROSA26 locus. 10 µg of DNA from each ES cell clone (all positively selected according to initial PCR screening) was digested with the restriction enzyme EcoRV (Roche) in a 100 µl reaction (total of 400 Units of enzyme over 24 hours). Digested samples were run on a 0.8 % agarose gel. Standard southern blot protocols were followed [207]. Briefly, gels were denatured for 1 hour in 1M NaOH and neutralised for 1 hour in buffer containing 1M Tris, 3M NaCl at pH 7.4. DNA was transferred onto Zeta-probe membrane (BIORAD) by upwards capillary transfer, and membranes were baked for 2 hours in an 80°C oven. A probe designed to target the ROSA26 locus (see Figure 5.5 for location) was labeled using the Megaprime Labeling Kit (Amersham) and hybridization was carried out in a formamide-based buffer. Correctly targeted clones (R26-TcH) were selected for secondary targeting depending on their ability to self-renew and differentiate. These properties were analysed by Alkaline Phosphatase staining (see Section 2.2.1) and EB formation (see Section 2.1.3.1).

2.8.3 Introduction of *Scl* cDNA under the Control of a Tetracycline Responsive Promoter

2.8.3.1 Modification of the Cloning Vector

To generate an exchange vector which would enable the introduction of *Scl* encoding cDNA under the control of the hCMV-1* tetracycline-responsive element within the ES cells, *Oct3/4* cDNA in the pEX-Oct3/4

exchange vector was excised and replaced with *Scl* cDNA, creating the *Scl* Exchange Vector targeting construct (see Figure 5.5).

To achieve this, the Oct3/4 transgene was excised by restriction enzyme digest using NotI and Xhol (New England Biosciences). The vector backbone was recovered by gel extraction, according to the supplied protocol (Qaigen). *Scl* cDNA was amplified by Phusion polymerase (see Section 2.7.1) and subsequently digested with Xhol. The *Scl* cDNA-Xhol fragment was ligated into the pEX-vector backbone using a Rapid Ligation Kit (Roche) and transformed into DH5-α competent bacterial cells (Invitrogen) using standard protocols. Correct insertion was assessed by PCR for *Scl* cDNA (see Section 2.7.5).

2.8.3.2 Lipofectamine Transfection

Selected R26-TcH clones were targeted with the *Scl* Exchange Vector construct. R26-TcH ES cells were co-transfected with 2 µg of *Scl* Exchange Vector and 5 µg of circular pCre-recombinase (see Figure 5.5 for strategy). Transfection was performed using Lipofecamine 2000 (Invitrogen) according to the manufacturer's protocol.

2.8.3.3 Selection of ES clones

Targeted ES clones were selected by puromycin resistance (1.5 µg/µl of Puromycin; Invitrogen) and hygromycin sensitivity (100 µg/µl of Hygromycin B). Puromycin resistant and Hygromycin B sensitive clones were screened by PCR to confirm correct recombination had occurred; genomic DNA was isolated from lysed ES cell clones, as described in Section 2.7.4.

PCR for amplification of SCL cDNA was performed as described in Section 2.7.5 (See Table 3). To ensure that selected clones correctly expressed the SCL protein, nuclear extracts were prepared from ES cell cultures, and western blots performed using anti-SCL antibodies (as described in Section 2.4.4).

Table One.
PCR Primers for Gene Expression Analysis

Gene	Forward 5'-3'	Reverse 5'-3'
18S	TTC CGA TAA CGA AAC GAG ACT CTG G	CTG AAC GCC ACT TGT CCC TCT AAG
Brachyury	TTG ATG CCA AAG AAA GAA ACG ACC	GGA ACA AGC CAC CCC CAT TG
Bglap I	GAG GAC CAT CTT TCT GCT CAC TCT	GAC ATG AAG GCT TTG TCA GAC TCA
Bgalp II	GCG CTA CCT TGG AGC TTC AG	CAT ACT GGT TTG ATA GCT CGT CAC A
Cdh5	TGC CCT GAA GAA CGA GGA CAG C	GCC CA TAC TTG ACC GTG ATG TTG G
Cdx2	CCA CAC TTG GGC TCT CCG AG	GCT GCT GCT TCT TCT TGA TTT TCC
Collagen II	CCG TCAG AGT ACC GAT CA	CAG GTC AGG TCA GCC ATT CA
Collagen X	AAG GAG TGC CTG GAC ACA AT	GTC GTA ATG CTG CTG CCT AT
Ebaf	GCT GGT CCG CTT TGC CTC G	TGG TGC TTC AGG GTC ACA GTC
Eomes	AAG TGG GTG ACC TGC GGC AAA G	TGT TAG GAG ATT CTG GGT GAA CG
Eto-2	CCACGGCTGCTTAAAGTGGT	GTCATTGCCAAATTGCTGTAGG
Fgf8	TGG AAG CAG AGT CCG AGT TCG C	CAG TCC TTG CCT TTG CCG TTG
Fli-1	ACG GAT TGA TGG AGA TTG AC	AGG AGA GGA CTT TTG TTG AGG
Flk-1	TGG TTG TGA ATG TCC CAC CCC	CCA TAC TGG TAG GAA TCC ATA GGC G
Flt4	CTC GCT CGG GAC ATC TAC AAA GAC	ATT TCA GAG GAA GTC GGG CAC TGC
FoxH1	ATG ACA AGC CCC CCT ACA C	GAC CTG ACG GAT AAT CTG AGC
Fst	CCG AGG AGG ATG TGA ACG ACA ATA C	ATT TTT TCC CAG GTC CGC AGT C
Gapdh	CAT GGC CTT CCG TGT TCC TA	CCT GCT TCA CCA CCT TCT TGA T
Gata2	CTA AGC AGA GAA GCA AGG CTC GC	GGC ACC ACA GTT GAC ACA CTC CC
Gata3	TTT ACC CTC CGG CTT CAT CCT CCT	TGC ACC TGA TAC TTG AGG CAC TCT
Gata4	CAA CCC TGG AAG ACA CCC CAA TC	TCC CGT CCC ATC TCG CCT C
Glt2	TCT TCC TGT GCC ATT TGC TGT TG	CCT GGG GTC CTC AGT CTT CTT TTC
Grb10	TGG AGG AAG CGA AGC ACA CG	AGAG GTG GAG GGG ACT TTG G
Gsc	GGA GGA GAA GGT GGA GGT CTG G	CGG CGA GGC TTT TGA GGA C
Hand2	CCG ACA CCA AAC TCT CCA AG	TCT CTG CTG CCT CCT TCT CCT TG
Hprt	TCC TCC TCA GAC CGC TTT TTG C	CAT CAT CGC TAA TCA CGA CGC TG
Ikaros	GGA TTC CTC GGT GCG GTT G	ACA AAG TCC AAA GGC TCT GGT TTC
Isl1	CAG CAT AGG CTT CAG CAA GAA CGA C	GCA GGC TAC ACA GCG GAA ACA CTC
Lyl-1	CTG ACA AAC CTG ACC ACA CAG CC	GGA CCC CAC GGA TAG AAT GCT AAC
Mef2c	AGA TAC CAA CAC ACC ACG CGC C	ATC CTT CAG AGA GTC GCA TGC GCT T
Meis1	CAG CAA ATC TAA CTG ACC AGC CCT C	TGA ACG AGT GGA TGC CGT GTC
Mesp1	TCC CAG GAAG CAG GAA ATG	TGC TGA AGA GCG GAG ATG AGG
Mesp2	CAA GAC TGG ACA CTG GAC ACA ATC	AGG CAG GGT TCT GGA GAC ACA G
Msx2	AGA TAC TCC CCG CCG CCC AG	GGT TGG TCT TGT GTT TCC TCA GG
MyoD	CAC TCC GGG ACA TAG ACT TGA CA	GTC GAA ACA CGG ATC ATC ATA GAA
Myogenin	GAC CCT ACA GAC GCC CAC AA	CAT ATC CTC CAC CGT GAT GCT
Nkx2.5	CCC AAG TGC TCT CCT GCT TTC C	GCT CTT TTT TAT CCG CCC GAG G
Nkx3.1	GTG AAC ATA ATC CAG GGG ACT TAG C	TTC TGT GGC TGC TTG GTG ACC
Osteocalcin	TCT GAC AAA GCC TTC ATG TCC	AAA TAG GAT ACC GTA GAT GCG
Pbx1	GCT AAC TCG CCC TCT ACT CCC AAC	CCG TCA CTG TAT CCT CCT GTC TGG
Pdgfra	TGT CTC TCT CAT CGT CCT GGT GGT C	TTT CAT ACC GTG GTT TCT GCT TCC
Pitx2	GTG TGG ACC AAC CTT ACG GAA GC	CGG CGA TTC TTG AAC CAA ACC C
Tbx5	TCT GCC CCC ACC TAA CCC ATA C	GTG CTC CGT GCT GGA ACA TTC
Tbx6	TCGCAGCCAATCCCTTTGC	CCCGCTCCCTCTTACAGTTTCTG
Tbx20	CGA GCC AAT GCC TTC TCC ATC	TCT CTG CTG CCT CCT TCT CCT TG
Tek/Tie2	GCG TCA ACA AGC ATC CTT CCT ACC	CAT TAT CTC ACA TTC CCT
Twist2	AGG TGC CGA AAG TCA CAG CC	CAG GTG CCG AAA GTC ACA GCC
Scl	CAC TAG GCA GTG GGT TCT TTG G	TCA CCC GGT TGT TGT TGG T
Skeletal Actin	CAC TGA GCG TGG CTA TTC CTT	CAC ATA GCA CAG CTT CTC TTT GAT GT
Sox9	CGT GGA CAT CGG TGA ACT GAG C	TTG GGT GGC AAG TAT TGG TCA AAC
Sox18	GCC CGT TTC CCA ATC CTC TGT C	CCA GTT GCT CTG TGC CCT CAA G

Table Two.
PCR Primers for ChIP Analysis

Genes	Forward 5'-3'	Reverse 5'-3'	Probe
Fli-1 +20	TTC CTT TTA TCA GCC CAA GTG A	CAA AGG GAA AAG GAA ACA TAT CA	ACT TCC TGT GGA CTG GCC CAG CTT C
FoxC1	TGG AAC CAG GGA GGA AAT CAT AG	TTG GAG AAA AGC GGG AGA GGT G	N/A
Gapdh	CAT CCATGA CAA CTT TGG TAT CGT	AAG GAC TCA TGA CCA CAG TCC ATG C	N/A
Gata2 -3	CGC GGA TCT GTG GTG GTA A	TTC AGA GCG CAC ACA ACA AAT	TGC ATT CGA TGG CGG GCG
Gata4 Enhancer	TGC CAA GAG GTG CTC ATT GAC C	AGG GGA AGC CGC AGA TAA GAA AC	N/A
Lyl-1 Promoter	GCA GCC AGA CCC TTA TCT GCA C	GCT TCC TCC CTC TTA CCC GCT	N/A
Nkx2.5	AGG GGA AGC CGC AGA TAA GAA AC	GCA CCA AGA AGG TCT GG	N/A
Rybf	GAT GCG AGG TTT CCT GGT GGA G	TGA GGA TGG CAT TGT GTG GCA C	N/A
Scl Promoter	GGG CGG GAG CGG GAT TAG	CCT TGG GGG TGT GCG TGG	N/A
Scl -16	GCA ATG AAC CTC CGA ACT GG	CGT CTA AGA AGG TGC CCA CAG	N/A
Scl +19	TTA AGC TCT CCC CAC TGG TTT G	CCT CCC CCA CTT CCT CTC TT	TGG GCT TTC TCT CCA CAC AAT AAC AGG ATG
Scl +40	TCC TAA ACC CTTGGT GCC TG	GAG CTG GCG ATA AGG AAG AGG	N/A
Sox9 Promoter	TGT CAG TTC AAG GTC GGC GTG	AGG CAA AGG CTG GCG GCA G	N/A

Table Three.
PCR Primers for Cell Line Generation

Name	Forward 5'-3'	Reverse 5'-3'
Scl cDNA	GCG AGA ATG ACG GAG CGG CCG CCG	ACT CCG GGG GCC AGC CCC ATC AGC
Scl cDNA + Xhol	GCG AGA CTC GAC ATG ACG GAG CGG CCG CCG	ACT CCG GGG GCC AGC CCC ATC AGC
Gata4 cDNA	CCG CAC TCT AGA ATG GAT TAT AAG GAT GAT GAT GAT AAG TAC CAA	CCT GAG TAT AGA TTA CGC GGT GAT TAT GTC CCC ATTG
tTA/i*Scl* Genotyping	AGA AGA GGC TGT GCT TTG T	TTC TCT TCC CAT GAA TTC G

3. Results Chapter One: SCL Expression During Early Mesoderm Development

3.1 Characterising Wild Type Mesoderm Populations in EBs

3.1.1 Flow Cytometry Protocol to Study Mesoderm Differentiation

Expression of cell surface proteins not only allows cells to respond to extra-cellular signalling pathways, thus promoting correct developmental differentiation, but also allows the characterisation of individual populations by flow cytometric analysis. For example, in the mouse embryo early mesoderm populations can be fractionated according to expression of two cell surface receptors, Flk-1 and PDGFRα [20-22, 95, 208]. *In vitro*, the mouse ES cell / embryoid body (EB) differentiation system recapitulates developmental processes occurring at the early stages of mouse embryonic development, including the formation of Flk-1/PDGFRα mesoderm populations (Figure 1.4) [208].

To study mesoderm development in our ES/EB system a flow cytometric protocol was developed using three cell surface markers; Flk-1, PDGFRα and E-Cadherin (Figure 3.1A). E-Cadherin expressing cells were excluded prior to analysis of Flk-1/PDGFRα populations to remove committed endoderm cells, as well as any E-Cadherin$^+$ PDGFRα^+ primitive mesendoderm cells [113, 209] (Figure 3.1A). Interestingly, only a small percentage of E-Cadherin expressing cells were identified, with expression peaking at 5% of total live cells in day 3 EBs (data not shown).

Out of these E-Cadherin$^+$ cells, more than 70% co-expressed PDGFRα, indicating the presence of early mesendoderm precursors rather than committed endoderm cells. This lack of cells committing to an endoderm fate reflects the nature of our *in vitro* ES/EB system. Our system is optimised for the study of mesoderm / blood formation (see Section 2.1.3.1) and therefore does not promote endoderm or ectoderm lineages. Unless otherwise stated, all Flk-1 and/or PDGFRα expressing cells described are negative for E-Cadherin expression.

3.1.2 Flk-1 and PDGFRα Expressing Mesoderm Populations in Wild Type EBs

To document the developmental kinetics of Flk-1/PDGFRα mesoderm formation in our ES/EB system, Flk-1 and PDGFRα expression was examined throughout the first 6 days of EB differentiation (Figure 3.1.B). Expression of both cell surface markers was first identified in day 2 EBs, with the occurrence of single positive populations, expressing either PDGFRα (single positive PDGFRα; SP-P) or Flk-1 (single positive Flk-1; SP-F) (Figure 3.1B). By day 3, an additional PDGFRα^+ Flk-1$^+$ (double positive; DP) population could be identified. All three Flk-1/PDGFRα populations expanded between days 3 and 4 of differentiation, whilst the double negative (DN) population decreased. By day 5/6, these dynamics changed, with the progressive loss of the DP population coinciding with an increase in DN cells (Figure 3.1B). The main expansion of Flk-1/PDGFRα expressing populations (SP-P, DP and SP-F) occurred between day 3 and 4 of EB differentiation (Figure 3.1B).

To understand the developmental relationship between these populations and the hierarchy of Flk-1/PDGFRα expression, SP-P, DP, SP-F, and DN populations were sorted from day 3.5 wild type EBs (Figure 3.2A) and re-aggregated to allow further differentiation for 24 hours. Re-aggregation of cells into EB-like structures is essential to provide the 3D environment and cell-cell interactions required for correct differentiation. After 24 hours of re-aggregation, cells populations were re-analysed Flk-1 and PDGFRα expression by flow cytometry (Figure 3.2B).

50% of DP cells retained expression of both cell surface markers. 31% of DP cells lost expression of Flk-1, becoming SP-P. 13% retained expression of Flk-1 but not PDGFRα, becoming SP-F and 9% of cells lost expression of both receptors (DN). This confirms previous reports that the DP populations give rise to both single positive populations [125].

61% of day 3.5 DN cells did not gain expression of either PDGFRα or Flk-1 after 24 hours additional differentiation. This suggests that the DN compartment contains cells that do not possess mesoderm potential.39% of DN cells became PDGFRα^{+}, of which 12% also expressed Flk-1 (becoming DP). This is in-line with previous ideas that mesoderm generation is an on-going process throughout embryonic development, and that PDGFRα expression is gained prior to Flk-1 expression [113].

The majority of SP-P cells (82%) retained expression of PDGFRα after 24 hours further differentiation, but did not acquire Flk-1 expression. 64% of SP-F cells retained Flk-1 expression, of which 15% also gained expression of PDGFRα (becoming DP), and 34% of SP-P cells lost expression of Flk-1 and became DN.

Table Four.
Expansion of Flk-1/PDGFRα Populations

Total cell numbers for Flk-1/PDGFRα immuno-phenotypic populations from $4x10^5$ Flk-1/PDGFRα sorted cell, after re-aggravation for 24 hours (See Figure 3.2).
(Values represent the mean number of cells from two independent experiments)

	Immuno-phenotype at Day 3.5 (sorted populations)			
Immuno-phenotype after 24 hours	SP-P	DP	SP-F	DN
SP-P	$35 x10^4$	$15 x10^4$	$1.3 x10^4$	$21 x10^4$
DP	$0.16 x10^4$	$20 x10^4$	$7.7 x10^4$	$0.8 x10^4$
SP-F	$0.36 x10^4$	$8 x10^4$	$17.4 x10^4$	$0.7 x10^4$
DN	$12 x10^4$	$2 x10^4$	$24 x10^4$	$28 x10^4$
Total Cell Number	$4.8 x10^5$	$4.5 x10^5$	$5.1 x10^5$	$5 x10^5$

Overall, this data indicated that both SP-P and SP-F populations arise from a common DP precursor population. Expression of Flk-1 in SP-F populations appears to be transient, which is consistent with Flk-1 being a temporary marker of blood development [20], resulting in DN haematopoietic cells by day 6 of EB differentiation (Figure 3.1B). In contrast, PDGFRα expression appears to be retained by SP-P populations as cells differentiate further. This reflects *in vivo* events, where Flk-1 expression is lost upon commitment to a blood cell fate, but retained by differentiating endothelial cells [20], whereas PDGFRα expression is retained by several mesoderm derived lineages [90]. The small percentage of PDGFRα expressing cells generated from SP-F cell populations (Figure 3.2B) most likely reflects the plasticity of these early precursor cells, which have been shown to share some degree of multi-potent behaviour and, under the appropriate conditions, are able to switch their cell fate program [125].

The expansion of SP-P and SP-F populations during the early stages of EB differentiation (day 3-4) is not necessarily coupled with a loss in DP and/or DN populations, despite a decrease in the percentage of cells expressing a DP or DN phenotype. Alternatively, DP and DN populations may remain unchanged, but contain cells which divide asymmetrically and generating one replacement cell and one differentiated SP-P or SP-F cell. For example, a sub-population of DP cells dividing asymmetrically may produce one DP cell and one SP-F cell. This would result in an expansion of the SP-F population without alternation of DP cell number, although the percentage of total cells with a DP immuno-phenotype is decreased.

3.2 Lineage Potential of Wild Type EB Mesoderm Populations

3.2.1 Cell Fate Potential of Day 3.5 Flk-1/ PDGFRα EB Cells

As a model of lateral mesoderm differentiation, cells were tested for their ability to generate hemangioblast-like blast colonies, containing haematopoietic and endothelial cells [119, 210]. The number of blast colony-forming cells (BL-CFCs / hemangioblast cells) was assessed by re-plating single cells in the presence of VEGF (Figure 3.3). Cardiac potential was measured by the ability of EBs to form spontaneously contracting colonies EBs under serum-free conditions (Figure 3.3). Paraxial mesoderm lineage potential was determined by re-plating cells with cytokines to promote the generation of chondrocytes or osteoblasts [106]. Chondrogenic nodules were identified by the production of specific cell surface proteoglycans, and osteogenic clusters were identified by calcium production (Figure 3.3).

To determine the *in vitro* cell fate potential of day 3.5 Flk-1/PDGFRα mesoderm populations (DP, SP-F, SP-P, DN; Figure 3.4A), sorted EB cells were re-plated under conditions to promote the differentiation of specific mesoderm lineages (Figure 3.3). Figure 3.4B shows a summary of the results.

Flk-1 expressing populations (SP-F and DP) gave rise to significantly more blast colonies (i.e. contained a higher frequency of BL-CFCs) than either SP-P or DN populations (Figure 3.4C). This is most likely due to Flk-1 allowing the cells to respond to VEGF signalling and undergo haematopoietic/endothelial differentiation. Day 3.5 SP-F populations generated relatively few chondrogenic nodules (Figure 3.4E) or calcium containing osteogenic clusters (Figure 3.4F).

Since the SP-F populations arise from DP populations (Figure 3.2), this suggests that DP populations contain mesoderm precursor cells which loose PDGFRα expression as they become more committed towards a lateral mesoderm lineage. In addition to blast colonies, DP populations gave rise to more chondrogenic nodules than SP-F or DN populations (Figure 3.4E). This provides further evidence that DP populations are multi-potent, containing progenitors for lateral mesoderm (SP-F) and paraxial mesoderm (SP-P). Furthermore, DP populations were able to generate significantly more functional cardiac colonies than SP-P or SP-F populations (Figure 3.4D). This not only identifies the presence of an Flk-1$^+$ PDGFRα^+ cardiac precursor within the day 3.5 DP population, but also classifies it as a tri-potent population that is able to give rise to all three mesoderm lineages.

SP-P cells gave rise to relatively few blast colonies (Figure 3.4C) or cardiac colonies (Figure 3.4D). Furthermore, SP-P populations were able to give rise to significantly more chondrogenic nodules (Figure 3.4E) and osteogenic clusters (Figure 3.4F) than any other population. This suggests SP-P populations are enriched for precursors of paraxial mesoderm lineages.

Overall, day 3.5 DN cells gave rise to relatively few mesoderm colonies, particularly haematopoietic and cardiac colonies. Again this reflects the presence of non-mesoderm cell types within the DN compartment. Interestingly, the DN cells were able to generate significantly more osteoblasts and chondrocytes (i.e. paraxial lineages) than blast colonies or cardiac colonies (Figure 3.4). This is most likely due to the ability of some DN cells to express PDGFRα upon further differentiation (Figure 3.2).

It is important to note that all four populations were able to generate all four mesoderm lineages to some degree. This reflects the plasticity and flexibility of early Flk-1/PDGFRα mesoderm populations, providing evidence that these early developmental cells are able to change their cell fate program in response to certain extra-cellular stimuli.

3.2.2 Mesoderm Gene Expression in Day 3.5 EBs

The phenotype of day 3.5 Flk-1/PDGFRα mesoderm populations was defined by gene expression analysis (Figure 3.5). SP-F cells were consistently found to express higher levels of lateral mesoderm associated genes (Figure 3.5A) and lower levels of paraxial mesoderm associated genes (Figure 3.5B) relative to DP and SP-P populations. Reversely, SP-P populations exhibited up-regulation of paraxial mesoderm genes (Figure 3.5B) and down-regulation of genes associated with haematopoietic, endothelial or cardiac development (Figure 3.5A and 3.5C).

DP populations were enriched for expression of some endothelial (Figure 3.5A; *Flk-1*, *Flt4*), haematopoietic (Figure 3.5A; *Lyl-1, Gata2*) and paraxial (Figure 3.5B; *Pdgfrα*, *Fst, Mesp2*) genes, as well as several cardiac genes (Figure 3.5C). In particular, expression of cardiac crescent and primary heart field genes (*Nkx2*.5, *Gata4*, *Mesp1*) [211] were up-regulated in DP populations compared to SP-P populations.

These findings are in agreement with results from *in vitro* cell fate experiments showing that in day 3.5 EBs; DP cells represent a population with potential for lateral, cardiac and paraxial mesoderm differentiation. SP-P cells are enriched for paraxial mesoderm progenitors and SP-F cells are enriched for lateral mesoderm progenitors (Summarised in Figure 3.4B).

Some cardiac genes were also up-regulated in SP-F populations, with similar expression levels as seen in DP populations (*Nkx2*.5, *Gata4*, *Sox18*; Figure 3.5C). This is consistent with *in vitro* cell fate data showing cells in day 3.5 SP-F populations possess increased cardiac potential compared to SP-P populations (Figure 3.4D). Curiously, expression of *Isl1*, a secondary heart field gene, was found to be preferentially expressed in SP-F populations (Figure 3.5C). Recent studies have suggested a role for Isl1 in cardiac endothelial cell development which precedes its role in cardiomyocyte development [97].

Since endothelial gene expression in day 3.5 EBs seems to be restricted to the SP-F lateral mesoderm associated, population (Figure 3.5A) this may account for the up-regulation of Isl1 expression in these cells. Alternatively, given the intertwined nature of mesoderm lineages during development, Isl1 and Flk-1 co-expression could mark an alternative source of cardiac endothelial/vascular cells ($Isl1^+$ $Flk\text{-}1^+$ cell populations give rise to all three cardiac lineages (cardiomyocyte, vascular smooth muscle, cardiac endothelium [94], either separate from, or down-stream of, a $Flk\text{-}1^+$ $PDGFR\alpha^+$ cardiac progenitor population.

3.2.3 Haematopoietic and Cardiac Potential of Day 6 EBs

By day 6 of differentiation, the Flk-1/PDGFRα immuno-phenotype exhibited by wild type EBs is drastically different to that observed at 3.5 (compare Figures 3.4A and 3.6A). This change is typified by the progressive loss of an Flk-1 ^{high} PDGFRα ^{high} DP population into SP-P and SP-F populations (note day 6 DP gated cells are either Flk-1 ^{low} PDGFRα ^{high}, or PDGFRα ^{low} Flk-1 ^{high}), and an increase in the DN population.

To investigate the cell fate potential of day 6 EB Flk-1/PDGFRα populations, expression of CD41 (the earliest marker of haematopoietic lineage commitment [28, 63] was used to determine haematopoietic potential, while *in vitro* assays were used to test for cardiac potential (Figure 3.6). Figure 3.6B shows a summary of the results. Flow cytometric analysis revealed that the majority of CD41^{+} cells were either SP-F or DN (Figure 3.6C). This suggests that early haematopoietic cells develop from the SP-F (i.e. from cells primed for a lateral mesoderm derived fate), with commitment to the blood lineage resulting in loss of expression of mesoderm/endothelial marks, such as Flk-1.

At day 3.5 of differentiation, the DP population is most significantly enriched for cardiac potential (Figure 3.4D). Cardiac potential of day 6 Flk-1/PDGFRα fractionated populations was measured by the ability to generate spontaneously contracting colonies (as in Figure 3.3). Day 6 SP-P populations were highly enriched for cardiac potential, compared to SP-F and DP cells (Figure 3.6D). This suggests that cardiac primed DP cells retain PDGFRα expression and loose Flk-1 expression as they further commit to

a cardiac lineage. This supports *in vivo* date showing Flk-1 is transiently expressed in pre-cardiac populations [20] while PDGFRα expression has been identified in the cardiac crescent [95]. DP populations also gave rise to more cardiac colonies than SP-F populations, presumably due to cardiac primed cells in the PDGFRα ^{high} Flk-1 ^{low} sub-population as differentiating cells progressively loose Flk-1 expression. Note that the cardiac potential of day 6 DN populations could not be tested as single cell sorted populations were unable to re-aggregate into EB-like structures. This is mostly likely due to the lack of cell-cell adhesion proteins expressed in the differentiated cell types within this population (approximately 25% of DN cells are $CD41^{+}$, i.e. blood cell progenitors).

3.3 Pattern of SCL Expression in Differentiating Wild Type EBs

3.3.1 SCL Expression in Differentiating EBs

Expression of the haematopoietic transcription factor SCL has previously been identified in Flk-1 expressing cells during EB differentiation [171] However, it is unknown whether SCL is expressed exclusively in cells already committed to the lateral mesoderm (i.e. SP-F cells) or if it is expressed in a subset of more immature PDGFRα expressing cells (i.e. DP populations).

To identify which mesoderm populations first express SCL and document SCL expression throughout Flk-1/PDGFRα mesoderm differentiation, an intracellular flow cytometry protocol was established (Figure 3.7A). Using this system, a small population of SCL expressing cells

were first observed between day 2.5 and day 3 of EB differentiation (Figure 3.7B). SCL expression increased with EB differentiation, with 36% of day 6 EB cells being SCL^+ (Figure 3.7B). The percentage of cells expressing SCL significantly increased between days 3 and 4 of differentiation (Figure 3.7B), coinciding with a general increase in Flk-1/PDGFRα expressing mesoderm populations (Figure 3.1B).

3.3.2 SCL Expression in Flk-1/PDGFRα Mesoderm Populations

SCL expression in Flk-1/PDGFRα fractionated mesoderm populations was determined through a combination of intra and extra-cellular Flow Cytometry (Figure 3.7A). Representative FACS plots showing co-expression of SCL with Flk-1 and/or PDGFRα are shown in Figure 3.8. The combined results from three independent EB differentiation experiments are shown in Figure 3.9B.

The first SCL expressing cells identified at day 3 of differentiation express Flk-1, but not PDGFRα (Figure 3.8). However, by day 3.25 of EB differentiation, a distinct SCL^+ $Flk\text{-}1^+$ $PDGFR\alpha^+$ (SCL^+ DP) population can be identified, accounting for 3% of total live EB cells. By day 3.5 of differentiation, the percentage of SCL^+ DP cells peaks at around 6% of total EB cells, together with an increase in SCL^+ $Flk\text{-}1^+$ $PDGFR\alpha^-$ (SCL^+ SP-F) populations. By day 4, the SCL^+ DP population is significantly reduced, with the vast majority of SCL^+ cells co-expressing Flk-1 only. This correlates with the loss of PDGFRα expression expected in DP cells which become further

committed to a lateral mesoderm lineage. A SCL^+ Flk-1 negative, PDGFRα negative population (SCL^+ DN) can first be identified in day 5 EBs, and accounts for 75% of all SCL^+ cells by day 6 (27% of total EB cells; Figure 3.8, bottom panel). This correlates with the time point during EB differentiation when cells start to gain CD41 expression (Figure 3.6C) and loose Flk-1 expression (Figure 3.8 bottom panels). Since SCL is a master regulator of blood development [1-4], this might suggest that the SCL^+ DP cells represent an early mesoderm population primed for haematopoietic differentiation, with SCL expression driving cells into lineage commitment.

3.3.3 *Scl* Gene Expression in Differentiating EBs

To further confirm our intracellular flow cytometry data with *Scl* mRNA levels, *Scl* expression in DP, SP-F, SP-P and DN sorted populations were quantified by real-time PCR (Figure 3.9A). The pattern of *Scl* mRNA expression was extremely similar to SCL protein expression (compare Figures 3.9A and 3.9B). Significant levels of *Scl* expression were first detected in day 3 DP populations. Between days 3 and 4 of EB differentiation *Scl* expression is down-regulated in DP populations, correlating with increased expression in SP-F cells. Likewise, between days 4 and 6 *Scl* expression is down-regulated in SP-F cells and increased in the DN population.

3.4 Summary

In the ES/EB system, there is an E-Cadherin negative mesoderm population which can be further fractionated, depending on Flk-1 and PDGFRα expression, into populations which differ in lineage potential. Gene expression analysis and *in vitro* cell fate assays show that co-expression of Flk-1 and PDGFRα (DP populations) marks an early developmental population containing progenitor cells capable of generating cells derived from lateral, paraxial and cardiac mesoderm lineages (Figures 3.4 and 3.5). During the early stages of EB differentiation (days 3-4), this DP population gives rise to SP-P and SP-F populations (Figure 3.2), which possess enhanced cell fate potential for lateral (blood cells and endothelium) and paraxial (cartilage and bone) lineages (Figures 3.4 and 3.5). Cardiac potential is mainly confined to Flk-1 expressing cells, in particular the Flk-1$^+$ PDGFRα^+ DP population (Figures 3.4 and 3.5). Expression of SCL is also confined to Flk-1 expressing cells, both in DP and SP-F populations (Figure 3.8 and 3.9).

As EB cells differentiate, their commitment to a chosen mesoderm lineage is marked by changes in their Flk-1/PDGFRα immuno-phenotype (Figure 3.1). By day 6 of EB differentiation, cardiac progenitors have lost Flk-1 expression, but retain PDGFRα expression (Figure 3.6). Lateral mesoderm cells (SP-F) committing to the blood lineage loose Flk-1 expression and gain expression of haematopoietic markers, such as CD41$^+$ (Figure 3.6). SCL expressing cells follow the same immuno-phenotypic trend, with the progressive loss of a SCL$^+$ DP population coupled with an increase in SCL$^+$ SP-F, and ultimately SCL$^+$ DN, populations (Figure 3.8 and 3.9).

Given the necessity of SCL expression during haematopoiesis, this suggests that SCL primes multi-potent DP cells early on during development for haematopoietic differentiation. The expression of SCL multi-potent DP populations may indicate a role for SCL in cell fate decisions that precede haematopoietic commitment, in particular, lateral mesoderm specification from a multi-potent progenitor. SCL is expressed within multi-potent DP populations between day 3 and 4 of EB differentiation, with 53% of DP SCL^+ cells expressing both Flk-1 and PDGFRα at day 3.25, and 26% at day 3.5 (Figure 3.8). This may indicate a role for SCL in cell fate decisions that precede haematopoietic commitment, in particular, lateral mesoderm specification from a multi-potent progenitor population.

4. Results Chapter Two: SCL is a Master Regulator of Mesoderm Specification

4.1 SCL Expression Affects Cell Fate Decisions in Flk-1 PDGFRα Populations

4.1.1 SCL Expression is Required for the Expansion of SP-F Populations

To investigate whether SCL expression influences the emergence of defined mesoderm populations, Flk-1/PDGFRα expression was examined in differentiating *Scl* $^{-/-}$ ES cells (Figure 4.1, middle panel). E-Cadherin^{+} cells were excluded prior to the analysis of Flk-1 and PDGFRα expression (Figure 3.1). It was noted that the percentage of E-Cadherin expressing *Scl* $^{-/-}$ cells did not significantly differ from that observed in wild type EBs, ranging from 2 to 6% at any given time point (data not shown). Like in wild type EBs, the majority of *Scl* $^{-/-}$ E-Cadherin^{+} cells also expressed PDGFRα (data not shown), thus corresponding to an early mesendoderm population which is unaffected in the absence of SCL.

Between days 2 and 3 of EB differentiation, *Scl* $^{-/-}$ EBs form DP, SP-P and SP-F populations were not significantly different from wild type (Figure 4.1 (compare top and middle panels) and Table 5). This is to be expected as few cells expressing SCL prior to day 3 of EB differentiation (Figures 3.8 and 3.9). However, by day 3.5 of EB differentiation, there was a clear block in the emergence of a *Scl* $^{-/-}$ SP-F population, with the SP-F cells accounting for only 6% of total *Scl* $^{-/-}$ EB cells, compared to 17% in wild type (Figure 4.1, (compare top and middle panels) and Table 5). Note that at this point in wild

Table Five.

Total Cell Numbers for Flk-1/PDGFRα Populations

Cell Numbers for Flk-1/PDGFRα expressing populations in differentiating EBs (See Figure 4.1), from an initial plating of $5x10^4$ ES cells.

Values are the mean from three independent ES cell differentiations.

		Total Cell Number	SP-P	DP	SP-F	DN
Day 2	Wild Type	$9.4 x10^4$	$0.37 x10^4$	$0.16 x10^4$	$0.19 x10^4$	$8.7 x10^4$
	Scl $^{-/-}$	$11.9 x10^4$	$0.64 x10^4$	$0.50 x10^4$	$0.38 x10^4$	$10.3 x10^4$
	Scl $^{RER/RER}$	$10.1 x10^4$	$0.50 x10^4$	$0.07 x10^4$	$0.19 x10^4$	9.36
Day 3	Wild Type	$23.4 x10^4$	$1.4 x10^4$	$0.56 x10^4$	$0.70 x10^4$	$20.0 x10^4$
	Scl $^{-/-}$	$24.5 x10^4$	$1.1 x10^4$	$0.51 x10^4$	$0.79 x10^4$	$22.1 x10^4$
	Scl $^{RER/RER}$	$23.0 x10^4$	$4.1 x10^4$	$2.5 x10^4$	$1.3 x10^4$	$15.1 x10^4$
Day 3.25	Wild Type	$35.7 x10^4$	$4.2 x10^4$	$2.5 x10^4$	$0.84 x10^4$	$28.0 x10^4$
	Scl $^{-/-}$	$26.3 x10^4$	$3.3 x10^4$	$1.9 x10^4$	$0.79 x10^4$	$20.2 x10^4$
	Scl $^{RER/RER}$	$32.2 x10^4$	$4.9 x10^4$	$6.8 x10^4$	$3.9 x10^4$	$16.7 x10^4$
Day 3.5	Wild Type	$45.1 x10^4$	$7.0 x10^4$	$12.3 x10^4$	$9.6 x10^4$	$16.3 x10^4$
	Scl $^{-/-}$	$39.3 x10^4$	$9.0 x10^4$	$8.9 x10^4$	$2.8 x10^4$	$18.5 x10^4$
	Scl $^{RER/RER}$	$40.3 x10^4$	$4.7 x10^4$	$10.5 x10^4$	$12.7 x10^4$	$12.4 x10^4$
Day 4	Wild Type	$55.8 x10^4$	$11.1 x10^4$	$13.3 x10^4$	$13.7 x10^4$	$17.8 x10^4$
	Scl $^{-/-}$	$48.2 x10^4$	$16.2 x10^4$	$8.8 x10^4$	$4.1 x10^4$	$19.1 x10^4$
	Scl $^{RER/RER}$	$59.2 x10^4$	$7.6 x10^4$	$14.9 x10^4$	$15.8 x10^4$	$20.7 x10^4$
Day 5	Wild Type	$90.2 x10^4$	$23.7 x10^4$	$18.0 x10^4$	$12.7 x10^4$	$35.9 x10^4$
	Scl $^{-/-}$	$87.5 x10^4$	$43.0 x10^4$	$25.0 x10^4$	$1.8 x10^4$	$18 x10^4$
	Scl $^{RER/RER}$	$82.5 x10^4$	$30.4 x10^4$	$21.3 x10^4$	$9.8 x10^4$	$21.9 x10^4$
Day 6	Wild Type	$100.5 x10^4$	$23.8 x10^4$	$13.5 x10^4$	$12.7 x10^4$	$50.5 x10^4$
	Scl $^{-/-}$	$90.3 x10^4$	$48.4 x10^4$	$15.1 x10^4$	$3.6 x10^4$	$23.2 x10^4$
	Scl $^{RER/RER}$	$92.8 x10^4$	$25.9 x10^4$	$19.8 x10^4$	$11.1 x10^4$	$36.0 x10^4$

type EB differentiation, SCL^+ DP and SCL^+ SP-F populations are expanding (Figures 3.8 and 3.9). This block in transition from a DP to SP-F phenotype was sustained throughout *Scl* $^{-/-}$ EB differentiation. There was no substantial increase in day 6 *Scl* $^{-/-}$ DN populations (Figure 4.1, middle panel). In fact, day 6 DN populations are greatly reduced in *Scl* $^{-/-}$ EBs (Figure 4.1 and Table 5). In wild type EBs, this increase in DN cells occurred as lateral mesoderm SP-F cells (including SP-F SCL^+ cells) began to lose Flk-1 expression, presumably as they commit to a $CD41^+$ haematopoietic fate. This indicated a block in lateral mesoderm specification in *Scl* $^{-/-}$ EBs.

4.1.2 SCL Specifies Flk-1$^+$ Lateral Mesoderm Lineages without Directly Binding DNA

Although SCL is a bHLH transcription factor which, historically was thought to acts though directly binding DNA, our lab has previously reported that SCL does not need to directly bind DNA to specify haematopoiesis [173, 187]. To investigate whether direct SCL DNA binding is necessary for SP-F lateral mesoderm specification, Flk-1/PDGFRα expression was analysed in EBs expressing a mutated version of SCL which is unable to bind DNA (*Scl* $^{RER/RER}$ ES cells) [187].

From day 3.25 onwards, *Scl* $^{RER/RER}$ EBs exhibited an Flk-1/PDGFRα expression pattern that is more comparable to a wild type than to a *Scl* $^{-/-}$ phenotype, although kinetics of mesoderm differentiation do differ slightly (Figure 4.1). Importantly, *Scl* $^{RER/RER}$ EBs were able to produce a SP-F population which expanded between days 3 and 4 of differentiation and decreased between days 5 and 6, coinciding with an enlarged DN population (Figure 4.1 and Table 5).

Flk-1/PDGFRα expression in *Scl* $^{RER/RER}$ EBs differed from wild type expression between day 3 and 3.25 in PDGFRα expressing populations, predominantly with an expansion of the SP-P population (Figure 4.1 and Table 5). This was coupled with a reduction in DN cells (Figure 4.1 and Table 5). However, since SCL is not expressed in these populations (Figure 3.8), this is most likely due to variance in differentiation kinetics as a result of sub-cloning during cell targeting.

Like in wild type EBs, SCL expression in day 3.5 *Scl* $^{RER/RER}$ EBs was confined to Flk-1^{+} populations (DP and SP-F; Figure 4.2A). The total number of *Scl* $^{RER/RER}$ SCL^{+} cells in day 3.5 EBs did not significantly differ from wild type, although the level of SCL expression within these cells was slightly down-regulated (Figure 4.2A, middle plots).

DN populations in day 6 *Scl* $^{RER/RER}$ EBs were reduced compared to wild type (Figure 4.1 (top and bottom panels; 44% in *Scl* $^{RER/RER}$ compared with 53% in wild type) and Table 5), as is the total number of SCL^{+} cells (30% in *Scl* $^{RER/RER}$ compared to 36% in wild type; Figure 4.2A). The pattern of SCL expression within SP-F and DN populations was unchanged (Figure 4.2A). The reduction in day 6 DN populations may reflect the decrease in SCL^{+} cells, thus a reduction in the number of DP/SP-F cells directed towards a haematopoietic fate. However, the percentage of DN cells expressing CD41^{+} is not decreased in *Scl* $^{RER/RER}$ EBs (Figure 4.2B) suggesting haematopoietic specification has not been disrupted.

4.1.3 Expansion of PDGFRα^+ Populations the Absence of SCL

DP populations formed normally in day 3 *Scl* $^{-/-}$ EBs and expand between days 3 and 5 at a rate similar to wild type (Figure 4.1, compare top and middle panels, and Table 5). Wild type DP populations were lost between day 5-6 of EB differentiation as SP-P and SP-F populations expand (Figure 4.1). In the absence of SCL, an increased percentage of cells retained a DP immuno-phenotype (30% in *Scl* $^{-/-}$ EBs compared to 20% wild type at day 5; Figure 4.1). Note that day 6 *Scl* $^{-/-}$ DP populations could be described as being PDGFRα^+ Flk-1^{low}, indicating DP cells are moving towards a SP-P phenotype.

The percentage of cells expressing PDGFRα expression increased in *Scl* $^{-/-}$ EBs from day 4 of EB differentiation onwards. This mainly relates to an expansion of in the SP-P compartment, which accounts for 46% of *Scl* $^{-/-}$ EB cells by day 6 of differentiation, compared to 22% of wild type EB cells (Figure 4.1). Although analysis of total EB cell numbers showed that *Scl* $^{-/-}$ EB clusters contained consistently less cells than wild type EBs, there is still a clear expansion of SP-P populations and loss of SP-F populations, which is apparent from day 3.5 onwards (Table 5). No noticeable change in levels or cell proliferation or apoptosis between day 3.5 wild type and *Scl* $^{-/-}$ cell populations was detected (Figure 4.3), therefore an expansion in SP-P populations may account for the loss of *Scl* $^{-/-}$ SP-F populations. This indicates that in the absence of SCL, multi-potent DP cells that would normally contribute towards lateral mesoderm lineages differentiate with an alternative, PDGFRα^+ associated, cell fate (i.e. a paraxial or cardiac mesoderm).

Surprisingly, day 6 SP-P populations were also expanded in *Scl* $^{RER/RER}$ EBs. By day 6 of EB differentiation 42% (±10) of *Scl* $^{RER/RER}$ EB cells expressed PDGFRα^+, compared to 34% (±11) in wild type (Figure 4.1). Overall, the total percentage of cells expressing PDGFRα was significantly increased day 6 of *Scl* $^{RER/RER}$ EBs, with a P-value = 0.037 (where n≥12). Like in *Scl* $^{-/-}$ EBs, this occurred at the expense of the DN population, which was reduced to 44% in *Scl* $^{RER/RER}$ EBs compared to 53% in wild type cells (Figure 4.1). Together this data indicates that SCL is involved not only in the promotion of Flk-1$^+$ lateral mesoderm lineage choice (which occurs in a DNA-binding independent manor), but also in the suppression of alternative PDGFRα^+ cell fates, possibly through a DNA-binding dependant mechanism.

4.1.4 Proliferation and Apoptosis are not affected in Flk-1/PDGFRα Populations in *Scl* $^{-/-}$ EBs

Cell cycle analysis and apoptosis assays were performed to determine whether the immuno-phenotypic differences observed between wild type and Scl $^{-/-}$ Flk-1/PDGFRα defined mesoderm populations were due to an increase in cell proliferation or death. Cell cycle and apoptosis flow cytometric analysis were performed by Dr Hedia Chagraoui.

Annexin V staining revealed that despite a block in the formation of day 3.5 SP-F populations in *Scl* $^{-/-}$ EBs (Figure 4.1), levels of apoptosis in *Scl* $^{-/-}$ DP and SP-F populations were comparable to wild type (Figure 4.3A). In addition, cell cycle analysis of day 3.5 wild type and *Scl* $^{-/-}$ EBs (performed using Brdu and 7AAD staining), showed that the increase in *Scl* $^{-/-}$ DP and SP-P populations was not a result of increased cell proliferation (Figure 4.3B).

Collectively, this suggests that the block in DP to SP-P transition exhibited by Scl $^{-/-}$ EBs is not a result of cellular apoptosis or a block in proliferation. Similarly, the expansion of *Scl* $^{-/-}$ SP-P populations is not due to increased proliferation or a block in apoptosis of paraxial or cardiac precursors. Therefore it can be concluded that mesoderm cells in early *Scl* $^{-/-}$ are directed down a different lineage fate pathway in the absence of SCL. This suggests that SCL direct acts to ensure correct cell fate decisions are made during mesoderm specification.

4.2 SCL is Essential for Correct Mesoderm Cell Fate Decisions

Given that SCL expression is important for directing cells towards a SP-F immuno-phenotype, while simultaneously suppressing cell differentiation into a SP-P immuno-phenotype, it is important to determine whether immuno-phenotypic changes correspond to actual changes in cell fate. To uncover which mesoderm lineages are affected by SCL expression, wild type, *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs were re-plated under conditions favouring differentiation into haematopoietic, chondrogenic, osteogenic, myogenic or cardiac lineages (as outlined in Figure 3.3).

4.2.1 SCL Expression is required for Lateral Mesoderm & Haematopoietic Commitment

In vitro assays to test lateral mesoderm potential showed that both wild type and *Scl* $^{RER/RER}$ EB cells were able to generate blast colonies, whereas *Scl* $^{-/-}$ cells were not (Figure 4.4A). The number and morphology of

blast colonies formed did not significantly differ between wild type and *Scl* $^{RER/RER}$ cells. Likewise, a similar number of EBs containing haemoglobin producing cells (Red EBs) formed from wild type and *Scl* $^{RER/RER}$ day 7 EBs. *Scl* $^{-/-}$ EBs did not generate any red EBs (Figure 4.4B). This confirms previous reports that *Scl* $^{-/-}$ cells lack any haematopoietic potential [2-4] while *Scl* $^{RER/RER}$ cells are able to generate primitive blood cells [173, 187].

4.2.2 Paraxial Mesoderm Lineage Potential of *Scl* $^{-/-}$ EBs

To determine whether the expansion of PDGFRα^+ cells in *Scl* $^{-/-}$ EBs (Figure 4.1) corresponds to a change in paraxial mesoderm potential, wild type and *Scl* $^{-/-}$ cells were tested for their ability to generate chondrocytes (Figure 4.5), osteoblasts (Figure 4.6) and myocytes (Figure 4.7)

Both wild type and *Scl* $^{-/-}$ EB cells generated chondrogenic nodules which were comparable in size and morphology (Figure 4.5A). However, significantly more chondrocytes were generated from *Scl* $^{-/-}$ cells compared to wild type. This is shown by an increase in the number of chondrogenic nodules formed (Figure 4.5B and 4.5C) and in expression of some cartilage-associated genes (*Collagen X* and *Sox9*; Figure 4.5D).

Importantly, in day 3.5 EBs, chondrogenic potential was enhanced in *Scl* $^{-/-}$ PDGFRα^+ cells (DP and SP-P; Figure 4.5C), with DP cells exhibiting the most significant increase in chondrogenic potential. This implies that, in absence of SCL, more multi-potent DP mesoderm progenitors are able to differentiate into a chondrogenic cell fate, suggesting a role for SCL in the suppression of some aspects of paraxial mesoderm specification.

Scl $^{-/-}$ did not exhibit a significantly enhanced ability to generate osteoblasts (Figure 4.6). *Scl* $^{-/-}$ EBs were able to generate cell clusters containing calcium deposits, a mark of osteogenic differentiation, which were similar in size, shape and number to wild type clusters (Figure 4.6A and 4.6B). In addition, expression of mature osteoblast markers in osteogenic cultures was not significantly altered in *Scl* $^{-/-}$ clusters (Figure 4.6D). However, upon re-plating of Flk-1/PDGFRα sorted cell populations, *Scl* $^{-/-}$ SP-P cells produced significantly less osteogenic clusters than wild type cells (Figure 4.6C). This is surprising as there is no obvious overall change in osteogenic potential (Figure 4.6B) which suggests SCL expression does not significantly impact osteoblast differentiation.

Although SCL is expressed in mature adult osteoclasts, and therefore may have a role in bone metabolism [170] previous studies have been unable to link SCL expression with bone development pathways expand [169]. Osteoblast development is closely linked to chondrogenic development (which is enhanced in *Scl* $^{-/-}$ EBs), with both cell types differentiating from a common progenitor *in vitro* [106]. Expression of pro-chondrogenic transcription factors inhibits osteoblast commitment [104, 106]. In particular, over-expression of Sox9 has been shown to direct paraxial mesoderm precursor cells into a chondrogenic fate, at the expense of osteogenic differentiation [104]. Interestingly, gene expression analysis comparing wild type and *Scl* $^{-/-}$ EBs (see Figure 4.12) revealed Sox9 to be over-expressed in day 3.5 Flk-1^{+} populations. Therefore the decrease in osteogenic potential in *Scl* $^{-/-}$ SP-P populations may be an indirect effect caused by increased chondrogenic potential in absence of SCL in DP multi-potent populations.

Myocytes generated from *Scl* $^{-/-}$ EBs formed with a similar morphology to those derived from wild type EBs (Figure 4.7A). Quantification of myocyte differentiation was performed by qPCR, comparing muscle gene expression (Figure 4.7B). *MyoD* and *Skeletal Actin* (mature myocyte genes) were expressed at equal levels in wild type and *Scl* $^{-/-}$ myogenic cultures, indicating unaltered levels of myogenesis in the absence of SCL. However, expression of *Myogenin* (the earliest marker of myocyte differentiation [110, 111] was up-regulated in *Scl* $^{-/-}$ cultures (Figure 4.7B). This may indicate that *Scl* $^{-/-}$ cultures did contain increased numbers of immature myocyte progenitors, although this needs to be confirmed using a quantitative functional assay.

Previous work has shown that forced over-expression of SCL in muscle progenitors blocks myocyte differentiation. SCL competes with MyoD (a bHLH protein essential for myocyte maturation) for its E-protein binding partners, resulting in repression of MyoD gene targets and a block in myocyte maturation [202, 203]. However, SCL is not normally expressed in muscle progenitor populations. An increase in Myogenin expression may be an indication of an increase in immature myocytes (i.e. Myogenin^{+} but negative for mature muscle markers such as MyoD and Skeletal Actin) in the absence of SCL. This needs to be confirmed with quantitative functional assays, but may reflect increased commitment to a muscle cell fate by multi-potent DP progenitors.

4.2.3 *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs have Increased Cardiac Potential

PDGFRα^+ populations in day 6, wild type EBs had more cardiac potential than SP-F and DN populations (Figure 3.7) and the SP-P compartment was expanded in day 6 *Scl* $^{-/-}$ EBs (Figure 4.1). Consequently, the cardiac potential of wild type and *Scl* $^{-/-}$ EBs was examined by re-plating EBs in serum-free conditions and quantifying the number of EB colonies able to generate cardiomyocytes (identified by spontaneous contraction; see Figure 3.3).

For analysis of cardiac potential, EBs generated using both the cell suspension method (Section 2.1.3.1) and hanging drop method (Section 2.1.3.2) were used (Hangning drop used in Figures 4.8A and 4.8B only). Spontaneous contracting colonies were first seen in wild type and *Scl* $^{-/-}$ cultures two days post re-plating under cardiac conditions (6 days differentiation from ES cells; Figure 4.8A). The percentage of colonies able to spontaneously contract increased with time, with significantly more contracting colonies observed in *Scl* $^{-/-}$ cultures from 4 day onwards (Figure 4.8A). After 6 days in serum-free conditions, 56% of *Scl* $^{-/-}$ colonies contracted, compared to 24% of wild type (Figure 4.8A and 4.8B). *Scl* $^{RER/RER}$ EBs also exhibited enhanced cardiac potential, with 50% of colonies able to spontaneously contract (Figure 4.8B). Note that size and morphology of cardiac colonies did not significantly differ between cell lines (Figure 4.8C).

Cardiac potential was also assessed by expression of cardiac Troponin T (cTnT), a cardiac specific contractile protein [102], by intracellular FACS analysis (Figure 4.9). Cardiac cultures generated from *Scl* $^{-/-}$ and

Scl $^{RER/RER}$ EBs consistently expressed a greater percentage of cTnT $^{+}$ cells (7% in *Scl* $^{-/-}$ colonies and 6% in *Scl* $^{RER/RER}$ colonies compared to 2% in wild type colonies), indicating an increase in cardiomyocyte formation, and therefore an increase in cardiac potential.

Re-plating day 3.5 Flk-1/PDGFRα sorted cell populations under cardiac inducing conditions revealed that Flk-1^{+} populations (DP and SP-F) in *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs gave rise to more cardiac colonies than Flk-1 negative populations (SP-P and DN) (Figures 4.8D). In addition, *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ day 3.5 DP and SP-F populations gave rise to significantly more cardiac colonies than their wild type counter parts (Figures 4.8D). As wild type EB cells differentiate, and DP populations disappear, cardiac potential became predominantly associated with SP-P populations by day 6 (Figures 3.6 and 4.10C). This was also true for *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ cells (Figure 4.10C), with the shift towards a SP-P cardiac population apparent from day 4.5 of EB differentiation (Figure 4.10B).

It has previously been reported that Flk-1^{+} cells which are capable of cardiac differentiation (contracting cardiomyoctyes) arise through a secondary wave of Flk-1^{+} expression (at EB day 4.25), while the primary wave of Flk-1^{+} cells (EB day 3.25) predominantly gives rise to haematopoietic cells (blast colonies) [19]. This second wave of Flk-1^{+} cells was isolated separately from the first by sorting Flk-1^{negative} cells from day 3.25 EBs and re-culturing them for 24 hours. We report same findings in our ES cell differentiation system, with day 3.25 wild type Flk-1^{+} being enriched for haematopoietic potential (assayed for by blast colony formation) compared to day 4.5 Flk-1^{+} cells (Figure 4.11B).

Cardiac potential (assayed for by contracting colony formation) was enriched in day 4.25 Flk-1^{+} cells, compared to day 3.25 wild type Flk-1^{+} cells (Figure 4.11C).Furthermore, cardiac potential appeared enhanced in *Scl* $^{-/-}$ day 4.25 Flk-1^{+} cells, which are able to generate 2 fold more contracting colonies compare to their wild type counterparts (Figure 4.11C). Neither wild type Flk-1^{+} nor *Scl* $^{-/-}$ Flk-1^{+} populations possessed cardiac potential at day 3.25 (Figure 4.11C), despite Flk-1^{+} populations in day 3.5 EBs being enriched for cardiac progenitors (Figure 4.8D and 4.10A). This again highlights how quickly mesoderm populations change with the first days of EB differentiation.

4.3 *Scl* $^{RER/RER}$ EBs Exhibit Aspects of Wild Type and *Scl* $^{-/-}$ Phenotypes

Scl $^{RER/RER}$ EBs exhibited a differentiation potential that partly reflects a wild type phenotype, and partly reflects a *Scl* $^{-/-}$ phenotype. Like wild type EBs, *Scl* $^{RER/RER}$ EB cells generated SP-F populations at day 3.5, and expanded DN populations at day 6 (Figure 4.1). Additionally, *Scl* $^{RER/RER}$ EB cells were able to differentiate into blast colonies (Figure 4.4A) and haemoglobin-producing cells (Figure 4.4B), whereas *Scl* $^{-/-}$ EBs could not. However, *Scl* $^{RER/RER}$ EBs also had an expanded day 6 SP-P population (Figure 4.1) and generated significantly more cardiac colonies than wild type EBs (Figures 4.8, 4.9 and 4.10), which is reminiscent of the *Scl* $^{-/-}$ phenotype.

To further characterise day 3.5 *Scl* $^{RER/RER}$ EBs, the expression of mesoderm associated genes was compared in day 3.5 Flk-1^{+} populations isolated from wild type, *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs (Figure 4.12).

Gene expression analysis showed lateral mesoderm associated genes to be down-regulated in *Scl*$^{-/-}$ Flk-1^{+} populations compared to wild type (*Lyl-1*, *Gata2*, *and Tie2*). Gene expression was not down-regulated in *Scl* $^{RER/RER}$ EBs, in-line with lateral mesoderm specification occurring in these cells. Several key developmental cardiac genes (*Gata4, FoxH1, MesP1, Hand2*) were up-regulated in both *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ Flk-1^{+} cells, with gene expression in *Scl* $^{RER/RER}$ cells equal to that in *Scl* $^{-/-}$ cells. Interestingly, expression of the chondrogenic development gene *Sox9* was significantly up-regulated in *Scl* $^{-/-}$ cells, but not in *Scl* $^{RER/RER}$ EBs.

Together, these results are in line with an increased potential for cardiac differentiation in both *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs. However, unlike *Scl* $^{-/-}$ EBs, *Scl* $^{RER/RER}$ EBs are able to generate lateral mesoderm. This suggests that increase in cardiac potential in *Scl* $^{-/-}$ EBs is not a default result caused by the lack of lateral mesoderm specification, but points towards a DNA-binding dependent role for SCL in the suppression of cardiac differentiation.

4.4 SCL Over-expression Favours Haematopoiesis over Cardiac Differentiation

4.4.1 SCL Over-expressing ES Cell Lines

To confirm an instructive role for SCL in Flk-1/PDFGRα mesoderm cell fate decisions, wild type and *Scl* $^{-/-}$ ES cell lines were modified using a vector designed to introduce ectopic SCL expression under the control of an Ef1α promoter (Figures 4.13A). SCL expression was introduced into wild type ES cells to create an SCL over-expressing cell line, and into *Scl* $^{-/-}$ ES

cells to determine whether re-introduction of SCL expression was able to rescue the *Scl* $^{-/-}$ cardiac phenotype (*Scl* $^{-/-}$/SCL ES cell line). ES clones were selected by Zeocin resistance, and screened for SCL expression by western blot (Figure 4.13B and 4.13C) as SCL is not endogenously expressed in undifferentiated ES cells. Introduction of the SCL expression construct into *Scl* $^{-/-}$ ES cells was performed by Dr Sarah Hoosdally (Figure 4.13C). The clone presented in Figure 4.13C was selected for analysis as it exhibited wild type levels of haematopoietic rescue in day 7 EBs, determined by flow cytometric analysis of Ter119 Expression (data not shown).

4.4.2 Haematopoietic Potential of SCL Over-expressing ES Cells

The Flk-1/PDGFRα immuno-phenotype of SCL over-expressing day 6 EBs was determined by flow cytometric analysis (Figure 4.14A). All clones were able to generate the four day 6 Flk-1/PDGFRα defined mesoderm population; with four SCL over-expressing ES clones (15, 20, 22 and 24) generating day 6 EBs significantly differing from wild type. These EBs had an expanded day 6 SP-F population (22-32% compared to 16% in wild type) coupled with a reduced SP-P compartment (8-15% compared to 20% in wild type) (Figure 4.14A). This implies a greater percentage of multi-potent DP cells have chosen to commit to a lateral mesoderm fate as opposed to a paraxial/cardiac mesoderm fate. Indeed, EBs from SCL over-expressing ES clones 15, 20 and 22 were able to generate significantly more blast colonies (Figure 4.15A) and haemoglobin-producing cells than wild type EBs (Figure

4.15B). In contrast, EBs generated from SCL over-expressing ES cells clones 4 and 16 exhibited a wild type Flk-1/PDGFRα immuno-phenotype (Figure 4.14A). Clone 4 EBs gave rise to wild type numbers of blast colonies (Figure 4.15A), but did produce more EBs containing red cells (Figure 15.B). Clone 16 EBs appeared to have reduced haematopoietic potential compared to wild type EBs (Figure 4.15).

Expression of SCL in SCL over-expressing cells did not differ between ES clones (Figure 4.13A). However, Western blotting for SCL expression in day 6 EBs revealed differences in the level of SCL expression (Figure 4.14B), with SCL expression in EBs from ES clones 15, 20 and 24 higher than 4 and 16. Targeting of ES cells was not directed; therefore the site of genome integration is unknown and most likely differs between clones. This may account for differences in SCL expression, as sites of genome integration can affect the efficiency of gene expression, especially upon EB differentiation when cells undergo vast transcriptional changes and chromatin re-arrangement. Differences in SCL expression explain the phenotypic differences, with only clones over-expressing SCL at a high enough level to provoke a phenotype.

4.4.3 Cardiac Potential of SCL Over-expressing ES Cells

EBs derived from SCL over-expressing clones 15, 20 and 24 had reduced SP-P populations. In addition, these SP-P populations lack PDGFR$\alpha^{+/high}$ cells present in wild type and *Scl* $^{-/-}$ EBs (Figure 4.14A). All SCL over-expressing EBs formed significantly less contracting colonies compared to wild type EBs (Figure 4.16).

In addition, intracellular FACS analysis did not identify cTnT expressing cells in day 6 SCL over-expressing EBs, or in SCL over-expressing cardiac cultures (data not shown). Not all clones exhibited a haematopoietic phenotype (Figure 4.15), but all showed decreased cardiac potential. This may indicate that the level of SCL over-expression affect the differentiation and blood and cardiac lineages differently. Overall, this indicates that forced SCL expression leads to increased haematopoietic potential coupled with a reduction in cardiac potential, possibly through re-direction of cardiac fated cells into lateral mesoderm lineages.

4.4.4 Phenotypic Rescue in *Scl* $^{-/-}$/SCL EBs

SCL expression was re-introduced into *Scl* $^{-/-}$ ES cells (*Scl* $^{-/-}$/SCL ES cells; Figure 4.13). Flow cytometry and *in vitro* cell fate assays were used to determine the degree of *Scl* $^{-/-}$ phenotypic rescue. Two *Scl* $^{-/-}$/SCL clones were analysed, results in Figures 4.17 and 4.18 show the combined results from both clones. Haematopoietic potential was restored in *Scl* $^{-/-}$/SCL EBs to levels comparable to wild type (Figure 4.17). Differentiating *Scl* $^{-/-}$/SCL ES cells formed wild type numbers of blast colonies (i.e. contained wild type numbers of BL-CFCs; Figure 4.17A) and were able to generate day 6 SP-F and CD41^{+} DN populations (Figure 4.16C). In addition, significantly more day 7 *Scl* $^{-/-}$/SCL EBs contain haemoglobin producing cells than wild type EBs (Figure 4.17B), plus day 6 *Scl* $^{-/-}$/SCL EBs contain more CD41 expressing cells than wild type (Figure 4.17C). This confirms re-introduction of SCL is sufficient to restore a wild type haematopoietic phenotype in *Scl* $^{-/-}$ cells. The increase in red EBs and CD41 expression in *Scl* $^{-/-}$/SCL EBs is

reminiscent of the SCL over-expressing phenotype and likely a result of higher levels of SCL expression in *Scl* $^{-/-}$/SCL cells in comparison to endogenous SCL.

Wild type cardiac potential was also restored in *Scl* $^{-/-}$/SCL EBs. *Scl* $^{-/-}$ EBs gave rise to 2 to 3 fold more contracting cardiac colonies than wild type (40% compared to 15%; Figure 4.18A), but only 5% of colonies formed from *Scl* $^{-/-}$/SCL EBs showed signs of spontaneous contraction (Figure 4.18). Consistent with this, *Scl* $^{-/-}$/SCL day 10 cardiac cultures expressed very low levels of cTnT (<2%; Figure 4.18B). Therefore it can be concluded that restoration of SCL expression (at ectopic levels, not endogenous) not only counteracts the *Scl* $^{-/-}$ phenotype, but reflects an SCL over-expression phenotype.

4.5 Summary

SCL expression is essential for SP-F lateral mesoderm specification from a DP multi-potent precursor population during EB differentiation (as shown in Section 3). In the absence of SCL expression (*Scl* $^{-/-}$) a SP-F lateral mesoderm-primed population does not emerge from early DP populations (Figure 4.1) resulting in the loss of haematopoietic potential (Figures 4.4, 4.11 and 4.12). This is not the result of increased apoptosis or a block in cell proliferation (Figure 4.3). Although it has previously been shown that SCL is essential for haematopoietic differentiation [2-4], its mechanism of action during early development was unknown. Results presented here indicate that SCL is in fact involved in developmental cell fate decisions in the mesoderm, and therefore essential at an earlier developmental time point than previously described.

Day 6 *Scl* $^{-/-}$ PDGFRα^{+} populations (i.e. paraxial and cardiac mesoderm precursors) where expanded compared to wild type (Figure 4.1). Again, this is not due to changes in cell proliferation (Figure 4.3). *In vitro* cell fate assays and gene expression analysis show that this correlates with enhanced potential for cardiac (Figure 4.8, 4.9, 4.10 and 4.12), chondrogenic (Figure 4.5, 4.12) and possibly myogenic, (Figures 4.7) differentiation. Most strikingly, *Scl* $^{-/-}$ EBs were able to generate 2-3 fold more functional cardiomyocytes than wild type EBs (Figure 4.8). This reflects previous work in the zebrafish model where knock-down of *Scl* expression causes a loss in blood formation coinciding with an expansion of the cardiac mesoderm [150].

Ectopic over-expression of SCL in wild type cells caused an increase in the percentage of DP cells committing to an SP-F lateral mesoderm fate (Figure 4.14), and therefore have an increase in haematopoietic potential (Figure 4.15). This expansion comes at the expense of commitment of DP cells to an SP-P cardiac mesoderm fate (Figure 4.14) shown functionally with a reduction in cardiomyocyte formation (Figures 4.18). Re-introduction of SCL expression into *Scl* $^{-/-}$ ES cells (*Scl* $^{-/-}$/SCL) was sufficient to rescue all aspects of the *Scl* $^{-/-}$ phenotype. *Scl* $^{-/-}$/SCL EBs generated SP-F and CD41^{+} DN populations (Figures 4.1 and 4.2) and gave rise to blast colonies and haemoglobin expressing cells (Figure 4.17), thus restoring normal haematopoiesis. In addition, cardiac potential was reduced in *Scl* $^{-/-}$/SCL EBs levels, shown by a decrease cardiomyocyte production and expression of cTnT (Figure 4.18).

Analysis of *Scl* $^{RER/RER}$ EBs confirmed that SCL does not need to directly bind DNA to specify lateral mesoderm commitment or promote the early stages of haematopoietic/endothelial differentiation (Figures 4.1, 4.2, 4.4 and 4.12). However, *Scl* $^{RER/RER}$ EBs also gives rise to an expanded day 6

SP-P population (Figure 4.1) and have the potential to generate more cardiac cells than wild type (Figures 4.8, 4.9, 4.10 and 4.12).

Taken together, these results suggest that SCL plays an instructive role in lateral mesoderm not only through specification of haematopoietic/ endothelial pathways, but also through suppression of cardiac differentiation. Analysis of *Scl* $^{RER/RER}$ EBs indicates suppression of cardiac differentiation by SCL occurs through a separate mechanism to that of haematopoietic specification. Furthermore, while SCL acts independently of direct DNA binding to promote lateral mesoderm associated cell fate, it may act through a DNA-binding dependent mechanism to suppress cardiac differentiation.

5. Results Chapter Three: An Inducible SCL Expression System

During early development, small temporal or spatial changes in the expression of key genes can alter cell fate decisions. As discussed in Sections 3 and 4, SCL is an important regulator of lineage commitment during mesoderm development. Expression of SCL differs within Flk-1/PDGFRα mesoderm compartments over relatively short periods of time, with the phenotype of Flk-1, PDGFRα and/or SCL expressing cells changing significantly within hours. To define the developmental window that critically requires SCL activity for lateral mesoderm specification, and/or cardiac lineage suppression, ES cell lines with inducible *Scl* expression (Tet-Off System; [212, 213]) were created through two stages. Firstly, knock-in of a tetracycline trans-activator (tTA) gene into the ROSA26 locus, and secondly, introduction of a tetracycline responsive promoter driving *Scl* cDNA transcription.

Inducible SCL cell lines were generated on a wild type background, in order to determine the effect of SCL over-expression, and on a *Scl* $^{-/-}$ background, to investigate if small, temporally controlled, doses of SCL are sufficient to rescue the *Scl* $^{-/-}$ haematopoietic and/or cardiac phenotype. Targeting strategy is outlined in Figures 5.1 and 5.5, and in Section 2.8.

5.1 tTA/ROSA26 Knock-In ES Cells

5.1.1 ES Cell Targeting

Tetracycline-controlled SCL expression was introduced into the ROSA26 locus of wild type and *Scl* $^{-/-}$ ES cells. The ROSA26 locus (Figure 5.1A) was targeted as it is expressed ubiquitously during embryonic development, both in adult mouse tissues and differentiating EBs [214-217]. Wild type and *Scl* $^{-/-}$ ES cells were targeted with the pROSA26-TcH vector (Figure 5.1B; supplied by **Dr** Shinji Masui, International Medical Centre of Japan, Tokyo, [213]) to introduce a tetracycline trans-activator (tTA) domain into the mouse genome by homologous recombination.

The pROSA26-TcH construct consisted of two overlapping fragments of the ROSA26 locus (2-3 kb each), thus targeting the ROSA26 allele between exons 1 and 2 and allowing expression of the tTA under the control of the ROSA26 promoter (ROSA26 transcription is initiated from Exon 1; Figure 5.1C). In addition to tTA domain, the construct introduced sequences encoding for an IRES-Venus and an hCMV-1 / hygromycin resistance cassette, flanked by LoxP sites.

The linearised construct was introduced into ES cells by electroporation. Hygromycin resistant ES cell colonies were screened by Southern blot to ensure that correct homologous recombination had occurred (Figure 5.2). The pROSA26-TcH targeting construct was designed to introduce a novel EcoRV digestion site into the ES cell genome (Figure 5.1C). Using a probe located in ROSA26 Exon 1, Southern blotting detected

an 11 Kb wild type fragment (Figure 5.1A), and a 2.3 Kb tTA/ROSA26 knock-in fragment (Figure 5.1C). 11.5% of ES cell clones tested were correctly targeted, containing one wild type allele and one knock-in allele (Figure 5.2).

5.1.2 Characterisation of tTA/ROSA26 ES Cells

The pluripotency and differentiation potential of correctly targeted wild type and *Scl* $^{-/-}$ tTA/ROSA26 knock-in ES cell lines was characterised to ensure the knock-in itself did not yield a change in phenotype.

The morphology and pluripotency of ES cell colonies was examined by alkaline phosphatase staining (Figure 5.3A and 5.4A). Pluripotent ES cell colonies, identified by a dark purple stain and tight colony morphology, were generated by all tTA/ROSA26 ES clones tested (examples indicated with black arrows). However, a high degree of differentiated cells were present in clones WT2, WT5 and WT27 (Figure 5.3A) and *Scl* $^{-/-}$ A12 (Figure 5.4A), identified by weak staining and the presence of fibroblast-like cells growing between tight ES colonies (indicated with red arrows).

The differentiation potential of wild type tTA/ROSA26 ES cell clones was assessed by their ability to generate day 7 EBs containing haemoglobin (red EBs) (Figure 5.3B and 5.3C). All four clones generated day 7 red EBs with wild type morphology (Figure 5.3B) and yield (Figure 5.3C).

Flk-1/PDGFRα flow cytometric analysis of day 4 EBs showed clones WT2 and WT5 exhibited a wild type immuno-phenotype, while clones WT12 and WT27 had enlarged DN populations, and fewer differentiating DP, SP-P or SP-F cells (Figure 5.3D). This could be due to WT12 and WT27 EB cells being less able to differentiate into Flk-1/PDGFRα mesoderm population, or a change in developmental kinetics causing mesoderm populations to emerge at later time points. Either way, these changes in differentiation potential are most likely due to sub-cloning of targeted ES cells, reflecting the heterogeneous nature of ES cell populations rather than as a result of cell targeting and/or tTA expression.

Overall it was concluded that wild type tTA/ROSA26 clone WT5 possessed the closest phenotype to parental wild type cells, and was subsequently used for secondary targeting. WT5 is referred to as WT R26 from this point on.

The differentiation potential of *Scl* $^{-/-}$ tTA/ROSA26 ES cell clones was determined based on the morphology of day 7 EBs (Figure 5.4B) and day 4 Flk-1/PDGFRα immuno-phenotype (Figure 5.4C). All clones generated day 7 EBs with a similar morphology to *Scl* $^{-/-}$ parental EBs (Figure 5.4B). Day 4 EBs from *Scl* $^{-/-}$ clone B26 expressed Flk-1 and PDGFRα at comparable levels to *Scl* $^{-/-}$ EBs (Figure 5.4C). Clones A12 and H8 clones seemed less able to differentiate into mesoderm populations (increase in DN cells) and clone A10 exhibited increase potential for generating a DP population. Again, these changes are mostly likely due to sub-cloning of targeted ES cells.

Scl $^{-/-}$ clone B26 was chosen for secondary targeting and is referred to as Scl $^{-/-}$ R26 from this point on.

5.2 ES Cell Line with Inducible SCL Expression

5.2.1 ES Cell Targeting

WT R26 and *Scl* $^{-/-}$ R26 ES cells were targeted with a vector designed to switch the ROSA26 hygromycin resistance cassette for one containing the *Scl* cDNA under the control of a tetracycline responsive (hCMV-1*) promoter and a PKG / puromycin resistance cassette (*Scl* Exchange Vector; Figure 5.5B). ES cells were co-transfected with the exchange vector and a cre-recombinase expressing vector to enable cleavage at LoxP sites, allowing the exchange and integration of *Scl* cDNA into the ROSA26 locus, 3' of the tTA.

Targeted ES cells which were both puromycin resistant (indicating the presence of the exchange vector within the genome) and hygromycin sensitive (indicating the excision of the hygromycin resistant cassette from the genome) were screened by PCR and Western blot (Figure 5.6). PCR screening confirmed the ES cell genotype, with the presence of the correct R26/i*Scl* genotype yielding a 1.8 Kb product (Primer location shown in Figure 5.5C, PCR products shown in Figure 5.6A). SCL expression was confirmed in correctly genotyped ES cells by Western blot (Figure 5.6B). 50% of correctly genotyped wild type ES cell clones tested expressed SCL (WT/i*Scl* ES cells), while 42% of *Scl* $^{-/-}$ ES cell clones expressed SCL (*Scl* $^{-/-}$/i*Scl* ES cells).

5.2.2 Control of SCL Expression in WT/i*Scl* and WT/i*Scl* $^{-/-}$ ES Cells

Correctly targeted WT/i*Scl* and *Scl* $^{-/-}$/i*Scl* ES cells should have inducible and reversible expression of SCL, depending on the presence or absence of tetracycline (or its more stable analogue, doxycycline). The Tet-Off system relies on the expression of a tetracycline trans-activator (tTA), in this case driven by transcription from ROSA26 Exon 1, which binds to a responder construct (consisting of a tetracycline responsive element (TRE) fused to a minimal hCMV-1* promoter), thus activating transcription of *Scl* cDNA (Figure 5.7A). However, in the presence of tetracycline/ doxycycline, the tTA is sequestered and transgene expression is inhibited (Figure 5.7B).

WT/i*Scl* ES cells 1, 4 and 5, and *Scl* $^{-/-}$/i*Scl* ES cells 17, 18 and 21 all expressed SCL in the absence of doxycycline (Figure 5.6B). In the presence of doxycycline, SCL expression was knocked down in *Scl* $^{-/-}$/i*Scl* 17, 18 and 21 ES cells (Figure 5.8A). SCL expression was restored upon removal of doxycycline from the culture medium; however this took up to 72 hours to occur (Figures 5.8A and 5.8B). In contrast, addition of doxycycline to the culture media knocked down SCL expression within 12 hours (Figure 5.8B). This implies that doxycycline-mediated inhibition of expression is very sensitive; therefore it is possible that the time delay prior to SCL expression after doxycycline removal from the culture medium is due to residual presence of doxycycline in the cells, in addition to the time required for gene transcription and translation.

A low level of SCL expression was detected in doxycycline treated ES cells (Figures 5.8). This phenomenon has previously been reported using a similar Tet-off expression system [217] and most likely due to presence of free tTA, and therefore may be abolished by increased doses of the antibiotic (not tested).

5.2.3 SCL Expression in WT/i*Scl* and *Scl*$^{-/-}$/i*Scl* EBs

To test whether ubiquitous expression of SCL in WT/i*Scl* and *Scl* $^{-/-}$ /i*Scl* ES cells resulted in an SCL over-expression phenotype (see Section 4), ES cells were differentiated, in the absence of doxycycline, under conditions to allow haemoglobin production (i.e. red EBs). WT/i*Scl* EBs generated red EBs in wild type numbers (50-60% red EBs; Figure 5.9A) and did not reflect a SCL over-expressing phenotype (70-80% red EBs; Figure 4.16B).

Scl $^{-/-}$/i*Scl* EBs, differentiated in the absence of doxycycline, did show some degree of haematopoietic potential (parental *Scl* $^{-/-}$ R26 EBs have no haematopoietic potential; Figure 5.9A). However, only 5-15% of EBs contained haemoglobin-expressing cells, compared to 60% in wild type (Figure 5.9A), and therefore were not reminiscent of the *Scl* $^{-/-}$ phenotypic rescue exhibited by *Scl* $^{-/-}$/SCL (70% red EBs; Figure 4.19).

WT/i*Scl* and *Scl* $^{-/-}$/i*Scl* clones expressed SCL at relatively high levels in ES cells (Figures 5.6 and 5.8). Analysis of SCL expression in WT/i*Scl* and *Scl* $^{-/-}$/i*Scl* day 7 EBs, generated in the absence of doxycycline, revealed that WT/i*Scl* EBs from clones 1, 4 and 5 expressed SCL at an endogenous level (Figure 5.9B) explaining their wild type phenotype.

Low levels of SCL expression was observed in *Scl* $^{-/-}$/i*Scl* EBs from clones 17, 18 and 21 (Figure 5.9B), accounting for the small percentage of EBs producing red cells.

To identify at which point in EB differentiation *Scl* expression is shut down, SCL expression was determined over the first 96 hours of *Scl* $^{-/-}$/i*Scl* 17 and *Scl* $^{-/-}$/i*Scl* 18 EB differentiation in the absence of doxycycline (Figure 5.9C). *Scl* $^{-/-}$ clones were used so that endogenous SCL expression was not mistaken for induced SCL expression. Down-regulation of SCL expression occurred between 24 and 36 hours of EB differentiation, and was sustained at a low level from then onwards (Figure 5.9C).

5.3 Summary

Based on the Tet-off system, ES cell lines were created, on wild type and *Scl* $^{-/-}$ backgrounds, with inducible SCL expression. SCL was expressed in ES cells in a doxycycline-responsive manner (Figure 5.8). Inducible ES cells expressed SCL in the absence of doxycycline, whereas gene expression was inhibited upon doxycycline treatment (Figure 5.8). Inhibition of SCL expression by doxycycline was rapid, with knock down occurring within 12 hours of treatment (Figure 5.8B). However, it took up to 72 hours to restore SCL expression following the removal of doxycycline (Figure 5.8A). In addition, in the absence of doxycycline, induced SCL expression was down-regulated in EBs within 48 hours of differentiation (Figure 5.9).

The aim of this system was to allow tight temporal regulation of SCL expression throughout EB differentiation to determine exactly when, and how long for, SCL is required for both the first steps of haematopoiesis and suppression of the cardiac lineage. Specification of blood and cardiac

mesoderm in EBs occurs within a tight temporal window between days 3 and 4 of differentiation, with endogenous SCL expression initiated at day 2.75 to 3 (see Sections 3 and 4). Therefore this inducible system was not adequate on two counts. Firstly, control of SCL expression in response to doxycycline treatment was not responsive enough to allow accurate detection of the temporal window at which SCL expression is required. Secondly, since inducible SCL expression is shut down at day 2 of EB differentiation, the system cannot be utilised to study SCL expression between days 3 and 4. For these reasons, this strategy for investigating the role of SCL in mesoderm specification was subsequently abandoned.

5.4 Discussion: Using the Tet-Off System

The Tet-off system is often used to control gene/protein expression as it allows activation and/or repression of gene transcription in a drug-mediated, reversible manner, depending on the presence or absence of tetracycline (or tetracycline homologue) [212]. The Tet-off system (where gene expression is repressed by tetracycline) is was chosen over the Tet-on system (where gene expression is activated by tetracycline) as background levels of transgene expression are oftern lower and a lower concentration of tetracycline is required for regulation [218, 219].

The Tet-off system was an attractive system to use to determine the temporal window where SCL is required during early development as it would allow short bursts of controlled SCL expression at key time points during EB differentiation. The Tet-off system has been previously reported to successfully control protein expression from the ROSA26 locus in ES cells and differentiating EBs [217, 220, 221]. Using a similar system to the one

described in Section 5, inducible protein expression has been reported in ES cells, in day 4 and 8 EBs, and in differentiated neuronal and endoderm lineages [217, 220, 221] . However, despite the similarities in targeting constructs and techniques used, in our system SCL expression was shut down within 48 hours of EB differentiation (Figure 5.9). SCL expression was rapidly repressed with the addition of doxycycline in ES cells (Figure 5.8), yet it took around 72 hours for SCL expression to be restored upon the removal of doxycycline (Figure 5.8). As ES cells differentiate, significant changes occur within short spaces of time. Consequently, the lag time between doxycycline removal and SCL expression would not allow tight controlled burst of SCL expression to be administered. Although studies have reported reversible, doxycycline-dependent, expression using the Tet-off system in ES cells/EBs, no detailed time course of drug administration versus expression has been published. Therefore it is unknown whether tighter temporal control is possible using this system.

Tetracycline-inducible expression systems relies on the use of the hCMV-1* tetracycline responsive promoter element. Studies comparing promoter activity in differentiating EBs have shown that CMV promoter activity is down-regulated as EB differentiate, but that CMV driven GFP expression could still be detected in day 14 EBs, but at a lower level than expression driven by a PGK or Ef1α promoter [222]. This actually makes using the CMV promoter a more attractive choice for driving SCL expression, as lower expression levels would reflect more endogenous SCL levels, as opposed to the ectopic over-expression of SCL that would be driven by PGK and Ef1α promoters. However, conflicting reports have also shown that the CMV promoter becomes completely inactive around day 4 of EB differentiation [222]. The discrepancy between these findings may be due to

the use of differences in EB culture conditions. Studies where gene/protein expression is sustained used the standard handing-drop method of EB generation followed by conditions to favour neuronal or endoderm lineages [217, 220, 222]. In contrast, studies indicating that the CMV promoter is deactivated in differentiating EBs are directed towards a mesoderm lineage (Figure 5.9) [223].

There has been some evidence that contaminating tetracyclines found in FCS can interfere with inducibility when using Tet-on or Tet-off systems. It is possible that the FCS selected for its mesoderm promoting qualities contains such contaminating tetracyclines, which inhibit gene activation in the absence of exogenous tetracycline treatment. This inhibition would presumably occur at low levels, and therefore only apparent once the CMV promoter activity is already partially down-regulated upon EB differentiation [222]. Using alternative serum-free techniques to differentiate the ES cells, such as the 2i/LIF system [132, 224] may allow use of our inducible system in differentiating EBs.

Alternative inducible systems include the 4-hydroxytamoxifen (4HT) [219] or the FKBP [225] systems. Both systems rely on the stabilisation and degradation of an already expressed protein in response to hormone or chemical treatment, as oppose to the Tet system which requires activation/repression of gene expression. This may allow tighter temporal control of SCL activity.

6. Results Chapter Four: Mechanisms of Action; Investigating SCL Gene Targets

During the earliest stages of EB differentiation, SCL^+ cells are located exclusively within Flk-1^+ populations (DP and SP-F; Figures 3.8 and 3.9). SCL^+ populations expand significantly between day 3 and 4 of EB differentiation, coinciding with the onset of DP population segregating into SP-P and SP-F populations (i.e. mesoderm specification; Figures 3.1 and 3.8). At day 3.5 of differentiation, the overall percentage of Flk-1^+ cells does not differ between wild type and *Scl* $^{-/-}$ EBs (Figure 6.1A). Despite this, the first significant block in SP-F cell emergence occurs in day 3.5 *Scl* $^{-/-}$ EBs (Figure 4.1) suggesting it is a critical time point for mesoderm cell fate decisions. Therefore, to begin to uncover the function and mechanism of action of SCL at the molecular level, high-throughput whole genome techniques were used to identify potential SCL target genes in day 3.5 Flk-1^+ EB populations.

6.1 Differential Gene Expression between Wild Type and *Scl* $^{-/-}$ EBs

6.1.1 Microarray Analysis

mRNA levels in Flk-1^+ day 3.5 wild type and *Scl* $^{-/-}$ EB derived cells were compared by microarray analysis to identify genes deregulated in the absence of SCL during mesoderm specification, thus generating a list of potential mesoderm gene targets (Figure 6.1; Table 6).

Microarray analysis revealed over 500 genes differentially expressed in day 3.5 *Scl* $^{-/-}$ Flk-1^{+} cells compared to their wild type counterparts (with a cut-off set at >1.3 fold difference in expression levels). These included genes involved in transcription, signalling, metabolism, transport and translation pathways, as well as tissue-specific genes (Figure 6.1aB; Table 6). 48% of differentially expressed genes were up-regulated in the absence of SCL, while 52% were down-regulated, indicating that SCL is involved in gene repression as well as gene activation in mesoderm populations.

Remarkably, 42% of the genes differentially regulated in *Scl* $^{-/-}$ cells encoded either a transcription factor or a signalling pathway protein (23% encoded transcription factors, 19% encoded signalling proteins; Figure 6.1aB), highlighting the importance of SCL in regulatory pathways. Gene ontology profiling revealed that genes encoding for DNA-binding proteins and proteins involved in transcriptional regulation were significantly over-represented (Figure 6.1b). The degree of perturbation of gene expression in *Scl* $^{-/-}$ EB cells ranged from 4.35 fold greater than wild type (*Gtl2*) to 50 fold less than wild type (*Grb10*). The majority of fold changes ranged from 3 fold greater to 4 fold less than wild type expression (Table 6), which is not unexpected as small changes in gene expression can equate to important differences in cell fate decisions during early development.

We have previously carried out microarray experiments comparing gene expression in CD71^{+} Ter119^{-} foetal liver pro-erythroblasts from E12.5 wild type and *Scl* $^{RER/RER}$ embryos [187].

Table Six.
Differential Gene Expression Exhibited by Flk-1^{+} *Scl* $^{-/-}$ EBs Compared to Wild Type
(Where wild type expression = 1)

Transcription Factors

Sox9	2.38	**Pml**	0.77	**Nkx3-1**	0.68
Cdx2	2.27	**Fbxo27**	0.77	**Fbxo15**	0.65
Gata3	1.75	**Zswim4**	0.76	**Zic2**	0.64
Sp6	1.75	**Pbx1**	0.75	**Ldb2**	0.63
Ddx5	1.69	**Phf6**	0.76	**Stat3**	0.63
Plac8	1.64	**Runx1**	0.75	**Tal2**	0.62
Dlx3	1.61	**Meis1**	0.73	**Zic3**	0.62
Twist2	1.59	**Rb1**	0.72	**Lhx5**	0.62
Bcl9l	1.54	**Mitf**	0.73	**Klf2**	0.61
Tbx20	1.52	**Smox**	0.73	**Pdzk1**	0.59
Dlx2	1.47	**Fli1**	0.71	**Pem**	0.59
Prrx2	1.47	**Dip3b**	0.72	**Gsc**	0.58
Msx2	1.47	**Ets2**	0.72	**Sp5**	0.57
Isl1	1.41	**Cbfa2t1h**	0.71	**Cbfa2t3h**	0.57
Phox2a	1.45	**Zfpn1a1**	0.71	**Fli1**	0.55
Foxh1	1.39	**Stat5a**	0.69	**Sox7**	0.44
Tcfl5	1.39	**Fbxo36**	0.68	**Gata2**	0.54
Rog	1.39	**Crabp1**	0.68	**Sox18**	0.36
Sertad1	1.37	**Etv4**	0.68	**Bcl11a**	0.51
Lisch7	1.35	**Pitx2**	0.68	**Lyl1**	0.43
Prr7	1.32	**Eomes**	0.67	**Hhex**	0.41
Zfp334	1.32	**Tcf15**	0.67	**Otx2**	0.44
Nfkbib	1.30	**Bcl11b**	0.65	**Zic3**	0.47
Zfp99	0.78	**Tbx6**	0.66	**Dppa4**	0.44
		Stat3	0.66	**Lhx1**	0.36

Signalling

Tgfb2	2.56	**Crlf1**	1.54	**Tnk1**	0.74
Krt1-18	2.17	**Rgs5**	1.49	**Leftb**	0.73
Map17	2.04	**Stard8**	1.49	**Agtrap**	0.71
Axin2	2.00	**Ptp4a2**	1.47	**Rab7l1**	0.71
Gna13	1.89	**Fgf13**	1.41	**Invs**	0.70
Rgs4	1.82	**Dab2ip**	1.39	**Kif3a**	0.70
Dok4	1.79	**Spint1**	1.35	**Wnt8a**	0.69
Ngfr	1.59	**Akt2**	1.33	**Egfl7**	0.68
Pdgfra	1.58	**Sh3md2**	1.33	**Mknk1**	0.67
Pdgfrl	1.47	**Il6st**	1.32	**Tnfaip2**	0.65
Sgk3	1.67	**Rgs16**	1.32	**Gmfg**	0.66
Dusp11	1.43	**Caml**	1.28	**Dusp4**	0.59
Pmvk	1.47	**Mapk9**	0.77	**Socs2**	0.54
Gprc5c	1.43	**Prkdc**	0.77	**Dact1**	0.54
Prkar1a	1.64	**Brip1**	0.76	**Egr1**	0.48
Rgs3	1.64	**Mapk8**	0.76	**Igf1**	0.43
Socs3	1.61	**Ptp4a3**	0.75	**Ptprb**	0.42
Krt2-8	1.56	**Arhgap22**	0.75	**Tek**	0.39
Slitl2	1.52	**Igtp**	0.74	**Cdh5**	0.27
Gpr44	1.56	**Mapkbp1**	0.74	**Ebaf**	0.25
		Gdf10	0.74		

Metabolism

Aars	1.28	**Siat10**	1.72	**Tpt1h**	0.76
Prss8	1.32	**Nos3as**	1.39	**Coasy**	0.75
Nans	1.32	**Pcsk7**	1.39	**Ogt**	0.75
Gus	1.33	**Atp10d**	1.43	**Osbpl2**	0.75
Sc5d	1.33	**Soat1**	1.41	**Alg6**	0.75
Cpt1b	1.45	**Hmgcs1**	1.39	**Gstt1**	0.75
Chst1	1.33	**B3gnt5**	1.39	**Aldh6a1**	0.74
Akr1b3	1.35	**Umps**	1.37	**Scly**	0.73
Ppic	1.35	**Hk2**	1.27	**Ddc**	0.74
Mgll	1.35	**Sgcb**	0.78	**Gstz1**	0.71

Insig1	1.37	**Pter**	0.78	**Mgst3**	0.70
		Galnt1	0.77	**Gst4**	0.61

Membrane Proteins

		Pvrl2	1.35	**Col4a2**	0.70
Map17	2.04	**Sema4b**	1.35	**Cd59a**	0.67
Gjb3	2.00	**Slc39a6**	1.33	**Adam23**	0.64
Tmc6	1.85	**Mbc2**	1.32	**Odz4**	0.65
Ankrd34	1.61	**Zdhhc3**	1.28	**Col23a1**	0.64
Slc39a11	1.54	**Cklfsf7**	1.30	**Itm2a**	0.62
Kctd12	1.41	**Stoml2**	1.33	**Dfy**	0.61
Rora	1.39	**Slc24a6**	1.30	**Gypc**	0.58
Dp1	1.39	**Itm2b**	0.76	**Cd79b**	0.39
Slc38a4	1.37	**Igsf4a**	0.75	**Slc1a3**	0.43
		Slc1a4	0.71	**Slc30a3**	0.33

Erythroid/Vasculogenesis

Unc5b	1.56
Epb4.1l3	0.75
Pink1	0.76
Lancl1	0.71
Alas2	0.68
Hbb-bh1	0.19
Ifit2	0.74
Myd116	0.75

Brain-Specific

Prss19	1.37
Nmb	1.33
Prss8	1.32
Lass2	1.30
Phf6	0.76
Pcp2	0.75
Pacsin3	0.74
Robo1	0.69
Prss12	0.66

Muscle-Specific

Acta1	2.04
Tnnc1	1.76
Tnnt3	1.47
Cpt1b	1.45
Mtmr4	1.32

Cardiac-Related

Tgfb2	2.56
Acta1	2.04
Axin2	2.00
Mmp9	1.76
Tpm1	1.75

Bone-Specific

Thbs3	1.33
Crtap	1.30

Sex-Specific

Tex2	1.33
Slp	0.78
Tex19	0.65

Cardiac-Related

Cxcl12	1.59
Pdgfra	1.58
Msx2	1.47
Nos3as	1.39
Clp1	1.33
Cxcr4	0.76
Wnt8a	0.69

Transport

Nxf7	1.33
Kdelr2	1.37
Kdelr3	1.61
Abcc4	0.75
Bbs2	0.74
Vamp4	0.74

Translation Factors

Eif2s3y	1.92
Arbp	1.45
Elovl1	1.41
Elovl4	1.39
Mrpl35	1.37
Tarbp2	1.32
Elovl6	1.35
Mrps18b	1.30
Mrpl22	1.28
Eif1ay	1.28

Mitochondria

Me2	1.30
Cyp7b1	1.33
Shmt2	1.37
Cyp4b1	0.73
Ucp1	0.63
Mmrn2	0.42

Transcription

Ctdspl	1.35
Sfrs2	1.33
Hist1h4j	0.78
Hist1h2bp	0.74
Stau2	0.70
Hist1h2ae	0.70

Oxidation

Prdx2	1.47
Ppargc1b	0.76
Gpx3	0.70

Cell Cycle

Psmd7	0.71
Psmb8	0.66
Psmb9	0.65
Rad51	1.30
Rad51c	0.75
Ccng1	0.62
Nek8	0.69
Hus1	0.75

Apoptosis

Faf1	1.39
Bcl2l11	1.39
Endog	1.30
Pawr	1.37
Pdcd4	0.62
Amid	0.71

Ubquitination

Smurf2	1.72
Usp29	1.69
Uchl5	1.47
Rnf19	**0.70**
Rnf31	0.70
Ube1l	0.63
Psmb3	0.75

In this study, although some transcription factors and signalling proteins were deregulated in *Scl* $^{RER/RER}$ cells, a large proportion of differentially expressed genes were found to be involved in red cell metabolism. In contrast, many more development-associated genes were deregulated in *Scl* $^{-/-}$ day 3.5 Flk-1^{+} EB cells. This highlights the different roles SCL has depending on when and where it is expressed. Whereas SCL plays a critical role in the maturation and maintenance of erythrocytes during definitive haematopoiesis, it appears to have a more developmental role in early stages of embryogenesis and be involved in mesoderm specification and patterning.

6.1.2 Validation & Analysis of Microarray Data

The differential expression of 30 genes of interest (chosen as they as have previously been assocated with mesoderm developmental pathways) was validated by real time qPCR, using the same material used for the microarray, plus new biological repeats (Figure 6.1C). With the exception of one gene (i.e. in 29/30 genes), values obtained by microarray and qPCR were extremely comparable, confirming that microarray results were reproducible and valid. The exception (*Isl1*) was shown to be up-regulated in *Scl* $^{-/-}$ cells by microarray analysis; yet reproducibly down-regulated in *Scl* $^{-/-}$ cells by qPCR validation (testes using two different sets of PCR primers). This highlights the importance of validating data generated by high-throughput techniques to avoid false positives.

6.1.2.1 Lateral Mesoderm Associated Genes

Unsurprisingly, many lateral mesoderm associated genes were deregulated in the absence of SCL. The differential expression of 10 genes, chosen due to their association with lateral mesoderm development, were validated by qPCR (Haematopoietic; *Eto-2*, *Hhex*, *Fli-1*, *Gata2*, *Lyl-1*, *Ikaros*, *Meis1* and *Pbx1*. Endothelial; *Cdh5*, *Tie2*). All genes were down-regulated in *Scl* $^{-/-}$ Flk-1^{+} cells compared to wild type (Figure 6.1C). Microarray analysis also identified 8 erythroid-specific or vasculogenesis-related genes to be down-regulated in *Scl* $^{-/-}$ Flk-1^{+} cells compared to wild type (Table 6), supporting the concept that SCL is essential for the emergence of lateral mesoderm derived lineages (Figures 4.1 and 4.2).

6.1.2.2 Cardiac Mesoderm Associated Genes

Microarray analysis identified 20 cardiac-related genes to be differentially expressed in the absence of SCL (Table 6). In accordance in an increased cardiac potential in *Scl* $^{-/-}$ EBs (Figure 4.8), some important cardiac transcription factors were confirmed to be up-regulated in *Scl* $^{-/-}$ cells (*Tbx20*, *FoxH1*; Figure 6.1C). Note that other cardiac transcription factors previously found to be up-regulated in day 3.5 *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs did not have representative probes on the microarray (*Mesp1*, *Gata4*, *Hand2*; Figure 4.12).

Some important cardiac transcription factors were not affected by the absence of SCL (*Tbx5*, *Nkx2.5*; Figure 6.1C) and others were down-regulated (*Sox18*, *Pitx2*; Figure 6.1C and Table 6). It is most likely that SCL does not act to suppress all cardiac transcription factors, but selectively represses the expression of only a few key mesoderm genes. It is important

to note that both direct and indirect targets of SCL will have been identified by microarray results and that the down-regulation of some cardiac-related genes may be a result of SCL perturbing other pathways, e.g. endothelial/endocardial lineages.

In addition to cardiac transcription factors, the expression of key signalling molecules involved in pathways known to affect cardiac differentiation were affected in *Scl* $^{-/-}$ cells. For example, TGFβ pathways have been linked to cardiac development in an agonistic manor [100, 137, 153] and *Tgfβ2* expression is significantly up-regulated in *Scl* $^{-/-}$ cells. In addition, the inhibition of canonical Wnt pathway during early development is essential to allow cardio-genesis to occur [226] and that the inhibition of Wnt signalling using antagonists (e.g. Dkk-1 or sFz8) increases cardiac differentiation [100, 153, 226, 227]. Consistent with this, *Wnt8a* (a canonical Wnt pathway ligand) was down-regulated in the absence of SCL. Furthermore, the Wnt/β-catenin inhibitor *Axin2* was up-regulated in *Scl* $^{-/-}$ populations.

6.1.2.3 Paraxial Mesoderm Associated Genes

Relatively few paraxial mesoderm related genes were deregulated in *Scl* $^{-/-}$ Flk-1^{+} populations. *Pdgfrα* gene expression was up-regulated by 1.58 fold *Scl* $^{-/-}$ Flk-1^{+} cells compared to wild type Flk-1^{+} cells (Table 6). Immuno-phenotypic analysis of day 3.5 Flk-1/PDGFRα mesoderm populations shows that this is due to an increase in the percentage of Flk-1 cells expressing PDGFRα in *Scl* $^{-/-}$ EBs (i.e. increased DP and decreased SP-F sub-populations; Figure 4.1) rather than increased PDGFRα expression in individual cells, thus reflecting a block in lateral mesoderm commitment. This

is an important point to consider when comparing gene expression in heterogeneous populations, as it is often unclear whether up-regulation gene expression is due to higher levels of gene expression, or due to an expansion of cells expressing a given gene. In addition, expression of *Pdgfrl* (PDGFR ligand; Table 6) was also up-regulated in *Scl* $^{-/-}$ Flk-1 populations, indicating increased signalling pathway activation in these cells.

Sox9 (an important chondrogenic gene) was the most up-regulated transcription factor identified by microarray analysis (2.38 fold up-regulated compared to wild type expression; Figure 6.1C and Table 6). Additionally, 5 muscle-associated genes and 2 bone-specific genes were up-regulated in *Scl* $^{-/-}$ populations (Table 6), supporting *in vitro* results suggesting that *Scl* $^{-/-}$ EBs have an increased potential for chondrogenic and myogenic differentiation (Figures 4.5 and 4.7).

6.1.2.4 Other Genes of Interest

In addition to its role in haematopoietic development, SCL is important for correct neuronal development [190-192]. In accordance with this, 9 genes associated with brain development were found to be differentially expressed in *Scl* $^{-/-}$ Flk-1^{+} populations. However, since our system has been specifically designed to generate and select for possible mesoderm precursors, it is that SCL affects the expression of many additional ectoderm-related genes that have not been identified here.

Interestingly, the most highly over-expressed and down-regulated genes (*Glt2* and *Grb10* respectively) were both maternally inherited imprinted genes. Both are non-coding transcripts, and function as RNAs to regulate developmental processes [228-230]. *Glt2* has been linked with cell

proliferation in mouse gestation and inhibition of angiogenesis [229, 230] . *Grb10* negatively regulates insulin/IGF-1 signalling and has been linked to foetal growth and glycogen synthesis [228, 231]. The vast deregulation of these genes in *Scl* $^{-/-}$ EBs opens up the possibility that SCL may have a regulatory role through pathways involving non-coding RNAs.

6.1.3 Lateral Mesoderm Gene Expression in Differentiating Wild Type, Scl $^{-/-}$ and *Scl* $^{RER/RER}$ EBs

The expression level of seven lateral, cardiac or paraxial mesoderm associated genes, which were differentially expressed between wild type and *Scl* $^{-/-}$ day 3.5 Flk-1^{+} EB cells, was determined by qPCR throughout day 2 to 6 of wild type, *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EB differentiation (Figures 6.2 to 6.4). Gene expression was also compared in day 3.5 DP, SP-F and SP-P mesoderm fractionated populations.

In depth analysis of *Lyl-1* and *Gata2* gene expression (haematopoietic transcription factors; see Section 1.2.4) confirmed that both genes were significantly down-regulated in *Scl* $^{-/-}$ EB cells from day 3.25 to 3.5 onwards (Figures 6.2A and 6.2B). In contrast, expression of *Lyl-1* and *Gata2* was not decreased in *Scl* $^{RER/RER}$ EBs at these time points. This confirms that specification of lateral mesoderm/haematopoietic populations does not occur in *Scl* $^{-/-}$ EBs, but also that this specification does not require direct SCL-DNA binding. However, *Scl* $^{RER/RER}$ EBs did down-regulate *Gata2* expression by day 5 of EB differentiation, with both *Gata2* and *Lyl-1* expression significantly reduced compared to wild type expression by day 6 (Figure 6.2A and 6.2B).

This suggests that, as development progresses, direct SCL-DNA binding is required to sustain correct haematopoietic development, reflecting the decrease in $CD41^+$ cells produced in *Scl* $^{RER/RER}$ *in vitro* (Figure 4.2) and supporting our previous *in vivo* work [187].

Detailed qPCR analysis of *Tie2* expression (an endothelial gene) showed that *Tie2* was also down-regulated in *Scl* $^{-/-}$ EBs from day 3.25 onwards (Figure 6.2C). Tie2 expression was down-regulated in day 3.5 SP-F populations, but not DP populations, indicating that Tie2 may only be affected by SCL in cells committed to a lateral mesoderm/endothelial cell fate. *Tie2* expression in *Scl* $^{RER/RER}$ EBs did not differ from wild type (Figure 6.2C); indicating further that a direct SCL-DNA interaction is not required for lateral mesoderm specification.

6.1.4 Cardiac Mesoderm Gene Expression in Differentiating Wild Type, Scl $^{-/-}$ and *Scl* $^{RER/RER}$ EBs

Microarray and qPCR analysis revealed that many cardiac-related genes are differentially expressed between wild type and *Scl* $^{-/-}$ Flk-1^+ populations (Figures 4.12, 6.1C and Table 6). Given that *Scl* $^{-/-}$ EBs have increased cardiac potential (Figures 4.8 to 4.10) detailed gene analysis was performed to determine the expression pattern of 3 important cardiac genes (*FoxH1*, *Gata4* and *MesP1,* all of which are up-regulated in *Scl* $^{-/-}$ EBs) in wild type, *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs (Figure 6.3). FOXH1, GATA4 and MESP1 are three transcription factors that play important regulatory roles in controlling cardiac gene expression in the early stages of cardiac development (see Section 1.3.4).

All three genes (*FoxH1*, *Gata4* and *MesP1)* showed some degree of over-expression in *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs, compared to wild type expression levels. *FoxH1* (Figure 6.3A) and *Gata4* (Figure 6.3B) were up-regulated in *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs from day 3.5 of differentiation onwards. In both cases, gene expression was increased in DP populations (i.e. cardiac precursor populations; Figures 3.4 and 4.8). Over-expression of *FoxH1* and *Gata4* was sustained throughout EB differentiation (Figures 6.3A and 6.3B). In contrast, *MesP1* expression was only up-regulated in *Scl* $^{-/-}$ and day 6 *Scl* $^{RER/RER}$ EBs after 5 days of differentiation (Figure 6.3C). However, in an independent experiment we have shown that *MesP1* is over-expressed in day 3.5 *Scl* $^{-/-}$ Flk-1^{+} populations (Figure 4.12). These discrepancies may be due subtle differences in *MesP1* expression in individual Flk-1/PDGFRα populations, which may be masked when looking at whole ES populations.

Overall, qPCR results provides molecular evidence of the increase in cardiac potential detected in *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs from day 3.5 onwards.

6.1.5 *Sox9* Expression in Differentiating Wild Type, *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs

The most up-regulated transcription factor in *Scl* $^{-/-}$ cells compared to wild type cells, as identified by microarray analysis, was *Sox9* (Figure 6.1C and Table 6). *Sox9* specifies cartilage development, and is expressed in early multi-potent mesenchymal precursors that generate osteoblasts and chondrocytes, as well as differentiating chondrocytes and mature cartilage tissue [232]. qPCR analysis confirmed *Sox9* over-expression in day 3.5 *Scl* $^{-/-}$ EBs (Figure 6.1C and 6.4), showing up-regulation in DP populations, as well

as in SP-P populations. This may indicate that an increased percentage of DP cells have committed to a paraxial mesoderm primed SP-P fate. Alternatively, high levels of expression of *Sox9* are being sustained in normal *Sox9* expressing mesenchymal precursor cells which usually down-regulate *Sox9* as they differentiate down a non-chondrogenic lineage. Increased *Sox9* expression is sustained throughout *Scl* $^{-/-}$ EB differentiation (Figure 6.4); presumably accounting for the increase in chondrogenic potential *Scl* $^{-/-}$ EBs exhibit *in vitro* (Figure 4.5).

Curiously, *Sox9* expression is down-regulated in *Scl* $^{RER/RER}$ EBs at day 4 of differentiation (Figure 6.4). This raises an interesting question; if in *Scl* $^{-/-}$ EB the increase in cardiac progenitors arises at the expense of haematopoiesis, what populations are lost in *Scl* $^{RER/RER}$ EBs where cardiac differentiation is enhanced but initial haematopoiesis is relatively normal? It may be that in the *Scl* $^{RER/RER}$ situation, commitment to other DP mesoderm derived lineages (i.e. paraxial mesoderm lineages) is decreased as cells choose alternative fates.

6.2 GATA4 in Cardiac and Haematopoietic Differentiation

The zinc-finger transcription factor GATA4 was found to be a significantly over-expressed cardiac transcription factor in day 3.5 *Scl* $^{-/-}$ EBs (Figure 6.3B). During early development, GATA4 is expressed in the cardiac mesoderm, cardiac crescent and developing primary heart field. GATA4 is known to interact with NKX2.5 and TBX5 in the cardiac crescent, thus regulating the primary heart field, and directly activates cardiac-associated target genes such as *Mef2C* and *cTnT* [99, 100, 233, 234]. Over-expression of GATA4 has been shown to enhance cardiogenesis both *in vivo* (using the

Xenous model) [235] and in cell lines [236, 237]. Furthermore, inducible Gata4 expression in ES/EB systems have shown that forced Gata4 expression during early mesoderm specification (Day 2-3 EBs) is sufficient to induce enhanced cardiac development [236, 238].

The importance of *Gata4* expression in the initiation of cardiac differentiation is further highlighted as it is one of the three key genes, along with *Mef2C* and *Tbx5,* which needs to be activated to re-programme fibroblast cells into functional cardiomyocytes [147]. In addition, GATA4 has been reported as being a pioneer factor, acting to open the chromatin structure in the cardiac loci, allowing the binding of other cardiac transcription regulators to bind and activate there target genes and leading to full activation of the cardiac programme [201].

GATA4 is also essential in the developing endoderm and regulates growth of endoderm-derived organs, including gut, liver and pancreas in zebrafish [239]. GATA4 is first expressed in the extra-embryonic visceral endoderm (VE), and *Gata4* $^{-/-}$ embryos are embryonic lethal around day 8.5 due to defective VE formation [133, 134]. In addition to its function in mesoderm and endoderm lineages, Gata4 expression has been linked to the emergence of both lineages from a common mesendoderm precursor in Xenopus [240].

Very few endoderm-primed cells are produced from EBs using our differentiation conditions and any endoderm or mesendoderm cells are excluded prior to analysis (see Section 3.1.1). Furthermore, SCL is not expressed in endoderm or mesendoderm populations (see Section 3.3.1). Therefore changes in GATA4 expression as a result of perturbed SCL expression are most likely associated with a cardiac lineage context.

For these reasons, GATA4 presents itself as a possible direct target gene of SCL. Although knock-down of GATA4 does not inhibit cardiac development, due to redundancy by GATA5 and GATA6 [241-243] over-expression of GATA4 does drive cardiac differentiation in cells that would normally chose a different cell fate [235-237]. Our hypothesis is that SCL may act to suppress GATA4 expression in cells destined for a lateral mesoderm fate. In the absence of SCL, uncommitted mesoderm precursors express *Gata4* abnormally and undergo cardiac differentiation.

6.2.1 GATA4 Expression in Wild Type, *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs

GATA4 presents itself as an appropriate marker of early cardiac progenitors in our system. Therefore, to further investigate links between blood and cardiac differentiation in the ES/EB system, and a possible link between GATA4 and SCL, GATA4 expression was analysed in differentiating wild type, *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs using an intracellular flow cytometry protocol.

Scl $^{-/-}$ and *Scl* $^{RER/RER}$ EBs contain a higher percentage of GATA4 expressing cells from day 3 of differentiation, with 10% of *Scl* $^{-/-}$ cells and 11% of *Scl* $^{RER/RER}$ cells being GATA4^{+} compared to 6% of wild type cells (Figure 6.5). By day 4 of differentiation, GATA4 expression increased dramatically in *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs, with 23% of *Scl* $^{-/-}$ cells and 22% of *Scl* $^{RER/RER}$ cells being GATA4^{+} compared to only 12% of wild type cells (Figure 6.5). Over-expression of GATA4 in *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs was sustained throughout days 5 and 6 of EB differentiation, with 34% of day 6

Scl $^{-/-}$ cells and 44% of *Scl* $^{RER/RER}$ cells expressing GATA4 in comparison to 19% in wild type (Figure 6.5 and Table 7). The rapid emergence of GATA4^{+} cells in *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs between days 3 and 4 of differentiation reflects their increased *in vitro* cardiac potential (Figures 4.8 to 4.10), suggesting that this is a critical time point for the specification of cardiac fated cells.

Co-expression of GATA4 with Flk-1 and/or PDGFRα was determined by flow cytometric analysis, to identify the mesoderm populations presenting an increase in GATA4^{+} ells (Figure 6.6). In all three cell lines (wild type, *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$) GATA4 was found to be first expressed in the DP multi-potent mesoderm population in day 3 and 3.5 EBs (Figure 6.6). At these time points, DP populations were enriched for cardiac precursors (Figures 3.4 and 4.8) and SCL^{+} cells with lateral mesoderm potential (Figures 3.4, 3.8 and 3.9). *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs contained significantly more GATA4^{+} DP cells than wild type EBs (Figure 6.6), consistent with their enhanced cardiac potential (Figures 4.8 to 4.10).

As differentiation continues, GATA^{+} cells became associated with the SP-P populations in all three cell lines (Day 4, 5 and 6 EBs; Figure 6.6). Day 6 *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs had an enlarged SP-P population compared to wild type (Figure 4.1). Indeed, it is the GATA4^{+} SP-P populations that was expanded in day 6 *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs, accounting for 20% of *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EB cells compared to 8% of wild type EB cells at day 6 (Figure 6.6 and Table 7). In addition, GATA4^{+} SP populations were significantly enlarged in *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs from day 4 onwards (Figure 6.6 and Table 7). Again this reflects the enhanced cardiac potential of these cells as cardiac precursors were enriched in SP-P populations in day 6 EBS (Figure 3.6).

Table Seven.

Percentage of Total Cells Co-expressing GATA4 with Flk-1 and/or PDGFRα

Values indicate the mean percentage of total live cells to co-express GATA4 with a Flk-1/PDGFRα immuno-phenotype, ± standard deviation, where n ≥ 3. The most significant populations are highlighted in blue (see Figure 6.6).

		Wild Type	*Scl* $^{-/-}$	*Scl* $^{RER/RER}$
Day 3	SP-P	1 ± 0.4	1 ± 0.4	2 ± 0.4
	DP	**4 ± 3**	**5 ± 4**	**7 ± 3**
	SP-F	1 ± 0.8	2 ± 2	1 ± 0.1
	DN	0.6 ± 0.2	1 ±0.7	1 ± 0.3
Day 3.5	SP-P	2 ± 2	3 ± 2	5 ± 3
	DP	**6 ± 3**	**15 ± 9**	**12 ± 3**
	SP-F	1 ± 1	3 ± 1	1 ± 0.2
	DN	2 ± 2	3 ± 1	4 ± 1
Day 4	SP-P	**3 ± 2**	**6 ± 1**	**11 ± 5**
	DP	**8 ± 1**	**11 ± 2**	**12 ± 1**
	SP-F	2 ± 0.5	1.5 ± 1	3 ± 2
	DN	1 ± 0.6	4 ± 1	4 ± 0.1
Day 5	SP-P	**9 ± 4**	**15 ± 2**	**18 ± 11**
	DP	**9 ± 5**	**16 ± 5**	**15 ± 2**
	SP-F	2 ± 0.5	3 ± 3	3 ± 1
	DN	4 ± 3	2 ± 0.3	4 ± 1
Day 6	SP-P	**8 ± 3**	**20 ± 9**	**18 ± 13**
	DP	**6 ± 5**	**8 ± 1**	**14 ± 4**
	SP-F	2 ± 1.5	1 ± 0.5	4 ± 2
	DN	5 ± 2.5	5 ± 4	5 ± 3

Increased GATA4 expression has been shown to enhance cardiac mesoderm specification [235-237]. Therefore, it is possible that increased levels of GATA4 expression in *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ DP mesoderm populations is the initiating molecular event that triggers aberrant differentiation towards a cardiac fate, resulting in an enlarged GATA4^{+} SP-P populations and increased cardiomyocyte differentiation.

6.2.2 GATA4 Expression in SCL Over-expressing EBs

EBs derived from ES cell lines that over-express SCL show a reduced ability to generate functional cardiomyocytes *in vitro* (Figure 4.16). Intracellular FACS was used to determine if this lack of cardiac potential was accompanied by a decrease in GATA4 expression.

GATA4 was expressed in 19% of day 6 wild type EBs (Figure 6.5), mainly in the SP-P population (Figure 6.6 and Table 7). GATA4 was over-expressed in day 6 *Scl* $^{-/-}$ EBs (Figure 6.5), most significantly in SP-P and DP populations (Figure 6.6 and Table 7). In contrast, day 6 SCL over-expressing EBs exhibited increased lateral mesoderm commitment (enlarged SP-F and DN populations) at the expense of the SP-P populations (Figure 6.7; top panels). This correlates with increased haematopoiesis (Figure 4.15) and decreased cardiac differentiation (Figure 4.16).

After day 6 of differentiation, SCL over-expressing EBs contain significantly less GATA4^{+} cells than wild type EBs; 7-9% in SCL over-expressing EBs compared to 19% in wild type EBs (Figure 6.7; middle panels), reflecting the decrease in cardiac potential. Interestingly, GATA4^{+}

cells in day 6 SCL over-expressing EBs are mainly Flk-1^{+} (Figure 6.7; bottom panels), whereas GATA4^{+} cells in wild type, *Scl* $^{-/-}$ and Scl $^{RER/RER}$ EBs are mainly PDGFRα^{+} (Figure 6.6).

Day 6 Scl$^{-/-}$/SCL EBs (where SCL is ectopically expressed on a Scl $^{-/-}$ background) expressed GATA4 in 9% of cells (compared to 19% in wild type and 34% in Scl $^{-/-}$ EBs; Figure 6.7) indicating that re-introduction of SCL prevents GATA4 over-expression. In addition, Scl $^{-/-}$/SCL EBs populations contain cells co-expressed GATA4 and Flk-1 (Figure 6.7). This not only suggests that SCL expression leads to down-regulation of GATA4, but that ectopic expression of SCL in remaining GATA4^{+} cardiac precursors is enough to re-direct cells into a SP-F/DN lateral mesoderm fate.

6.2.3 Creation of a GATA4 Over-expressing ES Cell Lines

GATA4 is an important transcription factor in cardiac development [100]. The absence of *Gata4* does not equate to a loss of cardiac development due to compensatory mechanism involving GATA5 and GATA6 [241, 242]. However, forced ectopic GATA4 expression has been shown to increase cardiomyocyte formation [235, 237].

To determine whether the over-expression of GATA4 documented in *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs was driving the increased in cardiac differentiation, wild type ES cells were electroporated with a vector designed to drive ectopic GATA4 expression (Figure 6.8A). Neomycin resistant ES clones were pooled into two populations, and genotyped by PCR to confirm integration of the targeting construct into the genome (Figure 6.8B). GATA4 protein expression was confirmed in these ES cell populations by Western blot (Figure 6.8C).

6.2.4 Cardiac Potential of GATA4 Over-expressing and *Gata4* $^{-/-}$ EBs

In vitro cell fate analysis for cardiac differentiation (see Figure 3.3) revealed that GATA4 over-expressing EBs were able to give rise to twice as many spontaneously contracting cardiac colonies compared to wild type EBs (40% in GATA4 over-expressing compared to 19% in wild type EBs; Figure 6.9A). This level of cardiac potential is similar to that exhibited by *Scl* $^{-/-}$ EBs (56% contracting colonies; Figure 4.8).

Analysis of cardiac potential in day 3.5 Flk-1/PDGFRα fractionated mesoderm populations showed that cardiac potential was increased in all four Flk-1/PDGFRα populations in GATA4 over-expressing EBs. This is presumably due to the ectopic nature of GATA4 expression from an Ef1α promoter, re-directing all types of mesoderm precursors to a cardiac path. However, significantly more cardiac colonies were still generated from DP and SP-F day 3.5 populations that SP-P and DN populations (Figure 6.9B), reflecting the increased cardiac potential of day 3.5 Flk-1^{+} populations seen in wild type EBs (Figure 3.4). *Gata4* $^{-/-}$ EBs generated wild type numbers of contracting colonies, indicating no loss in cardiac potential (Figures 6.9A and 6.9B), which is in line with previous studies showing that knock-down of GATA4 alone does not result in a significant cardiac phenotype due to redundancy by GATA5 and GATA6 [241-243].

6.2.5 Haematopoiesis in GATA4 Over-expressing and *Gata4* $^{-/-}$ EBs

To further explore links between cardiac and blood development, the haematopoietic potential of GATA4 over-expressing and *Gata4* $^{-/-}$ EBs was investigated using *in vitro* cell fate assays (see Figure 3.3). Significantly less blast colonies can be generated from day 3.5 GATA4 over-expressing EBs compared to wild type EBs (Figure 6.10A), due to decreased haematopoietic potential in Flk-1^{+} populations (DP and SP-F; Figure 6.10C). Despite low numbers, GATA4 over-expressing blast colonies have a normal size and morphology (Figure 6.10C). This implies that a smaller percentage of cells are committing to a lateral mesoderm lineage in GATA4 over-expressing EBs, but once cells have committed to a lateral mesoderm fate, differentiation into haematopoietic and endothelial cells continues normally. Furthermore, only 18% of GATA4 over-expressing EBs contained visible red cells compared to 63% in wild type; Figure 6.10B).

Although haemoglobin production in day 7 *Gata4* $^{-/-}$ EBs did not noticeably differ from wild type (Figure 6.10B), significantly more blast colonies were generated from day 3.5 *Gata4* $^{-/-}$ EB cells than wild type cells (Figure 6.10A). This is due to an increase in blast colonies forming from day 3.5 *Gata4* $^{-/-}$ DP populations (Figure 6.10C), indicating that this population contains more cells with BL-CFC potential than wild type DP populations. Altogether, this provides evidence for a function of GATA4 in repression of haematopoiesis and the existence of cross-antagonism between blood and cardiac differentiation during mouse development.

6.2.6 SCL Expression in GATA4 Over-expressing and *Gata4* $^{-/-}$ EBs

Since haematopoiesis is perturbed by changes in GATA4 expression, flow cytometric analysis was used to determine whether SCL expression is altered in GATA4 over-expressing and *Gata4* $^{-/-}$ EBs. SCL expression was found to be decreased in day 3.5 GATA4 over-expressing EBs; 5% in GATA4 over-expressing EBs compared to 22% in wild type EBs (Figure 6.11, top section).

SCL expression was decreased in both the DP and SP-F populations in day 3.5 GATA4 over-expressing EBs, correlating with the decreased blast colony forming potential exhibited by these cells (Figure 6.10C). Reduced expression of SCL in GATA4 over-expressing EBs was sustained throughout EB differentiation, with 16% of day 6 GATA4 over-expressing EBs expressing SCL compared to 36% in wild type (Figure 6.11, bottom section). SCL expression was decreased in SP-F and DN populations, accounting for defected haematopoiesis and haemoglobin production in day 7 GATA4 over-expressing EBs (Figure 6.10B).

Overall SCL expression was not altered in the absence of GATA4, with day 3.5 and day 6 *Gata4* $^{-/-}$ EBs expressing SCL at wild type levels (Figure 6.11). All SCL^+ cells in day 3.5 wild type ad *Gata4* $^{-/-}$ EBs co-expressed Flk-1 (SCL^+ SP-F). However, *Gata4* $^{-/-}$ EBs lacked a distinct SCL^+ DP population (Figure 6.11, top section). Instead, the majority of *Gata4* $^{-/-}$ SCL^+ cells were SP-F. Since DP cells become SP-F as they commit to a lateral mesoderm lineage (as shown in Section 3), this shift in SCL^+ cells from a DP to SP-F phenotype in *Gata4* $^{-/-}$ EB may indicate an acceleration in developmental kinetics in the absence of GATA4. This may account for the

increased number of blast colonies produced by day 3.5 *Gata4* $^{-/-}$ EB cells (Figure 6.10). By day 6 of differentiation, SCL expression is unaltered between wild type and *Gata4* $^{-/-}$ EBs.

Together these results suggest that ectopic expression of GATA4 results in increased cardiac differentiation at the expense of lateral mesoderm / haematopoietic commitment. The absence of GATA4 expression does not significantly alter either lineage, which is most likely due to compensatory mechanisms involving GATA5 and GATA6 [241-243]. Furthermore, SCL and GATA4 appear to be key indicators of early lineage commitment and may both act to specify and repress haematopoietic and cardiac fate decisions in an opposing manner (Figure 7.1).

6.3 Investigating Direct Gene Targets of SCL

Classically, transcription factors bind DNA in a sequence specific manner and recruit cofactors to form a complex that can alter chromatin structure and activate or repress target gene transcription. Although direct SCL-DNA binding is not necessary for the function of SCL in lateral mesoderm specification and the initial stages of haematopoietic development (Figures 4.1, 4.2 and 4.4) [173, 187], cardiac differentiation is enhanced in *Scl* $^{RER/RER}$ cells where SCL-DNA binding is inhibited (Figure 4.8 to 4.10), suggesting that SCL acts to repress cardiac development through a direct DNA interaction.

Previous reports have detected direct targets of SCL in erythroid and megakaryocyte cell lines and progenitor populations [67, 165, 166, 193, 194] and in T-ALL leukemic cell [197]. However, not surprisingly, SCL has been shown to regulate different target genes, and work through different

mechanisms, depending on the cellular context. This is highlighted by the differences in gene ontology of genes deregulated in *Scl* $^{-/-}$ Flk-1 EBs (microarray data; Figure 6.1 and Table 6) compared to our previous data from the analysis of erythroid progenitors [164] (discussed further in Section 7.7). Therefore, it is likely that SCL binds different genomic targets during mesoderm development and in mature blood cells.

To gain further mechanistic insight into how SCL controls gene expression in mesoderm patterning, chromatin immuno-precipitation (ChIP) experiments were setup to identify direct genomic targets of SCL. ChIP was, performed on Flk-1^{+} cells from day 3.5 wild type EBs, using an anti-SCL antibody [173], followed by either real-time qPCR or high-throughput Solexa (Illumina) sequencing.

6.3.1 SCL does not directly bind the GATA4 Proximal Promoter in Day 3.5 EBs

In the absence of SCL, more cells differentiated into a cardiomyocyte lineage (Figures 4.8 to 4.10). Several cardiac-related genes were found to be up-regulated in day 3.5 *Scl* $^{-/-}$ cells (Figure 4.12, 6.1, 6.3, 6.6 and Table 6). The most profound of which was *Gata4*, which was up-regulated from day 3 of EB differentiation onwards (Figures 4.12, 6.5 and 6.6). In addition, GATA4 expression was down-regulated upon over-expression of SCL (Figure 6.7), while over-expression of GATA4 recapitulated the *Scl* $^{-/-}$ cardiac phenotype (Figure 6.9). In addition, *Gata4* was also over-expressed in *Scl* $^{RER/RER}$ EBs (Figures 4.12, 6.3 and 6.6); therefore, we hypothesised that SCL may have a direct role in repressing *Gata4* expression in early EB cells.

To investigate this further, ChIP experiments were performed using conditions previously described to ChIP SCL-bound DNA fragments in primary erythroid cells [164] and analysed by real-time qPCR, using a set of primers designed around the *Gata4* transcriptional start site (TSS) (covering from 6kb upstream to 4 Kb downstream). SCL was not found bound to this region in day 3.5 EBs (Figure 6.12). Primers were also designed around know Gata4 enhancer sites [244, 245] SCL was not found bound at any of these genomic regions either (data not shown).

In addition to its role as DNA-binding transcription regulator, SCL has been shown to act as a repressor in a DNA-binding independent fashion by sequestering protein and inhibiting the formation of transcription regulating complexes [105, 202]. It was therefore attractive to hypothesise that SCL might sequester GATA4 in an “off-DNA” protein complex in order to inhibit its normal transcriptional functions. Co-IP techniques were used to determine if SCL and GATA4 exist within the same protein complex in day 3.5 EBs. SCL and GATA4 proteins could not be pulled down together, suggesting that they do not directly interact during early development (data not shown). Collectively, this strongly suggests that SCL does not inhibit cardiomyocyte differentiation through direct repression of GATA4.

6.3.2 Genome-Wide Mapping of Regions Bound by SCL

To determine the regions of DNA bound by SCL during mesoderm development, genome-wide mapping was performed using material from anti-SCL ChIP experiments (using day 3.5 Flk-1^{+} wild type EB cells) followed by high-throughput Solexa (Illumina) sequencing. Using Maq (Mapping and Assembly with Quality Programme), reads were mapped to the mouse genome (build m37) and uniquely mapped reads were counted for every nucleotide position. Using MIG (an in-house developed programme), peaks were determined and ranked according to peak intensity, number of reads, and greater than 15 fold enrichment over control-sample (no antibody ChIP). Using these parameters, 247 peaks were identified. The enrichment in number of reads in anti-SCL ChIP samples versus no-antibody control was not as profound as we have previously reported in red cells [164]; the most intense peak (*Rybp*) showing a fold enrichment of 45.61, and the overall mean fold enrichment being 17.98.

Gene ontology profiling revealed that SCL mainly binds near genes encoding transcription factors and proteins involved in transcriptional regulation (Figures 6.13a and 6.13b), providing evidence that SCL is an important regulator of the transcriptional mechanisms driving mesoderm specification. In day 3.5 EBs, SCL binds to haematopoietic regulators (e.g. *Lyl-1*, *Eto-2*, *Gata2*, *Nfe2*; Table 8), the expression of which are significantly down-regulated in *Scl* $^{-/-}$ EBs (Table 6). This implies that SCL is directly involved in transcriptional activation of these genes in order to induce a haematopoietic cell fate.

Table Eight.
Genes Associated with Genomic Regions Bound by SCL in Day 3.5 Wild Type EBs
(Ranked in order of fold enrichment over input control)

Transcription Factors	***Signalling***	***Transcription Processes***	***Metabolism***
Foxc1	Cntn3	Rybp	Ak3l1
Hey2	Dgkg	Esrrg	Iqgap2
Lyl-1	Pnoc	Grid2	Kcna6
Eto-2	Hdac4	Zfp300	Suclg1
Zfp317	Cd96	Akirin2	Gstm7
Nkx2.5	Fgfr2	Irx6	Nat12
Zfp280c	Npvf	Rpl19	23100A20Rik
Scl	Hnt	Zfpm2	Isoc1
Tbx5	Eltd1	Tmpo	Prnp
Tbx20	Fbxw7	Zfhx3	Pld1
Tbx3	Sod1	Tle3	Chi3l4
Myog	Olfr283	Cbx2	Nos1
Eomes	Accn2	Zfp169	Rabgap1l
Atxn7l3	Baiap2l	Pou2a1	Fahd2a
Gm13154	Cxcr4	Eif2c1	Fdxr
Foxf1a	Wnt5b	Calm2	Aloxe3
Hoxb8	Tnfrsf8	Jarid2	Scd4
Sp4	Tgfbr2	Smarc1	Rap2a
Gzmb	Hcrtr1	Trim69	Adrbk2
Gata2	Pdgfd	Tial1	Plcl2
Etv2	Ltbp1	Phb	Acsl6
Nfe2	Nxph4	Eif3b	
Foxk2	Prokr1	Ankrd7	
Snai3	Mc4r	Nfib	
E2f7			

Transport	***Development***	***Cell Cycle***	***Cell Adhesion***
Stxbp6	Bdnf	Nuf2	Comp
Kcnq4	Zfp217	Nsl1	Alcam
Slc2a9	Hfe	Ccnh	Cntn4
Steap3	Inhba	Cadm1	Olfm4
Spire1	Plcg2	Rhou	Ctnna2
Kcnmb4	Rb1cc1	Sh3bp4	Cdt1
Slc25a3	Chl1	Pik3r1	
Kctd11	Atg7	Cyr61	
Vamp3	Ar		
Trpm1	Etnk2		
Tomm7	Semaa		
Slc2a6	Ppp3r1		
Sec24d	Etl4		
Fam12b			
Ergic2			
Gsn			

Lineage Specific	***RNA-processing***	***Miscellaneous***	
Pln	Exosc2	Serpi3n	Serpinn
Ccrn4l	Raly	Crtam	Il22
Mtap1b	Eif2c1	Dusp10	Aurkc
Igfbp5	Nsun7	Prkacb	V1re10
Axin2	Fbl	Galnt2	Sh3bp4
Myo10		Ly75	Psd4
Chrdl2		Atic	Pla2g4d
F2rl1		Nudt7	Gdf3
Slc10a1		Psmb7	Epb4.3
Slc39a3		Dnaic2	Itch
Atp2b1		Stk39	Rnf180

ChIP sequencing did not find SCL bound within the vicinity of the Gata4 locus, indicating GATA4 is not likely to be a direct target of SCL in early EB cells. However, SCL does bind near several cardiac-associated genes (e.g. *Foxc1*, *Tbx20*, *Tbx5*, and *Nxk2.5;* Table 8). Analysis of genomic sequences bound by SCL revealed known SCL binding motifs (E-box and GATA motifs [161, 162, 164] under *Tbx20*, *Tbx5* and *Nkx2.5* peaks, providing further evidence that these regions are directly bound by SCL. During initial heart tube specification, TBX5 and NKX2.5 work together with GATA4 to regulate the primary heart field [86, 87, 97, 99, 238]. It would therefore be interesting to over-express and repress *Tbx5* and *Nxk2.5* in our ES/EB system, as this may result in cardiac, and potentially haematopoietic, defects similar to that of GATA4, thus recapitulating the *Scl* $^{-/-}$ phenotype. Note that SCL was not found bound to many cardiac genes which are up-regulated in *Scl* $^{-/-}$ EBs, implying that these are not direct gene targets of SCL (discussed further in Section 7.7). Therefore, to fully understand the function of SCL, further investigation is required to unearth the exact molecular pathways behind SCL mediated repression of cardiac fate. This will include repeating ChIP-sequencing using optimised conditions, such as changes in cross-linking reagents, protein beads and/or the concentration of salt washes (see Sections 6.3.3 and 7.9).

ChIP-sequencing showed SCL to bind DNA regions near genes involved in signalling pathways known to regulate mesoderm patterning during embryogenesis (e.g. *Fgfr2*, *Tgfbr2*, *Wnt5b*; Table 8). Increased levels of FGF signalling activate cardiac programmes and represses blood programmes in zebrafish anterior hemangioblasts [148]. Likewise, expression of the non-canonical Wnt ligand WNT5b has been shown to partially inhibit canonical Wnt/beta-catenin [246], which in turn has been

shown to inducing cardiac differentiation during early developmental stages [16, 89, 152, 226]. Therefore, it is possible that SCL acts to suppress cardiac commitment via transcriptional repression of the FGF receptor *Fgfr2* and /or the canonical-wnt pathway inhibitor *Wnt5b* (see Section 7.9.3).

Additionally, SCL was associated to genes coding for transcriptional inhibitors, such as histone deacetylases (*Hdac4*) and members of the polycomb complex (*Rybp*, *Jarid2*). HDAC4 is a critical transcriptional co-repressor which is active in a range of cell types, including cardiac lineages [247]. Furthermore, HDAC inhibition in EBs has been shown to cause increased spontaneous contractions (i.e. cardiomyocyte formation) and up-regulation of *Gata4*, *Mef2c*, *Nkx2.5* and cardiac actin mRNA expression chen11 [248]. Likewise, the polycomb complex has been linked to transcriptional repression and regulation of cell fate choice in many cellular contexts [249, 250]. Polycomb repressive protein complexes (PRC1 and PRC2) maintain gene expression patterns setup during early development by regulating chromatin structure. The YY1-binding protein RYBP is required for PRC1 and PRC2 recruitment [251], while JARID2 is a central component of PRC2 [249] and has been directly linked to cardiac differentiation via regulation of Notch signalling [252].

Overall, this data implicates SCL in a general role in controlling cellular transcriptional activity. Furthermore, it shows that SCL acts to specify mesoderm development through regulation of key transcription factor expression and/or signalling pathways (discussed further in Section 7.7).

6.3.3 ChIP Optimisation

In order to ensure that ChIP-sequencing experiments identified all the genomic regions bound by SCL, ChIP techniques were optimised at several stages. Optimisation was performed by ChIP followed by qPCR for genomic regions previous described as being bound by SCL in EBs (G*ata2* promoter and *Scl*+19 [253, 254]), regions identified from our initial ChIP-sequencing experiments (*Scl* promoter and *Lyl-1* promoter; Figure 6.13), and *Scl* -16 as a negative control.

ChIP techniques were optimised for cell number (Figure 6.14A), cross-linking reagents (Figure 6.14B) and beads used to pull down SCL-DNA complexes (Figure 6.14C) were tested. The use of low cells numbers ($5x10^5$ per ChIP) resulted in low enrichment levels compared to a control (no antibody) ChIP. Fold enrichment was amplified when using 10^6 cells, but there was no further improvement with increasing cell numbers (Figure 6.14A).

Fold enrichment of anti-SCL ChIP over a no-antibody control was not significantly altered at the G*ata2* promoter or *Scl*+19 regions by changes in cross-linking (formaldehyde (FA) versus EGS; Figure 6.14B) or by the use of different protein beads (Salmon sperm versus magnetic Dynabeads; Figure 6.14C). However, further optimisation performed using additional positive controls revealed that enrichment could be improved on some genomic regions by both EGS cross-linking and Dynabeads (Figures 6.14B and 6.14C). Therefore, repeating ChIP-sequencing, using optimised techniques (i.e. EGS cross-linking and Dynabeads), may uncover additional SCL gene targets (discussed further in Section 7.9).

6.4 Summary

High-throughput, whole-genome analysis (microarray and ChIP-sequencing) uncovered many potential gene targets for SCL in day 3.5 EBs (Tables 4 and 5). These include a variety of genes encoding transcription factors, signalling proteins, transcriptional repressors, and proteins involved in cell metabolism and transport. These differ somewhat from SCL target genes described in erythroid and leukemic cells [162, 164], although there was some overlap in haematopoietic gene targets (e.g. *Lyl-1*, *Gata2*, *Eto-2* and *Nfe2*). This indicates that SCL acts differently depending on its cellular context, and implicates SCL in a critical role during early development. Some haematopoietic gene loci were found to be bound by SCL in day 3.5 EBs (Table 8). Although direct SCL-DNA binding is not necessary for the transcriptional activity of SCL during haematopoietic specification (*Scl* $^{RER/RER}$ EB cells; Figures 4.1, 4.2, 4.12 and 6.3), this suggests that SCL works on the loci of DNA to activate haematopoietic gene transcription, most likely through recruitment via protein-protein interactions.

The expression of several cardiac genes were up-regulated with the loss of SCL expression (e.g. *FoxH1*, *Tbx20* and *Gata4*; Figures 6.1, 6.4 and Table 6). ChIP assays show that SCL binds the genome near the loci of many cardiac-related genes (e.g. *Tbx20*, *Tbx5* and *Nkx2.5*; Table 8), however, with the exception of *Tbx20*, there is no overlap between microarray and ChIP data (Figure 6.13bB). We have shown that direct SCL-DNA binding is required for SCL mediated repression of cardiac fate as *Scl* $^{RER/RER}$ EBs exhibit an increased potential for cardiomyocyte generation and over-express cardiac transcription factors (Figures 4.8 to 4.10, 4.12 and 6.3). However, this could be an indirect result of SCL-mediated regulation of upstream events.

Other potential SCL targets include genes associated with FGF, Wnt and TFGβ signalling pathways (Tables 4 and 5), all of which have been linked to embryonic development and cardiac differentiation. In addition, SCL may have a role in regulating the expression of transcriptional repressors, such as *Hdac*, *Rybp* and *Jarid2* (Table 8). Together, this data implicates SCL as a transcriptional regulator working through different molecular mechanisms to ensure correct cell fate during mesoderm specification.

Changes in GATA4 expression effected cardiac and haematopoietic development in a reciprocal manner to SCL, with over-expression of GATA4 driving cardiomyocyte generation at the expense of blood development. GATA4 is first expressed in day 3 DP populations, before becoming exclusively associated with SP-P populations (Figure 6.6). In addition, GATA4 is over-expressed in *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs (Figure 6.6) and repressed in SCL over-expressing EBs (Figure 6.7) in accordance with changes in cardiac potential (see Section 4). Over-expression of GATA4 results in increased cardiomyocyte production (Figure 6.9), coupled with a decrease in SCL expression (Figure 6.11) and haematopoietic development (Figure 6.10), thus mimicking the *Scl* $^{-/-}$ phenotype. Additionally, Gata4 $^{-/-}$ EBs have an increased ability to generate SCL^{+} lateral mesoderm populations (Figure 6.11) and haematopoietic /endothelial blast colonies (Figure 6.10).

Collectively, this suggests that antagonistic cross-talk exists between lateral and cardiac mesoderm during early stages of development, where networks of critical transcription factors act to specify a specific lineage while repressing alternative cell fates. SCL and GATA4 appear to be key transcription factors involved in these mechanisms.

7. Discussion

During embryonic development, multi-potent progenitor cells become committed to a specific pathway in response to establishment of cell specific transcription programmes [7, 9, 11]. Early embryonic cells exhibit a certain degree of plasticity and lineage flexibility, allowing them to switch their cell fate in response to changes in gene expression and/or extra-cellular signals [16, 17, 143, 145, 148] (see Section 1.6). This is illustrated by the potential to convert committed mesoderm progenitors into alternative cell fates upon forced changes in gene expression [147-151, 153, 208]. One of the mechanisms that safeguard patterning and cell fate decisions during development is the cross-antagonism between regulators of alternative cell fates. This ensures lineage specification and repression of alternative lineage pathways, thereby fixing a given cell fate.

The work described in Sections 3, 4 and 6 demonstrates the critical role played by the transcription factor SCL in such mechanisms during the early stages of haematopoietic development. Not only is SCL essential for the emergence of the first haematopoietic cells, but SCL is also required to repress non-haematopoietic lineages, ensuring correct cell fate decisions during the early phases of mesoderm specification. Furthermore, we have demonstrated that SCL activates haematopoietic programmes and represses cardiac differentiation via different molecular mechanisms. In addition, we have identified several blood and cardiac genes as SCL direct targets, thus opening the way towards molecular characterisation of the transcriptional mechanisms underlying mesoderm specification and patterning. In addition, we have shown that over-expression of GATA4 (a critical cardiac transcription factor) drives cardiomyocyte generation at the expense of blood

development. This further shows the importance of maintaining a fine balance in the levels of key transcription factors during critical stages of embryonic development.

7.1 Using the ES Cell System to Model Development

The ES/EB system provides an accessible and reliable system to study the early stages of development, although it is not without its limitations. Differentiating ES cells recapitulate the early stages of development; with mesoderm commitment occurring in the same temporal waves as in the embryo (see Section 1.5). The main drawback of the ES/EB system is the lack of spatial organisation, due to the lack of niche exposure to extra-cellular signalling that occurs in the embryo.

Choice of ES cell line and differentiation conditions affects EB development. Reflecting phenotypic variation between different mouse lines, differences exist between ES cell lines [255]. Therefore it is important to note that all genetically altered cell lines used in this study have a J1 wild type ES cell background, derived from the 129S4/Svjae mouse strain.

ES cell differentiation is affected by extra-cellular stimuli provided in culture medium. Modification of these conditions can induce different cell fate decisions. For example, our culture conditions are designed to enhance mesoderm, in particular haematopoietic, differentiation [115], as this is our developmental area of interest. Therefore, it is unsurprising that relatively few endoderm lineage precursors (identified by E-Cadherin expression) are specified (Section 3.1.1).

Previous studies have reported higher levels of E-Cadherin expression and endoderm formation within EBs, coupled with lower levels of Flk-1 expression. However, these studies use Activin, a cytokine known to activate endoderm differentiation, to induce ES cell differentiation [16, 17, 125, 141]. Likewise, studies using Flk-1 and PDGFRα as mesoderm markers have reported different percentages of cells contributing to DP, SP-P and SP-F lineages when ES differentiation is induced using Activin and BMPs [97, 125, 137].

Our culture conditions are advantageous to investigate mesoderm patterning as the full potential of cells to form mesoderm lineages can be exploited, allowing detailed study of the mechanisms that induce and maintain specific differentiation decision. However, it could be argued that we are inducing a bias into out differentiation assays, as illustrated by the lack of significant endoderm or ectoderm differentiation. Coupled with the lack of regionalisation of cellular compartments, this makes *in vivo* validation of our results essential.

7.2 Defining Mesoderm Populations

Within the mouse embryonic mesoderm, cells exhibit different lineage potential depending on regional and temporal organisation [7, 9, 13]. Mesoderm populations with distinct differentiation potentials can be fractionated on expression of specific cell surface markers, such as Flk-1 and PDGFRα [12, 18, 20-22]. These markers are also expressed in differentiating EBs, where they have been associated with lateral mesoderm [12, 19, 119, 122, 125, 171], paraxial mesoderm [106, 113, 125] cardiac mesoderm [19, 97, 122] precursor populations at given time points.

Here, for the first time, we have documented the dynamic changes in Flk-1 and PDGFRα expression that occur as EB cells differentiate. Furthermore; we link these immuno-phenotypic changes to lineage commitment, as outlined in (Figure 7.1). In addition, we show that expression of SCL and GATA4 identifies presumptive haematopoietic and cardiac progenitors within early mesoderm populations, thus providing an informative system to track and study the mechanisms governing blood and cardiac development.

7.2.1 Refining Mesoderm Patterning During EB Differentiation

Flk-1 expression has been associated with blood, endothelial and cardiac development in the ES/EB system, while PDGFRα expression has been linked to paraxial mesoderm and cardiac fates (as outlined in Figure 1.4). Detailed analysis of Flk-1 and PDGFRα expression and cell fate potential over a time course of EB differentiation (day 3 to 6), has shown that, in our system, the first DP populations (Flk-1$^+$ PDGFRα^+) are multi-potent and contain sub-populations of cells that give rise to lateral mesoderm (haematopoietic and endothelial), paraxial mesoderm (chondrogenic and osteogenic) and cardiac lineages (Figures 3.2, 3.4 and 3.5).

Early developmental events are mirrored during ES/EB differentiation. Indeed, during the early stages of EB differentiation, SP-F (Flk-1$^+$) and (PDGFRα^+) populations are primed for lateral lineages and paraxial lineages respectively (Figure 3.4). This reflects *in vivo* embryonic events, where Flk-1 is expressed in primitive streak derived cells destined for yolk sac and

intra-embryonic mesoderm lineages (haematopoietic and vascular endothelium) [12, 18, 49, 61, 62, 172] andPDGFRα expression marks primitive streak precursors of paraxial mesoderm lineages [13, 22]. As SP-F cells further commit to a haematopoietic lineage, they down-regulate Flk-1 expression (which is retained by endothelial lineages) and gain expression of early haematopoietic cell surface markers, such as CD41 (Figure 3.6). Again, this reflects *in vivo* events, where Flk-1 is only a transient marker of haematopoietic development [20, 172].

At early time points, cardiac potential was enriched in Flk-1$^+$ populations (Figure 3.4), in line with Flk-1 expression described on cardiac lineage precursors (endocardium and myocardium) *in vivo* [19, 20]. However, cardiac potential is restricted to Flk-1 negative cells in day 3.25 EBs (Figure 4.11) [19, 122]. This highlights the dynamic nature of the EB differentiation, as cardiac precursors must switch on Flk-1 expression within a 12 hour window, after haematopoietic and endothelial cells (summarised in Figure 7.1). As DP cells commit to a cardiac fate, Flk-1 expression is down-regulated (Figures 3.6 and 6.6). This supports *in vivo* data showing that PDGFRα is expressed by paraxial mesoderm precursors [13, 22] and by Nkx2.5$^+$ cardiac precursors in the cardiac crescent [21, 22, 95].

Although both paraxial and cardiac precursors undergo the same immuno-phenotypic transition (DP to SP-P), the kinetics of this differs. Day 3.5 SP-P cells have the potential to generate paraxial, but not cardiac, lineages (Figure 3.4), whereas at later time points cardiac progenitors are found almost exclusively in SP-P populations (Figures 3.6 and 6.6). Our data describes a link between cardiac and lateral mesoderm lineages during early stages of mesoderm development in the mouse, which reflects the cross-talk described between cardiac and hemangioblast development in the

zebrafish [148-150]. However, as cells become more differentiated, cardiac populations appear to become more closely related to paraxial lineages. This is in agreement with previous studies of EB differentiation which report that committed cardiac progenitors are still able to generate skeletal muscle and bone (i.e. paraxial mesoderm derived tissues), but not haematopoietic cells [97], while cells committed to a haematopoietic or endothelial cells (blast colonies) are not able to generate cardiomyocytes [119].

In conclusion, co-expression of Flk-1 and PDGFRα identifies a mesoderm population that contains precursors to lateral, paraxial and cardiac lineages. Importantly, these different progenitor populations transition through a DP population, in distinct temporal waves (summarised in Figure 7.1). As discussed in the following sections, it will be important to further fractionate the DP population to allow the isolation of distinct lateral, paraxial and cardiac mesoderm progenitor populations.

7.2.2 SCL Expression Marks Lateral Mesoderm-Primed Populations

The first SCL expressing cells have been associated with Flk-1$^+$ expression in the yolk sac blood island [12, 18, 172] and in differentiating ES cells (using a ES cell line with human CD4 expression knocked in to the SCL locus; *Scl* $^{+/hCD4}$ cells) [171]. However, co-expression of SCL and PDGFRα has not been previously described *in vitro* or *in vivo*.

Using intracellular flow cytometry and qPCR techniques, we have directly tracked the expression of SCL during EB differentiation to confirm that, in wild type EBs, the first SCL$^+$ cells are seen exclusively within Flk-1$^+$

cell populations (Figures 3.8 and 3.9). SCL expression was first detected in day 3 EBs, and steadily increases with EB differentiation up to day 6 (Figure 3.7 to 3.9). This is in agreement with results obtained from *Scl* $^{+/hCD4}$ cells (our unpublished data) [171]. Of note, Ismailoglu *et a*l report slightly different data; they describe a peak in *Scl* expression in day 4 EBs, after which *Scl* expression is down-regulated [185]. This discrepancy is most likely due to differences in ES cell lines and culture conditions.

Importantly, and for the first time, we show that SCL is expressed within a subset of Flk-1$^+$ PDGFRα^+ (DP) cells during early time points in EB differentiation (Figure 3.8). Throughout EB differentiation, SCL expression corresponds with Flk-1/PDGFRα populations which are enriched for haematopoietic and endothelial potential (Figures 3.4 and 3.6). This is consistent with SCL being associated with hemangioblast potential in both mouse and zebrafish [1, 3, 71, 171, 172, 175, 182].

In conclusion, we have shown that SCL is expressed in a more multi-potent mesoderm population than previously thought. SCL expression is sustained in these cells through haematopoietic specification, thus providing a system whereby the first presumptive blood cells can be identified and tracked throughout early development (Figure 7.1). Furthermore, this implicates SCL in a previously undetermined role in lateral mesoderm cell specification, prior to haematopoietic or endothelial lineage commitment.

It would be interesting to further define SCL$^+$ DP sub-population by identifying additional markers. RNA-sequencing of SCL/Flk-1/PDGFRα fractionated populations may uncover additional cell surface markers which can be used as an indication of specific lateral mesoderm lineages (i.e. haematopoietic, endothelial, vascular smooth muscle) versus cardiac and

paraxial primed cells. This would allow isolation of more pure, homogeneous, populations for molecular analysis.

7.2.3 GATA4 Expression Marks Cardiac Progenitors

GATA4 expression has been associated with both endoderm and cardiac development [133, 134, 236, 239]. In our ES/EB differentiation system, few cells are specified with endoderm potential and changes in GATA4 expression do not affect the percentage of EB cells committing to an endoderm fate (measured by E-Cadherin expression; data not shown). This therefore indicates that GATA4 expression is representative of cardiac potential. GATA4 is first expressed in a subset of DP cells at day 3 of EB differentiation, and is ultimately associated with SP-P populations (Figure 6.6). Expression of GATA4 correlates with enrichment in cardiac potential, suggesting that GATA4 marks cells primed for cardiac differentiation.

Previously, EB cardiac progenitors have been described as Brachyury$^+$ Flk-1$^+$ [19, 122], Flk-1$^+$ CXCR4$^+$ [256], Flk-1$^+$ PDGFRα^+ [137] and Flk-1$^+$ PDGFRα^+ MesP1$^+$ CXCR4$^+$ [97]. Our data confirms that cardiac potential arises from a subset of Flk-1$^+$ PDGFRα^+ mesoderm cells, and adds GATA4 as an early marker of cardiomyocyte fate in the mesoderm.

Whether GATA4 expression marks cells capable of generating other cardiac lineages (i.e. endocardium or cardiac vascular smooth muscle) is undetermined. *In vivo*, Flk-1 expression is lost upon cardiomyocyte differentiation, but is retained by endocardial-fated cells [94]. Therefore, GATA4$^+$ SP-P populations may specifically contain cells with a specific marker of cardiomyocyte fate. The expression pattern of GATA4 further

highlights the temporal waves of DP to SP-P transition exhibited by paraxial and cardiac mesoderm populations (summarised in Figure 7.1). This further illustrates the necessity to further fractionate mesoderm sub-populations according to distinct fates, thus allowing molecular analysis on pure, homogeneous populations.

7.3 SCL is required for Lateral Mesoderm Specification: New Insights into SCL Function

SCL is an important regulator of haematopoietic development and is absolutely required for the emergence of the first embryonic blood cells. *Scl* $^{-/-}$ ES cells are unable to generate haematopoietic cells *in vitro* (Figure 4.3) [1-3], reflecting the absence of yolk sac haematopoiesis observed in *Scl* $^{-/-}$ embryos [4, 5]. However, the action of SCL in early developmental lineages is poorly understood, as most studies have focused on SCL's role in erythrocyte and megakaryocyte development or in leukaemia.

SCL is first expressed within a cell population that contains progenitors for lateral, paraxial and cardiac lineages (DP population; Figure 3.8). In the absence of SCL (*Scl* $^{-/-}$ EBs), DP populations are unable to give rise to lateral mesoderm progenitors (Figures 4.1). This however, is not due increased cell death (Figure 4.3). Furthermore, lateral mesoderm populations are increased when SCL is over-expressed (Figure 4.14). This suggests an instructive role for SCL in cell fate decisions, where the loss of *Scl* expression results in lateral mesoderm precursors progressing down an alternative differentiation pathway. In agreement with this, expression of haematopoietic (e.g. *Lyl-1*, *Gata2*, *Ikaros*) and endothelial (e.g. *Tie2*, *Cdh5*) genes is significantly down-regulated in *Scl* $^{-/-}$ EBs (Table 6). SCL was also

found bound to the promoters of several haematopoietic regulators (e.g. *Lyl-1*, *Gata2*, *Nef2*) in day 3.5 wild type EBs (Table 8). In addition, SCL binds DNA regions near genes involved in signalling pathways know to regulate mesoderm patterning during embryogenesis (e.g. *Fgfr2*, *Tgfbr2*, *Wnt5b*, *Axin2*; Table 8). This opens up the possibility that SCL acts to specify mesoderm development through regulation of key transcription factor expression and/or signalling pathways.

Analysis of *Scl* $^{RER/RER}$ EBs confirms our previous findings that SCL does not directly bind DNA to specify early haematopoietic lineages (Figures 4.2 and 4.4) [173, 187]. We have extended this initial data and shown that SCL-DNA binding is dispensable for lateral mesoderm specification (Figure 4.1). ChIP experiments did find SCL bound to the promoters of several haematopoietic genes (Table 8), suggesting that SCL may be recruited to DNA through protein-protein interactions. Characterisation of SCL-containing multi-protein complexes in these cells will be an important step towards furthering our understanding of SCL mechanisms of action

Collectively this data shows that SCL is required at an earlier developmental time point that previously determined. It implicates the involvement of SCL in lateral mesoderm specification, prior to haematopoietic commitment, and further suggests that SCL works through a molecular mechanism which does not involve direct SCL-DNA binding.

7.4 SCL Negatively Regulates Cardiac Development

A single transcription factor can be involved in the specification of more than one lineage and acts through different molecular mechanisms depending on the cellular context [187]. In line with this, previous studies have shown that SCL is capable of acting as a repressor of cell differentiation (see Section 1.7.6). We have shown that during mouse ES cell differentiation, changes in SCL expression have an impact on non-haematopoietic programmes, particularly on cardiac development, implicating SCL as a repressor of cell fate.

We have documented this function in differentiating ES cells. (1) *Scl* $^{-/-}$ EBs contained an expanded cardiomyocyte progenitor population, which is not a result of increased cell proliferation (Figures 4.1, 4.3, 4.8 and 6.6). (2) Over-expression of SCL causes a decrease in cardiomyocyte progenitor specification (Figure 4.16). (3) Re-introduction of SCL expression in *Scl* $^{-/-}$ EBs counteracted the *Scl* $^{-/-}$ cardiac phenotype (Figure 4.17).

Analysis of *Scl* $^{RER/RER}$ ES cells showed that while SCL specifies commitment to a lateral mesoderm lineage through a mechanism independent of direct DNA binding, direct SCL-DNA binding is required for repression of cardiac differentiation (Figures 4.8 to 4.10). This provides strong evidence supporting the hypothesis that changes in SCL expression levels caused a change in cell fate. Indeed, the potential of *Scl* $^{RER/RER}$ EBs to generate normal lateral mesoderm populations but increased numbers of cardiomyocytes confirms that the *Scl* $^{-/-}$ cardiac phenotype is due to the loss of repression, and not an indirect result of the loss of haematopoietic specification. This directly implicated SCL in the regulation and control of mesoderm cell fate decisions. Moreover, it shows that SCL-driven

specification of lateral mesoderm and SCL-mediated repression of cardiac differentiation occurs through independent pathways.

The expression of several key cardiac genes is up-regulated in *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs (*Hand2*, *FoxH1*, *Mesp1*, *Tbx20*; Figure 4.12 and Table 6). ChIP-sequencing revealed SCL to be bound near several cardiac-associated genes (e.g.*Tbx20*, *Tbx5*, and *Nxk2.5;* Table 8) and genes involved in cardiac-related signalling pathways (e.g. *Wnt5b*, *Fgfr2*, *Axin2* and*Tgfbr2*; Table5). However, SCL was not found bound to many cardiac genes which are up-regulated in *Scl* $^{-/-}$ EBs (e.g. *Gata4*, *Hand2*, *and FoxH1*). One possibility is that SCL only actively regulates the expression of a few cardiac-related genes and/or signalling proteins which, in turn, regulate a cascade of down-stream events. Alternatively, our ChIP experiments may have failed to detect the whole spectrum of cardiac-related genes regulated by SCL. Therefore, further investigation is needed to unearth the exact molecular pathways behind SCL mediated repression of cardiac fate (see Section 7.9).

7.5 Reciprocal Antagonism between Cardiac Development and Haematopoiesis during Mouse Development

As discussed above, SCL plays a pivotal role in blood and cardiac lineage specification. Differentiation analysis of GATA4 over-expressing and *Gata4* $^{-/-}$ ES cells revealed that reciprocal cross-antagonism can be induced by perturbing the levels of this cardiac transcription factor.

7.5.1 GATA4 Expression Affects Cardiac and Blood Development

GATA4 expression can be used as a marker of cardiomyocyte potential, with expression of GATA4 being increased in *Scl* $^{-/-}$ EBs (Figure 6.6) and decreased in EBs over-expressing SCL (Figure 6.7), coinciding with gain and loss of cardiac differentiation potential. Using GATA4 over-expressing and *Gata4*$^{-/-}$ ES cells, we describe the reciprocal situation.

Over-expression of GATA4 induces ectopic cardiomyocyte generation (Figure 6.9) [237], which is not surprising given that GATA4 is a critical regulator of primary heart field specification [99, 133, 236, 239, 242, 243]. In contrast, loss of *Gata4* expression (*Gata4* $^{-/-}$ cells) does not perturb cardiac development (Figure 6.9), as compensatory mechanisms exist involving increased GATA5 and GATA6 expression to counteract loss of GATA4 expression [136, 241, 243].

More interestingly, over-expression of GATA4 resulted in fewer SCL^{+} cells (Figure 6.11). This maybe a result of repression of SCL expression in Flk-1^{+} GATA4^{+} cells, or an over-all loss of lateral mesoderm precursors in GATA4 over-expressing EBs. Cell fate assays showed that less commitment to a haematopoietic fate in GATA4 over-expressing EBs compared to wild type (Figure 6.10). This suggests that increased cardiac development occurs at the expense of haematopoiesis. *Gata4* $^{-/-}$ EBs generated haematopoietic cells to a wild type level (Figure 6.10). However, day 3.5 *Gata4* $^{-/-}$ EBs did exhibit an increased ability to generate colonies with haematopoietic and endothelial potential (Blast colonies; Figure 6.10). In addition, despite the overall percentage of *Gata4* $^{-/-}$ SCL^{+} cells not differing from wild type, a greater proportion of day 3.5 *Gata4* $^{-/-}$ SCL^{+} cells were SP-F, rather than DP,

in comparison to wild type (Figure 6.11). It is possible that in the absence of GATA4, DP cells are able to commit to a lateral mesoderm lineage (marked by the loss of PDGFRα expression) at an earlier time point.

The overall lack of a *Gata4* $^{-/-}$ haematopoietic phenotype reflects the lack of a *Gata4* $^{-/-}$ cardiac phenotype. Loss of *Gata4* and *Gata6* expression however results in severe cardiac defects [243, 257]. Therefore the haematopoietic potential of *Gata4* $^{-/-}$ *Gata6* $^{-/-}$ double knock-out ES cells should be examined to determine whether there is a phenotypic switch between cardiac and haematopoietic lineages.

7.5.2 Drawing Parallels with Zebrafish Development

Several studies have reported developmental antagonism between blood and cardiac programmes in zebrafish, highlighting how closely linked these lineages are during development. Indeed, alteration of extra-cellular signalling gradients (e.g. FGF) and levels of blood and cardiac transcription factors (e.g. *Cloche*, *Scl* or *Nkx2.5*) impacts on lineage decisions in the anterior lateral mesoderm (AML), which contains both cardiac and hemangioblast populations [148-150].

It is important to note that this anterior hemangioblast population does not exist in higher vertebrates, such as the mouse, as it is believed to have been recruited into the heart field during evolution [148, 258, 259]. In the mouse embryo, haematopoietic cells arise exclusively from the posterior mesoderm, while both the posterior (PLM) and anterior (ALM) mesoderm

contribute to haematopoiesis in the zebrafish. These evolutionary differences are in part due to changes in the spatial organisation of the embryonic endoderm. In the mouse, blood and vascular development occurs in proximal-posterior regions in response to cytokine signals from the visceral endoderm. In zebrafish, the ALM receives similar signals from the definitive endoderm [260-262].

Here, we have described a model for antagonism between blood and cardiac lineages during mouse development. This involves maintaining a fine balance in transcription factor expression and, mechanistically, parallels what has been described in the zebrafish. However, given the differences in cardiac and haematopoietic development in these two organisms, studying the molecular mechanisms that underlie cross-antagonism is likely to unveil differences between the two species. An example of this is the role that we describe for GATA4 in repressing haematopoietic specification in mouse development. This is in contrast to the zebrafish, where GATA4 is required earlier for the generation of both anterior hemangioblasts and cardiac tissue [262], while NKX2.5 is the cardiac transcription factor antagonising SCL [148]. This is likely to be reflective of differences in regulatory signals that govern the development in these two organisms, with the loss of zebrafish ALM specification in the absence of GATA4 being due to defects in endoderm signalling.

In contrast to the zebrafish, our data suggest that loss or gain of GATA4 does not appear to affect initial mesoderm patterning in mouse EBs, but ectopic expression of GATA4 does drive cardiomyocyte differentiation over haematopoietic and endothelial differentiation. This lack of early mesoderm phenotype may be due to the continued presence of increased GATA5 and GATA6 expression in these cells. Alternatively, it could reflect

the different molecular mechanisms controlling mesoderm specification in the zebrafish and mouse. Of note, *Gata4* expression has been detected in the mouse yolk sac blood islands, where it is thought to have a role in endoderm development [261, 263]. The ES/EB system reflects early developmental events during mouse embryogenesis (Figure 1.4), where the first presumptive blood and cardiac precursors are specified in the posterior region of the embryo, prior to their migration to extra-embryonic (haematopoietic) or to more anterior intra-embryonic (cardiac) regions. We believe that the blood/cardiac cross-antagonism seen in the ES/EB system most likely reflects cell fate plasticity of early primitive streak / posterior mesoderm populations, suggesting a link between cells destined for the yolk sac blood islands (first primitive haematopoietic cells) and the primary heart field. However, *in vivo* validation is needed to confirm this.

Collectively this provides evidence that the cross-antagonism mechanisms between cardiac and blood development, first described in zebrafish, are conserved throughout evolution and exist during mouse embryogenesis. This suggests that, although the embryonic locations of haematopoietic and cardiac tissues differ between species, the fundamental molecular mechanisms underlying lineage choice are conserved. For example, disruption of BMP signalling caused loss of *Scl* expression and haematopoietic defects during both Xenopus and mouse development [254, 264]. Therefore, it would be extremely interesting to further investigate which upstream events regulate expression of key transcription factors (i.e. *Scl* and *Gata4*) to determine if signalling pathways which are important for determining cardiac versus haematopoietic fate in the zebrafish are (such as FGF [148]) play similar roles during mesoderm patterning in mouse development.

7.6 SCL in Paraxial Mesoderm Development

In addition to an increase in cardiac potential, *Scl* $^{-/-}$ EB cells gave rise to significantly more chondrogenic nodules that wild type EB cells (Figure 4.5), but osteogenic and myogenic differentiation appeared largely unaffected (Figures 4.6 and 4.7). Previous studies in the zebrafish showed SCL over-expression to induce haematopoietic and vascular development in somitic paraxial mesoderm cells [151]. However, SCL has not previously been directly implicated in chondrogenic development, or as a repressor of paraxial mesoderm fate under normal expression condition.

Scl $^{-/-}$ EBs have an expanded SP-P population compared to wild type. This is first apparent around day 3.5 of EB differentiation, closely following the onset of SCL expression in wild type EBs. A substantial proportion of this is due to the expansion of *Scl* $^{-/-}$ GATA4^{+} SP-P cells (Figure 6.6), which account for the increase in cardiac potential exhibited by *Scl* $^{-/-}$ EBs. However, the increase in GATA4^{+} cells does not completely account for the total increase in SP-P cells; especially at earlier time points (compare Figures 4.1 and 6.6), suggesting increased specification of paraxial mesoderm progenitors.

7.6.1 Chondrogenic Differentiation

Scl $^{-/-}$ EBs were able to generate significantly more chondrogenic nodules compared to wild type EBs (Figure 4.5). This is particularly interesting as SCL has not been previously linked to chondrogenesis. Increased chondrogenic potential in *Scl* $^{-/-}$ EBs can be attributed to the over-expression of the critical chondrogenic regulator *Sox9* (Figures 4.12 and 6.4), which is known to drive cellular commitment to a chondrogenic fate

[104, 107, 232, 265]. However, ChIP-sequencing experiments did not reveal SCL to be bound to either the vicinity of the *Sox9* promoter, or to any other chondrogenic-related genes (Table 8). Also, *Sox9* is not over-expressed in *Scl* $^{RER/RER}$ EBs, where haematopoietic development is specified (Figures 4.3, 4.12 and 6.4). In fact, *Sox9* expression was down-regulated in day 6 *Scl* $^{RER/RER}$ EBs (Figure 6.4). This indicated that *Sox9* is not a direct target for SCL.

SCL may act to repress commitment to a chondrogenic fate through other mechanisms of action. SCL may act independently of direct DNA binding, through protein-protein binding to inhibit chondrogenic promoting pathways. This may occur through sequestering other transcription factors and preventing them from inducing the expression of chondrogenic genes. Therefore SCL would not be bound to chondrogenic genes, and EBs expressing a non-DNA binding version of SCL (*Scl* $^{RER/RER}$ ES cells) would not exhibit a chondrogenic phenotype. Alternatively, the increase in paraxial mesoderm precursors in *Scl* $^{-/-}$ EBs maybe an indirect effect of the loss of SCL expression, resulting from a switch in cells fate in the absence of lateral mesoderm specification.

To investigate this further, it would be informative to determine whether *Scl* $^{RER/RER}$ SP-P and DP EB populations have an enhanced capacity to generate chondrocytes. If so, it would be interesting to determine if SCL is required to directly bind and sequester proteins that induced chondrogenic differentiation. Furthermore, it is yet to be determined if there is an expansion of paraxial mesoderm tissue and/or chondrogenic differentiation in the absence of SCL *in vivo*.

7.6.2 Osteogenic Differentiation

Scl $^{-/-}$ ES cells give rise to wild type numbers of osteoblasts (Figure 4.6), with decreased potential of *Scl* $^{-/-}$ SP-P cells to generate calcium depositing colonies most likely due to cross-antagonism occurring between bone and cartilage development as a result of increased in *Sox9* expression (Figure 4.12). This is in agreement with previous studies which also suggest that osteoblast generation is not directly linked to SCL expression [169]. However, it has been shown that *Scl* $^{-/-}$ ES cells can not generate osteoclasts [170].

Osteoclasts differ from osteoblasts in function, phenotype and origin [266]. Osteoblasts are responsible for the formation of new bone and develop from the paraxial mesoderm and adult mesenchymal stem cells (MSCs; can generate bone, cartilage and brown adipose tissue[267]. In contrast, osteoclasts work to resorb old bone deposits. They are phenotypically similar to monocytes and macrophages, and are thought to arise from c-kit^{+} haematopoietic precursor cells, although little is known about the molecular processes involved [170, 266]. Therefore, the lack of osteoclast differentiation in *Scl* $^{-/-}$ EBs is most likely due to the loss of a common haematopoietic / osteoclast progenitor, and does not indicate a role for SCL in bone formation or paraxial mesoderm specification.

7.6.3 Myogenic Differentiation

Overall, myogenic activity appeared largely unaltered in the absence of SCL (Figure 4.7). SCL expression has previously been linked to muscle development, as ectopic expression of SCL in differentiating myocytes (which do not normally express SCL) blocks their terminal differentiation

program [202, 203]. However, this is not achieved through the re-direction of cell fate, but through SCL binding and sequestering important regulators of muscle differentiation.

Microarray analysis showed that some muscle-related genes were up-regulated in the absence of SCL (Table 6). Additionally, expression of Myogenin (an early marker of muscle lineage specification [110]) was up-regulated in *Scl* $^{-/-}$ myocyte cultures (Figure 4.7), suggesting that myocyte differentiation is enhanced in those cells. However, ChIP-sequencing experiments did not show SCL bound to muscle-associated genes (Table 8). Therefore is unlikely that SCL plays a direct role in the repression of myocyte specification, but that an increase in myogenesis is a result of an increased percentage of cells generically committing to a paraxial mesoderm fate in the absence of lateral mesoderm specification.

7.7 SCL Direct Gene Targets

ChIP-sequencing revealed that in early EB populations, SCL is mainly found bound in the vicinity of genes that encode transcription factors, signalling proteins and proteins involved in transcriptional repression (Table 8). These include several haematopoietic genes which are down-regulated in *Scl* $^{-/-}$ (Table 6). Previously we have reported that in erythroid progenitors SCL is recruited to haematopoietic genes even when direct SCL-DNA binding is inhibited (i.e. *Scl* $^{RER/RER}$ cells) [164], suggesting that SCL still works “on-DNA”, but is recruited via protein-protein interactions. Given that *Scl* $^{RER/RER}$ ES cells are capable of normal lateral mesoderm specification (Figures 4.1, 4.2 and 4.4), it can be concluded that SCL is recruited to

haematopoietic genes during mesoderm development via similar mechanisms.

The enhanced cardiac potential exhibited by *Scl* $^{RER/RER}$ EBs suggests that SCL directly binds DNA to repress cardiac differentiation. In accordance with this, we found several cardiac genes to be bound by SCL (Table 8). However, with the exception of *Tbx20*, these did not correlate with cardiac genes found to be differentially expressed in *Scl* $^{-/-}$ EBs (Table 6). Therefore, SCL may not be repressing cardiac pathways through direct repression of transcription factor expression.

FGF, TGF/BMP and Wnt signalling pathways constitute upstream events which control cardiac development [14, 116, 142, 148, 149, 152, 236]. ChIP-sequencing showed SCL to be bound near several genes involved in these pathways (*Fgfr2*, *Tgfbr2*, *Wnt5b*; Table 8), opening up the possibility that SCL acts to specify mesoderm development through regulation of key signalling pathways. In addition, SCL bound genes associated with transcriptional repression complexes, such as the polycomb complex (*Rybp*, *Jarid2*) which have also been linked to the regulation of cardiac development [252]. This implicates SCL in a broader role in controlling transcriptional activity.

7.8 Concluding Remarks

In conclusion, this study has identified potential gene targets of SCL during mesoderm patterning. We show that during mesoderm development SCL acts as a lateral mesoderm / haematopoietic inducing transcription factor, which is likely to be recruited to DNA via protein-protein interactions rather than direct DNA binding. In addition, our data suggests that SCL acts to repress cardiac differentiation by acting up-stream of cardiac transcription factor expression. This implicates SCL in a previously un-described role in the regulation of signalling pathways and/or transcriptional repressor complexes, and highlights the importance of maintaining a fine balance in transcription factor expression levels during early development.

It is unlikely that the ability of SCL to simultaneously activate and repress opposing lineages through different molecular mechanisms is unique among transcription factors. Instead, we believe that our work opens up the possibility to study alternative mechanisms of action of other essential transcription factors, thus providing a greater insight into the mechanisms regulating both normal development and disease.

7.9 Future Plans

7.9.1 Tracking Cell Fate Changes in the Absence of SCL

Analysis of cell cycle progression and apoptosis showed that levels of proliferation and cell death were not altered in *Scl* $^{-/-}$ mesoderm populations (Figure 4.3). This indicated that the loss of SP-F and DN CD41^{+} haematopoietic cells and gain of DP and SP-P GATA4^{+} cells in *Scl* $^{-/-}$ EBs is not due to increased cell death and proliferation. Instead, we believe that DP cells that would normally express SCL and become SP-F, undergo a fate change and become re-directed towards a SP-P cardiac lineage.

To prove this, we plan to utilise an ES cell line which expresses human CD4 from the Scl promoter, to track the fate of these cells in the absence of Scl expression. Using neomycin step-up on the previously described *Scl* $^{+/hCD4}$ knock in ES cells [171], we have isolated ES cell clones targeted at both loci; *Scl* $^{hCD4/hCD4}$ ES cells. These cells do not express SCL, but cells in which transcription is normally initiated from the *Scl* promoter will express hCD4. We believe that there will be a short window of time, before the *Scl* promoter is shut down (as cells switch to a non-haematopoietic fate) during which hCD4 will be expressed. Using FACS analysis, the fate of hCD4 expressing cells in DP population will be tracked to determine whether these cells do indeed become SP-P and/or express GATA4. In addition, these cells can be sorted and re-plated under cardiac inducing conditions to discover whether they exhibit an increased cardiac potential.

7.9.2 Determining the Temporal Window of SCL Activity

The Tet-off inducible SCL expression system described in Section 5 was not adequate to determine the temporal window in which SCL expression is absolutely required during mesoderm development (see Sections 5.4). There are however several other inducible expression systems which could be utilised to introduce short pluses of SCL expression in differentiating *Scl* $^{-/-}$ EB at different time points.

In the 4HT system, hormone ligand (usually oestrogen receptor; ER) is fused to a nuclear protein and sequesters it in the cytoplasm, causing effective inactivation. When administered, 4HT binds to the ER and induces translocation into the nucleus, thus allowing normal protein activity [219]. Alternatively, the FKBP system controls gene expression through chemical control of protein stability. A destabilising domain of the FKBP protein is fused to the protein of interest and causes its degradation. Upon treatment with a chemical inhibitor, this degradation is inhibited, thus providing a reversible expression [225]. Either of these systems may allow tighter temporal control of SCL activity to determine when SCL expression is most critical.

7.9.3 Further Investigation into SCL Direct Gene Targets

Through a combination of microarray and ChIP-sequencing experiments, we have identified many possible direct gene targets of SCL during early developmental stages (Tables 4 and 5). However, we do not yet have a clear understanding of the molecular pathways behind SCL-mediated repression of the cardiac lineage.

Binding of SCL to potential target genes will be verified by qPCR, focusing on cardiac-related transcription factors (*Tbx20*, *Tbx5*, *Nxk2.5*), signalling proteins (*Cxcr4*, *Fgfr2*, *Wnt5b*) and proteins associated with transcriptional repression, such as the Polycomb complex (*Rybp*, *Jarid2, Hdac4*). Differential expression of these genes in wild type, *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ EBs will be investigated by qPCR over a time course. Depending on their differential expression, gain and loss of function studies will be used to determine whether they are responsible for enhanced cardiomyocyte generation in Scl $^{-/-}$ EBs, thus helping build up a network of molecular mechanisms that underlie mesoderm patterning.

7.9.4 Uncovering Additional Direct Gene Targets of SCL

Although direct SCL-DNA binding is not required for the emergence of the blood cells in the embryo, we have shown that SCL required its DNA-binding activity to act as a repressor of the cardiac lineage. Furthermore, SCL may still act on DNA as part of a multi-protein complex during haematopoietic development, although it is not itself directly bound to DNA.

Initial ChIP experiments were conducted using only formaldehyde cross-linking to bind SCL to DNA. We have since optimised ChIP protocols using EGS cross-linking, which is normally used to ChIP co-factors that do not bind DNA directly, but are recruited to DNA by protein partners. ChIP experiments will be repeated using optimised conditions. Comparing ChIP-sequencing performed on material from wild type and *Scl* $^{RER/RER}$ EBs would help determine whether SCL is directly bound to DNA, or recruited off DNA via other transcriptional regulators. In addition, performing ChIP- sequencing at several time points during EB differentiation would give a clear idea of

when SCL activity is required. We hope to uncover more potential direct gene targets. We will also perform a de novo search to determine the DNA sequences to which SCL binds during early development, to reveal whether SCL binds to its classical E-box/GATA motifs, and/or novel genomic sequences.

To gain further mechanistic insights into how SCL works during early development, we will determine whether SCL is bound to DNA in its classical pentameric complex (see Figure 1.5C). ChIP will be performed using antibodies against core complex proteins (E2A, LMO2, LDB1, GATA1/2) to isolate regions of DNA potentially bound by both SCL and members of the pentameric complex. In addition, ChIP for SCL co-activators (e.g. p300, CBP), SCL co-repressors (e.g. ETO2, HACDs) and specific histone marks (e.g. H3K4Me as an active mark and H3K27Me / H3K9Me as repressive marks) will be performed to give an indication of the active or repressive nature of transcriptional control.

7.9.5 Action of SCL through DNA-Independent Mechanisms

It has been previously shown that SCL functions as a repressor off DNA by binding and sequestering other bHLH proteins (usually E-proteins) in T-ALL and myocyte development [202-204]. Therefore determination of SCL protein binding partners during mesoderm specification may give important insights into the molecular mechanisms involving SCL. We are currently using biochemical techniques, such as co-immuno precipitation (co-IP) and affinity purification, to pull-down SCL containing protein complexes in day 3 to 4 EBs, followed by mass spectrometry to identify SCL putative protein partners.

The ability of SCL to interact with other proteins is essential to its function [173]. For example, haematopoietic defects in *Scl* morphants cannot be rescued by mutant forms of SCL that cannot bind LMO2 [186]. *In vitro* differentiation of ES cells harbouring such mutations would determine whether specific protein-protein interactions required for SCL-driven haematopoietic specification are also required for SCL-mediated cardiac repression.

The absence of cardiac lineage repression in *Scl* $^{RER/RER}$ EBs, where SCL is unable to directly bind DNA, suggests that this repression occurs through a direct SCL-DNA binding mechanism. However, it is possible that the RER to AAA mutation introduced into the DNA-binding domain of SCL not only inhibits SCL-DNA binding, but may also incur an additional effect on the ability of SCL to bind certain protein partners. Performing additional co-IP and mass spectrometry using material from *Scl* $^{RER/RER}$ EBs would indicate whether protein-protein interactions are affected.

7.9.6 Verification in the Mouse Embryo

Due to the limitations of the ES/EB system (see Sections 1.5.2 and 7.1.1) it is important to confirm our finding *in vivo* in the mouse embryo. This will be done using embryos from wild type, *Scl* and *Scl* $^{RER/RER}$ mouse lines.

Using flow cytometric analysis techniques, we will determine the existence of a SCL$^+$ Flk-1$^+$ PDGFRα^+ cell population, either in the intra-embryonic primitive streak or extra-embryonic yolk sac mesoderm in wild type embryos. Furthermore, we will determine if there is an expansion of PDGFRα^+ populations, and/or GATA4$^+$ populations in *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$

embryos, both in the yolk sac and early intra-embryonic mesoderm regions. Aberrant expression of functional cardiac markers, such as cTnT, will be looked for in the yolk sac and intra-embryonic regions to determine if they exhibit abnormal cardiac potential in Scl $^{-/-}$ and *Scl* $^{RER/RER}$ embryos. Yolk sac and intra-embryonic cells will be cultured *ex vivo* under cardiac inducing conditions to determine if they exhibit increased cardiac potential compared to cells derived from wild type animals.

In addition, we will look at heart field defects in *Scl* $^{-/-}$ and *Scl* $^{RER/RER}$ embryos to see if the cardiac territory is expanded in the absence of SCL (to be determined using cTnT expression as an indicator of functional cardiomyocyte). However, underdeveloped vascular endothelial and endocardial structures in *Scl* $^{-/-}$ embryos causes cardiac defects as the lumen of the heart tube does not form properly. Consequently, looping is not initiated and the heart tube remains linear at the point of embryonic death [64]. In this instance, *Scl* $^{RER/RER}$ embryos provide a unique advantage as they survive longer than *Scl* $^{-/-}$ embryos, so that heart development can be studied at later stages.

References

1. D'Souza, S.L., A.G. Elefanty, and G. Keller, *SCL/Tal-1 is essential for hematopoietic commitment of the hemangioblast but not for its development.* Blood, 2005. **105**(10): p. 3862-70.
2. Porcher, C., et al., *The T cell leukemia oncoprotein SCL/tal-1 is essential for development of all hematopoietic lineages.* Cell, 1996. **86**(1): p. 47-57.
3. Robertson, S.M., et al., *A transitional stage in the commitment of mesoderm to hematopoiesis requiring the transcription factor SCL/tal-1.* Development, 2000. **127**(11): p. 2447-59.
4. Shivdasani, R.A., E.L. Mayer, and S.H. Orkin, *Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL.* Nature, 1995. **373**(6513): p. 432-4.
5. Robb, L., et al., *Absence of yolk sac hematopoiesis from mice with a targeted disruption of the scl gene.* Proceedings of the National Academy of Sciences of the United States of America, 1995. **92**(15): p. 7075-9.
6. Robb, L., et al., *The scl gene product is required for the generation of all hematopoietic lineages in the adult mouse.* The EMBO journal, 1996. **15**(16): p. 4123-9.
7. Baron, M.H., *Early patterning of the mouse embryo: implications for hematopoietic commitment and differentiation.* Experimental hematology, 2005. **33**(9): p. 1015-20.
8. Gardner, R.L., *Cell lineage and cell commitment in the early mammalian embryo.* Mead Johnson Symposium on Perinatal and Developmental Medicine, 1979(16): p. 18-24.
9. Tam, P.P. and R.R. Behringer, *Mouse gastrulation: the formation of a mammalian body plan.* Mech Dev, 1997. **68**(1-2): p. 3-25.
10. Lawson, K.A. and R.A. Pedersen, *Clonal analysis of cell fate during gastrulation and early neurulation in the mouse.* Ciba Foundation symposium, 1992. **165**: p. 3-21; discussion 21-6.
11. Kinder, S.J., et al., *The orderly allocation of mesodermal cells to the extraembryonic structures and the anteroposterior axis during gastrulation of the mouse embryo.* Development, 1999. **126**(21): p. 4691-701.
12. Huber, T.L., et al., *Haemangioblast commitment is initiated in the primitive streak of the mouse embryo.* Nature, 2004. **432**(7017): p. 625-30.
13. Lawson, K.A., J.J. Meneses, and R.A. Pedersen, *Clonal analysis of epiblast fate during germ layer formation in the mouse embryo.* Development, 1991. **113**(3): p. 891-911.
14. Kinder, S.J., D.A. Loebel, and P.P. Tam, *Allocation and early differentiation of cardiovascular progenitors in the mouse embryo.* Trends in cardiovascular medicine, 2001. **11**(5): p. 177-84.
15. Tam, P.P. and P.A. Trainor, *Specification and segmentation of the paraxial mesoderm.* Anatomy and embryology, 1994. **189**(4): p. 275-305.
16. Nostro, M.C., et al., *Wnt, activin, and BMP signaling regulate distinct stages in the developmental pathway from embryonic stem cells to blood.* Cell stem cell, 2008. **2**(1): p. 60-71.
17. Sumi, T., et al., *Defining early lineage specification of human embryonic stem cells by the orchestrated balance of canonical Wnt/beta-catenin, Activin/Nodal and BMP signaling.* Development, 2008. **135**(17): p. 2969-79.

18. Ueno, H. and I.L. Weissman, *Clonal analysis of mouse development reveals a polyclonal origin for yolk sac blood islands.* Developmental cell, 2006. **11**(4): p. 519-33.
19. Kattman, S.J., T.L. Huber, and G.M. Keller, *Multipotent flk-1+ cardiovascular progenitor cells give rise to the cardiomyocyte, endothelial, and vascular smooth muscle lineages.* Developmental cell, 2006. **11**(5): p. 723-32.
20. Ema, M., S. Takahashi, and J. Rossant, *Deletion of the selection cassette, but not cis-acting elements, in targeted Flk1-lacZ allele reveals Flk1 expression in multipotent mesodermal progenitors.* Blood, 2006. **107**(1): p. 111-7.
21. Takakura, N., et al., *PDGFR alpha expression during mouse embryogenesis: immunolocalization analyzed by whole-mount immunohistostaining using the monoclonal anti-mouse PDGFR alpha antibody APA5.* The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society, 1997. **45**(6): p. 883-93.
22. Kataoka, H., et al., *Expressions of PDGF receptor alpha, c-Kit and Flk1 genes clustering in mouse chromosome 5 define distinct subsets of nascent mesodermal cells.* Development, growth & differentiation, 1997. **39**(6): p. 729-40.
23. Ciruna, B.G. and J. Rossant, *Expression of the T-box gene Eomesodermin during early mouse development.* Mech Dev, 1999. **81**(1-2): p. 199-203.
24. Haar, J.L. and G.A. Ackerman, *A phase and electron microscopic study of vasculogenesis and erythropoiesis in the yolk sac of the mouse.* The Anatomical record, 1971. **170**(2): p. 199-223.
25. Moore, M.A. and D. Metcalf, *Ontogeny of the haemopoietic system: yolk sac origin of in vivo and in vitro colony forming cells in the developing mouse embryo.* British journal of haematology, 1970. **18**(3): p. 279-96.
26. Yoder, M.C., et al., *Murine yolk sac endoderm- and mesoderm-derived cell lines support in vitro growth and differentiation of hematopoietic cells.* Blood, 1994. **83**(9): p. 2436-43.
27. Ferkowicz, M.J. and M.C. Yoder, *Blood island formation: longstanding observations and modern interpretations.* Experimental hematology, 2005. **33**(9): p. 1041-7.
28. Ferkowicz, M.J., et al., *CD41 expression defines the onset of primitive and definitive hematopoiesis in the murine embryo.* Development, 2003. **130**(18): p. 4393-403.
29. Palis, J., et al., *Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse.* Development, 1999. **126**(22): p. 5073-84.
30. Silver, L. and J. Palis, *Initiation of murine embryonic erythropoiesis: a spatial analysis.* Blood, 1997. **89**(4): p. 1154-64.
31. Kingsley, P.D., et al., *Yolk sac-derived primitive erythroblasts enucleate during mammalian embryogenesis.* Blood, 2004. **104**(1): p. 19-25.
32. Fraser, S.T., J. Isern, and M.H. Baron, *Maturation and enucleation of primitive erythroblasts during mouse embryogenesis is accompanied by changes in cell-surface antigen expression.* Blood, 2007. **109**(1): p. 343-52.
33. McGrath, K.E., et al., *Circulation is established in a stepwise pattern in the mammalian embryo.* Blood, 2003. **101**(5): p. 1669-76.
34. Yoder, M.C., et al., *Characterization of definitive lymphohematopoietic stem cells in the day 9 murine yolk sac.* Immunity, 1997. **7**(3): p. 335-44.
35. Fraser, S.T., et al., *Definitive hematopoietic commitment within the embryonic vascular endothelial-cadherin(+) population.* Experimental hematology, 2002. **30**(9): p. 1070-8.

36. Goldie, L.C., et al., *Cell signaling directing the formation and function of hemogenic endothelium during murine embryogenesis.* Blood, 2008. **112**(8): p. 3194-204.
37. Nadin, B.M., M.A. Goodell, and K.K. Hirschi, *Phenotype and hematopoietic potential of side population cells throughout embryonic development.* Blood, 2003. **102**(7): p. 2436-43.
38. Cumano, A., F. Dieterlen-Lievre, and I. Godin, *Lymphoid potential, probed before circulation in mouse, is restricted to caudal intraembryonic splanchnopleura.* Cell, 1996. **86**(6): p. 907-16.
39. Cumano, A., et al., *Intraembryonic, but not yolk sac hematopoietic precursors, isolated before circulation, provide long-term multilineage reconstitution.* Immunity, 2001. **15**(3): p. 477-85.
40. Rhodes, B. and P. Anderlini, *Allogeneic peripheral blood stem cell collection as of 2008.* Transfus Apher Sci, 2008. **38**(3): p. 219-27.
41. Matsuoka, S., et al., *Generation of definitive hematopoietic stem cells from murine early yolk sac and paraaortic splanchnopleures by aorta-gonad-mesonephros region-derived stromal cells.* Blood, 2001. **98**(1): p. 6-12.
42. Samokhvalov, I.M., N.I. Samokhvalova, and S. Nishikawa, *Cell tracing shows the contribution of the yolk sac to adult haematopoiesis.* Nature, 2007. **446**(7139): p. 1056-61.
43. Esner, M., et al., *Smooth muscle of the dorsal aorta shares a common clonal origin with skeletal muscle of the myotome.* Development, 2006. **133**(4): p. 737-49.
44. Wasteson, P., et al., *Developmental origin of smooth muscle cells in the descending aorta in mice.* Development, 2008. **135**(10): p. 1823-32.
45. Medvinsky, A. and E. Dzierzak, *Definitive hematopoiesis is autonomously initiated by the AGM region.* Cell, 1996. **86**(6): p. 897-906.
46. de Bruijn, M.F., et al., *Hematopoietic stem cells localize to the endothelial cell layer in the midgestation mouse aorta.* Immunity, 2002. **16**(5): p. 673-83.
47. de Bruijn, M.F., et al., *Definitive hematopoietic stem cells first develop within the major arterial regions of the mouse embryo.* The EMBO journal, 2000. **19**(11): p. 2465-74.
48. Yokomizo, T., et al., *Three-dimensional imaging of whole midgestation murine embryos shows an intravascular localization for all hematopoietic clusters.* Blood, 2011. **117**(23): p. 6132-4.
49. Boisset, J.C., et al., *In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium.* Nature, 2010. **464**(7285): p. 116-20.
50. Cai, Z., et al., *Haploinsufficiency of AML1 affects the temporal and spatial generation of hematopoietic stem cells in the mouse embryo.* Immunity, 2000. **13**(4): p. 423-31.
51. Chen, M.J., et al., *Erythroid/myeloid progenitors and hematopoietic stem cells originate from distinct populations of endothelial cells.* Cell stem cell, 2011. **9**(6): p. 541-52.
52. North, T., et al., *Cbfa2 is required for the formation of intra-aortic hematopoietic clusters.* Development, 1999. **126**(11): p. 2563-75.
53. Davidson, A.J. and L.I. Zon, *The 'definitive' (and 'primitive') guide to zebrafish hematopoiesis.* Oncogene, 2004. **23**(43): p. 7233-46.
54. Warga, R.M., D.A. Kane, and R.K. Ho, *Fate mapping embryonic blood in zebrafish: multi- and unipotential lineages are segregated at gastrulation.* Developmental cell, 2009. **16**(5): p. 744-55.

55. Bertrand, J.Y., et al., *Definitive hematopoiesis initiates through a committed erythromyeloid progenitor in the zebrafish embryo.* Development, 2007. **134**(23): p. 4147-56.
56. Galloway, J.L. and L.I. Zon, *Ontogeny of hematopoiesis: examining the emergence of hematopoietic cells in the vertebrate embryo.* Current topics in developmental biology, 2003. **53**: p. 139-58.
57. Vogeli, K.M., et al., *A common progenitor for haematopoietic and endothelial lineages in the zebrafish gastrula.* Nature, 2006. **443**(7109): p. 337-9.
58. Laugwitz, K.L., et al., *Islet1 cardiovascular progenitors: a single source for heart lineages?* Development, 2008. **135**(2): p. 193-205.
59. Lam, E.Y., et al., *Live imaging of Runx1 expression in the dorsal aorta tracks the emergence of blood progenitors from endothelial cells.* Blood, 2010. **116**(6): p. 909-14.
60. Kissa, K. and P. Herbomel, *Blood stem cells emerge from aortic endothelium by a novel type of cell transition.* Nature, 2010. **464**(7285): p. 112-5.
61. Bertrand, J.Y., et al., *Haematopoietic stem cells derive directly from aortic endothelium during development.* Nature, 2010. **464**(7285): p. 108-11.
62. North, T.E., et al., *Runx1 expression marks long-term repopulating hematopoietic stem cells in the midgestation mouse embryo.* Immunity, 2002. **16**(5): p. 661-72.
63. Mikkola, H.K., et al., *Expression of CD41 marks the initiation of definitive hematopoiesis in the mouse embryo.* Blood, 2003. **101**(2): p. 508-16.
64. Elefanty, A.G., et al., *SCL expression in the mouse embryo detected with a targeted lacZ reporter gene demonstrates its localization to hematopoietic, vascular, and neural tissues.* Blood, 1999. **94**(11): p. 3754-63.
65. Kallianpur, A.R., J.E. Jordan, and S.J. Brandt, *The SCL/TAL-1 gene is expressed in progenitors of both the hematopoietic and vascular systems during embryogenesis.* Blood, 1994. **83**(5): p. 1200-8.
66. Chagraoui, H., et al., *SCL-mediated regulation of the cell-cycle regulator p21 is critical for murine megakaryopoiesis.* Blood, 2011. **118**(3): p. 723-35.
67. McCormack, M.P., et al., *A critical role for the transcription factor Scl in platelet production during stress thrombopoiesis.* Blood, 2006. **108**(7): p. 2248-56.
68. Hall, M.A., et al., *The critical regulator of embryonic hematopoiesis, SCL, is vital in the adult for megakaryopoiesis, erythropoiesis, and lineage choice in CFU-S12.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(3): p. 992-7.
69. Mikkola, H.K., et al., *Haematopoietic stem cells retain long-term repopulating activity and multipotency in the absence of stem-cell leukaemia SCL/tal-1 gene.* Nature, 2003. **421**(6922): p. 547-51.
70. Ciau-Uitz, A., F. Liu, and R. Patient, *Genetic control of hematopoietic development in Xenopus and zebrafish.* The International journal of developmental biology, 2010. **54**(6-7): p. 1139-49.
71. Gering, M., et al., *The SCL gene specifies haemangioblast development from early mesoderm.* The EMBO journal, 1998. **17**(14): p. 4029-45.
72. Lacaud, G., et al., *Runx1 is essential for hematopoietic commitment at the hemangioblast stage of development in vitro.* Blood, 2002. **100**(2): p. 458-66.
73. Wang, Q., et al., *Disruption of the Cbfa2 gene causes necrosis and hemorrhaging in the central nervous system and blocks definitive*

hematopoiesis. Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(8): p. 3444-9.
74. Yokomizo, T., et al., *Runx1 is involved in primitive erythropoiesis in the mouse.* Blood, 2008. **111**(8): p. 4075-80.
75. Okuda, T., et al., *AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis.* Cell, 1996. **84**(2): p. 321-30.
76. Swiers, G., M. de Bruijn, and N.A. Speck, *Hematopoietic stem cell emergence in the conceptus and the role of Runx1.* The International journal of developmental biology, 2010. **54**(6-7): p. 1151-63.
77. Lancrin, C., et al., *The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage.* Nature, 2009. **457**(7231): p. 892-5.
78. Tsai, F.Y. and S.H. Orkin, *Transcription factor GATA-2 is required for proliferation/survival of early hematopoietic cells and mast cell formation, but not for erythroid and myeloid terminal differentiation.* Blood, 1997. **89**(10): p. 3636-43.
79. Ling, K.W., et al., *GATA-2 plays two functionally distinct roles during the ontogeny of hematopoietic stem cells.* The Journal of experimental medicine, 2004. **200**(7): p. 871-82.
80. Liu, F., et al., *Fli1 acts at the top of the transcriptional network driving blood and endothelial development.* Current biology : CB, 2008. **18**(16): p. 1234-40.
81. Vicente, C., et al., *The role of the GATA2 transcription factor in normal and malignant hematopoiesis.* Critical reviews in oncology/hematology, 2011.
82. Giroux, S., et al., *lyl-1 and tal-1/scl, two genes encoding closely related bHLH transcription factors, display highly overlapping expression patterns during cardiovascular and hematopoietic ontogeny.* Gene expression patterns : GEP, 2007. **7**(3): p. 215-26.
83. Capron, C., et al., *The SCL relative LYL-1 is required for fetal and adult hematopoietic stem cell function and B-cell differentiation.* Blood, 2006. **107**(12): p. 4678-86.
84. Cai, C.L., et al., *Isl1 identifies a cardiac progenitor population that proliferates prior to differentiation and contributes a majority of cells to the heart.* Developmental cell, 2003. **5**(6): p. 877-89.
85. Buckingham, M., S. Meilhac, and S. Zaffran, *Building the mammalian heart from two sources of myocardial cells.* Nature reviews. Genetics, 2005. **6**(11): p. 826-35.
86. Snarr, B.S., C.B. Kern, and A. Wessels, *Origin and fate of cardiac mesenchyme.* Developmental dynamics : an official publication of the American Association of Anatomists, 2008. **237**(10): p. 2804-19.
87. Srivastava, D., *Making or breaking the heart: from lineage determination to morphogenesis.* Cell, 2006. **126**(6): p. 1037-48.
88. Lough, J. and Y. Sugi, *Endoderm and heart development.* Developmental dynamics : an official publication of the American Association of Anatomists, 2000. **217**(4): p. 327-42.
89. Klaus, A., et al., *Distinct roles of Wnt/beta-catenin and Bmp signaling during early cardiogenesis.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(47): p. 18531-6.
90. Zaffran, S., et al., *Right ventricular myocardium derives from the anterior heart field.* Circulation research, 2004. **95**(3): p. 261-8.

91. Hami, D., et al., *Zebrafish cardiac development requires a conserved secondary heart field.* Development, 2011. **138**(11): p. 2389-98.
92. Lazic, S. and I.C. Scott, *Mef2cb regulates late myocardial cell addition from a second heart field-like population of progenitors in zebrafish.* Developmental biology, 2011. **354**(1): p. 123-33.
93. Zhou, Y., et al., *Latent TGF-beta binding protein 3 identifies a second heart field in zebrafish.* Nature, 2011. **474**(7353): p. 645-8.
94. Moretti, A., et al., *Multipotent embryonic isl1+ progenitor cells lead to cardiac, smooth muscle, and endothelial cell diversification.* Cell, 2006. **127**(6): p. 1151-65.
95. Prall, O.W., et al., *An Nkx2-5/Bmp2/Smad1 negative feedback loop controls heart progenitor specification and proliferation.* Cell, 2007. **128**(5): p. 947-59.
96. Bondue, A. and C. Blanpain, *Mesp1: a key regulator of cardiovascular lineage commitment.* Circulation research, 2010. **107**(12): p. 1414-27.
97. Bondue, A., et al., *Defining the earliest step of cardiovascular progenitor specification during embryonic stem cell differentiation.* The Journal of cell biology, 2011. **192**(5): p. 751-65.
98. Saga, Y., et al., *MesP1 is expressed in the heart precursor cells and required for the formation of a single heart tube.* Development, 1999. **126**(15): p. 3437-47.
99. Maitra, M., et al., *Interaction of Gata4 and Gata6 with Tbx5 is critical for normal cardiac development.* Developmental biology, 2009. **326**(2): p. 368-77.
100. Sachinidis, A., et al., *Cardiac specific differentiation of mouse embryonic stem cells.* Cardiovascular research, 2003. **58**(2): p. 278-91.
101. von Both, I., et al., *Foxh1 is essential for development of the anterior heart field.* Developmental cell, 2004. **7**(3): p. 331-45.
102. Wang, Q., et al., *Comparative studies on the expression patterns of three troponin T genes during mouse development.* The Anatomical record, 2001. **263**(1): p. 72-84.
103. Rawls, A., J. Wilson-Rawls, and E.N. Olson, *Genetic regulation of somite formation.* Current topics in developmental biology, 2000. **47**: p. 131-54.
104. Akiyama, H., et al., *Interactions between Sox9 and beta-catenin control chondrocyte differentiation.* Genes & development, 2004. **18**(9): p. 1072-87.
105. Yang, L., et al., *Proper expression of helix-loop-helix protein Id2 is important to chondrogenic differentiation of ATDC5 cells.* The Biochemical journal, 2009. **419**(3): p. 635-43.
106. Kawaguchi, J., P.J. Mee, and A.G. Smith, *Osteogenic and chondrogenic differentiation of embryonic stem cells in response to specific growth factors.* Bone, 2005. **36**(5): p. 758-69.
107. Yano, F., et al., *The canonical Wnt signaling pathway promotes chondrocyte differentiation in a Sox9-dependent manner.* Biochemical and biophysical research communications, 2005. **333**(4): p. 1300-8.
108. Zhao, Q., et al., *Parallel expression of Sox9 and Col2a1 in cells undergoing chondrogenesis.* Developmental dynamics : an official publication of the American Association of Anatomists, 1997. **209**(4): p. 377-86.
109. Noda, M., *Transcriptional regulation of osteocalcin production by transforming growth factor-beta in rat osteoblast-like cells.* Endocrinology, 1989. **124**(2): p. 612-7.

110. Wright, W.E., D.A. Sassoon, and V.K. Lin, *Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD.* Cell, 1989. **56**(4): p. 607-17.
111. Sassoon, D., et al., *Expression of two myogenic regulatory factors myogenin and MyoD1 during mouse embryogenesis.* Nature, 1989. **341**(6240): p. 303-7.
112. Smith, A.G., et al., *Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides.* Nature, 1988. **336**(6200): p. 688-90.
113. Izumi, N., et al., *Dissecting the molecular hierarchy for mesendoderm differentiation through a combination of embryonic stem cell culture and RNA interference.* Stem cells, 2007. **25**(7): p. 1664-74.
114. Bain, G., et al., *Embryonic stem cells express neuronal properties in vitro.* Developmental biology, 1995. **168**(2): p. 342-57.
115. Keller, G., et al., *Hematopoietic commitment during embryonic stem cell differentiation in culture.* Molecular and cellular biology, 1993. **13**(1): p. 473-86.
116. Faloon, P., et al., *Basic fibroblast growth factor positively regulates hematopoietic development.* Development, 2000. **127**(9): p. 1931-41.
117. Fehling, H.J., et al., *Tracking mesoderm induction and its specification to the hemangioblast during embryonic stem cell differentiation.* Development, 2003. **130**(17): p. 4217-27.
118. Choi, K., *Hemangioblast development and regulation.* Biochemistry and cell biology = Biochimie et biologie cellulaire, 1998. **76**(6): p. 947-56.
119. Choi, K., et al., *A common precursor for hematopoietic and endothelial cells.* Development, 1998. **125**(4): p. 725-32.
120. Hirashima, M., et al., *Maturation of embryonic stem cells into endothelial cells in an in vitro model of vasculogenesis.* Blood, 1999. **93**(4): p. 1253-63.
121. Vittet, D., et al., *Embryonic stem cells differentiate in vitro to endothelial cells through successive maturation steps.* Blood, 1996. **88**(9): p. 3424-31.
122. Kouskoff, V., et al., *Sequential development of hematopoietic and cardiac mesoderm during embryonic stem cell differentiation.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(37): p. 13170-5.
123. Puceat, M., *Protocols for cardiac differentiation of embryonic stem cells.* Methods, 2008. **45**(2): p. 168-71.
124. Rohwedel, J., et al., *Muscle cell differentiation of embryonic stem cells reflects myogenesis in vivo: developmentally regulated expression of myogenic determination genes and functional expression of ionic currents.* Developmental biology, 1994. **164**(1): p. 87-101.
125. Sakurai, H., et al., *In vitro modeling of paraxial and lateral mesoderm differentiation reveals early reversibility.* Stem cells, 2006. **24**(3): p. 575-86.
126. Naujok, O., et al., *Changes in gene expression and morphology of mouse embryonic stem cells on differentiation into insulin-producing cells in vitro and in vivo.* Diabetes/metabolism research and reviews, 2009. **25**(5): p. 464-76.
127. Schroeder, I.S., et al., *Differentiation analysis of pluripotent mouse embryonic stem (ES) cells in vitro.* Methods in molecular biology, 2009. **530**: p. 219-50.
128. Lee, S.H., et al., *Efficient generation of midbrain and hindbrain neurons from mouse embryonic stem cells.* Nature biotechnology, 2000. **18**(6): p. 675-9.

129. Irion, S., et al., *Temporal specification of blood progenitors from mouse embryonic stem cells and induced pluripotent stem cells.* Development, 2010. **137**(17): p. 2829-39.
130. Lancrin, C., et al., *Blood cell generation from the hemangioblast.* Journal of molecular medicine, 2010. **88**(2): p. 167-72.
131. Chambers, I. and S.R. Tomlinson, *The transcriptional foundation of pluripotency.* Development, 2009. **136**(14): p. 2311-22.
132. Silva, J., et al., *Promotion of reprogramming to ground state pluripotency by signal inhibition.* PLoS biology, 2008. **6**(10): p. e253.
133. Kuo, C.T., et al., *GATA4 transcription factor is required for ventral morphogenesis and heart tube formation.* Genes & development, 1997. **11**(8): p. 1048-60.
134. Soudais, C., et al., *Targeted mutagenesis of the transcription factor GATA-4 gene in mouse embryonic stem cells disrupts visceral endoderm differentiation in vitro.* Development, 1995. **121**(11): p. 3877-88.
135. Molkentin, J.D., *The zinc finger-containing transcription factors GATA-4, -5, and -6. Ubiquitously expressed regulators of tissue-specific gene expression.* The Journal of biological chemistry, 2000. **275**(50): p. 38949-52.
136. Narita, N., M. Bielinska, and D.B. Wilson, *Cardiomyocyte differentiation by GATA-4-deficient embryonic stem cells.* Development, 1997. **124**(19): p. 3755-64.
137. Kattman, S.J., et al., *Stage-specific optimization of activin/nodal and BMP signaling promotes cardiac differentiation of mouse and human pluripotent stem cell lines.* Cell stem cell, 2011. **8**(2): p. 228-40.
138. Hidaka, M., W.L. Stanford, and A. Bernstein, *Conditional requirement for the Flk-1 receptor in the in vitro generation of early hematopoietic cells.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(13): p. 7370-5.
139. Eilken, H.M., S. Nishikawa, and T. Schroeder, *Continuous single-cell imaging of blood generation from haemogenic endothelium.* Nature, 2009. **457**(7231): p. 896-900.
140. Kyba, M., R.C. Perlingeiro, and G.Q. Daley, *HoxB4 confers definitive lymphoid-myeloid engraftment potential on embryonic stem cell and yolk sac hematopoietic progenitors.* Cell, 2002. **109**(1): p. 29-37.
141. Shiraki, N., et al., *Differentiation and characterization of embryonic stem cells into three germ layers.* Biochemical and biophysical research communications, 2009. **381**(4): p. 694-9.
142. Gadue, P., et al., *Wnt and TGF-beta signaling are required for the induction of an in vitro model of primitive streak formation using embryonic stem cells.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(45): p. 16806-11.
143. Dyer, M.A., et al., *Indian hedgehog activates hematopoiesis and vasculogenesis and can respecify prospective neurectodermal cell fate in the mouse embryo.* Development, 2001. **128**(10): p. 1717-30.
144. Monteiro, R., C. Pouget, and R. Patient, *The gata1/pu.1 lineage fate paradigm varies between blood populations and is modulated by tif1gamma.* The EMBO journal, 2011. **30**(6): p. 1093-103.
145. Graf, T. and T. Enver, *Forcing cells to change lineages.* Nature, 2009. **462**(7273): p. 587-94.
146. Takahashi, K. and S. Yamanaka, *Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors.* Cell, 2006. **126**(4): p. 663-76.

147. Yang, L., *From fibroblast cells to cardiomyocytes: direct lineage reprogramming.* Stem cell research & therapy, 2011. **2**(1): p. 1.
148. Simoes, F.C., T. Peterkin, and R. Patient, *Fgf differentially controls cross-antagonism between cardiac and haemangioblast regulators.* Development, 2011. **138**(15): p. 3235-45.
149. Molina, G., et al., *Zebrafish chemical screening reveals an inhibitor of Dusp6 that expands cardiac cell lineages.* Nature chemical biology, 2009. **5**(9): p. 680-7.
150. Schoenebeck, J.J., B.R. Keegan, and D. Yelon, *Vessel and blood specification override cardiac potential in anterior mesoderm.* Developmental cell, 2007. **13**(2): p. 254-67.
151. Gering, M., et al., *Lmo2 and Scl/Tal1 convert non-axial mesoderm into haemangioblasts which differentiate into endothelial cells in the absence of Gata1.* Development, 2003. **130**(25): p. 6187-99.
152. Schneider, V.A. and M. Mercola, *Wnt antagonism initiates cardiogenesis in Xenopus laevis.* Genes & development, 2001. **15**(3): p. 304-15.
153. Chen, V.C., et al., *Notch signaling respecifies the hemangioblast to a cardiac fate.* Nature biotechnology, 2008. **26**(10): p. 1169-78.
154. Lecuyer, E. and T. Hoang, *SCL: from the origin of hematopoiesis to stem cells and leukemia.* Experimental hematology, 2004. **32**(1): p. 11-24.
155. Murre, C., et al., *Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence.* Cell, 1989. **58**(3): p. 537-44.
156. Bernard, O., et al., *Two site-specific deletions and t(1;14) translocation restricted to human T-cell acute leukemias disrupt the 5' part of the tal-1 gene.* Oncogene, 1991. **6**(8): p. 1477-88.
157. Bernard, O., et al., *A third tal-1 promoter is specifically used in human T cell leukemias.* The Journal of experimental medicine, 1992. **176**(4): p. 919-25.
158. Calkhoven, C.F., et al., *Translational control of SCL-isoform expression in hematopoietic lineage choice.* Genes & development, 2003. **17**(8): p. 959-64.
159. Hsu, H.L., et al., *Enhancer-binding activity of the tal-1 oncoprotein in association with the E47/E12 helix-loop-helix proteins.* Molecular and cellular biology, 1991. **11**(6): p. 3037-42.
160. Massari, M.E. and C. Murre, *Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms.* Molecular and cellular biology, 2000. **20**(2): p. 429-40.
161. Hsu, H.L., et al., *Preferred sequences for DNA recognition by the TAL1 helix-loop-helix proteins.* Molecular and cellular biology, 1994. **14**(2): p. 1256-65.
162. Palii, C.G., et al., *Differential genomic targeting of the transcription factor TAL1 in alternate haematopoietic lineages.* The EMBO journal, 2011. **30**(3): p. 494-509.
163. Wadman, I.A., et al., *The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins.* The EMBO journal, 1997. **16**(11): p. 3145-57.
164. Kassouf, M.T., et al., *Genome-wide identification of TAL1's functional targets: insights into its mechanisms of action in primary erythroid cells.* Genome research, 2010. **20**(8): p. 1064-83.
165. Anderson, K.P., S.C. Crable, and J.B. Lingrel, *Multiple proteins binding to a GATA-E box-GATA motif regulate the erythroid Kruppel-like factor (EKLF) gene.* The Journal of biological chemistry, 1998. **273**(23): p. 14347-54.

166. Vyas, P., et al., *Different sequence requirements for expression in erythroid and megakaryocytic cells within a regulatory element upstream of the GATA-1 gene.* Development, 1999. **126**(12): p. 2799-811.
167. Schuh, A.H., et al., *ETO-2 associates with SCL in erythroid cells and megakaryocytes and provides repressor functions in erythropoiesis.* Molecular and cellular biology, 2005. **25**(23): p. 10235-50.
168. Pulford, K., et al., *Expression of TAL-1 proteins in human tissues.* Blood, 1995. **85**(3): p. 675-84.
169. Pimanda, J.E., et al., *Transcriptional link between blood and bone: the stem cell leukemia gene and its +19 stem cell enhancer are active in bone cells.* Molecular and cellular biology, 2006. **26**(7): p. 2615-25.
170. Yamane, T., et al., *Sequential requirements for SCL/tal-1, GATA-2, macrophage colony-stimulating factor, and osteoclast differentiation factor/osteoprotegerin ligand in osteoclast development.* Experimental hematology, 2000. **28**(7): p. 833-40.
171. Chung, Y.S., et al., *Lineage analysis of the hemangioblast as defined by FLK1 and SCL expression.* Development, 2002. **129**(23): p. 5511-20.
172. Ema, M., et al., *Combinatorial effects of Flk1 and Tal1 on vascular and hematopoietic development in the mouse.* Genes & development, 2003. **17**(3): p. 380-93.
173. Porcher, C., et al., *Specification of hematopoietic and vascular development by the bHLH transcription factor SCL without direct DNA binding.* Development, 1999. **126**(20): p. 4603-15.
174. Curtis, D.J., et al., *SCL is required for normal function of short-term repopulating hematopoietic stem cells.* Blood, 2004. **103**(9): p. 3342-8.
175. Patterson, L.J., M. Gering, and R. Patient, *Scl is required for dorsal aorta as well as blood formation in zebrafish embryos.* Blood, 2005. **105**(9): p. 3502-11.
176. Endoh, M., et al., *SCL/tal-1-dependent process determines a competence to select the definitive hematopoietic lineage prior to endothelial differentiation.* The EMBO journal, 2002. **21**(24): p. 6700-8.
177. Schlaeger, T.M., et al., *Tie2Cre-mediated gene ablation defines the stem-cell leukemia gene (SCL/tal1)-dependent window during hematopoietic stem-cell development.* Blood, 2005. **105**(10): p. 3871-4.
178. Lacombe, J., et al., *Scl regulates the quiescence and the long-term competence of hematopoietic stem cells.* Blood, 2010. **115**(4): p. 792-803.
179. Valtieri, M., et al., *Enforced TAL-1 expression stimulates primitive, erythroid and megakaryocytic progenitors but blocks the granulopoietic differentiation program.* Cancer research, 1998. **58**(3): p. 562-9.
180. Tijssen, M.R., et al., *Genome-wide analysis of simultaneous GATA1/2, RUNX1, FLI1, and SCL binding in megakaryocytes identifies hematopoietic regulators.* Developmental cell, 2011. **20**(5): p. 597-609.
181. Salmon, J.M., et al., *Aberrant mast-cell differentiation in mice lacking the stem-cell leukemia gene.* Blood, 2007. **110**(10): p. 3573-81.
182. Bussmann, J., J. Bakkers, and S. Schulte-Merker, *Early endocardial morphogenesis requires Scl/Tal1.* PLoS Genet, 2007. **3**(8): p. e140.
183. Liao, E.C., et al., *SCL/Tal-1 transcription factor acts downstream of cloche to specify hematopoietic and vascular progenitors in zebrafish.* Genes & development, 1998. **12**(5): p. 621-6.
184. Mead, P.E., et al., *SCL specifies hematopoietic mesoderm in Xenopus embryos.* Development, 1998. **125**(14): p. 2611-20.

185. Ismailoglu, I., et al., *Mesodermal patterning activity of SCL.* Experimental hematology, 2008. **36**(12): p. 1593-603.
186. Patterson, L.J., et al., *The transcription factors Scl and Lmo2 act together during development of the hemangioblast in zebrafish.* Blood, 2007. **109**(6): p. 2389-98.
187. Kassouf, M.T., et al., *Differential use of SCL/TAL-1 DNA-binding domain in developmental hematopoiesis.* Blood, 2008. **112**(4): p. 1056-67.
188. Visvader, J.E., Y. Fujiwara, and S.H. Orkin, *Unsuspected role for the T-cell leukemia protein SCL/tal-1 in vascular development.* Genes & development, 1998. **12**(4): p. 473-9.
189. Hayashi, S., et al., *Commitment and differentiation of stem cells to the osteoclast lineage.* Biochemistry and cell biology = Biochimie et biologie cellulaire, 1998. **76**(6): p. 911-22.
190. Muroyama, Y., et al., *Specification of astrocytes by bHLH protein SCL in a restricted region of the neural tube.* Nature, 2005. **438**(7066): p. 360-3.
191. Green, A.R., et al., *SCL is coexpressed with GATA-1 in hemopoietic cells but is also expressed in developing brain.* Oncogene, 1992. **7**(4): p. 653-60.
192. Bradley, C.K., et al., *The essential haematopoietic transcription factor Scl is also critical for neuronal development.* The European journal of neuroscience, 2006. **23**(7): p. 1677-89.
193. Lahlil, R., et al., *SCL assembles a multifactorial complex that determines glycophorin A expression.* Molecular and cellular biology, 2004. **24**(4): p. 1439-52.
194. Vitelli, L., et al., *A pentamer transcriptional complex including tal-1 and retinoblastoma protein downmodulates c-kit expression in normal erythroblasts.* Molecular and cellular biology, 2000. **20**(14): p. 5330-42.
195. Lecuyer, E., et al., *The SCL complex regulates c-kit expression in hematopoietic cells through functional interaction with Sp1.* Blood, 2002. **100**(7): p. 2430-40.
196. Goardon, N., et al., *ETO2 coordinates cellular proliferation and differentiation during erythropoiesis.* The EMBO journal, 2006. **25**(2): p. 357-66.
197. Kusy, S., et al., *NKX3.1 is a direct TAL1 target gene that mediates proliferation of TAL1-expressing human T cell acute lymphoblastic leukemia.* The Journal of experimental medicine, 2010. **207**(10): p. 2141-56.
198. Huang, S., et al., *p300 functions as a transcriptional coactivator for the TAL1/SCL oncoprotein.* Oncogene, 1999. **18**(35): p. 4958-67.
199. Huang, S., et al., *P/CAF-mediated acetylation regulates the function of the basic helix-loop-helix transcription factor TAL1/SCL.* The EMBO journal, 2000. **19**(24): p. 6792-803.
200. Nightingale, K.P., et al., *Histone acetylation facilitates RNA polymerase II transcription of the Drosophila hsp26 gene in chromatin.* The EMBO journal, 1998. **17**(10): p. 2865-76.
201. Cirillo, L.A., et al., *Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4.* Mol Cell, 2002. **9**(2): p. 279-89.
202. Goldfarb, A.N. and K. Lewandowska, *Inhibition of cellular differentiation by the SCL/tal oncoprotein: transcriptional repression by an Id-like mechanism.* Blood, 1995. **85**(2): p. 465-71.
203. Hofmann, T.J. and M.D. Cole, *The TAL1/Scl basic helix-loop-helix protein blocks myogenic differentiation and E-box dependent transactivation.* Oncogene, 1996. **13**(3): p. 617-24.

204. Hsu, H.L., et al., *Positive and negative transcriptional control by the TAL1 helix-loop-helix protein.* Proceedings of the National Academy of Sciences of the United States of America, 1994. **91**(13): p. 5947-51.
205. Li, E., T.H. Bestor, and R. Jaenisch, *Targeted mutation of the DNA methyltransferase gene results in embryonic lethality.* Cell, 1992. **69**(6): p. 915-26.
206. !!! INVALID CITATION !!!
207. Ausubel, F.M., *Short protocols in molecular biology : a compendium of methods from Current protocols in molecular biology*. 1989, New York ; Chichester: Wiley. xxii, 387 p.
208. Nishikawa, S.I., et al., *Progressive lineage analysis by cell sorting and culture identifies FLK1+VE-cadherin+ cells at a diverging point of endothelial and hemopoietic lineages.* Development, 1998. **125**(9): p. 1747-57.
209. Burdsal, C.A., C.H. Damsky, and R.A. Pedersen, *The role of E-cadherin and integrins in mesoderm differentiation and migration at the mammalian primitive streak.* Development, 1993. **118**(3): p. 829-44.
210. Kennedy, M., et al., *A common precursor for primitive erythropoiesis and definitive haematopoiesis.* Nature, 1997. **386**(6624): p. 488-93.
211. Olson, E.N. and D. Srivastava, *Molecular pathways controlling heart development.* Science, 1996. **272**(5262): p. 671-6.
212. Gossen, M. and H. Bujard, *Tight control of gene expression in mammalian cells by tetracycline-responsive promoters.* Proceedings of the National Academy of Sciences of the United States of America, 1992. **89**(12): p. 5547-51.
213. Masui, S., et al., *An efficient system to establish multiple embryonic stem cell lines carrying an inducible expression unit.* Nucleic acids research, 2005. **33**(4): p. e43.
214. Mao, X., Y. Fujiwara, and S.H. Orkin, *Improved reporter strain for monitoring Cre recombinase-mediated DNA excisions in mice.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(9): p. 5037-42.
215. Friedrich, G. and P. Soriano, *Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice.* Genes & development, 1991. **5**(9): p. 1513-23.
216. Zambrowicz, B.P., et al., *Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells.* Proceedings of the National Academy of Sciences of the United States of America, 1997. **94**(8): p. 3789-94.
217. Miyazaki, S., E. Yamato, and J. Miyazaki, *Regulated expression of pdx-1 promotes in vitro differentiation of insulin-producing cells from embryonic stem cells.* Diabetes, 2004. **53**(4): p. 1030-7.
218. Mizuguchi, H. and T. Hayakawa, *The tet-off system is more effective than the tet-on system for regulating transgene expression in a single adenovirus vector.* J Gene Med, 2002. **4**(3): p. 240-7.
219. Saunders, T.L., *Inducible transgenic mouse models.* Methods in molecular biology, 2011. **693**: p. 103-15.
220. Benninger, Y., et al., *Differentiation and histological analysis of embryonic stem cell-derived neural transplants in mice.* Brain pathology, 2000. **10**(3): p. 330-41.

221. Masui, S., et al., *Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells.* Nat Cell Biol, 2007. **9**(6): p. 625-35.
222. Wang, R., et al., *Promoter-dependent EGFP expression during embryonic stem cell propagation and differentiation.* Stem cells and development, 2008. **17**(2): p. 279-89.
223. Alexopoulou, A.N., J.R. Couchman, and J.R. Whiteford, *The CMV early enhancer/chicken beta actin (CAG) promoter can be used to drive transgene expression during the differentiation of murine embryonic stem cells into vascular progenitors.* BMC Cell Biol, 2008. **9**: p. 2.
224. Ying, Q.L., et al., *The ground state of embryonic stem cell self-renewal.* Nature, 2008. **453**(7194): p. 519-23.
225. Saez, E., et al., *Inducible gene expression in mammalian cells and transgenic mice.* Current opinion in biotechnology, 1997. **8**(5): p. 608-16.
226. Gessert, S. and M. Kuhl, *The multiple phases and faces of wnt signaling during cardiac differentiation and development.* Circulation research, 2010. **107**(2): p. 186-99.
227. Liu, Y., et al., *Sox17 is essential for the specification of cardiac mesoderm in embryonic stem cells.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(10): p. 3859-64.
228. Garfield, A.S., et al., *Distinct physiological and behavioural functions for parental alleles of imprinted Grb10.* Nature, 2011. **469**(7331): p. 534-8.
229. Gordon, F.E., et al., *Increased expression of angiogenic genes in the brains of mouse meg3-null embryos.* Endocrinology, 2010. **151**(6): p. 2443-52.
230. Sato, S., et al., *Methylation dynamics of IG-DMR and Gtl2-DMR during murine embryonic and placental development.* Genomics, 2011. **98**(2): p. 120-7.
231. Holt, L.J. and K. Siddle, *Grb10 and Grb14: enigmatic regulators of insulin action--and more?* The Biochemical journal, 2005. **388**(Pt 2): p. 393-406.
232. Wright, E., et al., *The Sry-related gene Sox9 is expressed during chondrogenesis in mouse embryos.* Nature genetics, 1995. **9**(1): p. 15-20.
233. Morrisey, E.E., et al., *GATA-4 activates transcription via two novel domains that are conserved within the GATA-4/5/6 subfamily.* The Journal of biological chemistry, 1997. **272**(13): p. 8515-24.
234. Grepin, C., et al., *A hormone-encoding gene identifies a pathway for cardiac but not skeletal muscle gene transcription.* Molecular and cellular biology, 1994. **14**(5): p. 3115-29.
235. Jiang, Y. and T. Evans, *The Xenopus GATA-4/5/6 genes are associated with cardiac specification and can regulate cardiac-specific transcription during embryogenesis.* Developmental biology, 1996. **174**(2): p. 258-70.
236. Holtzinger, A., G.E. Rosenfeld, and T. Evans, *Gata4 directs development of cardiac-inducing endoderm from ES cells.* Developmental biology, 2010. **337**(1): p. 63-73.
237. Grepin, C., G. Nemer, and M. Nemer, *Enhanced cardiogenesis in embryonic stem cells overexpressing the GATA-4 transcription factor.* Development, 1997. **124**(12): p. 2387-95.
238. Takeuchi, J.K. and B.G. Bruneau, *Directed transdifferentiation of mouse mesoderm to heart tissue by defined factors.* Nature, 2009. **459**(7247): p. 708-11.
239. Holtzinger, A. and T. Evans, *Gata4 regulates the formation of multiple organs.* Development, 2005. **132**(17): p. 4005-14.

240. Loose, M. and R. Patient, *A genetic regulatory network for Xenopus mesendoderm formation.* Developmental biology, 2004. **271**(2): p. 467-78.
241. Holtzinger, A. and T. Evans, *Gata5 and Gata6 are functionally redundant in zebrafish for specification of cardiomyocytes.* Developmental biology, 2007. **312**(2): p. 613-22.
242. Singh, M.K., et al., *Gata4 and Gata5 cooperatively regulate cardiac myocyte proliferation in mice.* The Journal of biological chemistry, 2010. **285**(3): p. 1765-72.
243. Peterkin, T., A. Gibson, and R. Patient, *Redundancy and evolution of GATA factor requirements in development of the myocardium.* Developmental biology, 2007. **311**(2): p. 623-35.
244. Rojas, A., et al., *Gata4 expression in lateral mesoderm is downstream of BMP4 and is activated directly by Forkhead and GATA transcription factors through a distal enhancer element.* Development, 2005. **132**(15): p. 3405-17.
245. Heicklen-Klein, A. and T. Evans, *T-box binding sites are required for activity of a cardiac GATA-4 enhancer.* Developmental biology, 2004. **267**(2): p. 490-504.
246. Kanazawa, A., et al., *Wnt5b partially inhibits canonical Wnt/beta-catenin signaling pathway and promotes adipogenesis in 3T3-L1 preadipocytes.* Biochemical and biophysical research communications, 2005. **330**(2): p. 505-10.
247. Parra, M. and E. Verdin, *Regulatory signal transduction pathways for class IIa histone deacetylases.* Current opinion in pharmacology, 2010. **10**(4): p. 454-60.
248. Chen, H., et al., *HDAC inhibition promotes cardiogenesis and the survival of embryonic stem cells through proteasome-dependent pathway.* Journal of cellular biochemistry, 2011.
249. Margueron, R. and D. Reinberg, *The Polycomb complex PRC2 and its mark in life.* Nature, 2011. **469**(7330): p. 343-9.
250. Schuettengruber, B. and G. Cavalli, *Recruitment of polycomb group complexes and their role in the dynamic regulation of cell fate choice.* Development, 2009. **136**(21): p. 3531-42.
251. Woo, C.J., et al., *A region of the human HOXD cluster that confers polycomb-group responsiveness.* Cell, 2010. **140**(1): p. 99-110.
252. Mysliwiec, M.R., E.H. Bresnick, and Y. Lee, *Endothelial Jarid2/Jumonji is required for normal cardiac development and proper Notch1 expression.* The Journal of biological chemistry, 2011. **286**(19): p. 17193-204.
253. Pimanda, J.E., et al., *The SCL transcriptional network and BMP signaling pathway interact to regulate RUNX1 activity.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(3): p. 840-5.
254. Pimanda, J.E., et al., *Gata2, Fli1, and Scl form a recursively wired gene-regulatory circuit during early hematopoietic development.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(45): p. 17692-7.
255. Kawase, E., et al., *Strain difference in establishment of mouse embryonic stem (ES) cell lines.* The International journal of developmental biology, 1994. **38**(2): p. 385-90.
256. Nelson, T.J., et al., *CXCR4+/FLK-1+ biomarkers select a cardiopoietic lineage from embryonic stem cells.* Stem cells, 2008. **26**(6): p. 1464-73.

257. Zhao, R., et al., *Loss of both GATA4 and GATA6 blocks cardiac myocyte differentiation and results in acardia in mice.* Developmental biology, 2008. **317**(2): p. 614-9.
258. Vincent, S.D. and M.E. Buckingham, *How to make a heart: the origin and regulation of cardiac progenitor cells.* Current topics in developmental biology, 2010. **90**: p. 1-41.
259. Simoes-Costa, M.S., et al., *The evolutionary origin of cardiac chambers.* Developmental biology, 2005. **277**(1): p. 1-15.
260. Belaoussoff, M., S.M. Farrington, and M.H. Baron, *Hematopoietic induction and respecification of A-P identity by visceral endoderm signaling in the mouse embryo.* Development, 1998. **125**(24): p. 5009-18.
261. Bielinska, M., et al., *Erythropoiesis and vasculogenesis in embryoid bodies lacking visceral yolk sac endoderm.* Blood, 1996. **88**(10): p. 3720-30.
262. Peterkin, T., A. Gibson, and R. Patient, *Common genetic control of haemangioblast and cardiac development in zebrafish.* Development, 2009. **136**(9): p. 1465-74.
263. Nemer, G. and M. Nemer, *Transcriptional activation of BMP-4 and regulation of mammalian organogenesis by GATA-4 and -6.* Developmental biology, 2003. **254**(1): p. 131-48.
264. Walmsley, M., A. Ciau-Uitz, and R. Patient, *Adult and embryonic blood and endothelium derive from distinct precursor populations which are differentially programmed by BMP in Xenopus.* Development, 2002. **129**(24): p. 5683-95.
265. Ikeda, T., et al., *Distinct roles of Sox5, Sox6, and Sox9 in different stages of chondrogenic differentiation.* Journal of bone and mineral metabolism, 2005. **23**(5): p. 337-40.
266. Olsen, B.R., A.M. Reginato, and W. Wang, *Bone development.* Annual review of cell and developmental biology, 2000. **16**: p. 191-220.
267. Porada, C.D., E.D. Zanjani, and G. Almeida-Porad, *Adult mesenchymal stem cells: a pluripotent population with multiple applications.* Current stem cell research & therapy, 2006. **1**(3): p. 365-9.