Ly-6A/E (Sca-1) Gene Regulatory Elements in Hematopoietic Stem Cells in Mouse

Ly-6A/E (Sca-1) genregulatie elementen in hematopoetische stamcellen in muizen

Proefschrift

ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van de
Rector Magnificus
Prof.dr.ir. J.H.van Bemmel
en volgens besluit van het College voor promoties

De openbare verdediging zal plaatsvinden op
woensdag 17 september 2003 om 11:45 uur

door

Xiaoqian Ma
麻晓迁
geboren te Wuhann, Hubei, P.R.China

Promotiecommissie

Promotoren:

Prof.dr. F.G. Grosveld
Prof.dr. E.A. Dzierzak

Overige leden:

Prof.dr. I.P. Touw
Dr. J.N.J. Philipsen
Dr. H.R. Delwel

Dit proefschrift kwam tot stand binnen de vakgroep Celbiologie en Genetica van de faculteit der Geneeskunde en Gezondheidswetenschappen van de Erasmus Universiteit Rotterdam. De vakgroep maakt deel uit het Medisch Genetisch Centrum Zuid –West Nederland.

To my parents

For Ting and Chongde

Contents

List of abbreviations:

7AAD	7-amino-actinomycin D
AGM	aorta-gonad-mesonephros
AML	acute myeloid leukemia
IFN	interferon
BMP	bone morphogenetic protein
CAFC	cobblestone area forming cell
CBFA2	core binding factor α2
CFU-C	colony forming unit-culture
CFU-S	colony forming unit-spleen
dpc	day of post coitum
FACS	fluorescence activated cell sorting
FCS	fetal calf serum
FDG	fluorescein di-(β-D-galactopyrannoside)
GFP	green fluorescent protein
GPI	glycophosphatidylinositol
hGH	human growth hormone
HSC	hematopoietic stem cell
HSS	hypersensitive site
IL	interleukin
IFN	interferon
ISRE	interferon sensitive response element
LTC-IC	long term culture initiating cell
LTR	long term repopulating
MEL	murine erythroid leukemia
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PE	phycoerythrin
PGK	phosphoglycerate kinase
PI	propidium iodide
SMA	single cell multipotential assay
P-Sp	para-aortic splanchnopleura
SCF	stem cell factor
TK	thymidine kinase
UGR	urogenital ridge
WGA	wheat-germ agglutinin

Chapter 1

General Introduction

Ly-6 gene and hematopoietic stem cells

Ontogeny and development of hematopoietic stem cells

The aim of this study

Scope of the thesis

1. *Ly-6* gene and hematopoietic stem cells

1.1 Characterization of hematopoietic stem cells

Hematopoietic stem cells (HSCs) are defined as cells that are able to self-renew and to multilineage reconstitute the hematopoietic system. Within this hierarchical system, HSCs differentiate along a pathway, giving rise to progenitor cells, such as lymphoid and myeloid progenitor cells, which have a limited ability to contribute to multilineage hematopoiesis. Further maturing progenitors, such as granulocyte or macrophage progenitors are committed and contribute to only a single lineage of cells (Keller, 1992). HSCs exist at very low frequencies in the adult bone marrow, representing approximately one out of 10^4-10^5 nucleated cells (about 0.05%) (Boggs et al., 1982) (Figure 1).

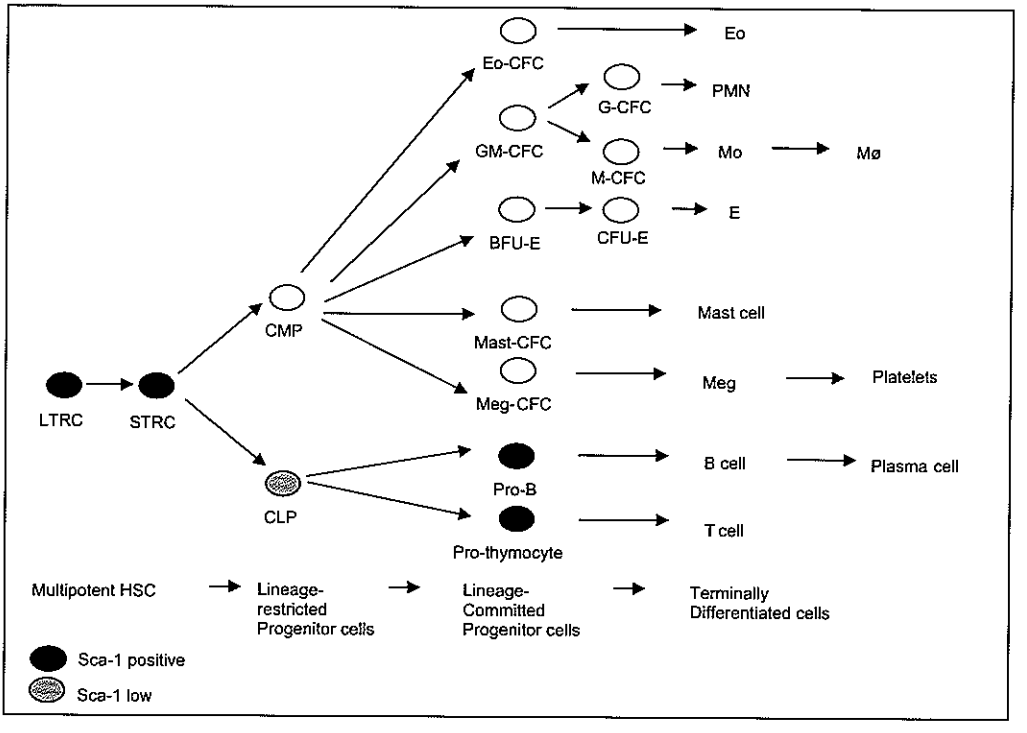

Figure 1. Schematic representation of the hematopoietic hierarchy. LTRC: long-term repopulating cell; STRC: short-term repopulating cell; CMP: common myeloid progenitor. CLP: common lymphoid progenitor. CFC: colony-forming cell; Eo: Eosinophil; GM: granulocyte/macrophage; G: granulocyte PMN: polymorphonuclear phagocyte; Mo: monocyte; MΦ: macrophage; BFU-E: burst-forming unit-erythroid; CFU-E: colony-forming unit-erythroid; E: erythrocyte; Meg: megakaryocyte; pro-B: progenitor B cell; HSC: hematopoietic stem cell.

The initial experiments to characterise the multilineage progenitor cells were done by Till and McCulloch in the early 60's (Till, 1961; Wu et al., 1967). Bone marrow cells were intravenously injected into irradiated adult mice, from which

macroscopic colonies on the spleen were formed after 8-11 days of injection, referred to as colony forming unit-spleen, CFU-S. The cells of CFU-S were composed of erythroid, myeloid and megakaryocytic lineages and these cells can also give rise to CFU-S in secondary irradiated transplanted recipient mice (Siminovitch L., 1963; Siminovitch L., 1964). Further experiments showed that there is some heterogeneity within the CFU-S cells of bone marrow. Colonies taken 12 days post injection of bone marrow had a greater potential to form secondary colonies than those taken 8 days post injection (Magli et al., 1982; Worton et al., 1969). However, neither of them were able to give rise to long term reconstitution (Jones et al., 1990; van der Loo et al., 1994). All these data together indicate that CFU-S assay is a useful tool for analysis of hematopoietic progenitor activity for myeloid, erythroid and megakaryocytic lineages, but not a suitable approach to test primary hematopoietic stem cells (Medvinsky et al., 1993; Moore and Metcalf, 1970).

In addition to *in vivo* CFU-S assays for progenitor cells, many *in vitro* culture systems have been developed for the study of hematopoietic potency. Generally for such analyses the hematopoietic cells are grown in semisolid media, like methylcellulose, in the presence of well-defined growth factors and/or on a stromal cell layer. Hematopoietic progenitors are then scored according to their ability to give rise to differentiated hematopoietic progeny (Metcalf and Nicola, 1984). This method, referred to as the colony forming unit-culture (CFU-C) assay, which is normally used for detecting only cells of myeloid and erythroid lineages, is still one of the most widely used *in vitro* assays. For detecting differentiation of the lymphoid lineage, other specific culture systems have been employed. For example, B lymphopoiesis is detected from putative progenitors using specific stromal cells and growth factor supplements in a co-culture system (Ogawa et al., 1988). T lymphopoiesis can be monitored by using T cell depleted fetal thymic explants as a stromal support to provide the source of growth factors and cellular signals for the specific proliferation/differentiation of T cell progenitors (Liu and Auerbach, 1991). In addition to the method described above, several multistep culture systems have been reported. For example, the single cell multipotential assay (SMA) (Godin et al., 1995b) is a two culture step assay, in which cells are first seeded at single cell dilution onto stromal cells in the presence of several interleukins and c-kit ligand (KL). After 10 to 15 days each clonally derived culture is assayed for multipotency by testing the cells in myeloid, B and T cell lineage differentiation cultures as described above. The long term culture initiating cell (LTC-IC) and cobblestone area forming cell (CAFC) assays (Dexter et al., 1977; Lemieux et al., 1995) are also co-culture systems, requiring bone marrow stroma layer, complex cocktail of growth factors and highly specific culture conditions. It should be taken into account that each different culture assay has its own significance and measures cells at different levels in the hematopoietic hierarchy. For example, the LTC-IC and CAFC assays provide an assessment of the hematopoietic progenitor content of a population but do not test the potential of single cells, while SMA assay provides information on the multilineage potential of a single cell.

The most stringent and widely accepted assay for hematopoietic stem cells is

to test the ability of such cells to give rise to long-term multilineage reconstitution of an irradiated adult recipient mouse (Abramson et al., 1977; Orlic and Bodine, 1994). Our current understanding about hematopoietic stem cells is based on studies using this type of *in vivo* transplantation assay. The crucial requirement for such transplantation experiments is to design a practical approach to discriminate donor-derived and recipient hematopoietic cells. These approaches include the use of different mouse strains displaying allelic differences in expression of certain cell surface antigens, like Thy-1.1 and Thy-1.2, or Ly-5.1 and Ly-5.2 (Spangrude et al., 1988), also the use of polymorphisms in hemoglobin (Down et al., 1991; Neben et al., 1993) or electrophoretically distinguishable enzymes, such as glucose phosphate isomerase (GPI) (Down et al., 1991; Harrison and Lerner, 1991) and phosphoglycerate kinase (PGK) (Neben et al., 1991). A highly sensitive method, from which signals can be amplified by the polymerase chain reaction (PCR), makes use of differences between donor and recipient at the DNA level. Previously, at the DNA level, the most often used discriminating methods were radiation induced chromosomal markers (Abramson et al., 1977). More recently, donor-recipient sex-mismatched transplantations (Jones et al., 1990), different transgenic markers (de Bruijn et al., 2000a; Miles et al., 1997) and retroviral transfection of donor stem cells (Dick et al., 1985; Keller et al., 1985; Lemischka et al., 1986) have been used together with PCR methods, resulting in an increase in the sensitivity and reliability of donor-marker detection.

To elucidate if an individual hematopoietic stem cell can repopulate all hematopoietic lineages and if clonality of hematopoietic reconstitution exists, various assays were established by using cells marked by the methods described above. The first experiment examining these questions involved the induction of cytological damage to donor bone marrow cells by irradiation. The radiation induced chromosomal markers allowed following the fate of these donor cells after transplantation into lethally irradiated recipient mice. In such experiments, progeny of the donor marked cells were detected in all hematopoietic lineages of the recipient mice, demonstrating clonal repopulation of the adult hematopoietic system by a single hematopoietic stem cell (Abramson et al., 1977; Wu et al., 1968).

The other approach of studying cell lineage and precursor/progeny in more detail is the use of replication defective recombinant retroviruses that harbour marker genes. The advantage of this method is the ability to avoid potentially damaging chromosomal changes (like chromosomal translocation) induced by irradiation (Dzierzak and Mulligan, 1988; Lemischka et al., 1986). Bone marrow or any test cell population can be infected with the replication-defective retrovirus, which will integrate randomly into the DNA of each infected cell. Given the size of the genome, it is likely that no two cells of those originally infected will have the same viral integration site. When cells from the original population divide, each specific integration pattern will be inherited; therefore the progeny of each starting cell can be clonally marked. Analysis of radiation chimaeras generated from retrovirally marked bone marrow demonstrated that a single cell could differentiate into all hematopoietic lineages and provide long term reconstitution (Dick et al., 1985; Keller et al., 1985; Lemischka et al., 1986).

1.2 Isolation and purification of hematopoietic stem cells

Because HSCs are present at a very low frequency in bone marrow, the development of highly efficient enrichment procedures was necessary. Hence, the powerful fluorescence activated cell sorting (FACS) technology was employed in the purification of HSCs through the use of fluorochrome-labeled antibodies specific for markers expressed on HSCs (Herzenberg and Sweet, 1976). To date, there is no single antibody identified that marks HSCs exclusively. However, through the combination of several different antibodies to cell surface markers, it is possible to achieve significant enrichment of HSCs from murine bone marrow (Spangrude, 1989) (Table 1).

Table 1. The mouse hematopoietic cell surface markers

Cell surface marker	*Major lineages of expression*	*Expression on HSCs/Multipotent progenitors*
Thy-1	T cells	Low
Sca-1	T cells	High
c-kit	Progenitors	Positive
Mac-1	Granulocytes, Monocytes	Negative or low
CD34	Progenitors	Positive
CD4	T cells	Negative or low
CD8	T cells	Negative
B220	B cells	Negative
Gr-1	Granulocytes	Negative
Ter119	Erythroid progenitors	Negative
H-2K	All hematopoietic cells	High
Ly-5	Most hematopoietic cells	Positive
CD27	Short term repopulating cell	Negative
CD31	Endothelial cells	Positive

Data compiled from (Aihara et al., 1986; Muller-Sieburg et al., 1986; Smith et al., 1991; Spangrude et al., 1988; Spangrude and Scollay, 1990) and (de Bruijn et al., 2002; North et al., 2002; Wiesmann et al., 2000).

It has been shown that mouse HSCs could be highly enriched among H-2K^{high} WGA^{+} fraction of bone marrow cells in 1984 (Visser et al., 1984). Shortly thereafter, monoclonal antibodies directly against the Thy-1 glycoprotein were also found as a marker for bone marrow HSCs isolation, and the bone marrow HSCs fall in the population of Thy-1^{-} and/or Thy-1^{low}. For higher enrichment of stem cells, a negative selection procedure was employed in which bone marrow was depleted of cells expressing markers of B cells (B220), T cells (CD4 and CD8), granulocytes (Gr-1) and myelomonocytic cells (Mac-1). Bone marrow cells expressing low levels of Thy-1 in the absence of mature lineage markers (Thy-1^{lo}, Lin^{-}) were 50-200 fold enriched in hematopoietic progenitor activity (Muller-Sieburg et al., 1986). This Thy-1^{lo}, Lin^{-} population has been further enriched using a monoclonal antibody E13-161.7 (Aihara et al., 1986) which reacts with antigen referred to as Sca-1. The Sca-1 marker is a member of the Ly-6 family and is also known as the Ly-6A/E antigen (Spangrude and Brooks, 1993; van de Rijn et al., 1989). The Thy-1^{lo}, Lin^{-}, Sca-1^{+} fraction of bone marrow from C57BL/Thy-1.1 mice represents approximately 0.05% of bone marrow cells and is approximately 2,000 fold enriched for HSC activity *in vitro* assays and

long term multilineage reconstitution assays (Li and Johnson, 1992; Smith et al., 1991; Spangrude et al., 1988). It is of considerable importance to choose certain mouse strains for HSC sorting studies, since both Thy-1 and Sca-1 exhibit strain specific polymorphisms, resulting in different antibody staining profiles in bone marrow of mice of different backgrounds. It has been shown that HSCs in Thy-1.2 strains of mice are both Thy-1^{lo} and Thy-1^{-}, while the vast majority of HSCs activity was found only in Thy-1^{lo} population in Thy-1.1 strains (instead of both Thy-1^{lo} and Thy-1^{-} in Thy-1.2 strains) (Spangrude and Brooks, 1992). Moreover, Sca-1 is not expressed on all marrow-repopulating cells in Ly-6E strains (only 25% of HSCs are Sca-1^{+}) but is expressed on virtually all marrow-repopulating cells in Ly-6A strains (Spangrude and Brooks, 1993). Despite the allelic differences, Sca-1 is now used routinely for HSC enrichment (Jurecic et al., 1993; Okada et al., 1992; Spangrude and Brooks, 1993), and it has been shown that the Thy-1^{lo} Lin^{-} Sca-1^{+} population from bone marrow of the C57BL/Ka-Thy1.1 strain includes almost all the detectable multipotent progenitors (Uchida and Weissman, 1992).

The Thy-1^{lo} Lin^{-} Sca-1^{+} bone marrow fraction can be further enriched by using c-kit receptor marker. Separation of Thy-1^{lo} Lin^{-} Sca-1^{+} bone marrow cells into c-kit expressing and non-expressing subsets demonstrated that only c-kit cells possessed CFU-S activity and were enriched in long-term repopulating activity (Ikuta and Weissman, 1992). Furthermore the analysis demonstrated that Sca-1 expression is associated with more immature bone marrow hematopoietic stem cells than is c-kit. During HSCs maturation and differentiation, they rapidly lose Sca-1 antigen from their cell surface, while the c-kit maker is maintained till the committed progenitor stage. Combining Sca-1 and c-kit antibodies provide a powerful system for HSC enrichment (Ikuta and Weissman, 1992; Okada et al., 1991; Okada et al., 1992). Nowadays, several other markers have also proven useful for enriching HSCs, such as CD34, AA4.1 and CD31 (de Bruijn et al., 2002; Jordan et al., 1995; North et al., 2002; Sowala et al., 1990).

1.3 Sca-1 and the *Ly-6* gene superfamily

Based on homology analysis of amino acid sequences, Ly-6A.2 and Ly-6E.1, also called Ly-6A and Ly-6E or Sca-1, were classified to the Ly-6 superfamily. So far, many murine members have been identified within this family, such as ThB, TSA-1, Ly-6F, Ly-6A/E,Ly-6C, Ly-6G, Ly-6B, etc (Gumley et al., 1995). The human homologous of murine Ly-6 family include CD59 (Williams et al., 1988), E48 (Brakenhoff et al., 1995), LY-6K (de Nooij-van Dalen et al., 2003), RIG-E (Mao et al., 1996), etc. Similar to the mouse Ly-6 family, many human LY-6 molecules are not only expressed in hematopoietic system, but also in keratinocytes (Brakenhoff et al., 1997; HogenEsch et al., 1993). The biological function of this family is not clear, and no ligand has been identified until recently. However three reports provide similar implications concerning the possible functions of this family of proteins. 1) The *Ly-6E* gene was found differentially expressed in mouse metastatic adenocarcinoma cells. 2) RIG-E can be induced by retinoic acid during the differentiation of acute promyelocytic leukemia cell. 3) LY-6K, another new member of this family, was identified as molecular marker for head-and-neck squamous cell carcinoma (Cohn et al., 1997; de Nooij-van Dalen et al., 2003; Mao et al., 1996). Taken together, these results suggest that the most likely biological functions of this family are

associated with cell-cell adhesion, signal transduction and/or cell proliferation. Indeed the first *Ly-6* gene *Ly-6E.1* was cloned from a tumor cell, the MethA fibrosarcoma cell line (LeClair et al., 1986). In mouse, using *Ly-6E* cDNA probe in genomic Southern blot revealed many hybridising bands, indicating the existence of a family of genes with high sequence homology (LeClair et al., 1986). Using the homology between *Ly-6* coding regions, the *Ly-6A* (Palfree et al., 1987; Reiser et al., 1988), *Ly-6C* (Palfree et al., 1988), *Ly-6F* and *Ly-6G* (Fleming et al., 1993), *ThB* (Gumley et al., 1992), *TSA-1* (MacNeil et al., 1993) genes have been cloned. The other two new *Ly-6* family members, the *Ly-6I* and *Ly-6M* genes were cloned in 2000 (Patterson et al., 2000; Pflugh et al., 2000). Gene mapping indicates that the Ly-6 locus is located on chromosome 15 and contains many homologous genes (Kamiura et al., 1992). The complexity and homology within this locus gives rise to much difficulty in the determination of how many functional related genes and/or how many pseudogenes exist.

All members of murine Ly-6 superfamily are low molecular weight (12-20 kDa) phosphatidylinositol (GPI) anchored cell surface glycoproteins and thus do not have a cytoplasmic domain. They are characterised by the presence of 10 cysteine residues (Gumley et al., 1992). These 10 conserved cysteine residues are equally spaced in the amino acid sequences of the Ly-6 superfamily and are predicted to form disulphide bridges which gives rise to a compact protein structure consisting of a tight, globular core with finger-like projections (Fleming et al., 1993). Such a predicted structure could be involved in ligand binding or cell-cell interactions (figure 2).

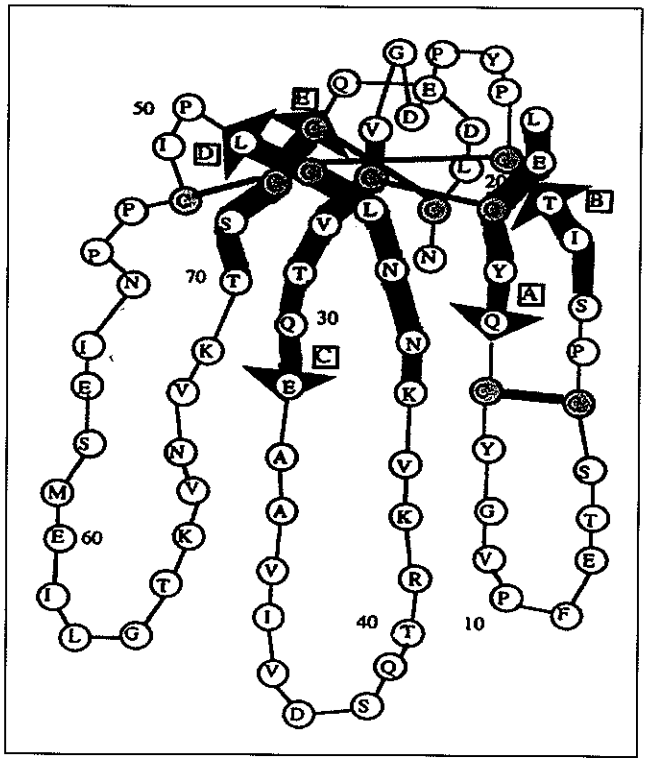

Figure 2. Structural model of Ly-6A molecule. The Ly-6A amino acid sequence was modeled on the crystal structure for α-bungarotoxin. The conserved cysteine residues are represented by a stippled background and putative disulfide bonds are indicated by filled bars. The thick arrows indicate regions of predicted β-sheet structure. The picture is adapted from review of Gumley (Gumley et al., 1995).

The expression pattern of the Ly-6 family of proteins in mouse is complex. During hematopoietic development, the expression patterns of the individual Ly-6 family members are tissue specific but quite often overlapping (see Table 2). The functional consequences of this complexity are not clear, but may reflect an important role for Ly-6 proteins in subtle modulation and/or differentiation of the hematopoietic system.

Table 2. Expression of Ly-6 antigens in hematopoietic system in mouse

Antigen	*thymus*	*Lymph Node*	*Spleen*	*Bone marrow*	*Activated Lymphocytes*
Ly-6A	+	++	++	+	++++
Ly-6E	+	+	+	+	+++
Ly-6B	-	+	+	++	-
Ly-6C	-	++	++	++	++
Ly-6M	n.d.	n.d.	n.d.	++	n.d.
Ly-6I	+	n.d.	+	++	n.d.
Ly-6G	-	-	-	+++	-
TSA-1	++++	-	-	+++	n.d.
ThB	++++	++	++	++	n.d.

Data updated from the thesis of A.Sinclair. Surface antigen expression within specific tissues is presented as follows: ++++=75-100%; +++=50-75%; ++=25-50; +=5-25%; -=less than 5%; n.d.=not done. (Fleming et al., 1993; Houlden et al., 1986; Kimura et al., 1984; Patterson et al., 2000; Pflugh et al., 2000; Tucek et al., 1992).

1.4 The murine *Ly-6A/E* gene

The *Ly-6A* and *Ly-6E* genes are strain-specific alleles within the multigene Ly-6 family and Ly-6A and Ly-6E protein expression within the hematopoietic system is complex (Kimura et al., 1984; Spangrude and Brooks, 1993; van de Rijn et al., 1989). The allelic genes differ by only three nucleotides in the coding sequence, resulting in two amino acid changes in the proteins (LeClair et al., 1986; Palfree et al., 1986; Palfree et al., 1987; Reiser et al., 1988; Rock et al., 1986). Both Ly-6E and Ly-6A proteins express the epitope recognized by the Sca-1 antibody. The allelic genes are interferon inducible, but the *Ly-6A* allele appears to be more widely and highly expressed (Kimura et al., 1984; Rock et al., 1986; Spangrude and Brooks, 1993). For example, Ly-6A strains of mice express this protein on 10-20% of adult thymocytes and 50-70% of peripheral T lymphocytes, while Ly-6E strains express the protein on 5-10% adult thymocytes and 10-15% of peripheral T lymphocytes. Similarly and more importantly, expression of these allelic products has been shown to differ on hematopoietic stem cells. Using the bone marrow repopulating assay as a readout, it was shown that Ly-6A strains of mice express Sca-1 on virtually all (99%) of hematopoietic stem cells while Ly-6E strains express Sca-1 on only 25% of these

cells (Spangrude and Brooks, 1993). These data indicate that complex developmental control mechanisms are likely to be involved in the regulation of both Ly-6A and Ly-6E expression, and *Ly-6A* gene regulatory elements are more active and/or responsible for high level of expression in HSCs. Further, the expression of the *Ly-6E* and *Ly-6A* genes is not only limited to the hematopoietic system. Expression was also found in keratinocytes and other tissues. The Ly-6A antigen can be found within the kidney and the vasculature of the heart, liver and brain (Reiser et al., 1988; van de Rijn et al., 1989). The analysis of *in situ* brain sections has demonstrated that the alloantigens exhibit extremely diverse expression patterns in this tissue. Ly-6A is limited to the vasculature and Ly-6E expression is in the hippocampal and midbrain regions (Cray et al., 1990) (Table 3).

Table 3. Comparison of surface alloantigen Ly-6E and Ly-6A expression patterns

Tissue	Ly-6E	Ly-6A
Lymph node	5%	60-70%
Spleen	10%	59%
Bone marrow	5%	5%
Activated lymphocytes	80-90%	100%
Thymus: CD4-CD8-	15%	67%
CD4+CD8+	2%	9%
CD4+CD8-	8%	39%
CD4-CD8+	30%	21%
Hematopoietic stem cell	25%	99%
Brain	Hippocampus, Midbrain	Vasculature

Data is presented as the percentage of cells within the particular tissue that express the antigens, unless stated. Data is adapted from Sinclair's thesis which is compiled from (Codias et al., 1989; Kimura et al., 1984; Spangrude and Brooks, 1993) and (Cray et al., 1990).

The regulation of *Ly-6A/E* gene expression is complicated. Previously, 5' cis-acting elements involved in regulating expression and responsiveness of the *Ly-6E* gene to interferons have been identified (Khan et al., 1990; Khan et al., 1993). In promoter deletion studies of the *Ly-6E* gene in mouse fibroblast transfectants, an interferon inducible (GAS) element has been localized between –1.76 and –0.9 kb and a purine rich sequence at –0.11 kb was found to be necessary for basal expression. Also, between –0.16 and –0.13 kb a putative repressor sequence was found. Through sequence comparisons, similar control elements are thought to be present in the *Ly-6A* promoter (McGrew and Rock, 1991). Interestingly, DNaseI hypersensitive site analysis has identified two sites in the promoter region at –1.2 and –0.1 kb which appear to coincide with the regulatory elements localized by the 5' deletional analysis (Sinclair and Dzierzak, 1993). Analysis of the 3' flanking regions of both the *Ly-6E* and *Ly-6A* alleles, by using MEL (endogenouse Ly-6A allele) and YAC-1 (endogenouse Ly-6E allele) cell lines, has localized six additional areas of DNaseI hypersensitivity +3.3, +4.9, +5.6, +6.7, +8.7 and +8.9 (Sinclair and Dzierzak, 1993). Furthermore, there was a rapid induction of DNaseI hypersensitive sites in the 3' flanking regions of the Ly-6A allele, particularly at +8.7 and +8.9. These γ-INF-induced hypersensitive sites were also found in MEL cells transfected a 14 kb Ly-6E

fragment (figure 3). Deletional studies of the *Ly-6E* 3' flanking sequences in murine erythroleukemia (MEL) cell transfectants revealed that the region containing these two most distal 3' HSS are responsible for high level, γ-interferon induced expression (Sinclair et al., 1996). These results suggest that *Ly-6A* transcriptional regulatory elements are generally similar to those of the *Ly-6E* allele, only subtle differences between the *Ly-6E* and *Ly-6A* alleles were observed in the hypersensitive site maps (Sinclair et al., 1996) (Figure 3).

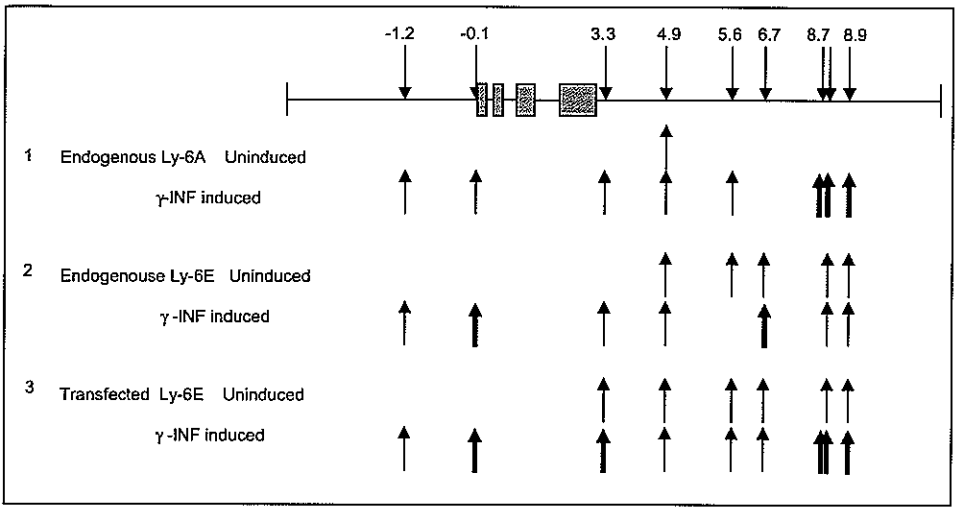

Figure 3. DNaseI hypersensitive sites in γ-IFN induced and uninduced hematopoietic cells. 1. MEL, mouse erythroid leukemia cell (Ly-6A allele), 2. YAC-1, mouse T cell (Ly-6E allele), 3. MEL cells transfected with 14 kb Ly-6E BamHI fragment, which is shown on the top of this figure. Regions that demonstrate hypersensitivity to DNaseI digestion are marked with arrows and numbered according to their location from the transcriptional start site. Exons are presented by gray boxes. The degree of hypersensitivity is presented by the bold arrows. The figure was redrawn from Sinclair et al. (Sinclair et al., 1996).

The further functional analyses of Ly-6A/E focussed on using different *Ly-6E* genomic fragments to drive *lacZ* gene expression in transgenic mice (Miles et al., 1997). The expression pattern of *Ly-6E/lacZ* transgene indicated that the *Ly-6E* genomic fragment containing the region approximately 9kb downstream from transcriptional start site including 3' hypersensitive sites at +8.7 and +8.9 is required for the distinct, restricted expression within adult hematopoietic stem cells and the embryonic hematopoietic tissue (AGM region). None of the transgenic lines, containing a construct in which the 3' region containing hypersensitive sites +8.7 and +8.9 were deleted, were found to express *lacZ* in a tissue-specific manner. Since Ly-6A strains of mice express Sca-1 on 99% hematopoietic stem cells and Ly-6E strains of mice express the protein on only 25% of hematopoietic stem cells, similar but subtly different gene regulatory mechanisms may exist between these two alleles. For purposes of molecular manipulation of hematopoietic stem cells, the use of the Ly-6A regulatory elements would appear to be of great advantage.

2. Ontogeny and development of hematopoietic stem cells

2.1 First site of HSC emergence in mouse embryos

For a long time it was believed that HSCs originate *in situ* in the yolk sac by differentiation of the mesoderm. Subsequently, they migrate and colonize the fetal liver during mid-gestation stages, and then finally home to the bone marrow before birth, where they reside throughout the life span of the animal (Houssaint, 1981; Johnson and Moore, 1975; Moore and Metcalf, 1970). However, this traditional paradigm of hematopoietic development has been strongly challenged by the studies of the origins of hematopoietic system in non-mammalian (Dieterlen-Lievre, 1975; Le Douarin, 1973) and in mammalian vertebrates (Muller et al., 1994).

Le Douarin and colleagues used the chicken-quail grafting approach to elucidate the origins of many adult tissues (Le Douarin, 1973). In this method, the nuclei of quail cells that possess a mass of heterochromatin can be used as a marker to distinguish quail cells from chicken cells. To study the embryonic origins of the avian adult hematopoietic system, interspecific chick/quail chimaeras were analysed by DNA staining after quail embryo bodies were engrafted onto intact chick yolk sacs (Dieterlen-Lievre, 1975; Le Douarin, 1973). Strikingly, it was shown that the adult hematopoietic system was generated from the quail embryo body, but not from the chick yolk sac. This result led to the proposal of a totally new concept of hematopoietic development.

Given the developmental analogy between different species, the studies of the hematopoietic system during mammalian embryonic development have been attracting more and more interest in recent years. During mouse embryonic development, gastrulation starts at 6-7 days post coitum (dpc), subsequently the three layers of ectoderm, mesoderm and definitive endoderm are formed and the intra/extra-embryonic territories become well defined. The yolk sac encompasses mesoderm and primitive endoderm while the embryo proper comprises all three layers formed during gastrulation. The intra-embryonic region derived from mesoderm and associated endoderm is called splanchnopleura. At 8 dpc, paired aortae appear in this region, and hence it is named PAS (para-aortic splanchnopleura) (Garcia-Porrero et al., 1995; Godin et al., 1993). Later at 9 dpc, the paired aortae fuse and primordia of the mesonephros and gonads become apparent at 10 dpc. Hence, the name AGM (aorta-gonads-mesonephros) was given to define this region at midgestation stage (Dzierzak and Medvinsky, 1995; Medvinsky et al., 1993; Muller et al., 1994). Many experiments have demonstrated that the PAS/AGM region contains hematopoietic activity, but the functional properties of the hematopoietic cells found in this region during mouse midgestation are quite different from those found in the yolk sac.

The first primitive erythrocytes can be observed within the blood islands of the yolk sac at 7.5 dpc (Russell, 1966). However, even at 7 dpc, the erythroid and granulocyte-macrophage progenitors from yolk sac can be detected by *in vitro* study (Moore and Metcalf, 1970). In addition, the T, B lymphoid lineages and multipotent progenitors with erythroid/myeloid and lymphoid potential have also been found in 8.5 dpc yolk sac by *in vitro* assay (Cumano et al., 1996; Godin et al., 1995a). The colony forming units-spleen (CFU-S) were able to be detected from 8 dpc, but no *in vivo* repopulating hematopoietic stem cells were found until 11 dpc

(Medvinsky et al., 1996; Medvinsky et al., 1993). Interestingly, long-term *in vivo* chimerism of erythroid and lymphoid lineages was detected when the 9 dpc yolk sac cells were injected transplacentally into early stage of embryos (Toles et al., 1989). It has also been shown that 9 dpc yolk sac cells could give rise to long-term, multi-lineage chimerism when injected into conditioned neonatal recipients (Yoder et al., 1997a; Yoder et al., 1997b). Nevertheless, long-term repopulation of lethally irradiated adult recipients is not found in the yolk sac until the embryonic stage of 11 dpc (Medvinsky and Dzierzak, 1996; Muller et al., 1994).

Compared to the yolk sac, multipotent progenitors for the erythroid and lymphoid lineages were able to be detected from PAS at 7.5 dpc, one day earlier than yolk sac (Cumano et al., 1996; Godin et al., 1995a). Although *in vivo* multipotent CFU-S are found in both yolk sac and PAS at 9 dpc, the abundance and potency of CFU-S from PAS are consistently greater than those from yolk sac (Medvinsky and Dzierzak, 1996; Medvinsky et al., 1996; Medvinsky et al., 1993). Furthermore, the HSCs capable of the full long-term, multilineage engraftment of lethal irradiated adult recipients are found in the PAS/AGM region beginning at 10 dpc, but not in yolk sac at this stage (Muller et al., 1994). Hence, the early appearance of multilineage progenitors, the number and potency of CFU-S and pluripotent HSCs in the PAS/AGM region are different from the yolk sac.

Since the blood circulation has already been set up between intra- and extra-embryonic tissues at 8.5 dpc, there is a possibility that the precursors of long term adult repopulating cells of the AGM region at 10 dpc may come from and/or depend on yolk sac. To exclude this possibility and to further confirm the hypothesis that the HSCs can be generated and/or expanded independently in PAS/AGM region, Medvinsky and colleagues have developed a explant culture system (Medvinsky and Dzierzak, 1996). The experiment is as follows: 10 dpc yolk sac and AGM were dissected and cultured individually on filters supported on metal mesh stands at the air-liquid interface. After 3 days, cultured yolk sac and AGM were collected separately, cell suspensions were made and transplanted to irradiated adult recipients. Interestingly, long term repopulating activity was only found from the AGM cell suspension but not from the yolk sac. And these AGM HSCs appear to possess the same potency as adult bone marrow stem cells, as they can give rise to high level long-term repopulation of all hematopoietic lineages of irradiated adult recipient mice. The same experiment has also been performed on 11 dpc yolk sac and AGM region and demonstrated that some HSC activity in yolk sac begins to be detectable whilst there was an increasing quantity of HSCs in AGM explants at this stage. Taking together, it is clear that the first adult repopulating HSCs are generated and/or expanded in the AGM region at 10 and 11 dpc, suggesting the intraembryonic PAS/AGM region is a site for potent definitive hematopoietic progenitors and HSC generation and expansion. More recently, it has been reported that HSCs activity was also detected in other major arteries such as umbilical and vitelline arteries at midgestation stage during the mouse development (de Bruijn et al., 2000b).

2.2 Localization and molecular markers of hematopoietic stem cells during mouse embryonic development

The PAS/AGM region is a very complex tissue structure composed of the dorsal aorta and surrounding mesenchyme, and laterally located genital ridges and mesonephros. Many studies have been performed to localize the first HSCs that

emerge in this region. In 1986, Cormier et al. demonstrated that cell clusters with hematopoietic potential within the AGM were localized within the mesoderm underlying and adhering to the ventral floor of the chick dorsal aorta (Cormier et al., 1986). Later on, morphology studies on tissue sections of human and murine embryos also revealed clusters of round cells adhering to the ventral floor of the dorsal aorta within AGM (Marshall et al., 1999; Tavian et al., 1996; Wood et al., 1997). Similar adherent clusters of cells have also been found in the proximal umbilical and vitelline arteries, which interconnect with the dorsal aorta (Tavian et al., 1999; Wood et al., 1997). In addition, our laboratory has performed functional studies in which the AGM region was dissected into aorta/mesenchyme and urogenital ridges. Subsequently, cell suspensions were made and transplanted to irradiated adult recipients. Although CFU-S_{11} were found in both aorta and urogenital ridges from 10-12 dpc, the frequency of CFU-S_{11} from 10-11 dpc aorta was higher than that of urogenital ridges from the same developmental stage (de Bruijn et al., 2000a). Most interestingly, HSC activity was only detected in 11 dpc dorsal aorta region and not in the same age urogenital ridge tissue. In addition, the experiment also showed that the vitelline and umbilical arteries contain HSC activity (de Bruijn et al., 2000b). All together, hematopoietic stem cell emergence occurs in close association with the major vasculature of the mouse embryo during midgestation.

With the aim to localize the first HSCs in mouse development more precisely, especially in AGM region, and to trace the cell lineages and cell differentiation during mouse development, the study of expression and function of hematopoietic markers has become a point of intense interest. Here, I focus on some important markers during mouse embryonic development. They include transcription factors GATA-2, Runx-1 (AML-1) and SCL, some endothelial markers such as CD31, Flk-1 and VE-cadherin, and some other markers often used in flow cytometric sorting and enrichment of HSCs and progenitors, like CD45, CD34, c-kit and Sca-1.

Many transcription factors are expressed in the AGM region and gene disruption studies have revealed important functions of at least two of these factors, GATA-2 and Runx-1, during hematopoietic development in the mouse embryo. GATA-2 and Runx-1 *null* mice showed embryonic lethality at E11.5 and E12.5 respectively. Furthermore, disruption of GATA-2 or AML-1 results in a complete block in the production of definitive hematopoietic lineages, leading to fetal liver anemia. Primitive hematopoiesis was found to be normal in AML-1 mutant embryos but some defects were found in GATA-2 mutant embryos (Castilla et al., 1996; Okuda et al., 1996; Tsai et al., 1994; Wang et al., 1996). The hematopoietic specific GATA-2 gene regulatory element has been identified and used to drive lacZ and GFP reporter genes *in vivo* (Minegishi et al., 1999). The reporter genes were detected specifically in the PAS/AGM region and *in situ* hybridization analysis showed that the reporter gene expression recapitulates the endogenous GATA-2 expression profile in this region. Moreover, flk-1, CD34, c-kit and CD45 were detected in GFP positive cells from the AGM region. By knock-in strategy, using the *LacZ* marker gene, Runx-1 (AML-1, Cbfa2) expression has been systematically followed during mouse hematopoietic ontogeny (North et al., 1999). Runx-1 expression was first detected in mesodermal cells in the yolk sac at 7.5 dpc and subsequently in both primitive erythrocytes and endothelial cells in the yolk sac at 8 dpc. At 8.5 dpc, its expression level decreases in primitive erythrocytes but remains high in a small proportion of endothelial cells and blood cells in yolk sac. Meanwhile, some other sites start to

express Runx-1, the sites include mesenchymal cells at the distal tip of the allantois, endothelial cells in the vitelline artery, and endothelial cells in the ventral part of paired dorsal aortae which later fuse to one artery and becomes part of AGM region. Between 9.5-11.5 dpc, Runx-1 is expressed in endothelial cells of vitelline and umbilical arteries, in endothelial and mesenchymal cells in the ventral portion of AGM region, in a small number of endothelial cells of yolk sac but in no other endothelial cells of the embryo. Interestingly, the clusters of Runx-1 positive blood cells attached to the ventral lumina of the dorsal aorta were detected by following the onset of endothelial cell expression. FACS analysis has shown about 35% Runx-1 positive cells in AGM region fall into c-kit$^+$CD34$^+$ population.

Another transcription factor, the T-cell leukemia oncoprotein tal-1/SCL is essential for both primitive (yolk sac) and definitive (intra-embryonic) blood formation. Lack of the *SCL* gene results in a complete block of blood formation in *null* embryos and is lethal between E8.5 and E10.5 within the period when the embryonic circulation is established (Shivdasani et al., 1995). Moreover, the *SCL*$^{-/-}$ murine embryonic stem cells fail to give rise to precursor or mature hematopoietic cells of any lineages, suggesting that SCL may be involved in blood specification at a pre-hematopoietic cell stage (Porcher et al., 1996). The gene regulatory fragment has been identified by DNaseI hypersensitive site mapping and further confirmed by a transgenic mouse approach (Sanchez et al., 1999). It was demonstrated that the 3' enhancer is essential for expression in endothelial cells and in some hematopoietic progenitors in multiple hematopoietic tissues during embryonic development, including the AGM region.

CD31, also called PECAM-1 (platelet/endothelial cell adhesion molecule) is expressed on all endothelial cells in the embryo, but not on angioblasts (Drake and Fleming, 2000). This antigen is also expressed on bone marrow HSCs (van der Loo et al., 1995) and on intra-aortic clusters (Garcia-Porrero et al., 1998). Flk-1, originally cloned from fetal liver hematopoietic cells, is expressed on endothelial cell progenitors and endothelial cells (Drake and Fleming, 2000; Matthews et al., 1991; Shalaby et al., 1997; Yamaguchi et al., 1993). It has been demonstrated that there is failure of blood island formation and vasculogenesis in flk-1 deficient mice (Shalaby et al., 1995). By knockin strategy, flk-1/lacZ expression was detected in intra-aortic hematopoietic clusters in flk-1$^{lacZ/+}$ AGM region by intracellular X-gal staining (Shalaby et al., 1997). VE-cadherin (vascular-endothelial cadherin) is expressed on endothelial cells in the dorsal aorta and major blood vessels and its expressing pattern is similar to CD31 during the mouse embryonic development (Drake and Fleming, 2000).

The pan-leukocyte marker CD45 has been shown to be expressed on the round cells in clusters associated with endothelial wall in chicken and human embryos (Jaffredo et al., 1998; Tavian et al., 1996). The membrane glycoprotein CD34, which is quite often used as a marker for adult bone marrow and human cord blood HSCs enrichment, has also been found to be expressed in the AGM region, particularly on the endothelial cells lining the dorsal aorta and associated hematopoietic cell clusters in human tissue (Marshall et al., 1999; Tavian et al., 1996). Combining CD34 and CD45, the hematopoietic clusters and their adjacent endothelial cells can be distinguished as CD34$^+$CD45$^+$ and CD34$^+$CD45$^-$ respectively (Garcia-Porrero et al., 1998; Marshall et al., 1999; Tavian et al., 1996; Wood et al., 1997).

The c-kit proto-oncogene encodes Kit, the tyrosine kinase receptor for stem cell factor (SCF). In the mouse, c-kit is allelic with the W locus, and it is widely used marker for enrichment of HSCs of adult bone marrow. In a mouse knockin experiment, the c-kit gene was inactivated and replaced by the *Escherichia coli lacZ* gene. In heterozygous embryos, expression of the lacZ gene was observed in hematopoietic progenitors of the blood islands of 8.5 dpc yolk sacs, and more interestingly also in 11.5 dpc AGM region. Expression in the AGM region was concentrated in the ventral part of the aorta in both endothelial and hematopoietic cells (Bernex et al., 1996). By using sorting and long-term transplantation techniques, it was found that c-kit marks all HSCs in both the AGM region and fetal liver in normal mouse embryos (Sanchez et al., 1996).

For many years, the Sca-1, GPI-linked cell surface glycoprotein has been used as a faithful marker of murine HSCs from various sources; i.e. adult bone marrow, fetal liver, and yolk sac (Huang and Auerbach, 1993; Spangrude et al., 1988). Sorting on the basis of the monoclonal antibody E13-161.7 specific for the Sca-1 epitope yields about a 100-fold enrichment of HSCs from adult bone marrow (Okada et al., 1992; Spangrude et al., 1988). Sca-1 is encoded by the strain specific allelic genes, *Ly-6E* and *Ly-6A* (Khan et al., 1990; Sinclair and Dzierzak, 1993; Stanford et al., 1992; van de Rijn et al., 1989) which are members of the large *Ly-6* gene family (Kamiura et al., 1992; LeClair et al., 1986). The relatively small size of the *Ly-6A/E* gene locus (14 kb) which may contain all the transcriptional regulatory elements of these alleles (Khan et al., 1990; Khan et al., 1993; Ma et al., 2001; McGrew and Rock, 1991; Sinclair et al., 1996; Sinclair and Dzierzak, 1993) has been used in directing *lacZ* marker gene expression in bone marrow HSCs in transgenic mice (Miles et al., 1997). In the Miles study, two constructs containing *Ly-6E* gene regulatory fragment, one 14 kb fragment containing 3'DNaseI hypersensitive sites +8.7/+8.9 and the other which deleted the segment containing the two hypersensitive sites, were used to drive *LacZ* gene for making transgenic mice. Several lines were established for each construct. Although the transgene copy number in some lines from deletion construct were comparable to those of non-deletion construct, tissue-specific *lacZ* transgene expression was only found in transgenic mice of non-deletion construct which contained 14 kb *Ly-6E* gene regulatory fragment. This transgenic line gave rise to hematopoietic and lymphoid cell expression in thymus, spleen, lymph node and bone marrow, and functional hematopoietic stem cells could be 100 fold enriched by FDG sorting of *lacZ* positive bone marrow cells. Furthermore, this line also gave rise to specific expression in embryonic hematopoietic tissue, in 11-12 dpc AGM region, especially in the location of pro/mesonephric tubules. However, X-gal stained, *lacZ* expression was not detected in cells around aorta and/or in the endothelial lining of the aorta, and the enrichment of hematopoietic stem cells from the AGM region was not successful by using the lacZ marker. According to the study of Spangrude and Brooks, only 25% hematopoietic stem cells express the *Ly-6E* gene and almost 100% hematopoietic stem cells express the *Ly-6A* gene (Spangrude and Brooks, 1993), it appears to be more advantageous to study the *Ly-6A* gene regulatory elements and use these regulatory elements to drive the expression of a marker gene so as to localize, isolate and manipulate hematopoietic stem cells from adult and embryonic hematopoietic tissues. My thesis is based on this idea.

3. The aim of this study

Hematopoietic stem cells are the clinically relevant cells involved in cell replacement therapies for leukemias and blood-related genetic diseases. Advances in such therapies, rely on our further understanding of the molecular, cellular and developmental properties of these cells so as to expand and manipulate them. During embryonic development HSC numbers increase spontaneously in the AGM region (Medvinsky and Dzierzak, 1996; Medvinsky and Dzierzak, 1998; Muller et al., 1994). Thus, this region offers a unique opportunity to examine the earliest generation of HSCs, particularly the signaling events and transcriptional regulation affecting HSCs.

The goals of my studies were 1) to examine the transcriptional regulatory elements of the Ly-6A/E gene which encodes one of the most widely used HSC markers, Sca-1 and 2) to use these regulatory elements direct expression of marker genes so as to localize and enrich for the earliest HSCs in the AGM region.

The Ly-6A/E gene was chosen since the Sca-1 protein expression pattern within the hematopoietic system is well characterized, because the gene is small (about 2 kb, with only 4 exons) and because some regulatory elements had already been described. In contrast other hematopoietic specific genes, c-kit, CD34 and AML-1, are much larger and information about their gene regulatory elements was limited at the outset of my project.

More specifically, the direct focus of my studies was to identify and isolate the definitive HSC control elements of the *Ly-6A* allele. Although the HSC specific control elements of the *Ly-6E* strain specific allele had been previously isolated and tested by *in vitro* and transgenic (with a *LacZ* marker) experiments (Miles et al., 1997; Sinclair et al., 1996), the isolation of HSCs from the AGM regions of *Ly-6E/lacZ* transgenics was not successful. The reasons for this failure could be: 1) The *Ly-6E* allele does not express highly in HSCs, as suggested by Spangrude and Brooks (Spangrude and Brooks, 1993); 2) The impact of variegated expression due to transgene integration site; 3) The multi-step isolation of HSCs from the AGM may influence the sensitivity of lacZ detection. Hence, my studies on the *Ly-6A* allele were aimed at more efficiently marking the HSCs in adult hematopoietic tissue and, more importantly, marking the first HSCs in the AGM region. In this thesis I present my studies in which I cloned the *Ly-6A* gene and related regulatory sequences. I identified by *in vitro* studies the *Ly-6A* gene fragment responsible for high-level hematopoietic cell expression and constructed the 14 kb expression cassette used in *in vivo* studies. Finally, I performed functional and localization studies on HSCs in *Ly-6A lacZ* and *GFP* transgenic adult mice and embryos.

4. Scope of the thesis

In Chapter 2, I describe studies on the cloning of the *Ly-6A* gene and gene expression studies in hematopoietic cells *in vitro*. 25 kb of the *Ly-6A* gene and flanking regulatory regions were isolated and constructs containing different DNaseI hypersensitive site fragments were transfected into MEL cells. The *Ly-6A* 14 kb sequences analogous to those in the *Ly-6E* allele were found responsible for high-level γ-IFN induced expression *in vitro*. It is also shown that the 3' distal *Ly-6A* fragment containing +8.7/+8.9 hypersensitive sites directs high level γ-IFN induced expression from a heterologous promoter, indicating that the 3' distal *Ly-6A* sequence contains a potent enhancer that could be useful for expression in hematopoietic stem cells *in vivo*.

In Chapter 3, I describe mouse transgenic studies in which the *in vivo* expression pattern of Ly-6A was examined. To study the specific expression pattern and hematopoietic regulation of the *Ly-6A* gene, a construct containing the 14 kb cassette from the genomic *Ly-6A* fragment was generated and used to drive the *lacZ* reporter gene in transgenic mice. Similar to *Ly-6E lacZ* transgenic mouse lines, we found that the *Ly-6A lacZ* transgene was expressed in the hematopoietic tissues in adult mice. Furthermore, functional hematopoietic stem cells from bone marrow can be enriched by sorting for β-galactosidase expressing cells. Interestingly, *Ly-6A lacZ* lines have a more wide embryonic expression pattern than the *Ly-6E lacZ* transgenic lines. Surprisingly, we did not find improvement of *lacZ* marker gene expression in hematopoietic cells when we compared *Ly-6A* to *Ly-6E* cassettes - not all HSCs are located within the lacZ-expressing cells and enrichment of embryonic HSCs by sorting FDG stained AGM cells was also not successful. Whether this is a result of the lacZ marker gene, the *Ly-6A/E* expression cassette, or the FDG substrate staining method is uncertain. Thus, the enhanced green fluorescent protein (GFP) gene was used within the *Ly-6A* cassette to further identify and localize HSCs in transgenic mice both in the adult and embryonic stage.

Ly-6A GFP transgenic mice were created and the transgene expression pattern has been studied both in adult mice (**Chapter 4**) and embryonic stage (**Chapter 5**). **In Chapter 4,** I show that the 14 kb *Ly-6A* expression cassette directs the transcription of the *GFP* marker gene in all function repopulating HSCs in the adult bone marrow and more than 100 fold enrichment of HSCs from bone marrow has been achieved by sorting *GFP* expressing cells. In addition, the *GFP* transgene expression pattern generally corresponded to that of endogenous Sca-1 protein in different hematopoietic tissues of adult mice. **In chapter 5,** instead of unsuccessful sorting HSCs by FDG staining AGM cells from *Ly-6A/E lacZ* transgenic embryos, enrichment of HSCs has been successfully performed by sorting GFP expressing cells. This *Ly-6A GFP* transgene marker is expressed in all functional HSCs in the midgestation aorta. Immunohistological study showed that GFP positive cells were specifically localized to the endothelial layer lining the wall of the dorsal aorta and the vitelline/umbilical arteries, suggesting that HSC activity arises within the endothelium of the vascular-structure during embryonic mid-gestation stage.

Taken together, the studies in this thesis demonstrate that the *Ly-6A* transgene expression cassette provides a useful tool for marking, isolation and localization of the first HSCs in the developing mouse. These transgenic mice could be useful in the characterization of the complete molecular program of these important cells.

References

Abramson, S., Miller, R.G. and Phillips, R.A. (1977) The identification in adult bone marrow of pluripotent and restricted stem cells of the myeloid and lymphoid systems. *J Exp Med*, **145**, 1567-79.

Aihara, Y., Buhring, H.J., Aihara, M. and Klein, J. (1986) An attempt to produce "pre-T" cell hybridomas and to identify their antigens. *Eur J Immunol*, **16**, 1391-9.

Bahrenberg, G., Brauers, A., Joost, H.G. and Jakse, G. (2000) Reduced expression of PSCA, a member of the LY-6 family of cell surface antigens, in bladder, esophagus, and stomach tumors. *Biochem Biophys Res Commun*, **275**, 783-8.

Bernex, F., De Sepulveda, P., Kress, C., Elbaz, C., Delouis, C. and Panthier, J.J. (1996) Spatial and temporal patterns of c-kit-expressing cells in WlacZ/+ and WlacZ/WlacZ mouse embryos. *Development*, **122**, 3023-33.

Boggs, D.R., Boggs, S.S., Saxe, D.F., Gress, L.A. and Canfield, D.R. (1982) Hematopoietic stem cells with high proliferative potential. Assay of their concentration in marrow by the frequency and duration of cure of W/Wv mice. *J Clin Invest*, **70**, 242-53.

Brakenhoff, R.H., Gerretsen, M., Knippels, E.M., van Dijk, M., van Essen, H., Weghuis, D.O., Sinke, R.J., Snow, G.B. and van Dongen, G.A. (1995) The human E48 antigen, highly homologous to the murine Ly-6 antigen ThB, is a GPI-anchored molecule apparently involved in keratinocyte cell-cell adhesion. *J Cell Biol*, **129**, 1677-89.

Brakenhoff, R.H., van Dijk, M., Rood-Knippels, E.M. and Snow, G.B. (1997) A gain of novel tissue specificity in the human Ly-6 gene E48. *J Immunol*, **159**, 4879-86.

Castilla, L.H., Wijmenga, C., Wang, Q., Stacy, T., Speck, N.A., Eckhaus, M., Marin-Padilla, M., Collins, F.S., Wynshaw-Boris, A. and Liu, P.P. (1996) Failure of embryonic hematopoiesis and lethal hemorrhages in mouse embryos heterozygous for a knocked-in leukemia gene CBFB-MYH11. *Cell*, **87**, 687-96.

Codias, E.K., Cray, C., Baler, R.D., Levy, R.B. and Malek, T.R. (1989) Expression of Ly-6A/E alloantigens in thymocyte and T-lymphocyte subsets: variability related to the Ly-6a and Ly-6b haplotypes. *Immunogenetics*, **29**, 98-107.

Cohn, M.A., Kramerov, D., Hulgaard, E.F. and Lukanidin, E.M. (1997) The differentiation antigen Ly-6E.1 is expressed in mouse metastatic tumor cell lines. *FEBS Lett*, **403**, 181-5.

Cormier, F., de Paz, P. and Dieterlen-Lievre, F. (1986) In vitro detection of cells with monocytic potentiality in the wall of the chick embryo aorta. *Dev Biol*, **118**, 167-75.

Cray, C., Keane, R.W., Malek, T.R. and Levy, R.B. (1990) Regulation and selective expression of Ly-6A/E, a lymphocyte activation molecule, in the central nervous system. *Brain Res Mol Brain Res*, **8**, 9-15.

Cumano, A., Dieterlen-Lievre, F. and Godin, I. (1996) Lymphoid potential, probed before circulation in mouse, is restricted to caudal intraembryonic splanchnopleura. *Cell*, **86**, 907-16.

de Bruijn, M.F., Ma, X., Robin, C., Ottersbach, K., Sanchez, M.J. and Dzierzak, E. (2002) Hematopoietic stem cells localize to the endothelial cell layer in the midgestation mouse aorta. *Immunity*, **16**, 673-683.

de Bruijn, M.F., Peeters, M.C., Luteijn, T., Visser, P., Speck, N.A. and Dzierzak, E. (2000a) CFU-S(11) activity does not localize solely with the aorta in the aorta- gonad-mesonephros region. *Blood*, **96**, 2902-4.

de Bruijn, M.F., Speck, N.A., Peeters, M.C. and Dzierzak, E. (2000b) Definitive hematopoietic stem cells first develop within the major arterial regions of the mouse embryo. *Embo J*, **19**, 2465-74.

de Nooij-van Dalen, A.G., van Dongen, G.A., Smeets, S.J., Nieuwenhuis, E.J., Stigter-van Walsum, M., Snow, G.B. and Brakenhoff, R.H. (2003) Characterization of the human Ly-6 antigens, the newly annotated member Ly-6K included, as molecular markers for head-and-neck squamous cell carcinoma. *Int J Cancer*, **103**, 768-74.

Dexter, T.M., Allen, T.D. and Lajtha, L.G. (1977) Conditions controlling the proliferation of haemopoietic stem cells in vitro. *J Cell Physiol*, **91**, 335-44.

Dick, J.E., Magli, M.C., Huszar, D., Phillips, R.A. and Bernstein, A. (1985) Introduction of a selectable gene into primitive stem cells capable of long-term reconstitution of the hemopoietic system of W/Wv mice. *Cell*, **42**, 71-9.

Dieterlen-Lievre, F. (1975) On the origin of haemopoietic stem cells in the avian embryo: an experimental approach. *J Embryol Exp Morphol*, **33**, 607-19.

Down, J.D., Tarbell, N.J., Thames, H.D. and Mauch, P.M. (1991) Syngeneic and allogeneic bone marrow engraftment after total body irradiation: dependence on dose, dose rate, and fractionation. *Blood*, **77**, 661-9.

Drake, C.J. and Fleming, P.A. (2000) Vasculogenesis in the day 6.5 to 9.5 mouse embryo. *Blood*, **95**, 1671-9.

Dzierzak, E. and Medvinsky, A. (1995) Mouse embryonic hematopoiesis. *Trends Genet*, **11**, 359-66.

Dzierzak, E.A. and Mulligan, R.C. (1988) Lineage specific expression of a human beta-globin gene in murine bone marrow transplant recipients. *Adv Exp Med Biol*, **241**, 41-3.

Fleming, T.J., O'HUigin, C. and Malek, T.R. (1993) Characterization of two novel Ly-6 genes. Protein sequence and potential structural similarity to alpha-bungarotoxin and other neurotoxins. *J Immunol*, **150**, 5379-90.

Garcia-Porrero, J.A., Godin, I.E. and Dieterlen-Lievre, F. (1995) Potential intraembryonic hemogenic sites at pre-liver stages in the mouse. *Anat Embryol (Berl)*, **192**, 425-35.

Garcia-Porrero, J.A., Manaia, A., Jimeno, J., Lasky, L.L., Dieterlen-Lievre, F. and Godin, I.E. (1998) Antigenic profiles of endothelial and hemopoietic lineages in murine intraembryonic hemogenic sites. *Dev Comp Immunol*, **22**, 303-19.

Godin, I., Dieterlen-Lievre, F. and Cumano, A. (1995a) B-lymphoid potential in pre-liver mouse embryo. *Semin Immunol*, **7**, 131-41.

Godin, I., Dieterlen-Lievre, F. and Cumano, A. (1995b) Emergence of multipotent hemopoietic cells in the yolk sac and paraaortic splanchnopleura in mouse embryos, beginning at 8.5 days postcoitus. *Proc Natl Acad Sci U S A*, **92**, 773-7.

Godin, I.E., Garcia-Porrero, J.A., Coutinho, A., Dieterlen-Lievre, F. and Marcos, M.A. (1993) Para-aortic splanchnopleura from early mouse embryos contains B1a cell progenitors. *Nature*, **364**, 67-70.

Gumley, T.P., McKenzie, I.F., Kozak, C.A. and Sandrin, M.S. (1992) Isolation and characterization of cDNA clones for the mouse thymocyte B cell antigen (ThB). *J Immunol*, **149**, 2615-8.

Gumley, T.P., McKenzie, I.F. and Sandrin, M.S. (1995) Tissue expression, structure and function of the murine Ly-6 family of molecules. *Immunol Cell Biol*, **73**, 277-96.

Harrison, D.E. and Lerner, C.P. (1991) Most primitive hematopoietic stem cells are stimulated to cycle rapidly after treatment with 5-fluorouracil. *Blood*, **78**, 1237-40.

Herzenberg, L.A. and Sweet, R.G. (1976) Fluorescence-activated cell sorting. *Sci Am*, **234**, 108-17.

HogenEsch, H., de Geus, B., Tielen, F. and Rozing, J. (1993) Constitutive expression of Ly-6.A2 on murine keratinocytes and inducible expression on TCR gamma delta+ dendritic epidermal T cells. *J Dermatol Sci*, **5**, 114-21.

Houlden, B.A., Hogarth, P.M. and McKenzie, I.F. (1986) Interrelationships of the "Ly-6 complex" antigens. *Immunogenetics*, **23**, 226-32.

Houssaint, E. (1981) Differentiation of the mouse hepatic primordium. II. Extrinsic origin of the haemopoietic cell line. *Cell Differ*, **10**, 243-52.

Huang, H. and Auerbach, R. (1993) Identification and characterization of hematopoietic stem cells from the yolk sac of the early mouse embryo. *Proc Natl Acad Sci U S A*, **90**, 10110-4.

Ikuta, K. and Weissman, I.L. (1992) Evidence that hematopoietic stem cells express mouse c-kit but do not depend on steel factor for their generation. *Proc Natl Acad Sci U S A*, **89**, 1502-6.

Jaffredo, T., Gautier, R., Eichmann, A. and Dieterlen-Lievre, F. (1998) Intraaortic hemopoietic cells are derived from endothelial cells during ontogeny. *Development*, **125**, 4575-83.

Johnson, G.R. and Moore, M.A. (1975) Role of stem cell migration in initiation of mouse foetal liver haemopoiesis. *Nature*, **258**, 726-8.Jones, R.J., Sharkis, S.J., Celano, P., Colvin, O.M., Rowley, S.D. and Sensenbrenner, L.L. (1987) Progenitor cell assays predict hematopoietic reconstitution after syngeneic transplantation in mice. *Blood*, **70**, 1186-92.

Jones, R.J., Wagner, J.E., Celano, P., Zicha, M.S. and Sharkis, S.J. (1990) Separation of pluripotent haematopoietic stem cells from spleen colony-forming cells. *Nature*, **347**, 188-9.

Jordan, C.T., Astle, C.M., Zawadzki, J., Mackarehtschian, K., Lemischka, I.R. and Harrison, D.E. (1995) Long-term repopulating abilities of enriched fetal liver stem cells measured by competitive repopulation. *Exp Hematol*, **23**, 1011-1015.

Jurecic, R., Van, N.T. and Belmont, J.W. (1993) Enrichment and functional characterization of Sca-1+WGA+, Lin-WGA+, Lin- Sca-1+, and Lin-Sca-1+WGA+ bone marrow cells from mice with an Ly-6a haplotype. *Blood*, **82**, 2673-83.

Kamiura, S., Nolan, C.M. and Meruelo, D. (1992) Long-range physical map of the Ly-6 complex: mapping the Ly-6 multigene family by field-inversion and two-dimensional gel electrophoresis. *Genomics*, **12**, 89-105.

Keller, G. (1992) Clonal analysis of hematopoietic stem cell development in vivo. *Curr Top Microbiol Immunol*, **177**, 41-57.

Keller, G., Paige, C., Gilboa, E. and Wagner, E.F. (1985) Expression of a foreign gene in myeloid and lymphoid cells derived from multipotent haematopoietic precursors. *Nature*, **318**, 149-54.

Khan, K.D., Lindwall, G., Maher, S.E. and Bothwell, A.L. (1990) Characterization of promoter elements of an interferon-inducible Ly- 6E/A differentiation antigen, which is expressed on activated T cells and hematopoietic stem cells. *Mol Cell Biol*, **10**, 5150-9.

Khan, K.D., Shuai, K., Lindwall, G., Maher, S.E., Darnell, J.E., Jr. and Bothwell, A.L. (1993) Induction of the Ly-6A/E gene by interferon alpha/beta and gamma requires a DNA element to which a tyrosine-phosphorylated 91-kDa protein binds. *Proc Natl Acad Sci U S A*, **90**, 6806-10.

Kimura, S., Tada, N., Liu-Lam, Y. and Hammerling, U. (1984) Studies of the mouse Ly-6 alloantigen system. II. Complexities of the Ly-6 region. *Immunogenetics*, **20**, 47-56.

Le Douarin, N.M. (1973) A biological cell labelling technique and its use in experimental embryology. *Dve. Bio.*, **30**, 217-230.

LeClair, K.P., Palfree, R.G., Flood, P.M., Hammerling, U. and Bothwell, A. (1986) Isolation of a murine Ly-6 cDNA reveals a new multigene family. *Embo J*, **5**, 3227-34.

Lemieux, M.E., Rebel, V.I., Lansdorp, P.M. and Eaves, C.J. (1995) Characterization and purification of a primitive hematopoietic cell type in adult mouse marrow capable of lymphomyeloid differentiation in long-term marrow "switch" cultures. *Blood*, **86**, 1339-47.

Lemischka, I.R., Raulet, D.H. and Mulligan, R.C. (1986) Developmental potential and dynamic behavior of hematopoietic stem cells. *Cell*, **45**, 917-27.

Li, C.L. and Johnson, G.R. (1992) Long-term hemopoietic repopulation by Thy-1lo, Lin-, Ly6A/E+ cells. *Exp Hematol*, **20**, 1309-15.

Liu, C.P. and Auerbach, R. (1991) In vitro development of murine T cells from prethymic and preliver embryonic yolk sac hematopoietic stem cells. *Development*, **113**, 1315-23.

Ma, X., Ling, K.W. and Dzierzak, E. (2001) Cloning of the Ly-6A (Sca-1) gene locus and identification of a 3' distal fragment responsible for high-level gamma-interferon-induced expression in vitro. *Br J Haematol*, **114**, 724-30.

MacNeil, I., Kennedy, J., Godfrey, D.I., Jenkins, N.A., Masciantonio, M., Mineo, C., Gilbert, D.J., Copeland, N.G., Boyd, R.L. and Zlotnik, A. (1993) Isolation of a cDNA encoding thymic shared antigen-1. A new member of the Ly6 family with a possible role in T cell development. *J Immunol*, **151**, 6913-23.

Magli, M.C., Iscove, N.N. and Odartchenko, N. (1982) Transient nature of early haematopoietic spleen colonies. *Nature*, **295**, 527-9.

Mao, M., Yu, M., Tong, J.H., Ye, J., Zhu, J., Huang, Q.H., Fu, G., Yu, L., Zhao, S.Y., Waxman, S., Lanotte, M., Wang, Z.Y., Tan, J.Z., Chan, S.J. and Chen, Z. (1996) RIG-E, a human homolog of the murine Ly-6 family, is induced by retinoic acid during the differentiation of acute promyelocytic leukemia cell. *Proc Natl Acad Sci U S A*, **93**, 5910-4.

Marshall, C.J., Moore, R.L., Thorogood, P., Brickell, P.M., Kinnon, C. and Thrasher, A.J. (1999) Detailed characterization of the human aorta-gonad-mesonephros region reveals morphological polarity resembling a hematopoietic stromal layer. *Dev Dyn*, **215**, 139-47.

Matthews, W., Jordan, C.T., Gavin, M., Jenkins, N.A., Copeland, N.G. and Lemischka, I.R. (1991) A receptor tyrosine kinase cDNA isolated from a population of enriched primitive hematopoietic cells and exhibiting close genetic linkage to c- kit. *Proc Natl Acad Sci U S A*, **88**, 9026-30.

McGrew, J.T. and Rock, K.L. (1991) Isolation, expression, and sequence of the TAP/Ly-6A.2 chromosomal gene. *J Immunol*, **146**, 3633-8.

Medvinsky, A. and Dzierzak, E. (1996) Definitive hematopoiesis is autonomously initiated by the AGM region. *Cell*, **86**, 897-906.

Medvinsky, A.L. and Dzierzak, E.A. (1998) Development of the definitive hematopoietic hierarchy in the mouse. *Dev Comp Immunol*, **22**, 289-301.

Medvinsky, A.L., Gan, O.I., Semenova, M.L. and Samoylina, N.L. (1996) Development of day-8 colony-forming unit-spleen hematopoietic progenitors during early murine embryogenesis: spatial and temporal mapping. *Blood* **87**, 557-66

Medvinsky, A. L., Samoylina, N. L., Muller, A. M., Dzierzak, E. A. (1993) An early pre-liver intraembryonic source of CFU-S in the developing mouse. *Nature* **364**, 64-67

Metcalf, D., Nicola, N. A. (1984) The regulatory factors controlling murine erythropoiesis in vitro. *Prog Clin Biol Res* **148**, 93-105

Miles, C., Sanchez, M.J., Sinclair, A. and Dzierzak, E. (1997) Expression of the Ly-6E.1 (Sca-1) transgene in adult hematopoietic stem cells and the developing mouse embryo. *Development*, **124**, 537-47.

Minegishi, N., Ohta, J., Yamagiwa, H., Suzuki, N., Kawauchi, S., Zhou, Y., Takahashi, S., Hayashi, N., Engel, J.D. and Yamamoto, M. (1999) The mouse GATA-2 gene is expressed in the para-aortic splanchnopleura and aorta-gonads and mesonephros region. *Blood*, **93**, 4196-207.

Moore, M.A. and Metcalf, D. (1970) Ontogeny of the haemopoietic system: yolk sac origin of in vivo and in vitro colony forming cells in the developing mouse embryo. *Br J Haematol*, **18**, 279-96.

Muller, A.M., Medvinsky, A., Strouboulis, J., Grosveld, F. and Dzierzak, E. (1994) Development of hematopoietic stem cell activity in the mouse embryo. *Immunity*, **1**, 291-301.

Muller-Sieburg, C.E., Whitlock, C.A. and Weissman, I.L. (1986) Isolation of two early B lymphocyte progenitors from mouse marrow: a committed pre-pre-B cell and a clonogenic Thy-1-lo hematopoietic stem cell. *Cell*, **44**, 653-62.

Neben, S., Anklesaria, P., Greenberger, J. and Mauch, P. (1993) Quantitation of murine hematopoietic stem cells in vitro by limiting dilution analysis of cobblestone area formation on a clonal stromal cell line. *Exp Hematol*, **21**, 438-43.

Neben, S., Redfearn, W.J., Parra, M., Brecher, G. and Pallavicini, M.G. (1991) Short- and long-term repopulation of lethally irradiated mice by bone marrow stem cells enriched on the basis of light scatter and Hoechst 33342 fluorescence. *Exp Hematol*, **19**, 958-67.

North, T.E., de Bruijn, M.F., Stacy, T., Talebian, L., Lind, E., Robin, C., Binder, M., Dzierzak, E. and Speck, N.A. (2002) Runx1 expression marks long-term repopulating hematopoietic stem cells in the midgestation mouse embryo. *Immunity*, **16**, 661-672.

North, T., Gu, T.L., Stacy, T., Wang, Q., Howard, L., Binder, M., Marin-Padilla, M. and Speck, N.A. (1999) Cbfa2 is required for the formation of intra-aortic hematopoietic clusters. *Development*, **126**, 2563-75.

Ogawa, M., Nishikawa, S., Ikuta, K., Yamamura, F., Naito, M. and Takahashi, K. (1988) B cell ontogeny in murine embryo studied by a culture system with the monolayer of a stromal cell clone, ST2: B cell progenitor develops first in the embryonal body rather than in the yolk sac. *Embo J*, 7, 1337-43.

Okada, S., Nakauchi, H., Nagayoshi, K., Nishikawa, S., Miura, Y. and Suda, T. (1991) Enrichment and characterization of murine hematopoietic stem cells that express c-kit molecule. *Blood*, **78**, 1706-12.

Okada, S., Nakauchi, H., Nagayoshi, K., Nishikawa, S., Miura, Y. and Suda, T. (1992) In vivo and in vitro stem cell function of c-kit- and Sca-1-positive murine hematopoietic cells. *Blood*, **80**, 3044-50.

Okuda, T., van Deursen, J., Hiebert, S.W., Grosveld, G. and Downing, J.R. (1996) AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. *Cell*, **84**, 321-30.

Orlic, D. and Bodine, D.M. (1994) What defines a pluripotent hematopoietic stem cell (PHSC): will the real PHSC please stand up! *Blood*, **84**, 3991-4.

Palfree, R.G., Dumont, F.J. and Hammerling, U. (1986) Ly-6A.2 and Ly-6E.1 molecules are antithetical and identical to MALA-1. *Immunogenetics*, **23**, 197-207.

Palfree, R.G., LeClair, K.P., Bothwell, A. and Hammerling, U. (1987) cDNA characterization of an Ly-6.2 gene expressed in BW5147 tumor cells. *Immunogenetics*, **26**, 389-91.

Palfree, R.G., Sirlin, S., Dumont, F.J. and Hammerling, U. (1988) N-terminal and cDNA characterization of murine lymphocyte antigen Ly- 6C.2. *J Immunol*, **140**, 305-10.

Patterson, J.M., Johnson, M.H., Zimonjic, D.B. and Graubert, T.A. (2000) Characterization of Ly-6M, a novel member of the Ly-6 family of hematopoietic proteins. *Blood*, **95**, 3125-32.

Pflugh, D.L., Maher, S.E. and Bothwell, A.L. (2000) Ly-6I, a new member of the murine Ly-6 superfamily with a distinct pattern of expression. *J Immunol*, **165**, 313-21.

Porcher, C., Swat, W., Rockwell, K., Fujiwara, Y., Alt, F.W. and Orkin, S.H. (1996) The T cell leukemia oncoprotein SCL/tal-1 is essential for development of all hematopoietic lineages. *Cell*, **86**, 47-57.

Reiser, H., Coligan, J., Palmer, E., Benacerraf, B. and Rock, K.L. (1988) Cloning and expression of a cDNA for the T-cell-activating protein TAP. *Proc Natl Acad Sci U S A*, **85**, 2255-9.

Rock, K.L., Yeh, E.T., Gramm, C.F., Haber, S.I., Reiser, H. and Benacerraf, B. (1986) TAP, a novel T cell-activating protein involved in the stimulation of MHC-restricted T lymphocytes. *J Exp Med*, **163**, 315-33.

Russell, E.S.a.B., S. E. (1966) *in Biology of the Laboratory Mouse*, 351-372.

Sanchez, M., Gottgens, B., Sinclair, A.M., Stanley, M., Begley, C.G., Hunter, S. and Green, A.R. (1999) An SCL 3' enhancer targets developing endothelium together with embryonic and adult haematopoietic progenitors. *Development*, **126**, 3891-904.

Sanchez, M.J., Holmes, A., Miles, C. and Dzierzak, E. (1996) Characterization of the first definitive hematopoietic stem cells in the AGM and liver of the mouse embryo. *Immunity*, **5**, 513-25.

Shalaby, F., Ho, J., Stanford, W.L., Fischer, K.D., Schuh, A.C., Schwartz, L., Bernstein, A. and Rossant, J. (1997) A requirement for Flk1 in primitive and definitive hematopoiesis and vasculogenesis. *Cell*, **89**, 981-90.

Shalaby, F., Rossant, J., Yamaguchi, T.P., Gertsenstein, M., Wu, X.F., Breitman, M.L. and Schuh, A.C. (1995) Failure of blood-island formation and vasculogenesis in Flk-1-deficient mice. *Nature*, **376**, 62-6.

Shivdasani, R.A., Mayer, E.L. and Orkin, S.H. (1995) Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL. *Nature*, **373**, 432-4.

Siminovitch L., M.E.A., Till J.E. (1963) The distribution of colony-forming ability of marrow cells among spleen colonies. *J.Cell.Comp.Physiol.*, **62**, 327-336.

Siminovitch L., T.J.E., McCulloch E.A. (1964) Decline in colony-forming ability of marrow cells subjected to serial transplantation into irradiated mice. *J.Cell.Comp.Physiol.*, **64**, 23-32.

Sinclair, A., Daly, B. and Dzierzak, E. (1996) The Ly-6E.1 (Sca-1) gene requires a 3' chromatin-dependent region for high-level gamma-interferon-induced hematopoietic cell expression. *Blood*, **87**, 2750-61.

Sinclair, A.M. and Dzierzak, E.A. (1993) Cloning of the complete Ly-6E.1 gene and identification of DNase I hypersensitive sites corresponding to expression in hematopoietic cells. *Blood*, **82**, 3052-62.

Smith, L.G., Weissman, I.L. and Heimfeld, S. (1991) Clonal analysis of hematopoietic stem-cell differentiation in vivo. *Proc Natl Acad Sci U S A*, **88**, 2788-92.

Sowala, H., Wunder, E. and Henon, P. (1990) Purification and characterisation of the CD34+ hematopoietic precursor cell population by flow cytometry. *Bone Marrow Transplant*, **5 Suppl 1**, 9-10.

Spangrude, G.J. (1989) Enrichment of murine haemopoietic stem cells: diverging roads. *Immunol Today*, **10**, 344-50.

Spangrude, G.J. and Brooks, D.M. (1992) Phenotypic analysis of mouse hematopoietic stem cells shows a Thy-1- negative subset. *Blood*, **80**, 1957-64.

Spangrude, G.J. and Brooks, D.M. (1993) Mouse strain variability in the expression of the hematopoietic stem cell antigen Ly-6A/E by bone marrow cells. *Blood*, **82**, 3327-32.

Spangrude, G.J., Heimfeld, S. and Weissman, I.L. (1988) Purification and characterization of mouse hematopoietic stem cells. *Science*, **241**, 58-62.

Spangrude, G.J. and Scollay, R. (1990) A simplified method for enrichment of mouse hematopoietic stem cells. *Exp Hematol*, **18**, 920-6.

Stanford, W.L., Bruyns, E. and Snodgrass, H.R. (1992) The isolation and sequence of the chromosomal gene and regulatory regions of Ly-6A.2. *Immunogenetics*, **35**, 408-11.

Tavian, M., Coulombel, L., Luton, D., Clemente, H.S., Dieterlen-Lievre, F. and Peault, B. (1996) Aorta-associated CD34+ hematopoietic cells in the early human embryo. *Blood*, **87**, 67-72.

Tavian, M., Hallais, M.F. and Peault, B. (1999) Emergence of intraembryonic hematopoietic precursors in the pre-liver human embryo. *Development*, **126**, 793-803.

Till, J.E.M. and McCulloch E.A. (1961) A direct measurement of the radiation sensitivityof normal mouse bonw marrow cells. *Radia. Res.*, **14**, 213-222.

Toles, J.F., Chui, D.H., Belbeck, L.W., Starr, E. and Barker, J.E. (1989) Hemopoietic stem cells in murine embryonic yolk sac and peripheral blood. *Proc Natl Acad Sci U S A*, **86**, 7456-9.

Tsai, F.Y., Keller, G., Kuo, F.C., Weiss, M., Chen, J., Rosenblatt, M., Alt, F.W. and Orkin, S.H. (1994) An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature*, **371**, 221-6.

Tucek, C.L., Godfrey, D.I. and Boyd, R.L. (1992) Five novel antigens illustrate shared phenotype between mouse thymic stromal cells, thymocytes, and peripheral lymphocytes. *Int Immunol*, **4**, 1021-30.

Uchida, N. and Weissman, I.L. (1992) Searching for hematopoietic stem cells: evidence that Thy-1.1lo Lin- Sca-1+ cells are the only stem cells in C57BL/Ka-Thy-1.1 bone marrow. *J Exp Med*, **175**, 175-84.

van de Rijn, M., Heimfeld, S., Spangrude, G.J. and Weissman, I.L. (1989) Mouse hematopoietic stem-cell antigen Sca-1 is a member of the Ly-6 antigen family. *Proc Natl Acad Sci U S A*, **86**, 4634-8.

van der Loo, J.C., Slieker, W.A., Kieboom, D. and Ploemacher, R.E. (1995) Identification of hematopoietic stem cell subsets on the basis of their primitiveness using antibody ER-MP12. *Blood*, **85**, 952-62.

van der Loo, J.C., van den Bos, C., Baert, M.R., Wagemaker, G. and Ploemacher, R.E. (1994) Stable multilineage hematopoietic chimerism in alpha-thalassemic mice induced by a bone marrow subpopulation that excludes the majority of day-12 spleen colony-forming units. *Blood*, **83**, 1769-77.

Visser, J.W., Bauman, J.G., Mulder, A.H., Eliason, J.F. and de Leeuw, A.M. (1984) Isolation of murine pluripotent hemopoietic stem cells. *J Exp Med*, **159**, 1576-90.

Wang, Q., Stacy, T., Binder, M., Marin-Padilla, M., Sharpe, A.H. and Speck, N.A. (1996) Disruption of the Cbfa2 gene causes necrosis and hemorrhaging in the central nervous system and blocks definitive hematopoiesis. *Proc Natl Acad Sci U S A*, **93**, 3444-9.

Wiesmann, A., Phillips, R.L., Mojica, M., Pierce, L.J., Searles, A.E., Spangrude, G.J. and Lemischka, I. (2000) Expression of CD27 on murine hematopoietic stem and progenitor cells. *Immunity*, **12**, 193-199.

Williams, A.F., Tse, A.G. and Gagnon, J. (1988) Squid glycoproteins with structural similarities to Thy-1 and Ly-6 antigens. *Immunogenetics*, **27**, 265-72.

Wood, H.B., May, G., Healy, L., Enver, T. and Morriss-Kay, G.M. (1997) CD34 expression patterns during early mouse development are related to

modes of blood vessel formation and reveal additional sites of hematopoiesis. *Blood*, **90**, 2300-11.

Worton, R.G., McCulloch, E.A. and Till, J.E. (1969) Physical separation of hemopoietic stem cells differing in their capacity for self-renewal. *J Exp Med*, **130**, 91-103.

Wu, A.M., Till, J.E., Siminovitch, L. and McCulloch, E.A. (1967) A cytological study of the capacity for differentiation of normal hemopoietic colony-forming cells. *J Cell Physiol*, **69**, 177-84.

Wu, A.M., Till, J.E., Siminovitch, L. and McCulloch, E.A. (1968) Cytological evidence for a relationship between normal hemotopoietic colony-forming cells and cells of the lymphoid system. *J Exp Med*, **127**, 455-64.

Yamaguchi, T.P., Dumont, D.J., Conlon, R.A., Breitman, M.L. and Rossant, J. (1993) flk-1, an flt-related receptor tyrosine kinase is an early marker for endothelial cell precursors. *Development*, **118**, 489-98.

Yoder, M.C., Hiatt, K., Dutt, P., Mukherjee, P., Bodine, D.M. and Orlic, D. (1997a) Characterization of definitive lymphohematopoietic stem cells in the day 9 murine yolk sac. *Immunity*, **7**, 335-44.

Yoder, M.C., Hiatt, K. and Mukherjee, P. (1997b) In vivo repopulating hematopoietic stem cells are present in the murine yolk sac at day 9.0 postcoitus. *Proc Natl Acad Sci U S A*, **94**, 6776-80.

Chapter 2

Cloning of the *Ly-6A* gene locus and identification of a 3' distal fragment responsible for high-level γ-interferon-induced expression in vitro

Xiaoqian Ma, Kam-Wing Ling and Elaine Dzierzak

British Journal of Haematology 2001; 114: 724-730

Chapter 2

Cloning of the *Ly-6A* (Sca-1) gene locus and identification of a 3′ distal fragment responsible for high-level γ-interferon-induced expression *in vitro*

XIAOQIAN MA, KAM-WING LING AND ELAINE DZIERZAK *Department of Cell Biology and Genetics, Erasmus University, Rotterdam, The Netherlands*

Received 28 March 2001; accepted for publication 15 May 2001

Summary. The *Ly-6A and Ly-6E* allelic genes encode the Sca-1 protein, which is one of the most widely used markers in haematopoietic stem cell isolation procedures. Identification of the specific gene regulatory elements that direct haematopoietic stem cell specific expression of Sca-1 is of current interest for purposes of stem cell manipulation. Both the *Ly-6E and Ly-6A* alleles have been examined for regions containing DNase I hypersensitive sites thought to be indicative of transcriptional regulatory elements. In these previous studies, the *Ly-6E* allele with its flanking regulatory sequences was cloned, and the region responsible for high-level γ-interferon (γ-IFN)-induced expression was localized to a 3′ distal sequence containing two strong DNaseI hypersensitive sites. Because the *Ly-6A* allele is thought to provide higher levels of expression in haematopoietic stem cells, we isolated over 25 kb of the *Ly-6A* gene and flanking regulatory regions. We show here that sequences analogous to those in the *Ly-6E* allele are responsible for high-level γ-IFN-induced expression *in vitro*. Furthermore, we show that this 3′ distal *Ly-6A* fragment directs high-level γ-IFN-induced expression from a heterologous promoter, suggesting that it is a potent enhancer that could be useful for expression in haematopoietic stem cells.

Keywords: Ly-6A, Sca-1, gene regulation, haematopoiesis.

The Sca-1 phosphatidylinositol-linked cell surface glycoprotein marker has been instrumental in the isolation of haematopoietic stem cells for biological studies (Spangrude *et al*, 1988; Okada *et al*, 1992). The Sca-1 protein is encoded by the strain-specific allelic genes *Ly-6E and Ly-6A* (van de Rijn *et al*, 1989; Khan *et al*, 1990; Stanford *et al*, 1992; Sinclair & Dzierzak, 1993), which are members of the multigenic *Ly-6* family (LeClair *et al*, 1986; Kamiura *et al*, 1992). The family consists of at least 18 highly homologous cross-hybridizing genes. However, studies using antibodies specific for some of the Ly-6 proteins have shown diverse patterns of tissue distribution (Kimura *et al*, 1984). Functional studies have been difficult (Bamezai & Rock, 1995; Stanford *et al*, 1997) because of the homologies of the genes and proteins and the suspected redundant/overlapping roles played by the members of this family.

Ly-6A and Ly-6E protein expression within the haematopoietic system is complex (Kimura *et al*, 1984; van de Rijn *et al*, 1989; Spangrude & Brooks, 1993). The allelic genes differ by only three nucleotides in the coding sequence, resulting in two amino acid changes in the proteins (LeClair *et al*, 1986; Palfree & Hammerling, 1986; Rock *et al*, 1986; Palfree *et al*, 1987; Reiser *et al*, 1988). Both Ly-6E and Ly-6A proteins express the epitope recognized by the Sca-1 antibody. The allelic genes are interferon inducible, but the *Ly-6A* allele appears to be more widely and highly expressed (Kimura *et al*, 1984; Rock *et al*, 1986; Spangrude & Brooks, 1993). For example, Ly-6A strains of mice express this protein on 10–20% of adult thymocytes and 50–70% of peripheral T lymphocytes, whereas Ly-6E strains express the protein on 5–10% adult thymocytes and 10–15% of peripheral T lymphocytes. Similarly, and more importantly, expression of these allelic products has been shown to differ on haematopoietic stem cells. Using the marrow-repopulating assay as a readout, it was shown that Ly-6A strains of mice express Sca-1 on virtually all (99%) haematopoietic stem cells, whereas Ly-6E strains express Sca-1 on only 25% of these cells (Spangrude & Brooks, 1993). Hence, for purposes of molecular manipulation of haematopoietic stem cells, the use of the Ly-6A regulatory elements would appear to be advantageous and the study of the regulatory elements leading to differences in expression patterns and levels of the two alleles is important.

From previous molecular studies, 5′ *cis*-acting elements

Correspondence: Professor E. Dzierzak, Department of Cell Biology and Genetics, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands. E-mail: dzierzak@ch1.fgg.eur.nl

involved in regulating expression and responsiveness of the *Ly-6E* gene to interferons have been identified (Khan *et al*, 1990, 1993). In promoter deletion studies of the *Ly-6E* gene in mouse fibroblast transfectants, an interferon-inducible (GAS) element has been localized between −1·76 and −0·9 kb, and a purine-rich sequence at −0·11 kb was found to be necessary for basal expression. Also, a putative repressor sequence was found between −0·16 and −0·13 kb. Through sequence comparisons, similar control elements are thought to be present in the *Ly-6A* promoter (McGrew & Rock, 1991). Interestingly, DNase I hypersensitive site (HSS) analysis has identified two sites in the promoter region at −1·2 and −0·1 kb, which appear to coincide with the regulatory elements localized by the 5′ deletional analysis (Sinclair & Dzierzak, 1993). Analysis of the 3′ flanking regions of both the *Ly-6E and Ly-6A* alleles has localized six additional areas of the DNase I hypersensitivity (+3·3, +4·9, +5·6, +6·7, +8·7 and +8·9) (Sinclair & Dzierzak, 1993). Deletional studies of the *Ly-6E* 3′ flanking sequences in murine erythroleukaemia (MEL) cell transfectants revealed that the region containing the two most distal 3′ HSS are responsible for high-level γ-interferon (γ-IFN)-induced expression (Sinclair *et al*, 1996). Furthermore, in transgenic mice, this 3′ distal region is required for expression in haematopoietic stem cells (Miles *et al*, 1997). These results together with the observation of the rapid induction of DNase I hypersensitive sites in the 3′ flanking regions of the *Ly-6A* allele, particularly at +8·7 and +8·9 (Sinclair *et al*, 1996), suggest that *Ly-6A* transcriptional regulatory elements are generally similar to those of the *Ly-6E* allele. However, some subtle differences between the *Ly-6E and Ly-6A* alleles were observed in the hypersensitive site maps (Sinclair *et al*, 1996).

Thus, with the long-term goal of determining what sequences direct high-level haematopoietic stem cell expression, we set out to investigate the 3′ distal sequences of the *Ly-6A* gene for transcriptional-enhancing activity in haematopoietic cell transfectants. We describe here the isolation of more than 25 kb of the *Ly-6A* gene and flanking regulatory regions. Deletional analyses performed on the 3′ sequences show that the most distal downstream region, analogous to that in the *Ly-6E* allele containing the +8·7 and +8·9 HSS, is responsible for high-level γ-IFN induced expression *in vitro* in haematopoietic cells. Furthermore, we show that this 3′ distal *Ly-6A* fragment directs high-level γ-IFN-induced expression from a heterologous promoter, suggesting that this transcriptional-enhancing element could be useful in stem cell-directed molecular manipulations.

MATERIALS AND METHODS

Cloning. Genomic DNA from mouse strain 129 was cut with *Sau*3A and ligated into EMBL-3 lambda phage DNA. The library consisted of $3{\cdot}5 \times 10^6$ independent phage clones and was a generous gift from Dr D. Meijer (Rotterdam, The Netherlands). Seven plates of 5×10^5 clones were blotted onto filters and screened with a random oligonucleotide P^{32}-labelled 760 bp *Eco*R1 fragment from the pCD2/Ly-6E.1 plasmid (LeClair *et al*, 1986) containing the cDNA for the *Ly-6E* gene. Restriction maps of the two overlapping clones XM1 and XM2 were generated by standard methods, and four subclones were generated in plasmid vector pUC19. For the generation of expression cassette Ly-6A14, a *Cla*1 site was created in the first untranslated Ly-6A exon using primers spanning a sequence from the 5′ *Spe*1 site, to the sequence in which the *Cla*1 site was generated (5′-CAAAAACTAGTAAAG GGCTGAGCA-3′ and 5′-TAATCGATGGTGTGAGGAGGGA-3′) and primers spanning the sequence from the created *Cla*1 site to the *Vsp*1 site (5′-CACACCATCGATTACTTCCTCTC- 3′ and 5′-CAATTAATAGACCCCCAATCACAT-3′).

Fragments of 240 bp and 210 bp, respectively, were amplified from XM2·2 DNA. These fragments were cut with *Spe*1 and *Cla*1 or *Vsp*1 and *Cla*1, respectively, and ligated into subclone XM2·2 in which the *Spe*1–*Vsp*1 fragment had been removed. Ly-6A14 was then generated from ligation of the following fragments: *Bam*H1–*Spe*1 of XM1·1, *Spe*1–*Hin*dIII of XM2·2 containing the *Cla*1 site, *Hin*dIII–*Eco*R1 of XM2·3, *Eco*R1–*Bam*H1 of XM2·4 and *Bam*H1 cut pPolyIII.

Deletion construct generation. PL2 was constructed from a 1·8 kb *Kpn*1–*Cla*1 fragment of the XM2·2 subclone, and ligated to the *hGH* gene in pϕGH (Selden *et al*, 1986). PL17 and PL10 were made by ligating a 5·4 kb *Eco*R1–*Kpn*1 and 7·4 kb *Eco*R1–*Bam*HI fragment, respectively, of Ly-6A14 downstream to the *hGH* gene in PL2. PL19 was made by ligating the 1·5 kb *Hin*dIII–*Kpn*1 fragment of XM2·4 to the TK promoter followed by the *hGH* reporter gene (Selden *et al*, 1986). For transfections, plasmid DNAs were linearized with Pst1 and, after agarose gel separation, fragments were purified using the Qiagen system according to manufacturers protocol.

MEL cells transfected populations. The MEL cell line C88 was grown in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal calf serum (FCS, pre-tested for growth of these cells), 10 U/ml of penicillin and 10 mg/ml of streptomycin. MEL cells were transfected by electroporation method (117 V, 1200 μF, 10 ms). For constructs PL10 and PL17, 80 μg of linearized plasmid was co-transfected with 5 μg of pMCINneo. For constructs PL2, PL19 and pTK, 50 μg of linearized plasmid was co-transfected with 5 μg of pMCINneo. Transfected cells were selected with 800 mg/ml for 14 d. Resulting populations were expanded and harvested for DNA preparation. Southern blot analysis was performed to test for gene copy number. Briefly, 10 μg of DNA was restricted with *Eco*R1, run on a 1% agarose gel and blotted on to filters. OLB P^{32}-labelled probes (0·8 kb *Sac*1–*Bgl*II hGH fragment and 1·2 kb *Kpn*1–*Xho*I Thy1 fragment) were used for filter hybridization. Phosphorimaging was used to determine the specific hybridizing signal for the transfected Ly-6A deletion constructs when normalized to the signal obtained for the single copy Thy1 gene control. Copy numbers of the populations varied from 1·3 to 32·5 copies per cell.

Human growth hormone (hGH) assay. After selection of each stably transfected population, 3×10^6 cells were pelleted in sterile falcon tubes, media removed and cells resuspended into 3 ml of fresh medium. Cells (500 μl) were added to each of six wells in a 24-well culture plate for examination of the basal and γ-IFN-induced production of

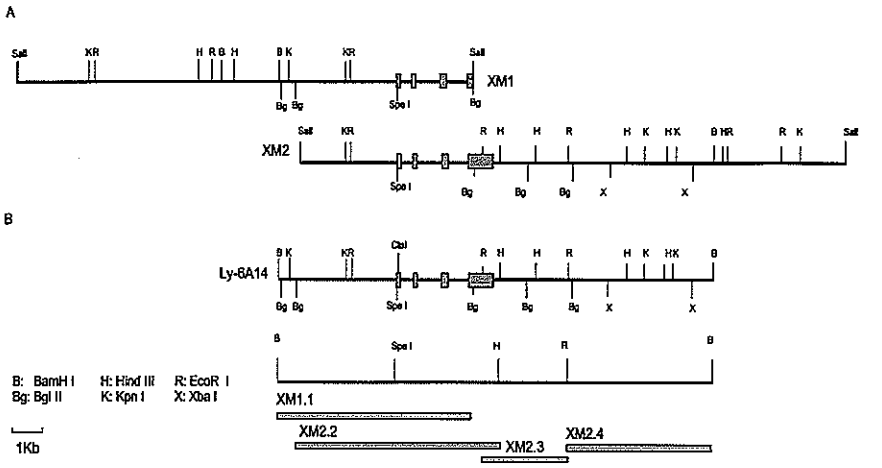

Fig 1. Restriction maps of *Ly-6A* phage clones, subclones and expression cassette. (A) Two overlapping phage clones XM1 and XM2 were isolated from a mouse strain 129 genomic library. The four exons of the *Ly-6A* gene are indicated by grey boxes. Restriction sites are indicated. (B) The restriction maps for the 14 kb Ly-6A14 expression cassette and the four subclones XM1·1, XM2·2, XM2·3 and XM2·4 are shown, along with the strategy for the generation of the expression cassette. The position of the PCR generated unique *Cla*1 cloning site is indicated.

A

XM2	TTT CTT GCA GAT TCT CAA AC
Ly-6E	--- --- --- --G --- --- --
Ly-6C.1	--- --- --- --A C-- --- --
Ly-6G.1	--- --- --- --A C-- --- --
Ly-6F.1	--- --- --- --A C-- --- --
TSA.1	--- --- --- --A A-G --- -C
ThB.2	--- --- --- --A -CT GGT G-

B

XM1	TTC TAT TTA GAC ACA GAA AAA A
Ly-6E	--- --- --- --A --- --- --- -
Ly-6C.1	-GT G-- -GG-GG T-T ATT --T T
Ly-6F.1	C-- --A --T--A --- AG- --- -
Ly-6G.1	--G GT- -C- -TG TAT -TG TGTG
TSA.1	AGA-GC-AC T-- --C TGG CTTC

Fig 2. Nucleotide sequences verifying that phage clones XM1 and XM2 represent the *Ly-6A* genomic locus. (A) Nucleotide sequence within the exon 4 region of phage clone XM2 and comparison with the *Ly-6E* allele and other members of *Ly-6* gene family. The sequence of phage clone XM2 is 100% homologous to the *Ly-6A* genomic sequence from GenBank, and there is a single nucleotide difference from the *Ly-6E* allele. (B) Nucleotide sequence 5′ to exon 1 from the *Spe*I site at 320 bp 5′ to the transcriptional start of *Ly-6A*. Alignment of phage clone XM1 nucleotide sequence with some other members of *Ly-6* family is shown. The sequence of phage clone XM1 in this region is 100% homologous to the *Ly-6A* genomic sequence from Genbank. The *Ly-6E* allele differs by one nucleotide base.

hGH. The final concentration of γ-IFN in culture was 100 U/ml. Cells were cultured for 48 h with no medium changes. Supernatants were harvested, diluted (so as to fall in the linear range of hGH detection) and triplet samples (200 μl each) were measured using the hGH enzyme-linked immunosorbent assay (ELISA) kit (Roch, Almere, The Netherlands) as per the assay conditions described by the manufacturer. The level of hGH in each sample was determined by comparison with a standard curve (manufacturer supplied hGH). After correction for transfectant copy number, the amount of hGH in pg/ml per copy was determined.

DNA sequencing. To confirm the identity of the two isolated overlapping phage clones, polymerase chain reaction (PCR) primers specific for exon 4 (5′-GTCCAGGTGCTGCCTCCATT-3′ and 5′-GGAAGCTCTGTTGTCCCTGC-3′) and 5′ to the first exon (5′-GGATGCAAGAGCCTACTGGGG-3′ and 5′-CACACC ACTCCTTACTTCCTCTC-3′) of the *Ly-6A* gene (sequences obtained from Genbank) were designed and used to amplify 215 bp and 960 bp fragments respectively. Sequence analysis was performed directly on the PCR products by using Big Dye Terminater Cycle Sequencing Kit and ABI 377 Automatic Sequencer (Applied Biosystems, Nieuwerkerke a/d Ijssel, The Netherlands). For sequencing of the region containing the +8·7 and +8·9 HSSs, a 1 kb *Kpn*I fragment from XM2·3 was subcloned into pBluescript. Common primers for the flanking T7 and T3 promoter sequences were used for amplification, followed by sequencing. The DNAMAN program was used for all sequence alignment and identity determination.

RESULTS

Cloning of the Ly-6A allele

A mouse strain 129 genomic DNA library was screened

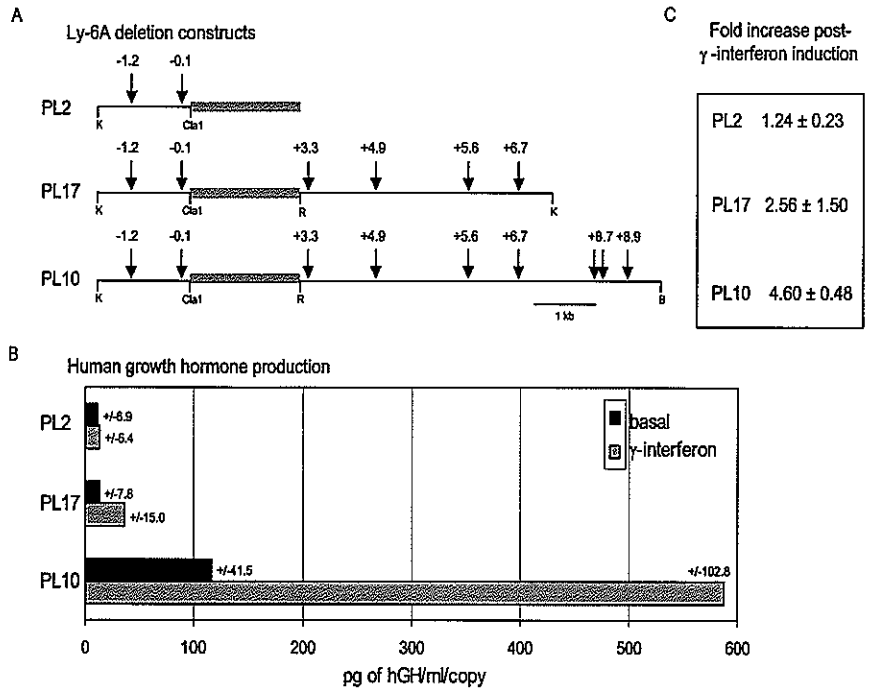

Fig 3. Basal and γ-IFN-induced expression of *hGH* by *Ly-6A hGH* reporter deletion constructs in MEL cell transfects. (A) Schematic maps of deletion constructs are shown with DNase I hypersensitive sites indicated by arrows. Position of each hypersensitive site from the transcriptional start site is noted. The *hGH* reporter gene is indicated by the grey rectangle. Restriction sites used for cloning fragments from the *Ly-6A* subclones are indicated; *Cla*1, K = *Kpn*, R = *Eco*RI, B = *Bam*HI. (B) Levels of hGH produced by MEL cell transfectants PL2, PL17 and PL10 (pg of hGH/ml of supernatant/gene copy) as measured using ELISA are indicated by the dark grey bars for basal and light grey bars for γ-IFN-induced expression. For each construct, at least two independently transfected populations were assayed two to six times. Average levels of hGH (pg/ml/copy) and standard deviations are shown. (C) Fold induction in expression of *hGH* from deletion constructs because of γ-IFN. Average values and standard deviations are shown.

with a 760 bp *Ly-6E* cDNA probe (LeClair *et al*, 1986). From this screening of over $3{\cdot}5 \times 10^6$ independent phage clones, 10 hybridizing clones were picked and rescreened. After two rounds of purification, the clones were characterized by restriction mapping. As shown in Fig 1, the restriction maps of two clones were found to be overlapping within the putative *Ly-6A* gene. Clone 1, XM1, contained 14·5 kb of genomic sequence, of which 12·2 kb of sequence was upstream of *Ly-6A* coding sequences. Clone XM2 was found to be 17·5 kb long and contained 3 kb of 5′ and 11·5 kb of 3′ sequence flanking the putative *Ly-6A* gene.

Because the *Ly-6* family of genes consists of numerous highly cross-hybridizing members, sequence analysis was performed to verify that both XM1 and XM2 represented the *Ly-6A* gene. The coding sequence in exon 4 that is unique for the *Ly-6A and Ly-6E* genes (Fig 2A) was examined in phage clone XM2. By comparison with exon 4 sequences of five of the closest related *Ly-6* gene family members, the sequence of the XM2 clone was confirmed to be that of the *Ly-6A* locus. Furthermore, this sequence of XM2 was identical to that of the *Ly-6E* gene except for one nucleotide. To verify that phage clone XM1 represented the *Ly-6A* gene, sequencing in the non-coding region 5′ to exon 1 was performed. The sequence of XM1 was found to be identical to the Genbank sequence of *Ly-6A*. As shown in Fig 2B, only one nucleotide differed from the *Ly-6E* sequence, whereas numerous other nucleotide differences were found in four other closely related *Ly-6* gene family members. Thus, the overlapping phage clones XM1 and XM2 represent the *Ly-6A* genomic locus.

Deletional analysis and localization of the Ly-6A gene regulatory sequences

Previous deletional analysis studies of *Ly-6E* genomic sequences demonstrated that a region located approximately 8–9 Kb downstream from the transcriptional start site, containing two strong DNase I hypersensitive sites, is responsible for the high-level transcription of a reporter gene in haematopoietic cells *in vitro* after γ-IFN induction (Sinclair *et al*, 1996). Furthermore, a 14 kb *Ly-6E* expression

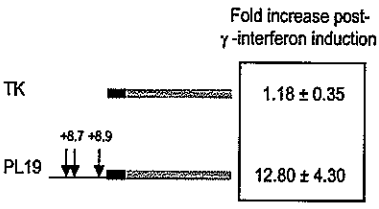

Fig 4. The 3′ distal fragment of the *Ly-6A* gene enhances γ-IFN-induced expression from the *TK* promoter. TK and PL19 constructs are shown. The *TK* promoter is indicated by the black rectangle and the *hGH* reporter gene by the grey rectangle. The average fold increase and standard deviation of *hGH* expression after γ-IFN induction is shown. Two or more independently transfected populations were tested two to six times.

cassette was previously found to yield high-level transcription in haematopoietic cells in transgenic mice (Miles *et al*, 1997). Hence, we engineered similar deletion constructs from the overlapping *Ly-6A* phage clones to examine γ-IFN-induced expression of the *hGH* reporter gene *in vitro* in haematopoietic cells.

The two overlapping clones XM1 and XM2 were subcloned (XM1·1, XM2·2, XM2·3 and XM2·4), and ligated to form a 14 kb fragment as per the strategy in Fig 1B and described in *Materials and Methods*. To expedite insertion of exogenous sequences, a unique *Cla*I cloning site was engineered in the first (untranslated) exon of *Ly-6A* using PCR amplification between the *Spe*I and *Vsp*1 sites in exon 1 of subclone XM2·2. As shown in Fig 3A, approximately 1·8 kb of 5′ flanking sequence containing the *Ly-6A* promoter (up to the engineered *Cla*I site) was ligated to the *hGH* reporter gene for each of the constructs PL2, PL17 and PL10. Approximately 5·4 kb of *Ly-6A* 3′ flanking sequence was ligated downstream of the *hGH* gene for construct PL17, and ≈7·4 kb of *Ly-6A* 3′ flanking sequence was used for construct PL10. The sequences used for PL10 included a region previously mapped and shown to contain two strong HSS +8·7 and +8·9 in a similar position to those in the *Ly-6E* locus. MEL cells were co-transfected with the deletion constructs and a neomycin-selectable marker gene. Independently transfected populations for each construct were selected in G418 for 14 d. The transfectant populations were tested for transgene copy number by Southern blot analysis and varied in copy number from 1·3 to 32·5 (average = 7·6 copies/cell).

Ly-6A 3′ distal flanking sequences promote high-level basal and γ-IFN-induced expression in MEL transfectants. MEL cell transfectants were seeded at the same density in suspension cultures in fresh medium in the presence or absence of γ-IFN and grown for 48 h. Supernatants were harvested and hGH levels were measured using ELISA. MEL populations transfected with the deletion constructs yielded low basal levels of hGH (Fig 3B). The PL2 construct, which contained only the 1·6 kb promoter region of the *Ly-6A* gene, produced 12 pg/ml/copy. Similarly, the PL17 transfectants produced an average basal expression of 13 pg/ml/copy. However, the PL10 construct containing the *Ly-6A*

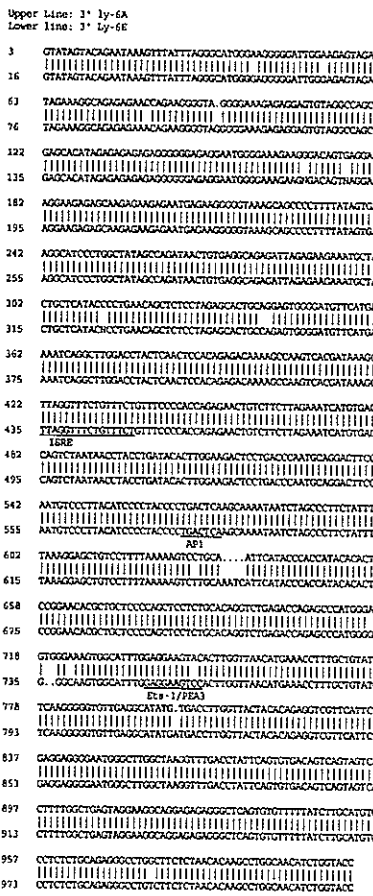

Fig 5. Nucleotide sequence comparisons of the 3′ distal 1 kb *Kpn*1–*Kpn*1 fragments of *Ly-6A* (top line) and *Ly-6E* genes (bottom line). The *Ly-6E* sequence was taken from Sinclair *et al*, 1996. Putative transcription factor consensus binding sites are underlined.

promoter and the complete 3′ sequence, including the region with the strong +8·7 and +8·9 HSS, yielded 10 times more hGH, 116 pg/ml/copy. Taken together, the data suggest that the *Ly-6A* promoter on its own does not promote abundant transcription, and that the *Ly-6A* 3′ distal flanking sequences promote higher levels of basal transcription from this promoter.

Supernatants from γ-IFN-induced cultures were also tested by hGH ELISA. As shown in Fig 3B, low level expression was observed both in PL2 and PL17 transfectants. When compared with the basal level of expression, the fold increase in expression caused by γ-IFN induction was 1·2 and 2·6 respectively (Fig 3C). However, after induction with γ-IFN, the PL10 MEL transfectant populations yielded over 587 pg/ml/copy of hGH expression,

representing a 4·6-fold increase from the basal level of hGH expression. Furthermore, a 41-fold difference was found between the induced levels of expression from only the promoter (PL2) and the PL10 construct with the 3′ sequences containing the distal HSS, suggesting that the *Ly-6A* 3′ distal sequences contain a potent γ-IFN-inducible transcriptional enhancer.

The 3′ distal Ly-6A 1·4-kb sequence enhances expression from a heterologous promoter and contains putative transcription factor binding sites. To further localize the 3′ sequences responsible for the high level of γ-IFN induced expression, an additional construct containing only the 1·4 kb 3′ *Ly-6A* sequence with the previously described strong +8·7 and +8·9 DNaseI hypersensitive sites was made. This sequence was cloned upstream to the *TK* promoter and the *hGH* reporter gene to yield PL19 (Fig 4). While MEL populations transfected with a control construct containing just the *TK* promoter gave a negligible 1·2-fold increase in hGH expression, PL19 transfectants produced high levels of hGH, representing an 12·8-fold increase. Thus, the *Ly-6A* 3′ distal sequences confer high levels of γ-IFN-induced expression on a heterologous promoter.

Previously, a 1 kb Kpn fragment of the *Ly-6E* flanking region containing two strong HSS was sequenced and shown to contain consensus binding sites for Myb, AP1 and Ets-1/PEA3 transcription factors and also an IFN stimulable responsive sequence (ISRE). We therefore examined the analogous region of the *Ly-6A* locus by DNA sequencing. As shown in Fig 5, the *Ly-6A* distal 3′ sequence showed a strong 98% identity to the *Ly-6E* sequence. There was complete identity in the putative ISRE, Myb and AP1 transcription factor binding sites. However, a single nucleotide difference was found in the putative binding site for the Ets-1/PEA3 factor. These data showing conservation of putative transcription factor binding sites, together with the deletion construct data showing enhanced reporter gene expression, suggest that these sequences play a role in high-level γ-IFN-induced haematopoietic cell expression.

DISCUSSION

The cloning of a *Ly-6A* locus has enabled us to begin investigations into the transcription regulatory elements that play a role in the expression of the cell surface Sca-1 glycoprotein marker on haematopoietic cells. Previously, several studies have demonstrated complex expression patterns of the *Ly-6A and Ly-6E* proteins on haematopoietic cells (Kimura *et al*, 1984; van de Rijn *et al*, 1989; Spangrude & Brooks, 1993). It was shown that Ly-6A strains of mice express Sca-1 on all haematopoietic stem cells rather than just 25% of Ly-6E strains (Spangrude & Brooks, 1993). These findings prompted us to clone the *Ly-6A* locus, make an 14 kb expression cassette with 3·5 kb of 5′ and 7 kb of downstream *Ly-6A* flanking sequence, and examine what sequences are necessary for high-level γ-IFN induced expression in haematopoietic cells. Because our previous strategy for the *Ly-6E* cassette has yielded high-level γ-IFN-induced expression *in vitro* (Sinclair *et al*, 1996) and correct developmental and cell lineage specific expression in transgenic mice (Miles *et al*, 1997), we reasoned that examination of the *Ly-6A* flanking regions may yield information concerning the differences in allele specific expression. This information would allow for faithful and specific expression of exogenous genes from the *Ly-6A* transcriptional control elements in haematopoietic stem cells for biological studies. Furthermore, compared with other haematopoietic stem cell-specific genetic markers such as *c-kit* (Gokkel *et al*, 1992) or *Runx-1* (Levanon *et al*, 2001), which span over 100 kb of sequence and contain many exons and introns, the relatively small size of the *Ly-6A* locus makes it an excellent candidate for such construction of an expression cassette.

Similar to previous results with the *Ly-6E* cassette, we have demonstrated that in these studies the *Ly-6A* regulatory elements that yield high-level expression after γ-IFN induction are located within a relatively small piece of DNA. The results of *Ly-6A* deletional analyses indicate that the 3′ distal fragment analogous to that in the *Ly-6E* locus is responsible for high-level γ-IFN-induced expression in MEL cells. Interestingly, both alleles show two strong hypersensitive sites in this region at +8·7 and +8·9 from the transcriptional start. However, after γ-IFN induction, the *Ly-6A* allele exhibits more intense HSSs and a doublet HSS at +8·7 compared with the *Ly-6E* allele. Thus, we examined whether these differences would be reflected in sequence changes in the 3′ distal regions. Sequencing analysis shows that, at least in the region containing the +8·7 and +8·9 HSS, the two alleles are very similar. Of the several consensus transcription factor binding sites located in this region, only one nucleotide difference was found in the putative Ets-1/PEA3 binding site in the *Ly-6A* compared with the *Ly-6E* allele. Because identity between these 3′ distal regions is 98%, further studies must be performed to determine the effects of these specific sequence changes on expression and transcription factor binding. Clearly, these issues are difficult to determine in *in vitro* analyses alone because allele-specific as well as developmental stage and cell-lineage-specific differences must be taken into account (Sinclair *et al*, 1996). Previously, cell lineage-specific differences in the appearance of +5·6 and +6·7 HSS have been found between MEL (*Ly-6A*) and YAC-1 (*Ly-6E*) cells. Allele-specific differences have also been found in the 5′ HSS that appear after γ-IFN induction. In the *Ly-6A* allele both the −1·2 and −0·1 are of the same intensity. However, in the *Ly-6E* allele a low intensity −1·2 HSS signal but a high intensity −0·1 HSS signal is observed (Sinclair & Dzierzak, 1993; Sinclair *et al*, 1996). Hence, it is likely that the sequences upstream of the *Ly-6A* and *Ly-6E* genes are also responsible for the allele-specific protein expression patterns and levels. In comparing 1607 bp upstream of the transcriptional start, 95·8% identity was found. Sequence changes do not occur in the GAS (interferon-sensitive) site, but other differences must be examined to determine the possible affects on expression and transcription factor binding.

Comparing *hGH* reporter levels in these studies with those previously published concerning the *Ly6E* allele, we observed a difference in both basal and γ-IFN-induced

expression. The basal levels of expression from the *Ly-6A* allele were approximately threefold higher and the γ-IFN-induced levels were about twofold higher. The significance of these differences would have to be confirmed by direct comparisons in the same experiments with *Ly-6E* transfectants. However, more direct relevant comparisons will be carried out in transgenic mice in whom the transcriptional activities of the two alleles can be examined in functional haematopoietic stem cells.

In summary, with the isolation of the *Ly-6A* gene and flanking regulatory regions, we have shown that 3′ distal flanking sequences are responsible for high-level γ-interferon-induced expression *in vitro*. Furthermore, we have shown that the 3′ distal *Ly-6A* fragment containing HSS 8·7 and 8·9 can be used to direct high-level γ-IFN-induced expression from a heterologous promoter. We are presently testing the Ly-6A14 cassette in transgenic mice in the hope that it will give high-level haematopoietic stem cell specific expression *in vivo*.

ACKNOWLEDGMENTS

The authors thank all the members of the laboratory for assistance and advice. This work was funded by the Netherlands Scientific Research Organization (NWO) 901-08-090, National Institutes of Health R01 DK51077, and the Leukemia Society of America 1034-94.

REFERENCES

Bamezai, A. & Rock, K.L. (1995) Overexpressed *Ly-6A.2* mediates cell-cell adhesion by binding a ligand expressed on lymphoid cells. *Proceedings of the National Academy of Sciences of the United States of America*, 92, 4294–4298.

Gekkel, E., Grossman, Z., Ramot, B., Yarden, Y., Rechavi, G. & Givol, D. (1992) Structural organization of the murine c-kit proto-oncogene. *Oncogene*, 7, 1423–1429.

Kamiura, S., Nolan, C.M. & Meruelo, D. (1992) Long-range physical map of the Ly-6 complex: mapping the Ly-6 multigene family by field-inversion and two-dimensional gel electrophoresis. *Genomics*, 12, 89–105.

Khan, K.D., Lindwall, G., Maher, S.E. & Bothwell, A.L. (1990) Characterization of promoter elements of an interferon-inducible Ly-6E/A differentiation antigen, which is expressed on activated T cells and hematopoietic stem cells. *Molecular Cell Biology*, 10, 5150–5159.

Khan, K.D., Shuai, K., Lindwall, G., Maher, S.E., Darnell, J.E. & Bothwell, A.L. (1993) Induction of the Ly-6A/E gene by interferon alpha/beta and gamma requires a DNA element to which a tyrosine-phosphorylated 91-kDa protein binds. *Proceedings of the National Academy of Sciences of the United States of America*, 90, 6806–6810.

Kimura, S., Tada, N., Liu-Lam, Y. & Hammerling, U. (1984) Studies of the mouse Ly-6 alloantigen system. II Complexities of the Ly-6 region. *Immunogenetics*, 20, 47–56.

LeClair, K.P., Palfree, R.G., Flood, P.M., Hammerling, U. & Bothwell, A. (1986) Isolation of a murine Ly-6 cDNA reveals a new multigene family. *EMBO Journal*, 5, 3227–3234.

Levanon, D., Glusman, G., Bangsow, T., Ben-Asher, E., Male, D.A., Avidan, N., Bangsow, C., Hattori, M., Taylor, T.D., Taudien, S., Blechschmidt, K., Shimizu, N., Rosenthal, A., Sakaki, Y., Lancet, D., Groner, Y. (2001). Architecture and anatomy of the genomic locus encoding the human leukemia-associated transcription factor, RUNX1/AML 1. *Gene*, 262, 23–33.

McGrew, J.T. & Rock, K.L (1991) Isolation, expression, and sequence of the TAP/Ly-6A.2 chromosomal gene. *Journal of Immunology*, 146, 3633–3638.

Miles, C., Sanchez, M.-J., Sinclair, A. & Dzierzak, E. (1997) Expression of the Ly-6E.1 (Sca-1) transgene in adult hematopoietic stem cells and the developing mouse embryo. *Development*, 124, 537–547.

Okada, S., Nakauchi, H., Nagayoshi, K., Nishikawa, S., Miura, Y. & Suda, T. (1992) In vivo and in vitro stem cell function of c-kit- and Sca-1-positive murine hematopoietic cells. *Blood*, 80, 3044–3050.

Palfree, R.G. & Hammerling, U. (1986) Biochemical characterization of the murine activated lymphocyte alloantigen Ly-6E.1 controlled by the Ly-6 locus. *Journal of Immunology*, 136, 594–600.

Palfree, R.G., LeClair, K.P., Bothwell, A. & Hammerling, U. (1987) cDNA characterization of an Ly-6.2 gene expressed in BW5147 tumor cells. *Immunogenetics*, 26, 389–391.

Reiser, H., Coligan, J., Palmer, E., Benacerraf, B. & Rock, K.L. (1988) Cloning and expression of a cDNA for the T-cell-activating protein TAP. *Proceedings of the National Academy of Sciences of the United States of America*, 85, 2255–2259.

Rock, K.L., Yeh, E.T., Gramm, C.E., Haber, S.I., Reiser, H. & Benacerraf, B. (1986) TAP, a novel T cell-activating protein involved in the stimulation of MHC-restricted T lymphocytes. *Journal of Experimental Medicine*, 163, 315–333.

Selden, R.F., Howie, K.B., Rowe, M.E., Goodman, H.M. & Moore, D.D. (1986) Human growth hormone as a reporter gene in regulation studies employing transient gene expression. *Molecular Cell Biology*, 6, 3173–3179.

Sinclair, A., Daly, B. & Dzierzak, E. (1996) The Ly-6E.1 (Sca-1) gene requires a 3′ chromatin-dependent region for high-level gamma-interferon-induced hematopoietic cell expression. *Blood*, 87, 2750–2761.

Sinclair, A.M. & Dzierzak, E.A. (1993) Cloning of the complete Ly-6E.1 gene and identification of DNase I hypersensitive sites corresponding to expression in hematopoietic cells. *Blood*, 82, 3052–3062.

Spangrude, G.J. & Brooks, D.M. (1993) Mouse strain variability in the expression of the hematopoietic stem cell antigen Ly-6A/E by bone marrow cells. *Blood*, 82, 3327–3332.

Spangrude, G.J., Heimfeld, S. & Weissman, I.L. (1988) Purification and characterization of mouse hematopoietic stem cells. *Science*, 241, 58–62.

Stanford, W.L., Bruyns, E. & Snodgrass, H.R. (1992) The isolation and sequence of the chromosomal gene and regulatory regions of Ly-6A.2. *Immunogenetics*, 35, 408–411.

Stanford, W.L., Haque, S., Alexander, R., Liu, X., Latour, A.M., Snodgrass, H.R., Koller, B.H. & Flood, P.M. (1997) Altered proliferative response by T lymphocytes of Ly-6A (Sca-1) null mice. *Journal of Experimental Medicine*, 186, 705–717.

van de Rijn, M., Heimfeld, S., Spangrude, G.J. & Weissman, I.L. (1989) Mouse hematopoietic stem-cell antigen Sca-1 is a member of the Ly-6 antigen family. *Proceedings of the National Academy of Sciences of the United States of America*, 86, 4634–4638.

Chapter 3

Expression of the *Ly-6A (Sca-1) lacZ* transgene in mouse hematopoietic stem cells and embryos

Xiaoqian Ma, Marella de Bruijn, Catherine Robin, Marian Peeters, John Kong-A-San, Ton de Wit, Cone Snoijs and Elaine Dzierzak.

British Journal of Haematology 2002; 116: 401-408

Expression of the *Ly-6A* (Sca-1) *lacZ* transgene in mouse haematopoietic stem cells and embryos

Xiaoqian Ma, Marella de Bruijn, Catherine Robin, Marian Peeters, John Kong-A-San, Ton de Wit, Corne Snoijs and Elaine Dzierzak *Department of Cell Biology and Genetics, Erasmus University, Rotterdam, The Netherlands*

Received 22 June 2001; accepted for publication 16 August 2001

Summary. The Sca-1 surface glycoprotein is used routinely as a marker for haematopoietic stem cell enrichment. Two allelic genes, *Ly-6A* and *Ly-6E*, encode this marker and appear to be differentially regulated in haematopoietic cells and haematopoietic stem cells. The Sca-1 protein has been shown to be expressed at a greater frequency in these cells from *Ly-6A* strains of mice. To study the specific expression pattern and haematopoietic regulation of the *Ly-6A* gene, we constructed a 14 kb cassette from a genomic *Ly-6A* fragment, inserted a *lacZ* reporter gene and created transgenic mice. We found that the *Ly-6A lacZ* transgene was expressed in the haematopoietic tissues and predominantly in the T-lymphoid lineage. Some expression was also found in the B-lymphoid and myeloid lineages. We demonstrated functional haematopoietic stem cell enrichment by sorting for β-galactosidase-expressing cells from the bone marrow. In addition, we found an interesting embryonic expression pattern in the AGM region, the site of the first haematopoietic stem cell generation. Surprisingly, when compared with data from *Ly-6E lacZ* transgenic mice, our results suggest that the *Ly-6A* cassette does not improve *lacZ* marker gene expression in haematopoietic cells.

Keywords: *Ly-6A/E*, Sca-1, haematopoietic stem cells, transgene, embryo.

Enrichment and characterization of the stem cells at the foundation of the haematopoietic hierarchy has relied on the Sca-1 phosphatidylinositol-linked cell surface glycoprotein marker (Spangrude *et al*, 1988). Through fluorescence-activated cell sorting using a monoclonal antibody specific for Sca-1, haematopoietic stem cells (HSC) can be enriched approximately 100-fold from adult bone marrow and, together with antibodies specific for other cell surface markers (i.e. Thy-1^{lo}, c-kit or depletion for cells with mature lineage markers), a greater than 1000-fold enrichment can be obtained (Spangrude *et al*, 1988; Okada *et al*, 1992).

The Sca-1 protein is encoded by the strain-specific allelic genes, *Ly-6E* and *Ly-6A* (van de Rijn *et al*, 1989; Khan *et al*, 1990; Stanford *et al*, 1992; Sinclair & Dzierzak, 1993), which are members of the multigenic *Ly-6* family (LeClair *et al*, 1986; Kamiura *et al*, 1992). The family consists of at least 18 highly homologous cross-hybridizing genes with diverse and overlapping patterns of expression (Kimura *et al*, 1984). Owing to the homologies of the *Ly-6* family of genes and proteins, and the suspected overlapping roles in cell adhesion played by members of this family, functional studies have been difficult (Bamezai & Rock, 1995; Stanford *et al*, 1997). Sca-1 protein expression is complex within *Ly-6A* and *Ly-6E* strains of mice (Kimura *et al*, 1984; van de Rijn *et al*, 1989; Spangrude & Brooks, 1993). The *Ly-6A* and *Ly-6E* genes differ only by three nucleotides in the coding sequence, resulting in two amino acid changes (LeClair *et al*, 1986; Reiser *et al*, 1988). Both gene products express the Sca-1 epitope recognized by the antibody E13-161-7 (LeClair *et al*, 1986; Palfree & Hammerling, 1986; Rock *et al*, 1986; Palfree *et al*, 1987; Reiser *et al*, 1988). Both genes are interferon inducible, but the *Ly-6A* allele appears to be more widely and highly expressed (Kimura *et al*, 1984; Rock *et al*, 1986; Spangrude & Brooks, 1993). Strains of mice with the *Ly-6A* gene express Sca-1 on 10–20% of adult thymocytes and 50–70% of peripheral T lymphocytes, while strains with the *Ly-6E* gene express Sca-1 on 5–10% of adult thymocytes and 10–15% of peripheral T lymphocytes. Similarly, *Ly-6A* strains of mice express Sca-1 on virtually all (99%) marrow repopulating cells, while *Ly-6E* strains express Sca-1 on only 25% of these cells (Spangrude & Brooks, 1993). Nonetheless, Sca-1 remains an important marker of HSCs and its gene

Correspondence: Professor Dr E. Dzierzak, Department of Cell Biology and Genetics, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands. E-mail: dzierzak@ch1.fgg.eur.nl

regulatory elements are of current research interest to direct expression to HSCs for potential therapeutic applications.

Previously, the *Ly-6E* transcriptional elements have been examined. Upstream *cis*-acting elements involved in regulating *in vitro* expression of *Ly-6E* have been identified (Khan *et al*, 1990, 1993) and sequence comparisons suggest that similar 5′ control elements are present in the *Ly-6A* promoter (McGrew & Rock, 1991). DNaseI hypersensitive site (HSS) mapping of both the *Ly-6E* and *Ly-6A* alleles show almost identical patterns (Sinclair & Dzierzak, 1993). Deletion studies using the *Ly-6E* and *Ly-6A* genes reveal that the region containing the two most distal 3′ HSS are responsible for high level, γ interferon-induced expression *in vitro* (Sinclair *et al*, 1996; Ma *et al*, 2001). Furthermore, this 3′ region is necessary in the context of a 14 kb *Ly-6E* expression cassette for high-level tissue-specific expression of a *lacZ* marker gene in transgenic mice (Miles *et al*, 1997). However, in such transgenic mice, it was found that some but not all HSCs can be sorted based on lacZ expression, suggesting that the *Ly-6E* expression cassette is not optimal for HSC expression *in vivo*. As subtle differences exist in HSS between the *Ly-6E* and *Ly-6A* alleles (Sinclair *et al*, 1996) and the *Ly-6A* gene product has been shown to be expressed in 100% of marrow repopulating cells, it was therefore of great interest to examine the *Ly-6A* sequences as a source for HSC-specific regulatory elements.

Thus, we cloned a *lacZ* reporter gene into a 14 kb *Ly-6A* gene cassette and generated transgenic mice. Here we present data from studies examining the differences in expression patterns and levels of *Ly-6A lacZ* transgene expression with that from a previously described *Ly-6E lacZ* transgene. In general, we found that the *Ly-6A lacZ* transgene is predominantly expressed in the cells of the T-lymphoid lineage. Moreover, we have shown that the *Ly-6A lacZ* transgene facilitates a > 100-fold enrichment of bone marrow HSCs. Surprisingly, while the *Ly-6A lacZ* transgene expression pattern in embryos was slightly more widespread than that of the *Ly-6E lacZ* transgene, the expected allele-specific differences in haematopoietic cell expression were not observed.

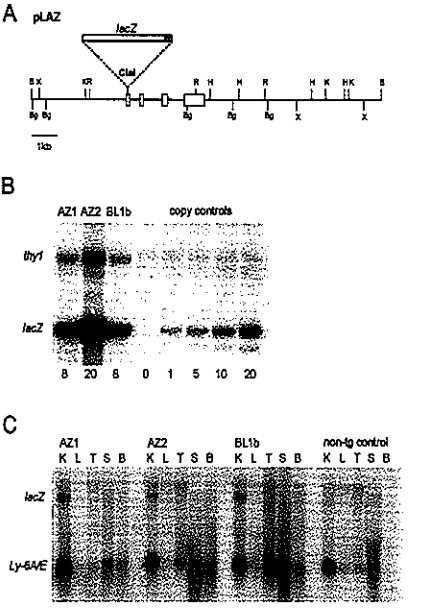

Fig 1. Generation and characterization of transgenic mouse lines. (A) Restriction map of the *Ly-6A lacZ* transgene in pLAZ. The *lacZ* marker gene (white rectangle) with SV40 poly A sequence (black rectangle) was cloned into the *Cla*1 site of the 14 kb *Ly-6A* cassette and this fragment was used to generate the AZ1 and AZ2 transgenic mouse lines. B = *Bam*H1, Bg = *Bgl*II, K = *Kpn*, R = *Eco*R1, H = *Hind*III, X = *Xba*. (B) Southern blot of transgenic mouse DNA. Hybridization with the *Thy1* probe (DNA normalization control) and *lacZ* probe (transgene) was performed and the signal detected by phosphorimaging was compared with plasmid copy controls for the determination of transgene copy number (indicated below each lane) in AZ1 and AZ2 transgenic lines. (C) Northern blot of total RNA from tissues of *Ly-6A lacZ* transgenic lines. Hybridization was performed with a *lacZ* probe for transgene expression and *Ly-6A/E*-specific probe for endogenous gene expression. Tissues of a BL1b and a non-transgenic control are also shown. K = kidney, L = lymph node, T = thymus, S = spleen, B = bone marrow.

MATERIALS AND METHODS

Constructs and transgenic mice. The 14 kb Ly-6A cassette (Ly-6A14) was constructed as described previously (Ma *et al*, 2001). The *lacZ* gene in p610ZA (gift of D. Meijer) was modified, converting a 3′ *Sma*I site to a *Nar*I site using oligonucleotide adaptors. The 3·6 kb *lacZ* NarI fragment was cloned into Ly-6A14 to generate pLAZ.

Fertilized (C57BL/10 × CBA)F1 oocytes were microinjected with a 17·6-kb NotI fragment containing the *Ly-6A lacZ* gene from pLAZ (Fig 1A). This fragment was gel purified for removal of all vector sequences. Positive founder animals were bred with (C57BL/10 × CBA)F1 mice and lines were maintained as heterozygotes. The C57BL/10 strain contains an endogenous *Ly-6A* allele and the CBA strain contains an endogenous *Ly-6E* allele. (C57BL/10 × CBA)F1 mice co-dominantly express both alleles. Southern blot analysis of tail DNA was used to identify founder transgenic mice.

DNA and RNA analysis. Genomic DNA (5–10 μg) for Southern blot analysis (Miles *et al*, 1997) was digested using BamH1 and electrophoresed through 1% agarose/Tris, acetate, EDTA gels prior to transfer to Hybond-N membranes. Transgene copy number controls were generated by addition of appropriate amounts of pLAZ to non-transgenic genomic DNA. Filters were probed with lacZ and *Thy-1* gene fragments. Normalization for DNA content of each lane was performed after phosphorimage analysis of *Thy-1* signal. Copy number was determined subsequently by comparing the *lacZ* signal obtained from the transgenic mice with that of the plasmid controls on the linear portion of the standard curve.

Total cellular RNA for Northern blot analysis was prepared using the lithium chloride/urea method and 5–15 μg was fractionated on 1% agarose/formaldehyde gels (Fraser *et al*, 1990) prior to transfer to Hybond-N membranes. Filters were probed with *lacZ* and *Ly-6E* cDNA fragments.

Probes used for hybidization to Southern or Northern filters were labelled using the random oligonucleotide priming procedure incorporating [^{32}P]-ATP. The fragments used were as follows: 1·1 kb *Bam*H1–*Eco*RV *lacZ* containing fragment from p610ZA; 1·2 kb *Xba*1–*Nru*1 *Thy-1* gene fragment from pD7 (Spanopoulou *et al*, 1988); 761 bp *Eco*R1 *Ly-6E* cDNA fragment from pLy6·1–2R (LeClair *et al*, 1986). After hybridization, filters were washed to a stringency of 0·2 × saline sodium citrate (SSC)/0·1% sodium dodecyl sulphate (SDS) and exposed to a phosphorimager screen for quantification using IMAGEQUANT software.

Genomic DNA (200 ng) from the peripheral blood of transplant recipients was analysed using polymerase chain reaction (PCR) with oligonucleotide primers for myogenin-specific sequences: (myo1) 5′-TTACGTCCATCGTGGACAGC-3′ and (myo2) 5′-TGGGCTGGGTGTTAGTCTTA-3′; and for *lacZ*-specific sequences: (lacZ1) 5′-GCGACTTCCAGTTCAACATC-3′ and (lacZ2) 5′-GATGAGTTTGGACAAACCAC-3′.

DNA was subjected to an initial 5 min denaturation at 94°C followed by 30 cycles of denaturation (5 s at 94°C), annealing (30 s at 60°C), elongation (30 s at 72°C). Serial dilutions of blood DNA from a transgenic animal were used as a control to evaluate the levels of donor cell reconstitution in transplanted mice.

β-galactosidase and antibody staining. For analysis of β-galactosidase expression in transgenic bone marrow, thymus, spleen and lymph node, 10^6 cells were suspended in 100 μl of prewarmed phosphate-buffered saline (PBS) with 5% fetal calf serum (FCS) prior to loading with 100 μl of 2 mmol/l fluorescein di-(β-D-galactopyranoside) (FDG) in H_2O. Cells were incubated at 37°C for 60 s. The uptake was stopped by the addition of 2 ml of ice-cold PBS with 5% FCS and the reaction was allowed to proceed for 1–3 h on ice in the dark. Propidium iodide (PI, 1 μg/ml) or 7-amino-actinomycin D (7AAD, 2·5 μg/ml; Pharmingen, Alphen a/d Rijn, The Netherlands) was used to exclude dead cells. A FACSCAN and FACSVANTAGE SE (Becton-Dickinson, Alphen a/d Rijn, The Netherlands) were used for analysis and sorting.

Sca-1, CD4, CD8, B220 and Mac-1 antibodies were direct phycoerythrin (PE) conjugates (Pharmingen). Briefly, after 1–2 h of FDG staining, 10^6 cells were stained with antibody, incubated on ice for 30 min and washed three times in cold PBS with 5% FCS.

Whole embryos were isolated into ice-cold PBS, fixed in 1 ml of X-gal fix (1% formaldehyde, 0·2% gluteraldehyde) at 4°C for 1 h and stained overnight at room temperature in 1 mg/ml X-gal (Sigma, Zwijndrecht, The Netherlands). After staining, embryos were dehydrated through increasing concentrations of ethanol in ice-cold PBS and mounted in paraffin wax. Sections (6–10 μm) were cut onto APES (3-aminopropyltriethoxysilane, Sigma, Zwijndrecht, The Netherlands)-coated microscope slides and dried overnight at room temperature. Slides were dewaxed in Histoclear and rehydrated through decreasing concentrations of ethanol before standard counterstaining with haematoxylin-eosin and mounting.

Bone marrow transplantation. Donor transgenic bone marrow cells for transplantations were FDG, Sca-1 and Hoechst 33258 stained *ex vivo* in PBS with 5% FCS. FACS-sorted cells were counted, diluted and suspended in a final volume of 500 μl of PBS for intravenous injection into the tail vein of male (C57BL/10 × CBA)F1 mice. On the day of transfer, the recipients were exposed to a split dose (3 h interval) of 900 rad irradiation from a ^{137}Cs source. Adult (C57BL/10 × CBA)F1 spleen cells (2×10^5) were co-injected with the donor cells to promote short-term survival. All recipients were housed in filter-top isolators and received 1·6 g/l neomycin in drinking water for at least 1 month. Peripheral blood was taken at 1 and 4 months post transplantation for analysis.

RESULTS

The Ly-6A lacZ transgene is expressed in adult mice

The *Ly-6A* gene was previously cloned and analysed for *in vitro* expression in haematopoietic cells. A genomic expression cassette containing a distal 3′ fragment with strong DNaseI hypersensitive sites (Sinclair & Dzierzak, 1993) was found to yield high level, γ interferon-induced expression (Ma *et al*, 2001). To determine if this cassette could be used to express exogenous genes in haematopoietic stem cells *in vivo*, we inserted a *lacZ* marker gene into an engineered *Cla*1 site in the first untranslated exon of the *Ly-6A* gene (Fig 1A; Ma *et al*, 2001). Two transgenic mouse lines were produced with the *Ly-6A lacZ* construct: AZ1 and AZ2. Southern blotting of DNA from these established mouse lines was compared with DNA from a previously generated *Ly-6E lacZ* transgenic line (BL1b) which, in the homozygous state, contains eight copies of this allelic transgene. Figure 1B shows that AZ1 contains eight copies and AZ2 contains > 20 copies of the *Ly-6A lacZ* transgene in the hemizygous state. Northern blot analysis was performed on RNA derived from various haematopoietic and non-haematopoietic tissues of these transgenic lines (Fig 1C). High level *Ly-6A lacZ* transgene expression was found in the kidney of both the AZ1 and AZ2 transgenic lines and was similar to that observed in the BL1b transgenic line. Other tissues, such as the bone marrow, spleen and thymus, show little or undetectable expression. No expression was found in a non-transgenic littermate control. In general, the tissue-specific expression pattern followed closely the transcription of the endogenous *Ly-6A/E* gene. Interestingly, the higher copy AZ2 line showed equivalent levels of expression to the AZ1 and BL1b lines using this analysis. Thus, the specific expression pattern of the *Ly-6A lacZ* transgene was similar in both AZ1 and AZ2 adult tissues and was consistent with the general pattern in several lines of *Ly-6E lacZ* transgenic mice (Miles *et al*, 1997) including BL1b.

Ly-6A lacZ transgene is expressed in haematopoietic cells

Although Northern blot analysis showed little expression in haematopoietic tissues, a more sensitive method,

FDG-FACS, was performed on cells from thymus, spleen, lymph node and bone marrow of transgenic mice to detect β-galactosidase expression. To determine if allelic-specific differences in transgene expression could be observed, the two *Ly-6A lacZ* transgenic lines were analysed and compared with the *Ly-6E lacZ* BL1b transgenic line. As shown in the representative FACS histograms in Fig 2, no FDG-positive cells were found in the tissues of a non-transgenic control mouse, while both AZ1 and AZ2 mouse lines expressed the *Ly-6A lacZ* transgene in all four haematopoietic tissues. When the FACS-FDG profiles of the *Ly-6A lacZ* tissues were then compared with those of the *Ly-6E lacZ* transgenic line BL1b, similar percentages of FDG-positive cells were observed in all four tissues.

To determine in which adult haematopoietic lineages the *Ly-6A lacZ* transgene expresses, we performed FDG-FACS analysis together with antibodies specific for T-lymphoid, B-lymphoid and myeloid cells. Table I shows the percentages of CD4-, CD8-, B220- and Mac-1-positive cells in the FDG^+ fraction of bone marrow, spleen, thymus and lymph node cells. As expected, predominant transgene expression was found in the T-lymphoid lineage, with some expression in the B-lymphoid and myeloid lineages. In addition, the percentages of FDG^+ cells of the different lineages found in the bone marrow, spleen and thymus of *Ly-6E lacZ* and *Ly-6A lacZ* transgenic adults were similar. Slight differences were found in the bone marrow CD4 and Mac-1 subsets, probably the result of low sample numbers. Taken together, these results strongly suggest that the lineage distribution of *lacZ* marker expression is not different for the *Ly-6E* and *Ly-6A* allelic transgene cassettes.

The Ly-6A lacZ transgene marks functional haematopoietic stem cells in adult bone marrow

As the Ly-6A (Sca-1) protein is used extensively for the enrichment of HSCs from the bone marrow of adult mice and the *Ly-6A lacZ* transgene is expressed in 5–6% of adult bone marrow cells, we determined, using limiting dilution transplantation analysis, whether HSC activity was enriched in the FDG^+ population. To begin these studies, we first examined what percentage of bone marrow cells were positive for transgene and endogenous Sca-1 expression. The FACS plots in Fig 3A show the distribution and percentages of negative, double-positive and single-positive cells found in representative *Ly-6A lacZ* AZ1 and AZ2 transgenic bone marrow. The percentage of cells within each of the four quadrants was similar between AZ1 and AZ2 as well as BL1b (not shown). While some cells expressed both markers, not all FDG^+ cells were $Sca\text{-}1^+$ and vice versa. Thus, regulation of transgene expression overlapped but did not completely recapitulate endogenous *Ly-6A/E* gene regulation.

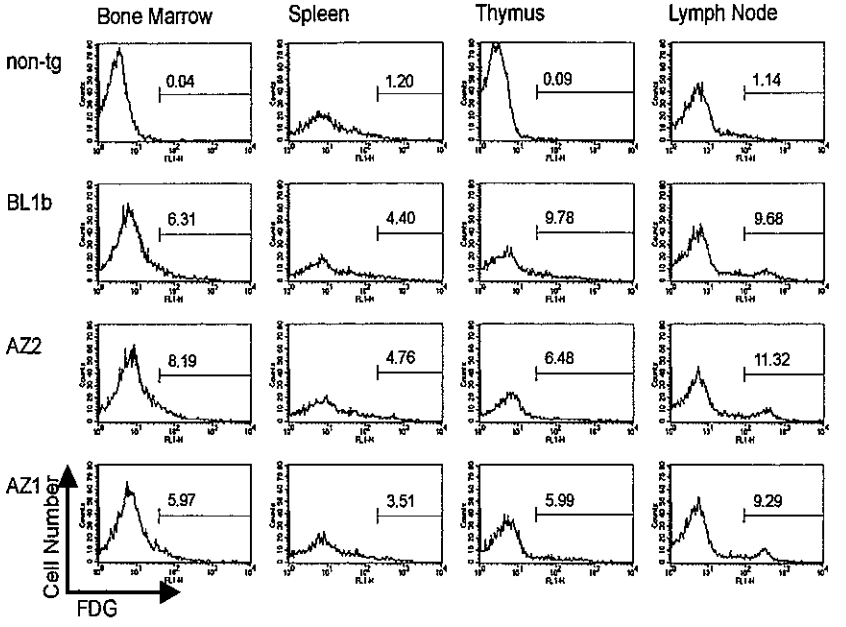

Fig 2. Representative fluorescein di-(β-D-galactopyranoside) fluorescence-activated cell sorting (FDG-FACS) analysis of *lacZ* transgene expression in haematopoietic tissues. Bone marrow, spleen, thymus and lymph node cells from control non-transgenic, BL1b, AZ1 and AZ2 age-matched male transgenic mice were stained with the FDG substrate and analysed using flow cytometry. Histograms show levels of fluorescence intensity on a logarithmic scale (abscissa) and number of cells (ordinate). Percentages of FDG-positive cells are indicated.

Table 1. Subsets of haemato/lymphoid cells found in the FDG$^+$ fractions of *Ly-6E lacZ and Ly-6A lacZ* transgenic mice.

		Mean percentage of subset in total FDG$^+$ population (SD)			
Tissue	Transgene	CD4	CD8	B220	Mac-1
Bone marrow	Ly-6E lacZ	28·2*	14·2	20·8	33·1
	Ly-6A lacZ	46·5	19·4	22·8	12·5*
Spleen	Ly-6E lacZ	42·9	29·0	16·5	5·8
	Ly-6A lacZ	49·2 (17·4)	38·7 (13·9)	12·5 (5·7)	6·7 (3·7)
Thymus	Ly-6E lacZ	68·7	37·8	ND	ND
	Ly-6A lacZ	64·9 (13·3)	38·8 (18·1)	ND	ND
Lymph node	Ly-6E lacZ	55*	35·3*	3·5*	ND
	Ly-6A lacZ	58·3 (15·5)	39·2 (12·9)	3·6 (3·4)	ND

*Only one experiment performed.

Cell suspensions were stained with the FDG substrate and specific antibodies against the indicated cell lineage markers. At least 2×10^4 cells were examined. For the *Ly-6E lacZ* results, BL1b and BL19 transgenic adult mice were examined (see Miles *et al*, 1997). For the *Ly-6A lacZ* results, AZ1 and AZ2 transgenic adult mice were used. Numbers in brackets (SD) are the standard deviation (three experiments performed). ND = not done.

To test for the presence of HSCs in each of the phenotypically described populations, AZ1 and AZ2 bone marrow cells were sorted based on FDG and Sca-1 staining and injected in varying doses into irradiated adult recipients. At 4 months post transplantation, the recipient mice were tested for donor cell haematopoietic engraftment. As shown in Fig 3B, the combined results of two independent experiments show the highest enrichment of HSCs in the sorted Sca-1$^+$FDG$^-$ and Sca$^+$FDG$^+$ cells (as few as 100 sorted cells yield repopulation). Some enrichment was also observed in the sorted Sca-1$^-$FDG$^+$ cells (2×10^4 cells yield repopulation). In contrast, the Sca-1$^-$FDG$^-$ population of bone marrow was greatly decreased in HSC activity, requiring greater than $1–5 \times 10^5$ cells for repopulation. Unsorted control bone marrow was found to be at least five times more efficient than the Sca-1$^-$FDG$^-$ sorted bone marrow. When these transplantation data were compared with equivalent sorting and transplantation data from *Ly-6E lacZ* transgenic mice (Miles *et al*, 1997; and data not shown), no clear quantitative difference was found between *Ly-6A lacZ* and *Ly-6E lacZ* transgenics in bone marrow HSC activity enriched by FDG sorting.

Embryonic expression of the Ly-6A lacZ transgene in the AGM region is similar to that of the Ly-6E lacZ transgene

The expression of the *Ly-6A lacZ* transgene in the haematopoietic lineages and the HSCs of the adult mouse led us to examine the specific expression pattern of this transgene during development. Localization of β-galactosidase expression by X-gal staining could indicate the first site(s) of HSC appearance within the embryo. At E11, the expression pattern of *Ly-6A lacZ* was limited to the embryo body, with no expression in the yolk sac. The most striking X-gal staining was in the caudal tail region and the limb buds of AZ1 and AZ2 embryos (Fig 4A). The caudal expression pattern along the dorso-ventral axis in the AZ1 and AZ2 lines was slightly more widespread than in BL1b embryos. However, the antero-posterior limit of expression in all three lines was confined to the posterior area containing the hindgut. The high-level limb bud expression was specific to the AZ1 and AZ2 lines and was not observed in the BL1b line. Furthermore, limb bud expression was not observed in other *Ly-6E lacZ* transgenic lines (Miles *et al*, 1997). Thus, in mid-gestational mouse embryos the *Ly-6A lacZ* transgene was differentially expressed compared with the *Ly-6E lacZ* transgene.

Histological sectioning and staining was performed to determine in which embryonic tissues the *Ly-6A lacZ* transgene was expressed. In transverse sections from the truncal region of E11 AZ1, AZ2 and BL1b transgenic embryos, intense blue staining was observed in the epithelial cells lining the tubules of the pronephros and mesonephros (Fig 4B). The staining pattern was identical between all three lines. As the dorsal aorta and the surrounding mesenchyme have been found to be the only area with the AGM region containing functional HSCs (de Bruijn *et al*, 2000), we carefully examined the transverse sections for β-galactosidase activity at the site. No X-gal staining was found in the dorsal aorta or surrounding mesenchyme in any of the E11 sections examined from AZ1, AZ2 or BL1b transgenic lines. As the counterstaining may obscure the weak β-galactosidase signal from this area, we also examined transverse sections stained only with X-gal. While the pro/mesonephros showed high level β-galactosidase expression, not even weak X-gal staining was observed in the dorsal aorta or surrounding mesenchyme (data not shown). More sensitive FACS analysis verified this result, strongly suggesting that AGM HSCs are negative or beneath the limits of detection for *Ly-6A lacZ* transgene expression.

DISCUSSION

The results of the studies presented here demonstrate that the *Ly-6A lacZ* transgene is transcribed consistently in a lineage-specific manner. This 14 kb cassette produced high

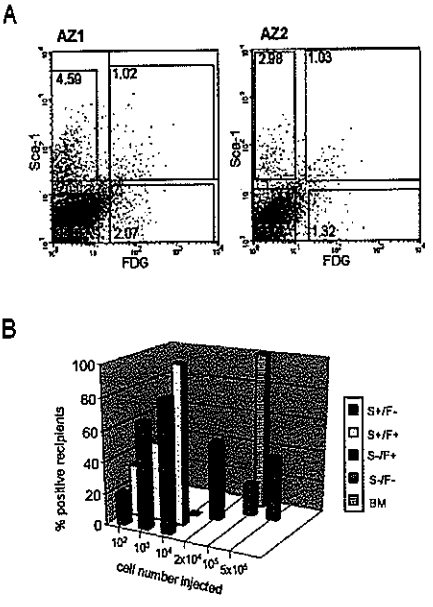

Fig 3. Repopulation of adult recipients using flow cytometric-sorted bone marrow cells from *Ly-6A lacZ* transgenic bone marrow. (A) Fluorescence-activated cell sorting (FACS) plots and sorting gates of bone marrow cells from AZ1 and AZ2 transgenic mice used for limiting dilution transplantation experiments. Adult bone marrow cells were stained with FDG and Sca-1 and sorted into four populations. Percentages of cells in each of the quadrants is indicated. (B) Limiting dilution repopulation frequency of sorted FDG and Sca-1 stained bone marrow from *Ly-6A lacZ* bone marrow. AZ1 and AZ2 bone marrow was sorted into double-negative, double-positive and single-positive populations and injected into irradiated adult recipients in limiting numbers. Engraftment by donor cells was tested at greater than 4 months post transplantation and mice found to be more than 10% repopulated in the peripheral blood (as determined by *lacZ* semiquantitative polymerase chain reaction) were considered positive. The percentage of positive recipients is plotted on the ordinate and the number of sorted cells transplanted is plotted on the abscissa. Coded vertical bars represent the various sorted cell populations injected into the recipient mice. S = Sca-1, F = FDG and BM = whole unsorted bone marrow.

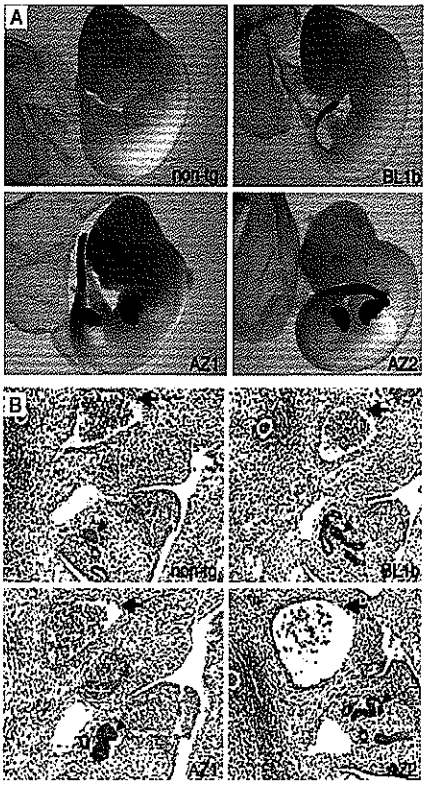

Fig 4. *Ly-6A lacZ* expression in E11 transgenic embryos. (A) Whole embryos. E11 non-transgenic, BL1b, AZ1 and AZ2 embryos were stained with the X-gal substrate to detect β-galactosidase expression. Staining is observed in the limb buds and caudal regions of *Ly-6A lacZ* embryos. No staining is observed in the yolk sac of transgenic embryos or the control. (B) AGM transverse sections. E11 embryos (as above) were sectioned after X-gal staining to reveal the β-galactosidase expression pattern. Counterstaining was performed after transverse sectioning. Expression is observed in the pro/mesonephric tubules (arrow head). The arrow indicates the ventral wall of the dorsal aorta which appears negative for β-galactosidase expression.

levels of *lacZ* transcripts in the kidney as previously observed in *Ly-6E lacZ* transgenic mice and recapitulated the endogenous *Ly-6A/E* gene transcription expression pattern in adults. Although we examined only two *Ly-6A lacZ* transgenic lines, the levels of transcription of the transgene appeared to be identical between the AZ1 line, carrying eight copies of the transgene, and the AZ2 line, which had 20 copies of the transgene. Furthermore, the transcriptional levels were similar to that of the BL1b *Ly-6E lacZ* transgenic line which carries eight transgene copies (in homozygous animals). Thus, in the context of the *lacZ* reporter gene, the *Ly-6A* cassette does not appear to direct copy number dependent expression.

Both lines of *Ly-6A lacZ* transgenic adult mice express β-galactosidase similarly in haematopoietic cells. As expected from the previous results of flow cytometric analysis with the Sca-1 antibody, the *Ly-6A lacZ* transgene is expressed in all haematopoietic organs. Similar percentages of FDG-positive cells were found in the bone marrow, thymus, spleen and lymph nodes of the *Ly-6A lacZ* transgenic mice and corresponded to the percentages found in *Ly-6E lacZ* transgenic tissues. The predominant lineage

expressing the transgene is the CD4 subset of T cells. Also, cells of the CD8 subset, B and myeloid lineages were positive for transgene expression. Again, these data on the *Ly-6A lacZ* transgenic mice correspond well with the percentages of haematopoietic subsets positive for *Ly-6E lacZ* transgene expression.

Surprisingly, we did not observe the allele-specific differences noted by previous Sca-1 FACS analysis of the different allelic mouse strains and the percentages of FDG^+ haematopoietic cells were always slightly less than $Sca\text{-}1^+$ cell percentages.

Indeed, in the *Ly-6A lacZ* transgenic mice we found FDG expression in some but not all functional adult repopulating HSCs. Flow cytometric sorting of FDG and Sca-1 double-stained bone marrow showed that not all adult HSCs were in the FDG fraction. While almost all HSCs are in the Sca-1 fraction, equal numbers of HSCs are found in the FDG^+ and FDG^- fractions. We observed this same distribution in the *Ly-6E lacZ* transgenic mice. The incomplete overlap in FDG and Sca-1 staining in bone marrow may be owing to the following: (1) Ly-6A/E molecules are surface GPI-linked glycoproteins, while β-galactosidase is cytoplasmic. Thus, the kinetics of protein production as well as protein half-life could be vastly different. (2) Not all the appropriate transcriptional control elements are contained within the 14kb *Ly-6A/E* cassettes or, more likely, position effect variegation has occurred. (3) The cell permeability to FDG is inefficient. Either the entry of FDG into the cells is suboptimal or there is a loss through leakage, or both. (4) The *lacZ* gene is bacterial in origin and may be constrained in its expression in mice. For example, β-galactosidase production may reach a physiological threshold with higher levels being toxic and, thus, levels appear to be limited in the *Ly-6A lacZ* transgenic mice. At this time it is unclear which of these possibilities is responsible for suboptimal transgene expression. However, we have made several lines of transgenic mice in which mammalian genes such as the *tal-1* transcription factor (unpublished observations) and the *Bcl-2* antiapoptotic gene (unpublished observations) have been inserted into these cassettes. Both genes have been found to be expressed in haematopoietic cells.

Unlike the adult, differences in *Ly-6A lacZ* and *Ly-6E lacZ* expression were observed in transgenic embryos. The consistent expression in the limb buds and dorsal-caudal tail of the *Ly-6A lacZ* but not *Ly-6E lacZ* embryos strongly suggests that the proper regulatory elements are present, at least for these tissues. However, no functional HSCs have been isolated from either *Ly-6A lacZ* or *Ly-6E lacZ* AGMs. Sectioning and staining of this region (together with preliminary data with a *Ly-6A GFP* transgene) suggests that β-galactosidase expression is not high enough to yield an enrichment of HSCs from the AGM region.

Despite incomplete expression of the *Ly-6A lacZ* transgene in $Sca\text{-}1^+$ cells, this transgene cassette does lead to faithful expression in some HSCs. For manipulation of HSCs *in vivo* and *in vitro* and for localization of HSCs within the whole animal, the *Ly-6A* cassette appears at present to be the best transgene construct, outside of targeting a marker gene by homologous recombination in embryonic stem cells. The clear advantage in the use of *Ly-6A* sequences for regulated expression in HSCs is the relative size of this gene and, particularly important, the 3′ distal 1 kb regulatory sequence, compared with other genes encoding proteins expressed in HSCs. The genes encoding HSC marker proteins c-kit and AML-1 (Gokkel *et al*, 1992; Levanon *et al*, 2001) span over 100 kb of sequence and contain many exons and introns, thus making identification of regulatory elements difficult. Thus, the further dissection of the regulatory elements of the *Ly-6A* gene expression cassette should lead to the generation of retroviral vectors for efficient transduction of and expression in HSCs.

ACKNOWLEDGMENTS

We sincerely thank all the members of the laboratory for help with experiments, critical comments and suggestions regarding this work. We also thank the EDC staff for assistance with animal breeding and care. This work was funded by Netherlands Scientific Research Organization 901–08–090, National Institutes of Health R01 DK51077, Netherlands Cancer Society EUR1999–65, La Ligue Nationale Contre le Cancer (CR) and the European Community QLK-CT-1999–00020.

REFERENCES

Bamezai, A. & Rock, K.L. (1995) Overexpressed Ly-6A.2 mediates cell-cell adhesion by binding a ligand expressed on lymphoid cells. *Proceedings of the National Academy of Sciences of the United States of America*, 92, 4294–4298.

de Bruijn, M.R.T.R., Speck, N.A., Peeters, M.C.E. & Dzierzak, E. (2000) Definitive hematopoietic stem cells first emerge from the major arterial regions of the mouse embryo. *EMBO Journal*, 19, 2465–2474.

Fraser, P., Hurst, J., Collis, P. & Grosveld, F. (1990) DNaseI hypersensitive sites 1, 2 and 3 of the human beta-globin dominant control region direct position-independent expression. *Nucleic Acids Research*, 18, 3503–3508.

Gokkel, E., Grossman, Z., Ramot, B., Yarden, Y., Rechavi, G. & Givol, D. (1992) Structural organization of the murine *c-kit* proto-oncogene. *Oncogene*, 7, 1423–1429.

Kamiura, S., Nolan, C.M. & Meruelo, D. (1992) Long-range physical map of the Ly-6 complex: mapping the Ly-6 multigene family by field-inversion and two-dimensional gel electrophoresis. *Genomics*, 12, 89–105.

Khan, K.D., Lindwall, G., Maher, S.E. & Bothwell, A.L. (1990) Characterization of promoter elements of an interferon-inducible Ly- 6E/A differentiation antigen, which is expressed on activated T cells and hematopoietic stem cells. *Molecular and Cellular Biology*, 10, 5150–5159.

Khan, K.D., Shuai, K., Lindwall, G., Maher, S.E., Darnell, J.E. & Bothwell, A.L. (1993) Induction of the *Ly-6A/E* gene by interferon alpha/beta and gamma requires a DNA element to which a tyrosine-phosphorylated 91-kDa protein binds. *Proceedings of the National Academy of Sciences of the United States of America*, 90, 6806–6810.

Kimura, S., Tada, N., Liu-Lam, Y. & Hammerling, U. (1984) Studies of the mouse Ly-6 alloantigen system. II. Complexities of the Ly-6 Region. *Immunogenetics*, 20, 47–56.

LeClair, K.P., Palfree, R.G., Flood, P.M., Hammerling, U. & Bothwell, A. (1986) Isolation of a murine *Ly-6* cDNA reveals a new multigene family. *EMBO Journal*, 5, 3227–3234.

Levanon, D., Glusman, G., Bangsow, T., Ben-Asher, E., Male, D.A., Avidan, N., Bangsow, C., Hattori, M., Taylor, T.D., Taudien, S., Blechschmidt, K., Shimizu, N., Rosenthal, A., Sakaki, Y., Lancet, D. & Groner, Y. (2001) Architecture and anatomy of the genomic locus encoding the human leukemia-associated transcription factor RUNX1/AML1. *Gene*, 262, 23–33.

Ma, X., Ling, K.-W. & Dzierzak, E. (2001) Cloning of the *Ly-6A* (Sca-1) gene locus and identification of the 3′ distal fragment responsible for high level γ interferon-induced expression *in vitro*. *British Journal of Haematology*, 114, 724–730.

McGrew, J.T. & Rock, K.L. (1991) Isolation, expression, and sequence of the *TAP/Ly-6A.2* chromosomal gene. *Journal of Immunology*, 146, 3633–3638.

Miles, C., Sanchez, M.-J., Sinclair, A. & Dzierzak, E. (1997) Expression of the *Ly-6E.1* (Sca-1) transgene in adult hematopoietic stem cells and the developing mouse embryo. *Development*, 124, 537–547.

Okada, S., Nakauchi, H., Nagayoshi, K., Nishikawa, S., Miura, Y. & Suda, T. (1992) *In vivo* and *in vitro* stem cell function of c-kit- and Sca-1-positive murine hematopoietic cells. *Blood*, 80, 3044–3050.

Palfree, R.G. & Hammerling, U. (1986) Biochemical characterization of the murine activated lymphocyte alloantigen Ly-6E.1 controlled by the Ly-6 locus. *Journal of Immunology*, 136, 594–600.

Palfree, R.G., LeClair, K.P., Bothwell, A. & Hammerling, U. (1987) cDNA characterization of an Ly-6.2 gene expressed in BW5147 tumor cells. *Immunogenetics*, 26, 389–391.

Reiser, H., Coligan, J., Palmer, E., Benacerraf, B. & Rock, K.L. (1988) Cloning and expression of a cDNA for the T-cell-activating protein TAP. *Proceedings of the National Academy of Sciences of the United States of America*, 85, 2255–2259.

van de Rijn, M., Heimfeld, S., Spangrude, G.J. & Weissman, I.L. (1989) Mouse hematopoietic stem-cell antigen Sca-1 is a member of the Ly-6 antigen family. *Proceedings of the National Academy of Sciences of the United States of America*, 86, 4634–4638.

Rock, K.L., Yeh, E.T., Gramm, C.F., Haber, S.I., Reiser, H. & Benacerraf, B. (1986) TAP, a novel T cell-activating protein involved in the stimulation of MHC-restricted T lymphocytes. *Journal of Experimental Medicine*, 163, 315–333.

Sinclair, A., Daly, B. & Dzierzak, E. (1996) The *Ly-6E.1* (Sca-1) gene requires a 3′ chromatin-dependent region for high-level gamma-interferon-induced hematopoietic cell expression. *Blood*, 87, 2750–2761.

Sinclair, A.M. & Dzierzak, E.A. (1993) Cloning of the complete *Ly-6E.1* gene and identification of DNase I hypersensitive sites corresponding to expression in hematopoietic cells. *Blood*, 82, 3052–3062.

Spangrude, G.J. & Brooks, D.M. (1993) Mouse strain variability in the expression of the hematopoietic stem cell antigen Ly-6A/E by bone marrow cells. *Blood*, 82, 3327–3332.

Spangrude, G.J., Heimfeld, S. & Weissman, I.L. (1988) Purification and characterization of mouse hematopoietic stem cells. *Science*, 241, 58–62.

Spanopoulou, E., Giguere, V. & Grosveld, F. (1988) Transcriptional unit of the murine *Thy-1* gene: different distribution of transcription initiation sites in brain. *Molecular and Cellular Biology*, 8, 3847–3856.

Stanford, W.L., Bruyns, E. & Snodgrass, H.R. (1992) The isolation and sequence of the chromosomal gene and regulatory regions of *Ly-6A.2*. *Immunogenetics*, 35, 408–411.

Stanford, W.L., Haque, S., Alexander, R., Liu, X., Latour, A.M., Snodgrass, H.R., Koller, B.H. & Flood, P.M. (1997) Altered proliferative response by T lymphocytes of *Ly-6A* (Sca-1) null mice. *Journal of Experimental Medicine*, 186, 705–717.

Chapter 4

The *Ly-6A (Sca-1) GFP* transgene is expressed in all adult mouse hematopoietic stem cells

Xiaoqian Ma, Catherine Robin, Katrin Ottersbach, Elaine Dzierzak

Stem Cells 2002; 20: 514-521

The *Ly-6A* (Sca-1) *GFP* Transgene is Expressed in all Adult Mouse Hematopoietic Stem Cells

XIAOQIAN MA,[a,b] CATHERINE ROBIN,[b] KATRIN OTTERSBACH,[b] ELAINE DZIERZAK[b]

[a]Pathology Department and [b]Department of Cell Biology and Genetics, Erasmus University, Rotterdam, Netherlands

Key Words. *Hematopoietic stem cells · Sca-1 · Ly-6A GFP · Transgenic mouse*

ABSTRACT

The Sca-1 cell surface glycoprotein is used routinely as a marker of adult hematopoietic stem cells (HSCs), allowing a >100-fold enrichment of these rare cells from the bone marrow of the adult mouse. The Sca-1 protein is encoded by the *Ly-6A/E* gene, a small 4-exon gene that is tightly controlled in its expression in HSCs and several hematopoietic cell types. For the ability to sort and localize HSCs directly from the mouse, we initiated a transgenic approach in which we created *Ly-6A* (Sca-1) green fluorescent protein (*GFP*) transgenic mice. We show here that a 14-kb *Ly-6A* expression cassette directs the transcription of the *GFP* marker gene in all functional repopulating HSCs in the adult bone marrow. A >100-fold enrichment of HSCs occurred by sorting for the *GFP*-expressing cells. Furthermore, as shown by fluorescence-activated cell sorting and histologic analysis of several hematopoietic tissues, the *GFP* transgene expression pattern generally corresponded to that of Sca-1. Thus, the *Ly-6A GFP* transgene facilitates the enrichment of HSCs and presents the likelihood of identifying HSCs in situ. *Stem Cells 2002;20:514-521*

INTRODUCTION

During adult stages, the hematopoietic system is constantly renewed from rare hematopoietic stem cells (HSCs) harbored in the bone marrow. HSCs can be retrospectively identified based on their functional repopulation properties observed only by transplantation into hematopoietic-depleted adult recipients. The properties that define HSCs are long-term, high-level repopulation of all hematopoietic lineages and the ability to self-renew [1, 2].

Using the transplantation assay to identify stem cells, HSCs have been enriched and characterized by flow cytometric sorting using a wide range of antibodies detecting cell-surface markers. The established surface marker profile of adult HSCs is high expression of c-kit and Sca-1, low expression of Thy-1, and an absence (or very low expression) of mature lineage markers, including CD3, CD4, and CD8 for T lymphocytes, B220 for B lymphocytes, Mac-1 for macrophages, and Gr-1 for granulocytes [3-5]. This sorting procedure has been verified by many laboratories and has been used to sort HSCs to relative homogeneity [6]. However, it is still unclear whether the use of antibody-mediated cell sorting has any activating effect on HSCs, since, for example, in lymphocytes, antibody-mediated cross-linking of some cell-surface proteins results in cell proliferation and/or differentiation.

Correspondence: E. Dzierzak, Ph.D., Department of Cell Biology and Genetics, Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, Netherlands. Telephone: 31-10-408-7169; Fax: 31-10-408-9468; e-mail: dzierzak@ch1 fgg.eur.nl *Received May 29, 2002; accepted for publication July 11, 2002.*

Sca-1, a glycoprotein-I-linked cell-surface glycoprotein, was one of the first epitopes used for antibody-mediated enrichment of murine HSCs. *Spangrude et al.* [3] used the Sca-1-specific monoclonal E13-161.7 antibody [7] for sorting HSCs from adult murine bone marrow to yield about a 100-fold enrichment of these cells [3, 4]. Since then, others have used Sca-1 as a marker for enrichment of embryonic day 14 (E14) fetal liver and E11 yolk sac HSCs [8]. While the Sca-1 antibody is useful for enrichment of viable HSCs, immunostaining of tissue sections, particularly from embryos, has been difficult. Hence, to more specifically localize Sca-1$^+$ cells in embryonic tissues and adult bone marrow, a transgenic approach with the gene encoding Sca-1 has been taken.

The Sca-1 epitope is encoded by the strain-specific *Ly-6E/A* allelic gene [9]. Examination of the specific expression pattern of *Ly-6E/A* has been difficult [9-11], since the *Ly-6* gene family consists of at least 18 highly homologous cross-hybridizing genes [12, 13]. However, it was found that the Ly-6A and Ly-6E proteins differ in their expression pattern on marrow-repopulating cells, with *Ly-6A* strains of mice expressing Sca-1 on 99% and *Ly-6E* strains expressing Sca-1 on 25% of such cells [11]. Nonetheless, the Sca-1 marker has proven useful in identifying HSCs, and the small size of the *Ly-6A/E* locus and flanking transcriptional regulatory elements make it useful in directing marker gene expression in HSCs in transgenic mice.

The transcriptional regulatory elements and 5′ and 3′ DNaseI hypersensitivity patterns of *Ly-6E* and *Ly-6A* have been examined and appear to be similar [14-17]. Deletional studies show that the 3′ flanking regions contain two hypersensitive sites, which are most likely responsible for high-level, γ-interferon-induced expression in vitro [18, 19]. This distal flanking element is also responsible for high-level, tissue-specific expression of a *lacZ* marker gene in transgenic mice [20, 21]. HSCs can be enriched from adult bone marrow based on *lacZ* transgene expression [20, 21]. However, not all HSCs are located within the lacZ-expressing population. Whether this is a result of the *lacZ* marker gene, the *Ly-6A/E* expression cassette, or the fluorescein di-β-D-galactopyranoside (FDG) substrate staining method [22] is uncertain.

Thus, to further examine, identify, and localize HSCs in the adult mouse, we created transgenic mice with a green fluorescent protein (*GFP*) gene within the context of the *Ly-6A* expression cassette. We report here that all bone marrow HSCs were positive for expression of GFP from the *Ly-6A* transgene. *Ly-6A GFP* was also expressed in other Sca-1$^+$ hematopoietic cells, and histological analyses revealed an interesting pattern of expression of GFP within the thymus, bone marrow, and spleen cells, suggestive of hematopoietic and endothelial cell expression. Thus, the *Ly-6A GFP* transgene marker facilitated the enrichment of functional adult repopulating HSCs and their identification in situ.

Materials and Methods

Transgenic Mice

The 14-kb Ly-6A cassette (pLy-6A14) was constructed as described previously [18, 21]. The enhanced GFP (EGFP) gene AgeI/AflII fragment containing simian virus 40 polyA in pEGFP-N1 (Clontech; Alphen aan den Rijn, The Netherlands; http://www.clontech.com/index.shtml) was modified, converting 3′ and 5′ AgeI/AflII sites to NarI sites using oligonucleotide adaptors. The 970-bp EGFP NarI fragment was cloned into pLy-6A14 to generate pLAG.

Fertilized (C57BL/10 × CBA)F_1 oocytes were microinjected with a 15-kb BamH1 fragment containing the *Ly-6A GFP* gene from pLAG (Fig. 1A). This fragment was gel purified for removal of all vector sequences. Positive founder animals were bred with (C57BL/10 × CBA)F_1 mice, and lines were maintained as heterozygotes. Polymerase chain reaction (PCR) analysis of tail DNA was used to identify founder transgenic mice. Animals were housed according to institutional guidelines, with free access to food and water. Animal procedures were carried out in compliance with the Standards for Humane Care and Use of Laboratory Animals.

DNA and RNA Analysis

Genomic DNA (5-10 μg) for Southern blot analysis [20] was digested with BamH1 and electrophoresed through 1% agarose/Tris, acetate, EDTA gels prior to transfer to Hybond-N membranes. Transgene copy number controls were generated by addition of appropriate amounts of pLAG to nontransgenic genomic DNA. Filters were probed with *GFP* and *Thy-1* gene fragments. Normalization for DNA content of each lane was performed after phosphorimaging analysis of *Thy-1* signal. Copy number was determined subsequently by comparing the *GFP* signal obtained from the transgenic mice with that of the plasmid controls on the linear portion of the standard curve.

Total cellular RNA for Northern blot analysis was prepared using the lithium chloride/urea method, and 5-15 μg was fractionated on 1% agarose/formaldehyde gels [23] prior to transfer to Hybond-N membranes. Filters were probed with *GFP* and *Ly-6E* cDNA fragments. Probes used for hybridization to Southern or Northern filters were labeled by a random oligonucleotide priming procedure, incorporating ^{32}P ATP. The fragments used were as follows:

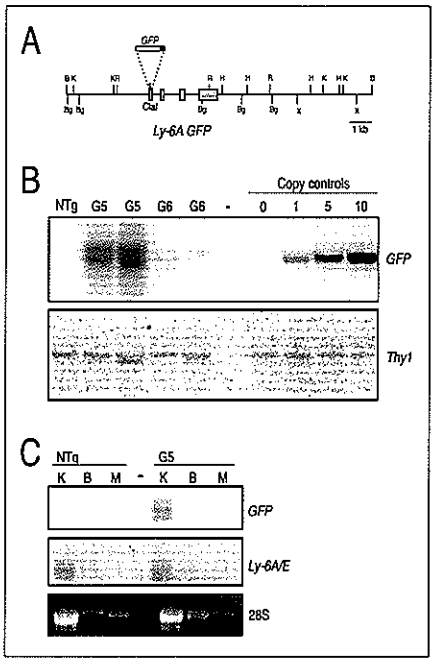

Figure 1. Ly-6A GFP *transgene construct and transgenic mouse characterization. A) A schematic drawing of the 14-kb BamHI fragment of the* Ly-6A *gene and the insertion of the* EGFP *gene into the ClaI cloning site. Restriction sites are indicated. B = BamH1, BG = BglII, K = KpnI, R = EcoR1, H = HindIII, and X = XbaI. B) Southern blot analysis of DNA from two* Ly-6A GFP *transgenic lines, G5 and G6, and a nontransgenic control (NTg). Copy controls are nontransgenic DNA with* Ly-6A GFP *plasmid DNA added at 1, 5, and 10 copy equivalents. After hybridization with* GFP *and* Thy-1 *gene probes, the intensity of signal was determined by phosphorimaging. C) Northern blot analysis of total RNA from kidney (K), bone marrow (B), and muscle (M) of an adult G5 transgenic and a nontransgenic (NTg) littermate. Hybridization was performed with a* GFP *gene and* Ly-6E *cDNA probes, and the 28S RNA signal was used for RNA quantitation.*

(970 bp PCR EGFP products from pLAG); 1.2 kb XbaI-NruI *Thy-1* gene fragment from pD7 [24]; 761 bp EcoR1 *Ly-6E* cDNA fragment from pLy6.1-2R [13]. After hybridization, filters were washed to a stringency of 0.2 × standard saline citrate/0.1% SDS and exposed to a phosphorimager screen for quantitation using Imagequant software.

Genomic DNA (200 ng) from the peripheral blood of transplant recipients was analyzed by PCR using oligonucleotide primers for GAPDH-specific sequences, (GAPDH1) 5′CTTCACCACCATGGAGAAGG 3′ and (GAPDH2) 5′CCACCCTGTTGCTGTAGCC 3′, and for *GFP*-specific sequences, (Ly6GFP) 5′GACAGAACTTGCCACTGTGC 3′ and (GFP) 5′AAGAAGATGGTGCGCTCCTG 3′.

DNA was subjected to an initial 5 minute denaturation at 94°C followed by 30 cycles of denaturation (1 minute at 94°C), annealing (2 minutes at 55°C), and elongation (2 minutes at 72°C). Serial dilutions of blood DNA from a transgenic animal were used as a control to evaluate the levels of donor cell reconstitution in transplanted mice.

Analysis of Long-Term Multilineage Repopulating Activity

Bone marrow cells were isolated from the femurs and tibias of adult transgenic mice and sorted for GFP expression on a FACSVantage (Becton Dickinson; San Jose, CA; http://www.bd.com), and sorted adult bone marrow cells (in limiting dilution) were assayed for the presence of HSCs by intravenous transfer into irradiated adult recipients, as described [25, 26]. Briefly, (C57BL/10 × CBA)F_1 male recipients were exposed to a split dose of 900 rad of gamma-irradiation from a ^{137}Cs source. Recipient mice were bled at 1 and >4 months after transfer and analyzed for percentage donor contribution by donor marker-specific PCR on DNA isolated from peripheral blood [20, 25, 26]. Reconstitution was evaluated by ethidium bromide staining of agarose gels, and in some cases, by Southern blot hybridization, as described previously [25-27]. To test for long-term multilineage hematopoietic repopulation, genomic DNA was isolated from peripheral blood, thymus, lymph node, sorted splenic B and T cells, and bone marrow myeloid and lymphoid cells. Percentage donor-cell contribution was analyzed by PCR, and in some cases, by Southern blot hybridization and phosporimaging.

Fluorescent Antibody Surface Staining and Flow Cytometry

All antibodies used in flow cytometric sorting and analysis were obtained from PharMingen (Alphen aan den Rijn, The Netherlands; http://www.bdbiosciences.com/pharmingen). The monoclonal antibodies used were directly conjugated with either phycoerythrin (PE) or biotin and included: PE-anti-c-kit, PE-anti-Sca-1, PE-anti-CD4, PE-anti-CD8, PE-anti-B220, PE-anti-Mac, and biotinylated CD31 and CD34. Single-cell suspensions were prepared as described [21]. After incubation with specific antibodies for 30 minutes on ice, cells were washed twice and incubated with PE-conjugated streptavidin (Caltag Laboratories; Burlingame, CA; http://www.caltag.com) when required. Labeled cells were finally washed twice and filtered through a nylon mesh screen prior to sorting. One μg/ml Hoechst 33258 or 2 μg/ml 7 amino-actinomycin D were added to identify dead cells. To determine the background levels, cells were stained with fluorochrome-conjugated immunoglobulin isotype controls from Pharmingen. During the entire staining procedure, phosphate-buffered saline (PBS) containing 10% fetal calf serum and penicillin/

streptomycin was used. Cells were sorted using a FACSVantage SE (Becton-Dickinson); the purity of the sorted cells ranged from 89%-98%. Analyses were performed on a FACScan.

Cryosectioning and Histology

Tissues (kidney, bone marrow, spleen, lymph nodes, thymus, and brain) from adult *Ly-6A GFP* mice were isolated and fixed for 1 hour at room temperature in 2% paraformaldehyde/PBS. Tissues were equilibrated in 20% sucrose/PBS overnight at 4°C, quick frozen in tissue tek, and stored in liquid nitrogen until cryosectioning. 10-μm thick sections were placed on glass slides, air-dried and mounted with vectashield (Vector Laboratories; Burlingame, CA; http://www.vectorlabs.com).

Results

Generation of *Ly-6A GFP* Transgenic Mice

To provide a means by which HSCs could be enriched without the use of antibodies or FDG (β-galactosidase substrate) and to facilitate in vivo HSC localization, we generated transgenic mice expressing the *GFP* reporter gene under the control of the *Ly-6A* gene regulatory sequences. Since our previous results with *Ly-6A* transgenic embryos showed incomplete expression of the inserted *lacZ* marker in bone marrow HSCs [20, 21], we sought to alleviate a possible experimentation problem (due to cell permeabilization necessary for FDG substrate entry into these cells) by using the *GFP* reporter gene. The enhanced *GFP* gene was inserted into the ClaI cloning site of the first untranslated exon of *Ly-6A* (Fig. 1A) and was flanked by 4 kb of upstream and 10 kb of downstream genomic sequences containing regulatory elements necessary for high-level specific expression in adult bone marrow HSCs. Transgenic mice were generated, and mice from line G5, containing eight copies of transgene (hemizygous), were used in all the studies described here (Fig. 1B). As expected from other published studies [20, 21], Northern blot analysis of adult tissues revealed high-level transgene expression in the kidney (Fig. 1C). Low-level transgene expression was found in the bone marrow, and limited or no expression was found in muscle tissue. Thus, the *Ly-6A GFP* transgene produced a restricted expression pattern similar to that found previously in *Ly-6A lacZ* and *Ly-6E lacZ* transgenic mice [20, 21].

Ly-6A GFP Expression in Tissues and Subsets of Hematopoietic Cells of Adult Transgenic Mice

Previously, it was shown, by immunostaining of sectioned adult hematopoietic tissues, that Sca-1 is expressed on some cells of the thymus, spleen, and lymph nodes [28]. Sca-1 is also expressed in the brain vasculature and the cortical tubules of the kidney [9, 29]. Hence, histologic sections from the hematopoietic tissues of *Ly-6A GFP* adult mice were examined for transgene expression under a fluorescence microscope. As shown in Figure 2, GFP signal was detected in some of the cells of the kidney (D), bone marrow (E), spleen (F), lymph nodes (G), thymus (H), and brain (I). Within the kidney,

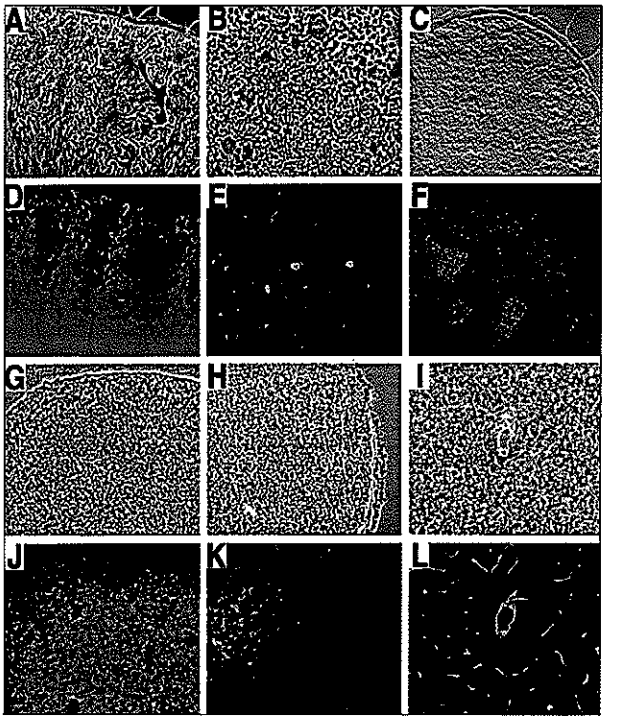

Figure 2. Histologic sectioning and microscopy of* Ly-6A GFP *hematopoietic tissues. *Brightfield and fluorescent images, respectively, of kidney (A and D), bone marrow (B and E), spleen (C and F), lymph node (G and J), thymus (H and K), and brain (I and L). Tissues were cryosectioned at a thickness of 10 microns. Images B, E, H, K, I, and L were taken at 10× magnification. All other images were taken at 4× magnification.*

the cortical tubules were highly positive. The bone marrow showed highly GFP$^+$ endothelial cells in the vasculature and other scattered positive cells (presumably, HSCs, some progenitors, and stromal cells). In the peripheral hematopoietic tissues, the GFP expression pattern appeared to be strongest in the lymphoid areas of the spleen, in the lymph nodes, and in the medulla of the thymus (with a few subcapsular thymus cells also GFP$^+$). Finally, a vascular endothelial pattern of GFP expression was found in the brain. Thus, the general expression pattern of the *Ly-6A GFP* transgene was identical to the published Sca-1 immunostaining patterns [9, 28, 29].

To determine the percentage of GFP$^+$ cells in each of the tissues, flow cytometric analysis was performed. All hematopoietic tissues tested, i.e., bone marrow, thymus, lymph node, and spleen, contained GFP$^+$ cells (Table 1). The lymph node contained the highest percentage (80%) of cells expressing the GFP marker, while the bone marrow and thymus contained less than 20% of GFP$^+$ cells. Sca-1 antibody staining of cells from each of the hematopoietic tissues showed an overlap in the expression of GFP with the endogenous Sca-1 protein. Approximately 50% or more of Sca-1$^+$ cells in bone marrow, thymus, spleen, and lymph node were also GFP$^+$ (data not shown). In the lymph node, thymus, and spleen, 40% or greater of GFP$^+$ cells were Sca-1$^+$. However, in the bone marrow, a much lower percentage of GFP$^+$ cells were Sca-1$^+$, suggesting that GFP was expressed more highly or was more readily detectable than the Sca-1 antigen in the progenitors and stem cells of this tissue.

To determine the distribution of the GFP phenotype within hematopoietic lineages, antibodies directed against progenitor/stem cell markers and mature lineage markers were used in flow cytometric analyses (Table 1). In the bone marrow, cells of all mature lineages, T cell (CD4 and CD8), B cell (B220), and myeloid (Mac-1), and also progenitor/stem cells (c-kit, CD34, CD31) were GFP$^+$. Monocytes/granulocytes and their precursors (Mac-1$^+$) represented the largest percentage (72%) of bone marrow cells expressing GFP. In the thymus and lymph nodes, high percentages of T lymphoid cells, most notably CD4 cells, were found to be GFP$^+$. Cells of the myeloid lineage and the B lymphoid lineage in the spleen were also GFP$^+$. Thus, the *Ly-6A GFP* transgene was expressed to some degree in all hematopoietic lineages but appeared to be predominantly expressed in mature T cells in the lymphoid organs and in myeloid and immature hematopoietic progenitor/stem cells in the bone marrow.

Ly-6A GFP Expression Marks All Functional Bone Marrow HSCs

To determine whether the *Ly-6A GFP* transgene was expressed in functional adult repopulating HSCs, we performed transplantation experiments with flow cytometric-sorted cells from adult bone marrow. A representative fluorescence profile of bone marrow cells sorted on the basis of endogenous Sca-1 and GFP expression is shown in Figure 3A. In this profile, at least five populations of cells are apparent; most predominant are the Sca-1$^-$GFP$^-$, Sca-1$^-$GFP^{lo}, and Sca-1$^-$GFP^{hi}, and less frequent are Sca-1$^+$GFP^{lo+neg} and Sca-1$^+$GFP^{hi}. These five populations were sorted and injected into irradiated adult recipients at limiting dilution to test for long-term, high-level, multilineage repopulation. At greater than 4 months posttransplantation, peripheral blood DNA was tested for the presence of the *Ly-6A GFP* transgene marker. Only recipients showing greater than 10% donor cell repopulation using a semiquantitative PCR method for the *GFP* transgene were considered positive for HSC repopulation. The results of three independent sorting and transplantation experiments show that all HSC activity was found within the GFP^{hi} fractions (Fig. 3B). These animals were high-level, multilineage engrafted as determined by donor marker PCR analysis of DNA from hematopoietic tissues and sorted populations of B, T, and myeloid cells (Fig. 3D). No HSC activity was found in the GFP^{lo} or GFP$^-$ fractions. Although one animal receiving 5×10^5 Sca-1$^-$GFP$^-$ cells was positive, this recipient was only 10% engrafted by the donor cells, suggesting no enrichment for HSCs compared with the whole bone marrow control. As expected, most of the HSC activity was highly enriched in the Sca-1$^+$ fraction of the GFP^{hi} cells. As few as 1,000 Sca-1$^+$ GFP^{hi} cells resulted in reproducibly high-level multilineage repopulation (Fig. 3B). Limited HSC activity was also found in the Sca-1$^-$ GFP^{hi} fraction, but only when 10^5 cells were transplanted (Fig. 3C). These results demonstrate that GFP expression has no adverse effects on adult bone marrow HSCs and can be used to enrich for these cells by approximately 100-fold. Finally, these results, in combination with the expression

Table 1. Hematopoietic lineage marker expression on GFP$^+$ cells within adult tissues

	Bone marrow	Thymus	Lymph node	Spleen
% GFP$^+$	15.4	10.8	80.7	26.3
Marker	% marker$^+$ cells within GFP$^+$ population			
Sca-1	13.3	40.6	65.0	69.4
c-kit	9.7	nd	nd	nd
CD34	8.5	nd	nd	nd
CD31	35.0	nd	nd	nd
CD4	8.9	83.1	62.1	41.8
CD8	3.0	40.9	43.7	25.4
B220	16.2	nd	22.8	48.5
Mac-1	72.0	nd	nd	16.6

Abbreviation: nd = not done.

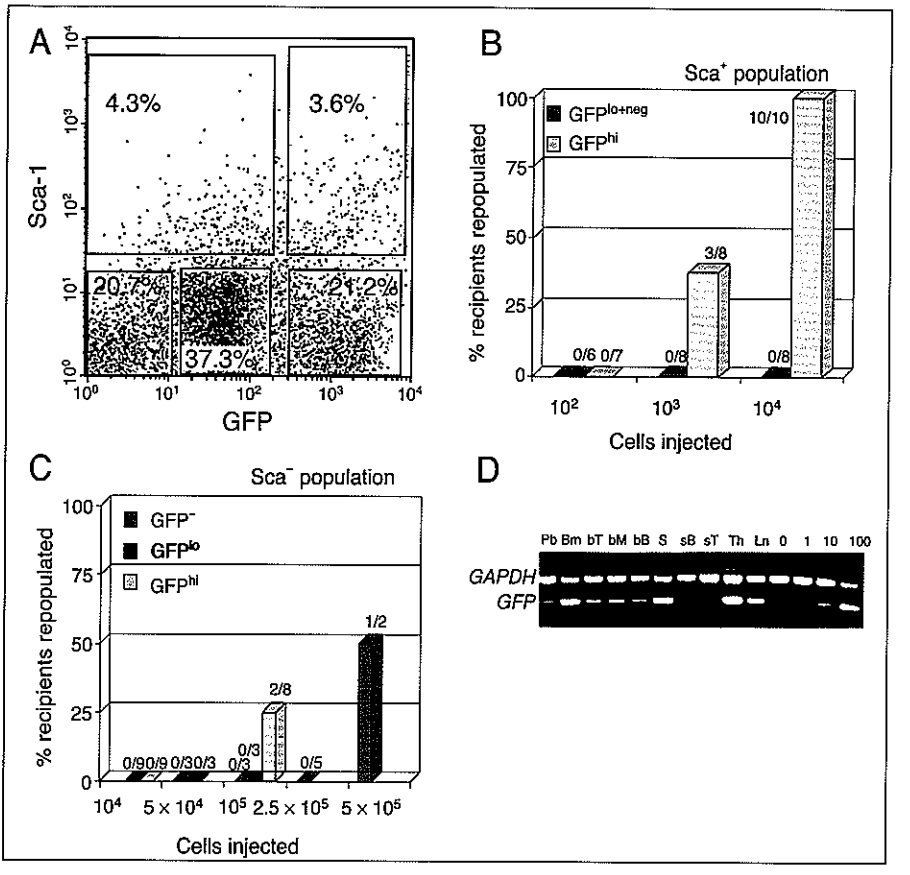

Figure 3. A) Fluorescence profiles and gates used for sorting GFP⁺ and Sca-1 antibody stained Ly-6A GFP *transgenic adult bone marrow are shown. The percentages of bone marrow cells in each sorted population are indicated. B) Transplantation data from limiting dilution transplantation of Sca-1⁺ cells within the GFP$^{low+neg}$ and GFP^{high} gates. The X axis shows the number of cells injected per recipient. The bars represent the percentage of recipients found positive for donor repopulation at greater than 4 months posttransplantation. The number of recipients repopulated/total number of recipients is indicated for each bar. C) Transplantation data from limiting dilution transplantation of Sca-1⁻ cells within the GFP⁻, GFP^{low}, and GFP^{high} gates. The axis and bar indications are as described for (B). Note the higher cell numbers transplanted. D) Hematopoietic multilineage analysis of a recipient repopulated with Sca-1⁺GFP^{high} cells. At greater than 4 months posttransplantation, DNA was isolated from peripheral blood (Pb), bone marrow (Bm), bone marrow T cells (bT), bone marrow myeloid cells (bM), bone marrow B cells (bB), spleen (S), spleen B cells (sB), spleen T cells (sT), thymus (Th), and lymph nodes (Ln). PCR analysis was performed using oligonucleotide primers for* GFP *(bottom band) and* GAPDH *(top band). 0, 1, 10, and 100 were the controls indicating percentage donor cell engraftment. Transplantation results were obtained from three independent experiments.*

observed in tissue sections, suggest that the GFP marker is more highly expressed by HSCs than is Sca-1.

Discussion

We have shown here that the *Ly-6A GFP* transgene faithfully marks all HSCs in adult bone marrow. This is in contrast to previous transgenic mice made with the *Ly-6A* and *Ly-6E* expression cassettes directing the transcription of the *lacZ* marker gene [20, 21]. In those mice, some HSCs did not express *lacZ*. It is possible that those transgenic HSCs were inefficiently labeled by FDG. Not all HSCs may be sufficiently permeabilized so as to allow entry of

FDG and, thus, appeared to be negative for *lacZ* expression. In addition, too much permeability may allow leakage of the substrate and again result in no or low FDG signal. Alternatively, it may be possible that expression of the bacterial *lacZ* is regulated in transgenic mice (for purposes such as viability) so that only low-level expression is allowed in HSCs. Despite the incomplete isolation of bone marrow HSCs from *Ly-6E lacZ* and *Ly-6A lacZ* transgenic mice, the *GFP* marker in the *Ly-6A* cassette was optimally expressed in all bone marrow HSCs and allowed the enrichment of all HSCs. Thus, the *Ly-6A* cassette appears to contain all the necessary transgene regulatory elements for HSC-specific expression. And when used together in a transgene context, the *Ly-6A* cassette and the *GFP* marker are highly advantageous in maintaining a viable, unmanipulated pool of HSCs in the absence of enzyme substrate addition or antibody staining.

Is Sca-1 surface glycoprotein limiting on the surface of some HSCs and hematopoietic cells? The expression profile of *Ly-6A GFP* in adult tissues and hematopoietic cells is reminiscent of the Sca-1 profile [9, 10, 28, 29], particularly in mature hematopoietic cells of T lymphoid, B lymphoid, and myeloid lineages. However, our fluorescence-activated cell sorting profiles suggest that not all GFP^+ cells were $Sca\text{-}1^+$. Indeed, the fluorescence signal produced by GFP was more intense. This is most likely due to the fact that, within the G5 transgenic line of mice, there are eight copies of the transgene, compared with the normal diploid copy number of the endogenous *Ly-6A/E* gene encoding Sca-1. Furthermore, unlike the cell-surface localization of Sca-1, which requires extra processing steps, such as transport through the plasma membrane, glycosylation, and GPI linkage, GFP is an easily expressed cytoplasmic protein. Thus, GFP may be a better marker than Sca-1 in low-expressing cells of the various hematopoietic tissues, particularly HSCs.

Although we have not directly compared limiting dilution transplantation of Sca-1-sorted HSCs with GFP-sorted HSCs, the general enrichment gained by *Ly-6A GFP* sorting is about 100-fold. This is similar to the Sca-1 enrichment levels established in the published literature [3, 10]. While differences do not appear to exist on the level of bone marrow HSC enrichment, we are examining whether differences exist between antibody-mediated sorting and GFP sorting in the intraembryonic tissue-generating HSCs at the earliest stages of mouse development. This earliest site, the aorta-gonads-mesonephros (AGM) region, generates the first adult repopulating HSCs at E10.5 [27]. Interestingly, in related studies, we reproducibly isolated only 50% of HSCs from the AGM region using the Sca-1 antibody, but could isolate 100% of AGM HSCs using *Ly-6A GFP* marker expression [30]. This supports the notion that the *Ly-6A GFP* marker is more efficiently expressed and identifies the $Sca\text{-}1^{low}$-expressing population of HSCs as they are being generated within the embryo. This may be particularly useful in future studies of precursor-progeny relationships and possible lineage relationships of HSCs with endothelial cells.

Further to this point, when we used a *lacZ* marker gene in the *Ly-6A* cassette, we found high-level β-galactosidase expression in the mesonephros of the E11 AGM but no expression in the dorsal aorta, the site where the first AGM HSCs are generated [21]. HSC sorting procedures based on the FDG substrate for β-galactosidase were unsuccessful. However, we have found GFP-expressing cells lining the walls of the dorsal aorta in *Ly-6A GFP* embryos, and these cells contained all the adult repopulating HSC activity [30]. While both *lacZ* and *GFP* marker gene expressions were found in the mesonephros, only the *GFP* reporter was optimally expressed in AGM HSCs. Thus, *GFP* expression from the *Ly-6A* cassette in the G5 line of transgenic mice serves as an excellent tool for HSC enrichment and for the localization of the HSCs in the adult as well as the embryo. Further transgenic studies will focus on the deletion mapping of HSC-specific transcriptional regulatory elements so as to express exogenous genes only in HSCs. Combined with inducible expression, the *Ly-6A* transgenesis will allow for a better understanding of HSC regulation and migration during each precise stage of development and adult hematopoietic maintenance.

ACKNOWLEDGMENTS

The authors would like to thank *Nancy Speck, Marella de Bruijn, Marian Peeters, Kam-Wing Ling*, and all the members of the lab for help; *John Kong-A-San* for microinjections; *Corne Snoijs* for flow cytometry; and *Danielle Zondervan-Bakker* and *Henk Dronk* for animal care. This work was funded by Netherlands Scientific Research Organization 901-08-090, National Institutes of Health R01 DK51077, La Ligue Nationale Contre le Cancer (*C.R.*), the European Community QLK-CT-1999-00020, and The Wellcome Trust (*K.O.*).

REFERENCES

1 Lemischka IR. Clonal, in vivo behavior of the totipotent hematopoietic stem cell. Semin Immunol 1991;3:349-355.

2 Spangrude GJ, Smith L, Uchida N et al. Mouse hematopoietic stem cells. Blood 1991;78:1395-1402.

3 Spangrude GJ, Heimfeld S, Weissman IL. Purification and characterization of mouse hematopoietic stem cells. Science 1988;241:58-62.

4 Okada S, Nakauchi H, Nagayoshi K et al. In vivo and in vitro stem cell function of c-kit- and Sca-1-positive murine hematopoietic cells. Blood 1992;80:3044-3050.

5 Morrison SJ, Wandycz AM, Hemmati HD et al. Identification of a lineage of multipotent hematopoietic progenitors. Development 1997;124:1929-1939.

6 Uchida N, Aguila HL, Fleming WH et al. Rapid and sustained hematopoietic recovery in lethally irradiated mice transplanted with purified Thy-1.1^{lo} Lin^{-}Sca-1^{+} hematopoietic stem cells. Blood 1994;83:3758-3779.

7 Aihara Y, Buhring HJ, Aihara M et al. An attempt to produce "pre-T" cell hybridomas and to identify their antigens. Eur J Immunol 1986;16:1391-1399.

8 Huang H, Auerbach R. Identification and characterization of hematopoietic stem cells from the yolk sac of the early mouse embryo. Proc Natl Acad Sci USA 1993;90:10110-10114.

9 van de Rijn M, Heimfeld S, Spangrude GJ et al. Mouse hematopoietic stem-cell antigen Sca-1 is a member of the Ly-6 antigen family. Proc Natl Acad Sci USA 1989;86:4634-4638.

10 Kimura S, Tada N, Liu-Lam Y et al. Studies of the mouse Ly-6 alloantigen system. II. Complexities of the Ly-6 region. Immunogenetics 1984;20:47-56.

11 Spangrude GJ, Brooks DM. Mouse strain variability in the expression of the hematopoietic stem cell antigen Ly-6A/E by bone marrow cells. Blood 1993;82:3327-3332.

12 Kamiura S, Nolan CM, Meruelo D. Long-range physical map of the Ly-6 complex: mapping the Ly-6 multigene family by field-inversion and two-dimensional gel electrophoresis. Genomics 1992;12:89-105.

13 LeClair KP, Palfree RG, Flood PM et al. Isolation of a murine Ly-6 cDNA reveals a new multigene family. EMBO J 1986;5:3227-3234.

14 Khan KD, Lindwall G, Maher SE et al. Characterization of promoter elements of an interferon-inducible Ly- 6E/A differentiation antigen, which is expressed on activated T cells and hematopoietic stem cells. Mol Cell Biol 1990;10:5150-5159.

15 Khan KD, Shuai K, Lindwall G et al. Induction of the Ly-6A/E gene by interferon alpha/beta and gamma requires a DNA element to which a tyrosine-phosphorylated 91-kDa protein binds. Proc Natl Acad Sci USA 1993;90:6806-6810.

16 McGrew JT, Rock KL. Isolation, expression, and sequence of the TAP/Ly-6A.2 chromosomal gene. J Immunol 1991;146:3633-3638.

17 Sinclair AM, Dzierzak EA. Cloning of the complete Ly-6E.1 gene and identification of DNase I hypersensitive sites corresponding to expression in hematopoietic cells. Blood 1993;82:3052-3062.

18 Ma X, Ling K-W, Dzierzak E. Cloning of the Ly-6A (Sca-1) gene locus and identification of a 3′ distal fragment responsible for high-level gamma-interferon-induced expression in vitro. Br J Haematol 2001;114:724-730.

19 Sinclair A, Daly B, Dzierzak E. The Ly-6E.1 (Sca-1) gene requires a 3′ chromatin-dependent region for high-level gamma-interferon-induced hematopoietic cell expression. Blood 1996;87:2750-2761.

20 Miles C, Sanchez M-J, Sinclair A et al. Expression of the Ly-6E.1 (Sca-1) transgene in adult hematopoietic stem cells and the developing mouse embryo. Development 1997;124:537-547.

21 Ma X, de Bruijn M, Robin C et al. Expression of the Ly-6A (Sca-1) lacZ transgene in mouse haematopoietic stem cells and embryos. Br J Haematol 2002;116:401-408.

22 Nolan GP, Fiering S, Nicolas JF et al. Fluorescence-activated cell analysis and sorting of viable mammalian cells based on beta-D-galactosidase activity after transduction of *Escherichia coli* lacZ. Proc Natl Acad Sci USA 1988;85:2603-2607.

23 Fraser P, Hurst J, Collis P et al. DNaseI hypersensitive sites 1, 2 and 3 of the human beta-globin dominant control region direct position-independent expression. Nucleic Acids Res 1990;18:3503-3508.

24 Spanopoulou E, Giguere V, Grosveld F. Transcriptional unit of the murine *Thy-1* gene: different distribution of transcription initiation sites in brain. Mol Cell Biol 1988;8:3847-3856.

25 Dzierzak E, de Bruijn M. Isolation and analysis of hematopoietic stem cells from mouse embryos. In: Klug C, Jordan C, eds. Methods in Molecular Medicine: Hematopoietic Stem Cell Protocols. Totowa, New Jersey: The Humana Press Inc., 2002:1-14.

26 Muller AM, Medvinsky A, Strouboulis J et al. Development of hematopoietic stem cell activity in the mouse embryo. Immunity 1994;1:291-301.

27 Medvinsky A, Dzierzak E. Definitive hematopoiesis is autonomously initiated by the AGM region. Cell 1996;86:897-906.

28 Spangrude GJ, Aihara Y, Weissman IL et al. The stem cell antigens Sca-1 and Sca-2 subdivide thymic and peripheral T lymphocytes into unique subsets. J Immunol 1988;141:3697-3707.

29 Cray C, Keane RW, Malek TR et al. Regulation and selective expression of Ly-6A/E, a lymphocyte activation molecule, in the central nervous system. Brain Res Mol Brain Res 1990;8:9-15.

30 de Bruijn MF, Ma X, Robin C et al. Hematopoietic stem cells localize to the endothelial layer in the midgestation mouse aorta. Immunity 2002;16:673-683.

Chapter 5

Hematopoietic stem cells localize to the endothelial cell layer in the midgestation mouse aorta

Marella de Bruijn*, Xiaoqian Ma*, Catherine Robin*, Katrin Ottersbach, Maria-Jose Sanchez, and Elaine Dzierzak.

Immunity 2002; 16: 673-683

*** These authors contributed equally to this work**

Hematopoietic Stem Cells Localize to the Endothelial Cell Layer in the Midgestation Mouse Aorta

Marella F.T.R. de Bruijn,[1,4] Xiaoqian Ma,[1,4] Catherine Robin,[1,4] Katrin Ottersbach,[1] Maria-Jose Sanchez,[2] and Elaine Dzierzak[1,3]
[1]Department of Cell Biology and Genetics
Erasmus University
Rotterdam
Netherlands
[2]Department of Hematology
Cambridge University
CIMR Centre
Cambridge
United Kingdom

Summary

The emergence of the first adult hematopoietic stem cells (HSCs) during mammalian ontogeny has been under intense investigation. It is as yet unresolved whether these first HSCs are derived from intraembryonic hemangioblasts, hemogenic endothelial cells, or other progenitors. Thus, to examine the spatial generation of functional HSCs within the mouse embryo, we used the well-known HSC marker, Sca-1, and a transgenic approach with an *Ly-6A* (Sca-1) *GFP* marker gene. Our results show that this transgene marker is expressed in all functional HSCs in the midgestation aorta. Immunohistology of aorta-gonads-mesonephros (AGM) regions show that GFP$^+$ cells are specifically localized to the endothelial layer lining the wall of the dorsal aorta but not to the mesenchyme, strongly suggesting that HSC activity arises within a few cells within the endothelium of the major vasculature.

Introduction

During adult stages, the vertebrate hematopoietic system is constantly renewed from hematopoietic stem cells (HSCs) harbored in the bone marrow. These HSCs are characterized by functional repopulation properties elaborated after transplantation into adult recipients depleted for endogenous hematopoietic activity. The defining characteristics of HSCs are long-term, high level repopulation of all hematopoietic lineages and the ability to self-renew (Lemischka, 1991; Spangrude et al., 1991). During ontogeny, the first adult-type HSCs arise in the aorta-gonads-mesonephros (AGM) region, as shown by direct transplantation of AGM cells into adult recipients (Muller et al., 1994) or by transplantation of such cells after AGM explant culture (Medvinsky and Dzierzak, 1996). Further investigation of HSC emergence within the embryo has demonstrated their presence in the regions of major vasculature: the vitelline and umbilical arteries and the dorsal aorta (de Bruijn et al., 2000). Furthermore, within all mammalian and nonmammalian vertebrates analyzed, clusters of hematopoietic cells have been observed along the vitelline and umbilical arteries as well as the ventral wall of the dorsal aorta (Garcia-Porrero et al., 1995; Shalaby et al., 1997; Tavian et al., 1999; Wood et al., 1997), suggesting that hematopoietic cell emergence occurs in close association with the major vasculature of the embryo.

Indeed, many years ago, based on microscopic observations of yolk sac blood islands, it had been proposed that there exists a common mesodermal precursor cell, the hemangioblast, for hematopoietic and endothelial lineages (Murray, 1932). Gene targeting experiments have yielded some insight into the relationship between these lineages and have shown that the Flk-1 receptor tyrosine kinase, for example, is required for development of both the endothelial lineage and the hematopoietic lineage (Shalaby et al., 1997). However, unlike the coordinated emergence of hematopoietic and endothelial cells in the yolk sac, the dorsal aorta is formed before the emergence of adult repopulating HSCs, suggesting a slightly different type precursor or hemangioblast. Immunohistochemical analyses of the AGM region reveal overlap in the expression of hematopoietic and endothelial markers in the clusters of cells that appear to be emerging from the ventral wall of the dorsal aorta (Marshall and Thrasher, 2001). The most interesting expression pattern is exhibited by the Runx1 (previously known as Cbfa2 or AML1) transcription factor (North et al., 1999), which has been shown to be required for HSC activity in the AGM but does not appear to affect the major vasculature or early yolk sac hematopoiesis (Cai et al., 2000; Wang et al., 1996; Okuda et al., 1996; Mukouyama et al., 2000). At embryonic day 10 (E10), *Runx1* expression is found in the hematopoietic clusters and endothelial cells lining the walls of the vitelline and umbilical arteries and the ventral wall of the dorsal aorta (North et al., 1999). Additionally, expression is found in some mesenchymal cells underlying the ventral endothelial cells of the dorsal aorta. Taken together, these studies suggest that adult repopulating HSCs are derived from precursors within the hematopoietic clusters, endothelium, and/or underlying mesenchyme, and only the use of further markers will yield a precise identification of the HSC precursor. Recently, flow cytometric sorting of Runx1 positive cells together with combinations of antibodies recognizing other marker proteins on hematopoietic, endothelial, and mesenchymal cells has supported a mesenchymal and/or endothelial origin for HSCs (North et al., 2002 [this issue of *Immunity*]).

For many years, the Sca-1, GPI-linked cell surface glycoprotein has been used as a faithful marker of murine HSCs from various sources; i.e., adult bone marrow, fetal liver, and yolk sac (Spangrude et al., 1988; Huang and Auerbach, 1993). Sorting on the basis of the monoclonal antibody E13-161.7 specific for the Sca-1 epitope yields about a 100-fold enrichment of HSCs from adult bone marrow (Okada et al., 1992; Spangrude et al., 1988). Sca-1 is encoded by the strain specific allelic genes, *Ly-6E* and *Ly-6A* (Khan et al., 1990; Sinclair and Dzierzak, 1993; Stanford et al., 1992; van de Rijn et al.,

[3]Correspondence: dzierzak@ch1.fgg.eur.nl
[4]These authors contributed equally to this work.

1989), which are members of the large *Ly-6* gene family (Kamiura et al., 1992; LeClair et al., 1986). Sca-1 protein expression is complex within *Ly-6A* and *Ly-6E* strains of mice (Kimura et al., 1984; Spangrude and Brooks, 1993; van de Rijn et al., 1989). *Ly-6A* strains of mice express Sca-1 on virtually all (99%) of marrow-repopulating cells from the adult bone marrow, while *Ly-6E* strains express Sca-1 on only 25% of these cells (Spangrude and Brooks, 1993). Despite the allelic differences, the Sca-1 marker should be useful in identifying and localizing the first HSCs as they emerge in the AGM region. The relatively small size of the *Ly-6A/E* gene locus (14 kb) which contains all the transcriptional regulatory elements of these alleles (Khan et al., 1990, 1993; McGrew and Rock, 1991; Sinclair and Dzierzak, 1993; Ma et al., 2001; Sinclair et al., 1996) has been useful in directing *lacZ* marker gene expression in bone marrow HSCs in transgenic mice (Miles et al., 1997; Ma et al., 2002).

Thus, to identify the lineage of cells from which the first functional HSCs emerge in the AGM region, we performed studies using the Sca-1 surface glycoprotein marker and the enhanced green fluorescent protein (GFP) marker within the context of an *Ly-6A* transgene in midgestation mouse embryos. We report here that while only some AGM HSCs express the Sca-1 protein, all AGM HSCs are positive for expression of GFP from the *Ly-6A GFP* transgene. Immunohistological analyses reveal expression of the GFP marker within a single layer of cells lining the wall of the dorsal aorta and the vitelline/umbilical arteries. Together with functional data demonstrating the presence of HSC activity exclusively in the GFP positive fraction of the aorta region, our results strongly suggest that the first HSCs within the midgestation embryo are localized within the endothelial cell layer lining the wall of the dorsal aorta.

Results

The Sca-1 Antigen Is Expressed on Some AGM HSCs

To examine whether the Sca-1 antigen is expressed on AGM HSCs, we sorted Sca-1$^+$ cells from E11 AGM cell suspensions and performed transplantation experiments to test for functional hematopoietic repopulation of irradiated adult recipient mice by the donor cells. A representative fluorescence profile (Figure 1A) shows that 3.8% (range 1.95%–3.9%) of E11 AGM cells are Sca-1$^+$. At greater than 4 months posttransplantation, recipients were tested for the presence of donor cells by PCR of peripheral blood DNA (donor cell marker is a *lacZ* or human β-*globin* transgene). When transplanted at 1 embryo equivalent (ee) of sorted cells per recipient, adult repopulating HSCs were found in both the Sca-1$^+$ and Sca-1$^-$ fractions (Figure 1B) (donor engraftment ranging from 15% to 38%). This result may reflect a low or negative expression of Sca-1 by some HSCs as they emerge in the E11 AGM.

Previously, we had shown that all E11 AGM HSCs are c-kit$^+$ (Sanchez et al. 1996). To further subfractionate AGM HSCs, we sorted E11 AGM cells on the basis of c-kit and Sca-1 antigen expression. In these experiments, we confirm that all functional adult repopulating HSCs are c-kit$^+$ and that they are divided into a Sca-1$^+$ and a Sca-1$^-$ population (data not shown). Thus, the c-kit marker can provide a further enrichment for E11 AGM HSCs.

The Ly-6A GFP Transgene Marks All Functional HSCs from the E11 AGM Region

To study HSC generation and localization within the midgestation mouse embryo, we generated transgenic mice expressing the *GFP* reporter gene under the control of the *Ly-6A* gene regulatory sequences. The enhanced *GFP* gene was inserted into the first untranslated exon of the *Ly-6A* gene (Figure 1C) and was flanked by 4 kb of upstream and 10 kb of downstream genomic sequences containing regulatory elements shown to be necessary for high level specific expression in transgenic mice (Ma et al., 2002).

Since all adult repopulating bone marrow HSCs were found to express the *Ly-6A GFP* transgene, yielding up to a 100-fold enrichment of HSCs (data not shown), we explored whether transgene expression could be used to enrich for HSCs from the AGM region. Sorting experiments were performed on *Ly-6A GFP* E11 AGM cells. Approximately 1.7% (range 1.5%–2.0%) of AGM cells are GFP$^+$ (Figure 1D). Sorted GFP$^+$ and GFP$^-$ cells were transplanted at cell doses of 2, 1, and 0.3 ee (average of 3 × 10^3 GFP$^+$ cells/ee). At greater than 4 months posttransplantation, >10% donor cell engraftment (ranging from 15% to 70%) was found in recipients receiving either 1 or 2 ee of GFP$^+$ cells (Figure 1E). No donor-derived repopulation was found with 0.3 ee of GFP$^+$ cells or any dose of GFP$^-$ cells. Thus, unlike the Sca-1 marker, *Ly-6A GFP* transgene expression marks all HSCs in the AGM region.

All E11 AGM HSCs Are Localized to the GFP$^+$ Population in the Aorta Region

While previous E11 AGM subdissection studies demonstrated that HSCs are present in the region of the aorta with its surrounding mesenchyme (de Bruijn et al., 2000), we tested whether such cells could be further enriched based on GFP expression. We performed flow cytometric sorting and transplantation of GFP$^+$ cells from the aorta-mesenchyme subregion of E11 *Ly-6A GFP* embryos. The aorta-mesenchyme contained on average 1.9% (1.4% to 3.0% GFP$^+$ cells; see Figure 2A, left panel). Long-term adult repopulating HSC activity (Figure 2A, right panel) was found in the GFP$^+$ fraction of the aorta-mesenchyme when 1 or 2 ee of GFP$^+$ cells was transplanted (average of 1.6 × 10^3 GFP$^+$ cells/aorta-mesenchyme). No donor cell engraftment was found when the GFP$^-$ fraction was transplanted. Engraftment of recipients, as measured by peripheral blood DNA PCR for the donor cell *GFP* marker at greater than 4 months posttransplantation, ranged from 11% to 100% (data not shown).

To determine whether donor cell hematopoietic engraftment was multilineage, flow cytometric analysis was performed on peripheral blood cells. Representative data of one recipient at 2 months posttransplantation is shown in Figure 2C (bottom panels). GFP expression was found in B220$^+$, CD4$^+$/8$^+$, and Mac-1$^+$ fractions, thus revealing the donor derivation of 35% of the monocytes, 18% of the B cells, and 5.3% of the T

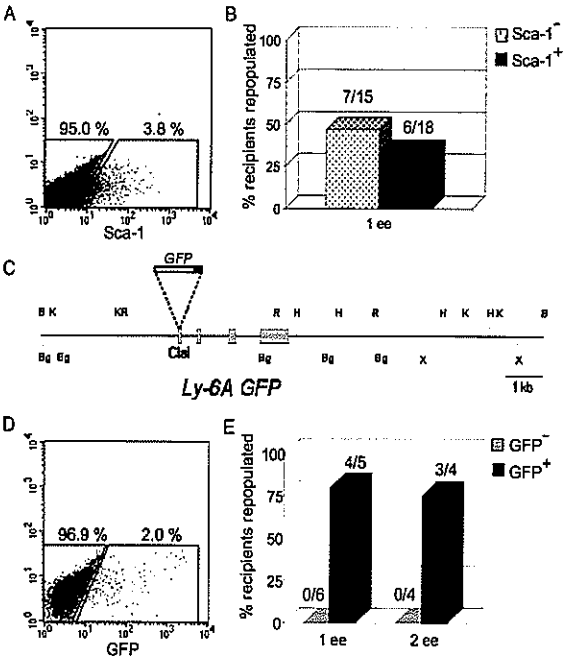

Figure 1. Characterization of E11 AGM Adult Repopulating HSCs for Expression of the Sca-1 Antigen or *Ly-6A GFP* Transgene

(A) E11 AGM region cells were stained with the Sca-1 PE antibody and sorted into Sca-1$^+$ and Sca-1$^-$ fractions. Percentages in each fraction are indicated.

(B) One embryo equivalent of each fraction of Sca-1$^+$ (1.0×10^4) and Sca-1$^-$ (25.3×10^4) cells was injected into adult recipient mice, and after 4 months peripheral blood DNA was examined for the presence of donor cells. The number of repopulated animals/total number of animals transplanted is indicated above each column (percentage of recipients repopulated is indicated on *y* axis). Only those animals with >10% donor cells are considered positive. Engraftment of recipients by donor-derived cells as measured by peripheral blood DNA PCR ranged from 15%–38% for Sca-1$^+$ sorted cells and 16%–36% for Sca-1$^-$ sorted cells. Results are from four independent experiments.

(C) *Ly-6A GFP* transgene construct. The 14 kb BamHI fragment of the *Ly-6A* gene and the insertion of the enhanced *GFP* gene into the ClaI cloning site is shown. Restriction sites are indicated. B, BamH1; Bg, BglII; K, Kpn1; R, EcoR1; H, HindIII; X, XbaI.

(D) Flow cytometric profile for E11 *Ly-6A GFP* transgenic AGM cells, with percentages of cells within the sorting gates indicated.

(E) Long-term repopulation of recipients (>4 months posttransplantation) receiving 1 or 2 ee of either GFP$^+$ or GFP$^-$ sorted E11 AGM cells. One ee of sorted E11 GFP$^-$ AGM cells is on average 3×10^3 cells, and one ee of GFP$^-$ cells is 17.0×10^4. The number of repopulated animals/total number of animals transplanted is indicated above each column (*y* axis shows percentage of recipients repopulated). Only those recipients with greater than 10% donor positive cells are considered positive. Engraftment ranged from 15%–70% as measured by peripheral blood DNA PCR for the donor cell GFP marker. Results are from four independent experiments.

cells as compared to the control transgenic (top panels). Furthermore, at greater than 4 months posttransplantation, donor marker DNA PCR analysis on two other recipients transplanted with GFP$^+$ aorta-mesenchyme cells revealed repopulation in all hematopoietic tissues and lineages tested (Figure 2D). Engraftment ranged from 21% to 100%. Thus, all long-term, multilineage adult repopulating HSCs from the E11 aorta-mesenchyme can be isolated and enriched within the GFP$^+$ fraction.

The Temporal and Spatial Expression Pattern of Ly-6A GFP Is Correlated with Sites of HSC Generation

The general expression pattern of the *Ly-6A GFP* transgene could be visualized with fluorescence and confocal microscopy. When marker transgene expression was examined in the late E11 AGM region, we found highly fluorescent GFP$^+$ cells in the mesonephric tubules and the Wolffian/Mullerian ducts (Figures 3A and 3B), as we had found previously for the *Ly-6A lacZ* transgene (Ma et al., 2002). However, in contrast to *Ly-6A lacZ* transgenic embryos, we also found punctate GFP expression in the cells along the wall of the dorsal aorta (Figures 3A and 3C). Confocal microscopy revealed the GFP-expressing cells to be dispersed along the circumference and length of the E11 dorsal aorta (Figures 3A and 3C). Compared to the mesonephric tubules in the urogenital ridges (UGR) (Figure 3D), GFP fluorescence of dorsal aorta cells appeared to be of lower intensity (Figure 3C). However, as revealed in transverse sections through the E11 (43 somite pairs [sp]) dorsal aorta (Figure 3F), only a single layer of cells interspersed along the circumference of the aorta was found to be GFP$^+$ (Figure 3E). No GFP expression was found in the mesenchymal tissue directly surrounding the dorsal aorta. GFP expression in the dorsal aorta was found beginning at E9 (24 sp stage) and is confined to only a very few cells in the ventral wall (data not shown). Expression in the aorta becomes more widespread, first along the ventral and then along the dorsal wall at E10 and E11, respectively, and continues until at least E14. In contrast to the dorsal aorta, no GFP expression was observed in the whole yolk sac at E9 or E10 (34 sp stage; Figures 3G and 3H). However, at late E11 (48 sp stage; Figures 3I and 3J) GFP$^+$ cells could be found in the yolk sac but only on very few of the large vessels. Expression in the yolk sac vessels continues at least until E14. Most interestingly, high level GFP expression was found in both the vitelline (Figures 3K and 3L) and umbilical arteries (data not shown) at E10 (38 sp stage). GFP expression in these vessels begins at E9 (24 sp stage) and continues at least until E14 (Figures 3M and 3N). Of the vitelline vessels, high level GFP expression appears to be confined to the artery.

Other sites in the midgestational *Ly-6A GFP* embryo

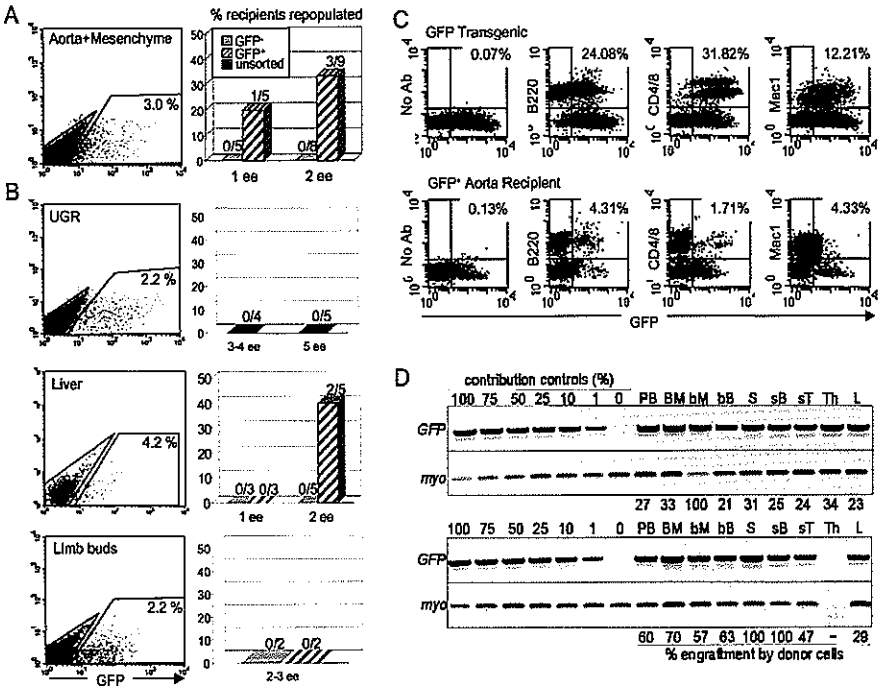

Figure 2. Long-Term Adult Repopulation by GFP-Sorted Cells from E11 Tissues

(A and B) In the left panels flow cytometric profiles and sorting gates are shown for (A) E11 aorta-mesenchyme and (B) UGR, liver, and limb bud cells. The percentage of GFP$^+$ cells for each tissue is indicated. In the right panels the results of transplantation of these sorted cells is shown. The *x* axis of the bar charts indicates the number of embryo equivalents of each sorted cell population that was transplanted. One ee of GFP$^+$ cells from the aorta-mesenchyme is on average 1.6×10^3 (range 1.3 to 2.1×10^3), and one ee of GFP$^-$ cells is 9×10^4 (range $8–10 \times 10^4$). The bars represent the percentage of recipients found positive for donor repopulation at >4 months posttransplantation. The number of recipients repopulated/total number of recipients is indicated for each bar. Donor cell engraftment, as measured by peripheral blood DNA PCR ranged from 11% to 100%.

(C) Flow cytometric analysis of peripheral blood from a normal adult *Ly-6A GFP* transgenic control and a GFP$^+$ aorta cell transplant recipient. GFP fluorescence is on the *x* axis, and phycoerythrin fluorescence is on the *y* axis. Directly conjugated phycoerythrin B220-, CD4-, CD8-, and Mac-1-specific antibodies were used for blood cell staining. Percentage of double-positive cells is indicated in the upper right quadrants.

(D) Hematopoietic multilineage analysis of two recipient mice repopulated with GFP$^+$ aorta-mesenchyme cells. At >4 months posttransplantation, DNA was isolated from peripheral blood (PB), bone marrow (BM), bone marrow myeloid cells (bM), bone marrow B cells (bB), spleen (S), spleen B cells (sB), spleen T cells (sT), thymus (Th), and lymph nodes (L). PCR analysis was performed using oligonucleotide primers for *GFP* (top band) and myogenin (*myo*) (lower band) genes. One hundred, seventy-five, fifty, twenty-five, ten, one, and zero are the contribution controls indicating percentage donor cell engraftment. Percentage engraftment by donor cells (indicated below each gel) was calculated by ImageQuant. Transplantation data are from two independent experiments.

were found to express GFP and include the liver, hindgut, and limb buds (data not shown). These GFP$^+$ tissues were tested for HSC activity. Flow cytometry (Figure 2B, left panels) revealed that these tissues contained between 2% and 4.2% GFP$^+$ cells. As shown in Figure 2B (right panels), no HSCs (GFP$^+$ or GFP$^-$) were found in the E11 UGRs or the limb buds. However, HSCs of the E11 liver were exclusively in the GFP$^+$ fraction. Thus, all HSCs are GFP$^+$ but not all GFP$^+$ cells are HSCs.

Ly-6A GFP Expression Is Localized to the Endothelial Cell Layer of the Midgestation Dorsal Aorta

Immunohistochemical staining was performed to establish what cell lineage(s) the GFP-expressing cells in the dorsal aorta may represent. Transverse sections through the trunkal regions of E11 *Ly-6A GFP* embryos were stained using antibodies to CD31 (PECAM-1), a marker expressed on all endothelial cells and on intraaortic hematopoietic clusters (Cai et al., 2000; Drake and Fleming, 2000; Garcia-Porrero et al., 1998), and Sca-1. Low magnification of a transverse section through the trunkal region of an E11 *Ly-6A GFP* embryo stained with CD31 shows a single layer of CD31$^+$ (red; Cy5) endothelial cells lining the entire circumference of the dorsal aorta (Figures 4A and 4B). GFP expression appears to be colocalized within a few of these endothelial cells. High magnification of the ventral area within the box shows single green and red fluorescence signals

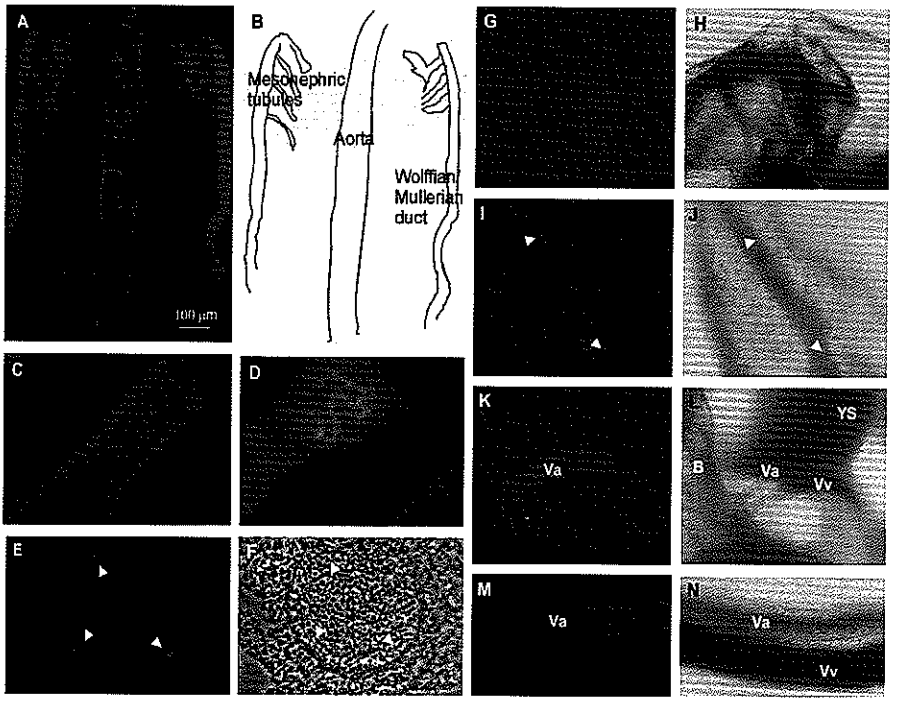

Figure 3. Fluorescent Microscopic Analysis of Whole AGM and Tissue Sections from *Ly-6A GFP* Embryos

(A) Confocal microscopic analysis of whole late E11 AGM tissue at 10× magnification. Twenty-four scanned images of this AGM region were obtained at 10 µm intervals. The scanned image in this figure is the 12th optical slice of 24 with a thickness of approximately 1 µm. All 24 images of the Z stack of this AGM were used to build the three-dimensional projection shown at www.eur.nl/fgg/ch1 (see Cell Biology).

(B) Schematic drawing of the GFP expression region of the E11 AGM shown in (A). GFP positive areas include the aorta, mesonephric tubules, and Wolffian/Mullerian ducts.

(C and D) GFP fluorescent image of a portion of the E11 aorta (C) and mesonephric tubules (D) at 20× magnification, 1 µm thickness.

(E and F) Fluorescent (E) and bright field (F) images of a transverse section of the AGM region of a 43 somite pair (sp) transgenic embryo (E11) showing the dorsal aorta. Dorsal side is up and ventral is down. A single-cell layer of interspersed GFP$^+$ cells is found only around the circumference of the dorsal aorta and is indicated by the arrowheads. Magnification 10×.

(G and H) Fluorescent (G) and bright field (H) images of a yolk sac from a 34 sp transgenic embryo (E10). No fluorescent signal is detected in the yolk sac, particularly the vessels. Magnification 4×.

(I and J) Fluorescent (I) and bright field (J) images of a yolk sac from a 48 sp transgenic embryo (E11). GFP signal is found in the cells lining the large vessel. Arrowheads indicate the vessel walls.

(K and L) Fluorescent (K) and bright field (L) images of the vitelline vessels as they emerge from the body wall (B) of a 38 sp transgenic embryo (E10). Strong GFP signal is observed in the vitelline artery (Va) with some low signal in the underlying vein (Vv). The fluorescent negative yolk sac (YS) is shown in the background. Magnification 4×.

(M and N) Fluorescent (M) and bright field (N) images of the vitelline vessels from an E14 transgenic embryo. The endothelial cells of the vitelline artery (Va) are strongly GFP positive. Magnification 4×.

(Figures 4E and 4F). An overlay of both signals (Figure 4G) reveals GFP expression (green) within the cytoplasm and the red signal of the CD31 antibody on the surface of a pair of double-positive cells (see arrow, Figure 4G). The two adjacent cells are located within the single layer of cells lining the aorta. A third double-positive cell is seen further to the right. Unfortunately, while we examined many of these transverse sections to evaluate the GFP expression status, we were unable to find any hematopoietic clusters. Thus, from the results presented here, GFP expression is specifically localized to a few of the cells within the single-cell layer of endothelial cells lining the dorsal aorta, and these cells express the CD31 endothelial/hematopoietic cluster marker.

Finally, the expression of GFP and endogenous Sca-1 was examined by fluorescent microscopy on transverse sections through the trunkal region of E11 *Ly-6A GFP* embryos. As shown in Figure 4C (GFP expression) and D (GFP and Sca-1 expression), Sca-1 expression clearly overlaps with the GFP signal in the mesonephric tubules. However, no (or undetectable) Sca-1 expression was found in the GFP$^+$ cells lining the walls of the dorsal

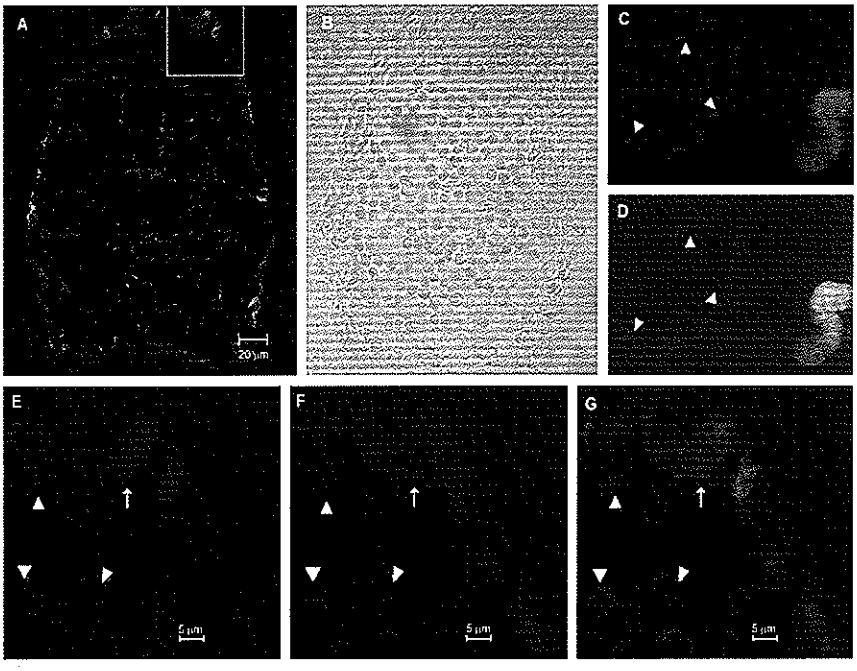

Figure 4. Fluorescence and Immunostaining Confocal Microscopic Analysis of AGM Tissue Sections from *Ly-6A GFP* Embryos

(A and B) Fluorescence microscopic image (A) and bright field image (B) of a transverse section through the dorsal aorta of an E11 (43 somite pair) embryo, showing immunofluorescent staining with an antibody against CD31 (red) and GFP transgene signal (green). Part of the ventral wall of the aorta (marked by a box) is shown at a higher magnification in (E) GFP signal, (F) CD31 Cy5 signal, and (G) an overlay of GFP and CD31 signals. The arrow indicates two adjacent cells that coexpress the *Ly-6A GFP* transgene and CD31. Arrowheads indicate some of the autofluorescent erythrocytes.

(C) Transverse section through an E11 AGM region showing a single layer of GFP$^+$ cells (green) around the dorsal aorta (and strongly fluorescent mesonephric tubules). Ventral is to the upper right.

(D) Sca-1 (red) and GFP immunofluorescence on section as shown in (C). Overlap in expression of Sca-1 and GFP occurs in the mesonephric tubules, but no (or undetectable) overlap is found in aorta cells. Size bars are included for (A), (E), (F), and (G), and (C) and (D) are at a 40× magnification.

(E–G) For (E), (F), and (G) (CD31 immunostaining), optical slice eight from 18 optical scans of approximately 1 μm thickness at 0.5 μm intervals is shown. All 18 optical slices, respectively, were used to build the three-dimensional projection shown at www.eur.nl/fgg/ch1 (see Cell Biology).

aorta, suggesting that either expression or the sensitivity of detection is more limited for the endogenous Sca-1 antigen than for GFP fluorescence.

GFP$^+$ Aorta Cells Express Additional Endothelial and Hematopoietic Markers

Since we have shown that hematopoietic and/or endothelial markers such as c-kit, CD31, CD34, and VE-cadherin are expressed on functional adult repopulating AGM HSCs from normal embryos (Sanchez et al., 1996; North et al., 2002), we performed flow cytometric analysis with antibodies specific for these markers and for Sca-1 on cells from E11 *Ly-6A GFP* aorta-mesenchyme. Flow cytometric profiles for UGRs are provided for comparison. As shown in Figure 5, a large percentage of GFP$^+$ cells in the UGRs are Sca-1$^+$ (68%) as expected from the immunostaining results showing high coexpression of Sca-1 and GFP in the mesonephric tubules. However, in the aorta-mesenchyme fewer GFP$^+$ cells coexpress Sca-1$^+$ (19%). This is in accordance with the immunohistochemical and transplantation results and suggesting that cell surface Sca-1 antigen expression is limited on cells of the aorta and particularly on HSCs.

We next examined the expression of c-kit, CD34, CD31, and VE-cadherin on cells from these subregions. Interestingly, all of these markers have been shown to be expressed in hematopoietic clusters as well as the endothelial cells lining the dorsal aorta in midgestation embryos (Bernex et al., 1996; Wood et al., 1997; Drake and Fleming, 2000; Garcia-Porrero et al., 1998; Cai et al., 2000; North et al., 2002). In the aorta-mesenchyme, c-kit$^+$ cells make up 37%–51% of the GFP$^+$ population. As all AGM HSCs are known to express c-kit, these results strongly suggest that HSCs are in this double-positive population. As expected from the endothelial localization of GFP$^+$ cells by microscopy, many GFP$^+$

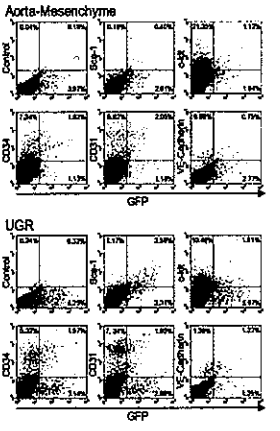

Figure 5. Representative FACS Profiles for E11 Aorta-Mesenchyme and UGR Subregions

Cells from these two AGM subregions were obtained from *Ly-6A GFP* transgenic embryos and were stained with antibodies specific for Sca-1, c-kit, CD34, CD31, and VE-cadherin antigens. Three to five times 10^4 cells were analyzed for GFP (*x* axis) and surface antigen (*y* axis) expression on a FACS Vantage SE. Dead cells were gated out from the analysis. Percentages of cells in each quadrant are indicated and are representative of the profiles obtained from two independent analyses.

cells in the aorta-mesenchyme are also CD34$^+$, CD31$^+$, and VE-Cadherin$^+$ (63%, 64%, and 21% of all GFP$^+$ cells, respectively). In combination with our data that all normal AGM HSCs express c-kit, CD34, CD31, and VE-Cadherin (Sanchez et al., 1996; North et al., 2002), these data indicate that the first HSCs are among the GFP$^+$ cells in the aorta that coexpress these endothelial/hematopoietic markers.

Early E10 AGM Explants Contain GFP$^+$ Multipotential/Lymphoid Progenitors

Previously, we have been unable to obtain high level, multilineage hematopoietic repopulation from directly transplanted AGM regions (Muller et al., 1994) or AGM explants (Medvinsky and Dzierzak, 1996) before the 35 somite pair stage. The appearance of a few GFP$^+$ cells in the AGM region of *Ly-6A GFP* transgenic embryos beginning at the 24 somite pair stage (late E9), together with the finding that not all E11 AGM GFP$^+$ cells are HSCs, prompted us to test these early appearing GFP$^+$ cells for other hematopoietic progenitor activity. We cultured AGM explants from early E10 *Ly-6A GFP* embryos (31 to 34 somite pairs) for 3 days, sorted the GFP$^+$ and GFP$^-$ cells, and transplanted these cells into irradiated adult recipients (Figure 6A). At 4 months posttransplantation, only those animals (two out of two) transplanted with GFP$^+$ cells were positive for donor cell engraftment. Peripheral blood DNA PCR revealed levels of engraftment of 1%–2% in these animals. Multilineage analysis of hematopoietic tissues was performed at 6 months posttransplantation for one animal. Engraftment of donor-derived cells was found mainly in the peripheral

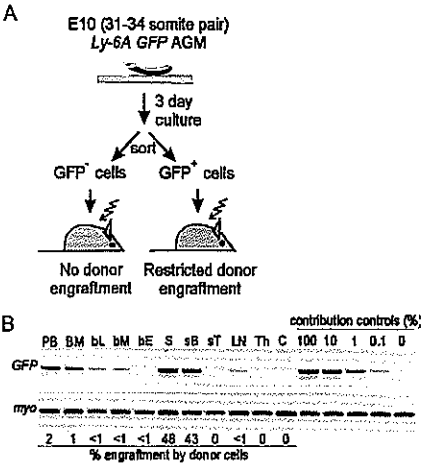

Figure 6. Hematopoietic Engraftment by Sorted GFP$^+$ Cells after Explant Culture of Early E10 *Ly-6A GFP* AGMs

(A) Experimental protocol. Twelve transgenic AGMs of the 31 to 34 somite pair stage were cultured as explants for 3 days and harvested, and the cells were sorted on the basis of GFP transgene expression. Irradiated adult recipient mice were injected with either 3 ee or 1.5 ee of GFP$^+$ (2.6×10^4 cells/ee) or GFP$^-$ (3.5×10^5 cells/ee).

(B) DNA PCR analysis, for the donor cell *GFP* transgene marker and the myogenin (*myo*) gene normalization control, showing low level, restricted hematopoietic engraftment of a recipient mouse at 6 months after injection of 3 ee of GFP$^+$ cells. Cells were obtained from each of the hematopoietic tissues indicated and from lineage-specific sorts performed using antibodies directed against B cells, T cells, monocytes/granulocytes, and erythroid cells. PB, peripheral blood; BM, bone marrow; bL, bone marrow lymphoid; bM, bone marrow myeloid; bE, bone marrow erythroid; S, spleen; sB, spleen B cells; sT, spleen T cells; LN, lymph nodes; Th, thymus; and C, control transplanted with nontransgenic cells. Contribution controls show signal expected from varying percentages of *GFP* transgene containing DNA. Percentage engraftment is indicated below each lane. Quantitation was performed using ImageQuant.

blood, bone marrow, and spleen (Figure 6B). After sorting cells from the bone marrow and spleen with antibodies specific for myeloid, erythroid, and B and T lymphoid lineages, a predominance of donor-derived cells in the B lymphoid lineage was found. Only very low levels (<1%) of myeloid and erythroid lineage cells were found, suggesting that this animal was long-term repopulated with a multipotent/lymphoid progenitor. Thus, the early expression of GFP in the AGM region (before E10.5) may represent cells that are the predecessors of high level, multilineage repopulating HSCs or unrelated progenitors.

Discussion

Ly-6A GFP Marks Definitive HSC Emergence in the Dorsal Aorta and Vitelline and Umbilical Arteries

Unlike the Sca-1 surface glycoprotein, flow cytometric sorting based on GFP expression reliably enriches for all AGM HSCs. This difference is due most likely to the limiting nature of Sca-1 protein on the surface of these

HSCs compared to a more intense fluorescence signal produced by GFP. Indeed, the transgenic mice contain eight copies of the transgene as compared to the normal diploid copy of the endogenous *Ly-6A/E* gene encoding Sca-1. Furthermore, GFP is a cytoplasmic protein not requiring the processing steps necessary to get the Sca-1 protein properly displayed on the cell surface (i.e., glycosylation, transport through the plasma membrane, and GPI linkage). Thus, the *Ly-6A GFP* transgene appears to be an optimally expressed reporter in AGM HSCs and is an excellent marker for HSC isolation from the embryo.

Flow cytometric analyses show that there are approximately 1600 GFP$^+$ cells within the E11 aorta-mesenchyme. While in transplantation studies all E11 AGM aorta-mesenchyme HSCs are GFP$^+$, it is clear that not all 1600 GFP$^+$ aorta-mesenchyme cells are functional HSCs. Outside the aorta subregion, we also found a significant percentage of GFP$^+$ cells in the limb buds, tail, and UGRs. In the E11 UGRs, we show here and in a previous study (de Bruijn et al., 2000) that there are no HSCs even at high cell numbers transplanted. However, considering that HSCs can be detected in E12 UGRs or explant cultured E11 UGRs (de Bruijn et al., 2000), further studies must be performed to determine whether these HSCs are GFP$^+$. Recent reports on the multipotentiality of stem cells for lineages outside the expected lineage (reviewed in Wulf et al., 2001 and Morrison, 2001) led us to analyze the hematopoietic potential of the other GFP$^+$ sites. If indeed the *Ly-6A GFP* transgene marks more totipotent stem cells, these cells in the developing limb and tail might be able to contribute to adult hematopoietic repopulation. In support of this, Godin and colleagues (1999) found some hematopoietic colony-forming activity in limb bud tissues. When we sorted and transplanted high ee of GFP$^+$ limb bud cells, we found no hematopoietic repopulation. Similarly, we have transplanted high ee of cells from the tail and found no detectable adult hematopoietic repopulation (our unpublished data). However, given the low frequency of stem cells and the limited number of mice transplanted, additional studies are warranted.

Enrichment of Aorta HSCs as They Emerge during Midgestation

At E11 there is at least one adult repopulating HSC in 1600 GFP$^+$ aorta cells. Our FACS analysis data and transplantation results with sorted AGM cells using conventional antibody-mediated enrichment suggest that further enrichment can be achieved. For example, of the 1600 GFP$^+$ aorta cells, a smaller number are also positive for c-kit, CD31, CD34, or VE-cadherin, markers previously used for enrichment of AGM HSCs (Sanchez et al., 1996; North et al., 2002). Consistent with our transplantation data for the aorta-mesenchyme in which not all HSCs are Sca-1$^+$, more GFP$^+$ cells express c-kit than Sca-1. Since it has already been shown that all AGM HSCs are c-kit$^+$ and we have shown here that all aorta HSCs are GFP$^+$, a 50% enrichment can be achieved, and only 800 c-kit$^+$GFP$^+$ cells should be required for complete multilineage adult recipient repopulation. Three or four color sorting using c-kit and GFP expression together with CD34 and/or CD31 may allow an even greater enrichment of aorta HSCs. In addition, close observation of the GFP fluorescence profile of E11 aorta-mesenchyme cells reveals two populations of cells: a GFP^{high}-expressing population representing on average 0.3% of total cells and a GFP^{low}-expressing population representing 2.1% of total cells. As a means of achieving greater enrichment of aorta HSCs, we have begun to examine which population contains adult repopulating HSCs. Our preliminary results show adult repopulating HSCs within the GFP^{high} population. In the best instance this may represent a further 10-fold enrichment that, when put together with c-kit sorting, may yield one HSC in 80 c-kit$^+$GFP^{high} cells.

Specific Endothelial Localization of HSCs in the Midgestation Embryo

The temporal and spatial expression pattern of the *Ly-6A GFP* transgene within the expected hematopoietic sites of the midgestation mouse embryo is highly suggestive of expression in HSCs. Indeed, in our functional transplantation and fluorescence microscopy studies we have shown that all aorta-mesenchyme HSCs are GFP$^+$ and that they are localized to the single-cell layer lining the dorsal aorta. Furthermore, GFP expression is found in the vitelline/umbilical arteries, which together with the aorta, are the other sites within the embryo in which HSCs first appear at mid-E10. At later time points, GFP expression is found in the E11 liver (Figure 2) and E11 yolk sac (Figure 3), again coincident with the appearance of HSCs as determined in previous studies (Muller et al., 1994; Medvinsky and Dzierzak, 1996). We have shown here that all HSCs in the E11 liver are GFP$^+$, and in other related studies (Ma et al., unpublished data) we have found that all bone marrow HSCs in adult transgenic mice are also GFP$^+$. Thus, in the various spatially and temporally distinct sites tested, the *Ly-6A GFP* transgene faithfully marks all HSCs.

The identification of the direct cellular precursors of HSCs is of great importance in understanding how the adult hematopoietic system is generated. Transverse sections through the E11 dorsal aorta reveal GFP$^+$ cells only in the single layer of CD31$^+$ endothelial cells. No GFP$^+$ cells were found within the mesenchymal tissue immediately surrounding and flanking the dorsal aorta. Indeed, functional adult HSCs at this stage are CD31$^+$ (North et al., 2002). Thus, our data strongly suggest that HSCs are generated within the single layer of endothelial cells lining the wall of the dorsal aorta and possess both endothelial and hematopoietic cell surface characteristics.

At early E10, GFP expression in the aorta is limited to the ventral endothelium, the area in which many investigators have demonstrated emerging hematopoietic clusters. On the basis of our data we cannot make the distinction between the aortic endothelium or the hematopoietic clusters as the site of HSC emergence since we have not been able to find any clusters in sections from numerous GFP$^+$ embryos analyzed. Despite our inability to find clusters, if such cells were GFP$^+$ we would postulate that either these cells, the GFP$^+$ cells within the endothelial layer, or both cell populations are the functional adult repopulating HSCs. Indeed, our recent findings using combinations of antibodies discern-

ing hematopoietic, endothelial, and mesenchymal cells in the wild-type E11 AGM region demonstrate that the HSCs exhibit not only a hematopoietic but also an endothelial surface phenotypic profile (North et al., 2002). Taken together, these results strongly suggest that adult repopulating HSC activity normally arises within the endothelial and/or hematopoietic cluster cells and not in the mesenchymal cell layer of the dorsal aorta. It is of interest to note that North and colleagues revealed a population of adult-type HSCs in *Runx1* haploinsufficient embryos with a mesenchymal surface phenotype, which is absent from normal embryos, indicating that a full dose of *Runx1* is required for the normal development of the adult repopulating HSCs in the aorta (North et al., 2002).

Finally, we have shown by explant culture of early E10 AGM regions and subsequent GFP sorting and transplantation that an earlier restricted hematopoietic progenitor is marked by the *Ly-6A GFP* transgene. Similar low level, adult repopulating cells in the AGM region have been recently reported by Cumano and colleagues (2001), and others have reported neonatal repopulating hematopoietic cells in the AGM (Yoder et al., 1997) at time points earlier than E10.5 onset of HSC function. Here we have observed the appearance of GFP expression in the AGM region and the vitelline/umbilical arteries at times earlier than E10.5. Hence, these *Ly-6A GFP*-expressing cells may represent the direct precursors to the high level adult repopulating cells or alternatively represent the differentiation products of the same or similar endothelial cohort that gives rise to the HSCs. Interestingly, an abundance of GFP^{low}-expressing cells were found in the explant cultures of early E10 AGM regions, and future sorting and transplantion experiments will examine whether these restricted low level repopulating cells are correlated with GFP expression levels.

In summary, we have shown here by transplantation and immunofluorescence analyses of midgestational *Ly-6A GFP* transgenic mouse AGM regions that all HSCs in the aorta are GFP^+ and that GFP-expressing cells are localized to the endothelial layer of the dorsal aorta but not the underlying mesenchyme. This localization can now lead to an identification of the immediate precursors to HSCs and a further understanding of the molecular signals necessary for the emergence of these somatic stem cells.

Experimental Procedures

Transgenic Mice and Embryo Generation

Timed matings were set up between *Ly-6A-GFP*, BL1b (Miles et al., 1997), or line 72 (Strouboulis et al., 1992) transgenic male mice and wild-type (C57BL/10 × CBA)F, females. The day of vaginal plug detection is day 0. For embryo generation, the transgene was always inherited through the male parent to avoid contribution by genetically marked maternal cells in transplantation assays. Transgenic embryos were typed on the basis of their GFP expression using an Olympus IX70 fluorescent microscope. Animals were housed according to institutional guidelines, and procedures were carried out in compliance with the Standards for Humane Care and Use of Laboratory Animals.

Cell Preparations

Pregnant mice were killed by cervical dislocation. Embryos were isolated from the uterus, and AGMs were dissected and subsequently subdissected into the aorta with its surrounding mesenchyme and the UGRs using 27 gauge needles (Becton Dickinson, San Jose, CA). Tissues were either incubated directly with collagenase (type I, Sigma; final concentration 0.12% v/v) for 1 hour at 37°C in phosphate-buffered saline (PBS) supplemented with 10% fetal calf serum (FCS) and penicillin/streptomycin or cultured as explants for 3 days prior to collagenase digestion (Medvinsky and Dzierzak, 1996). Cells were dispersed, washed, resuspended in PBS+FCS, sorted, and transplanted intravenously into irradiated mice. Viable cell numbers were determined with a Bürker hemocytometer using trypan blue. Cells were kept on ice at all times subsequent to collagenase digestion until the time of injection.

Analysis of Long-Term Multilineage Repopulating Activity

Embryonic cell suspensions were assayed for the presence of definitive HSCs by intravenous transfer into irradiated adult recipients (Dzierzak and de Bruijn, 2002; Muller et al., 1994; Miles et al., 1997; Medvinsky and Dzierzak, 1996). In brief, (C57BL/10 x CBA)F, male recipients were exposed to a split dose of 900 rad of γ irradiation from a ^{137}Cs source. Adult (C57BL/10 x CBA)F, spleen cells (2×10^5/mouse) were coinjected with the embryonic cells to promote survival. Recipient mice were bled at 1 and >4 months after transfer and analyzed for donor contribution by donor marker-specific PCR on peripheral blood DNA (Dzierzak and de Bruijn, 2002; Miles et al., 1997; Muller et al., 1994). Reconstitution was evaluated by ethidium bromide staining of agarose gels and in some cases by Southern blot hybridization as described (Dzierzak and de Bruijn, 2002; Medvinsky and Dzierzak, 1996; Muller et al., 1994). To test for long-term multilineage hematopoietic repopulation, genomic DNA was isolated from peripheral blood, thymus, lymph node, splenic B and T cells, and bone marrow myeloid, erythroid, and lymphoid cells. Percentage donor-cell contribution was analyzed by ImageQuant using the *myo* gene as the DNA normalization control.

Fluorescent Antibody Surface Staining and Flow Cytometry

All antibodies used in flow cytometric sorting and analysis were obtained from Pharmingen. The monoclonal antibodies used were directly conjugated with either phycoerythrin (PE) or biotin, or unconjugated, and included PE-anti-c-kit (clone 2B8), PE-anti-Sca-1 (Ly-6A/E; clone E13-161.7), PE-anti-CD4 (clone GK1.5), PE-anti-CD8a (clone 53-6.7), PE-anti-B220 (CD45R; clone RA3-6B2), PE-anti-Mac-1 (CD11b; clone M1/70), biotinylated CD31 (PECAM-1; clone MEC 13.3), CD34 (RAM34), and unconjugated purified VE-Cadherin (CD144; clone 11D4.1). Single-cell suspensions were prepared as described above. After incubation with specific antibodies for 30 min on ice, cells were washed twice and incubated with PE-conjugated streptavidin or PE-conjugated goat anti-rat IgG (mouse-absorbed) (Caltag Laboratories, Burlingame, CA) when required. Cells were washed twice and filtered through a 45 μm nylon mesh screen prior to sorting. Hoechst 33258 (1 μg/ml) or 2 μg/ml 7AAD was added to identify dead cells. To determine the background levels, cells were stained with appropriate immunoglobulin isotype controls (Pharmingen). During the entire staining procedure, PBS containing 10% FCS and penicillin/streptomycin was used. Cells were sorted/analyzed using a FACS Vantage SE (Becton-Dickinson). The purity of the sorted cells ranged from 80%–98%. For phenotypic analysis of stained whole blood cell samples, suspensions were treated for 10 min with FACS lysis solution (Becton-Dickinson) to eliminate erythrocytes and washed before analysis.

Immunohistochemical Staining and Confocal Microscopy

Ly-6A GFP transgenic embryos were fixed in 2% paraformaldehyde/PBS for 30 min at room temperature, equilibrated in 20% sucrose/PBS overnight at 4°C, quick frozen in Tissue Tek, and stored at −70°C until use. After cryosectioning, tissues were treated in 100% cold acetone for 10 min, washed with PBS (0.05% tween), blocked with 0.5% BSA and 50% v/v Avidin D block solution (Vector Laboratories Inc., Burlingame, CA) for 15 min, washed three times, blocked with Biotin blocking solution (Vector Laboratories, Inc.) for 15 min, and washed three times. Subsequently, sections were incubated with biotinylated antibodies for CD31 or Sca-1 (Pharmingen) at room temperature for 1 hr, washed three times, incubated with the detection reagent avidin TexasRed (Vector Laboratories Inc.) or Streptavi-

din-Cy5 (Rockland, Gilbertsville, PS), washed three times, dehydrated in ethanol (from 70% to 100%), and mounted with vectashield (Vector Laboratories, Burlingame, CA). For whole mount confocal microscopy, AGM tissues were carefully dissected from E11 *Ly-6A GFP* embryos, kept in PBS (containing 10% FCS), and directly observed.

DNA Analysis

Genomic DNA (200 ng) from the peripheral blood of transplant recipients was analyzed by PCR using oligonucleotide primers for *myogenin*, 5'TTACGTCCATCGTGGACAGC3' and 5'TGGGCTGGGTGT TAGTCTTA3', GFP 5'GACAGAACTTGCCACTGTGC 3' and 5'AAG AAGATGGTGCGCTCCTG 3', human β-*globin* 5'CTTCAGGTTCCC AGTGAGGATG3' and 5'GCTCCCTAAGGGGTAAAGAGTG3', and *lacZ* 5'GCGACTTCCAGTTCAACATC3' and 5'GATGAGTTTGGACA AACCAC3'.

DNA was subjected to an initial 5 min denaturation at 94°C followed by 30 cycles of denaturation (1 min at 94°C), annealing (2 min at 58°C [β-*globin*] or 55°C [*lacZ* and *GFP*]), and elongation (2 min at 72°C). Serial dilutions of blood DNA from a transgenic animal were used as a control to evaluate the levels of donor cell reconstitution in transplanted mice.

Acknowledgments

The authors would like to thank Nancy Speck for her constant enthusiasm and contributions to this work. We also would like to thank Karin van der Horn, Marian Peeters, Kam-Wing Ling, and all the members of the lab for help; John Kong-A-San for microinjections; Corne Snoijs for flow cytometry; Gert van de Cappelen for confocal microscopy; and Danielle Zondervan-Bakker, Henk Dronk, and the EDC. This work was funded by Netherlands Scientific Research Organization 901-08-090, National Institutes of Health R01 DK51077, La Ligue Nationale Contre le Cancer (C.R.), and European Community QLK-CT-1999-00020.

Received: September 24, 2001
Revised: April 3, 2002

References

Bemex, R., De Sepulveda, P., Kress, C., Elbaz, C., Delouis, C., and Panthier, J.-J. (1996). Spatial and temporal patterns of *c-kit*-expressing cells in $W^{lacZ}/+$ and W^{lacZ}/W^{lacZ} mouse embryos. Development *122*, 3023–3033.

Cai, Z.L., de Bruijn, M., Ma, X., Dortland, B., Luteijn, T., Downing, J.R., and Dzierzak, E. (2000). Haploinsufficiency of AML1/CBFA2 affects the temporal and spatial generation of hematopoietic stem cells in the mouse embryo. Immunity *13*, 423–431.

Cumano, A., Ferraz, J.C., Klaine, M., Di Santo, J.P., and Godin, I. (2001). Intraembryonic, but not yolk sac hematopoietic precursors, isolated before circulation provide long-term multilineage reconstitution. Immunity *15*, 477–485.

de Bruijn, M.R.T.R., Speck, N.A., Peeters, M.C.E., and Dzierzak, E. (2000). Definitive hematopoietic stem cells first emerge from the major arterial regions of the mouse embryo. EMBO J. *19*, 2465–2474.

Drake, C.J., and Fleming, P.A. (2000). Vasculogenesis in the day 6.5 to 9.5 mouse embryo. Blood *95*, 1671–1679.

Dzierzak, E., and de Bruijn, M. (2002). Isolation and analysis of hematopoietic stem cells from mouse embryos. In Methods in Molecular Medicine: Hematopoietic Stem Cell Protocols, C. Klug and C. Jordan, eds. (Totowa, NJ: The Humana Press Inc.), pp. 1–14.

Garcia-Porrero, J.A., Godin, I.E., and Dieterlen-Lievre, F. (1995). Potential intraembryonic hemogenic sites at pre-liver stages in the mouse. Anat. Embryol. (Berl.) *192*, 425–435.

Garcia-Porrero, J.A., Manaia, A., Jimeno, J., Lasky, L.L., Dieterlen-Lièvre, F., and Godin, I.E. (1998). Antigenic profiles of endothelial and hemopoietic lineages in murine intraembryonic hemogenic sites. Dev. Comp. Immunol. *22*, 303–319.

Godin, I., Garcia-Porrero, J.A., Dieterlen-Lievre, F., and Cumano, A. (1999). Stem cell emergence and hemopoietic activity are incompatible in mouse intraembryonic sites. J. Exp. Med. *190*, 43–52.

Huang, H., and Auerbach, R. (1993). Identification and characterization of hematopoietic stem cells from the yolk sac of the early mouse embryo. Proc. Natl. Acad. Sci. USA *90*, 10110–10114.

Kamiura, S., Nolan, C.M., and Meruelo, D. (1992). Long-range physical map of the Ly-6 complex: mapping the Ly-6 multigene family by field-inversion and two-dimensional gel electrophoresis. Genomics *12*, 89–105.

Khan, K.D., Lindwall, G., Maher, S.E., and Bothwell, A.L. (1990). Characterization of promoter elements of an interferon-inducible Ly-6E/A differentiation antigen, which is expressed on activated T cells and hematopoietic stem cells. Mol. Cell. Biol. *10*, 5150–5159.

Khan, K.D., Shual, K., Lindwall, G., Maher, S.E., Damell, J.E., and Bothwell, A.L. (1993). Induction of the Ly-6A/E gene by interferon alpha/beta and gamma requires a DNA element to which a tyrosine-phosphorylated 91-kDa protein binds. Proc. Natl. Acad. Sci. USA *90*, 6806–6810.

Kimura, S., Tada, N., Liu-Lam, Y., and Hammerling, U. (1984). Studies of the mouse Ly-6 alloantigen system. II. Complexities of the Ly-6 region. Immunogenetics *20*, 47–56.

LeClair, K.P., Palfree, R.G., Flood, P.M., Hammerling, U., and Bothwell, A. (1986). Isolation of a murine Ly-6 cDNA reveals a new multigene family. EMBO J. *5*, 3227–3234.

Lemischka, I.R. (1991). Clonal, in vivo behavior of the totipotent hematopoietic stem cell. Semin. Immunol. *3*, 349–355.

Ma, X., Ling, K.-W., and Dzierzak, E. (2001). Cloning of the Ly-6A (Sca-1) gene locus and identification of a 3' distal fragment responsible for high level γ-interferon-induced expression in vitro. Br. J. Haematol. *114*, 724–730.

Ma, X., de Bruijn, M., Robin, C., Peeters, M., Kong-A-San, J., Snoijs, C., and Dzierzak, E. (2002). Expression of the Ly-6A (Sca-1) lacZ transgene in mouse hematopoietic stem cells and embryos. Br. J. Haematol. *116*, 401–408.

Marshall, C.J., and Thrasher, A.J. (2001). The embryonic origins of human hematopoiesis. Br. J. Haematol. *112*, 838–850.

McGrew, J.T., and Rock, K.L. (1991). Isolation, expression, and sequence of the TAP/Ly-6A.2 chromosomal gene. J. Immunol. *146*, 3633–3638.

Medvinsky, A., and Dzierzak, E. (1996). Definitive hematopoiesis is autonomously initiated by the AGM region. Cell *86*, 897–906.

Miles, C., Sanchez, M.-J., Sinclair, A., and Dzierzak, E. (1997). Expression of the Ly-6E.1 (Sca-1) transgene in adult hematopoietic stem cells and the developing mouse embryo. Development *124*, 537–547.

Morrison, S.J. (2001). Stem cell potential: can anything make anything? Curr. Biol. *11*, R7–R9.

Mukouyama, Y.-S., Chiba, N., Hara, T., Okada, H., Ito, Y., Kanamaru, R., Miyajima, A., Satake, M., and Watanabe, T. (2000). The AML1 transcription factor functions to develop and maintain hematogenic precursor cells in the embryonic aorta-gonad-mesonephros region. Dev. Biol. *220*, 27–36.

Muller, A.M., Medvinsky, A., Strouboulis, J., Grosveld, F., and Dzierzak, E. (1994). Development of hematopoietic stem cell activity in the mouse embryo. Immunity *1*, 291–301.

Murray, P. (1932). The development *in vitro* of the blood of the early chick embryo. Proc. R. Soc. Lond. B Biol. Sci. *11*, 497–521.

North, T., Gu, T.-L., Stacy, T., Wang, Q., Howard, L., Binder, M., Marin-Padilla, M., and Speck, N. (1999). Cbfa is required for the formation of intraaortic hematopoietic clusters. Development *126*, 2563–2575.

North, T.E., de Bruijn, M.F.T.R., Stacy, T., Talebian, L., Lind, E., Robin, C., Binder, M., Dzierzak, E., and Speck, N.A. (2002). Runx1 expression marks long-term repopulating hematopoietic stem cells in the midgestation mouse embryo. Immunity *16*, this issue, 661–672.

Okada, S., Nakauchi, H., Nagayoshi, K., Nishikawa, S., Miura, Y., and Suda, T. (1992). In vivo and in vitro stem cell function of c-kit- and Sca-1-positive murine hematopoietic cells. Blood *80*, 3044–3050.

Okuda, T., van Deursen, J., Hiebert, S.W., Grosveld, G., and Down-

ing, J.R. (1996). AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. Cell *84*, 321–330.

Sanchez, M.J., Holmes, A., Miles, C., and Dzierzak, E. (1996). Characterization of the first definitive hematopoietic stem cells in the AGM and liver of the mouse embryo. Immunity *5*, 513–525.

Shalaby, F., Ho, J., Stanford, W.L., Fischer, K.D., Schuh, A.C., Schwartz, L., Bernstein, A., and Rossant, J. (1997). A requirement for Flk1 in primitive and definitive hematopoiesis and vasculogenesis. Cell *89*, 981–990.

Sinclair, A.M., and Dzierzak, E.A. (1993). Cloning of the complete Ly-6E.1 gene and identification of DNase I hypersensitive sites corresponding to expression in hematopoietic cells. Blood *82*, 3052–3062.

Sinclair, A., Daly, B., and Dzierzak, E. (1996). The Ly-6E.1 (Sca-1) gene requires a 3' chromatin-dependent region for high level gamma-interferon-induced hematopoietic cell expression. Blood *87*, 2750–2761.

Spangrude, G.J., and Brooks, D.M. (1993). Mouse strain variability in the expression of the hematopoietic stem cell antigen Ly-6A/E by bone marrow cells. Blood *82*, 3327–3332.

Spangrude, G.J., Heimfeld, S., and Weissman, I.L. (1988). Purification and characterization of mouse hematopoietic stem cells. Science *241*, 58–62.

Spangrude, G.J., Smith, L., Uchida, N., Ikuta, K., Heimfeld, S., Friedman, J., and Weissman, I.L. (1991). Mouse hematopoietic stem cells. Blood *78*, 1395–1402.

Stanford, W.L., Bruyns, E., and Snodgrass, H.R. (1992). The isolation and sequence of the chromosomal gene and regulatory regions of Ly-6A.2. Immunogenetics *35*, 408–411.

Strouboulis, J., Dillon, N., and Grosveld, F. (1992). Developmental regulation of a complete 70-kb human beta globin locus in transgenic mice. Genes Dev. *6*, 1857–1864.

Tavian, M., Hallais, M.F., and Peault, B. (1999). Emergence of intraembryonic hematopoietic precursors in the pre-liver human embryo. Development *126*, 793–803.

van de Rijn, M., Heimfeld, S., Spangrude, G.J., and Weissman, I.L. (1989). Mouse hematopoietic stem-cell antigen Sca-1 is a member of the Ly-6 antigen family. Proc. Natl. Acad. Sci. USA *86*, 4634–4638.

Wang, Q., Stacy, T., Binder, M., Marín-Padilla, M., Sharpe, A.H., and Speck, N.A. (1996). Disruption of the Cbfa2 gene causes necrosis and hemorrhaging in the central nervous system and blocks definitive hematopoiesis. Proc. Natl. Acad. Sci. USA *93*, 3444–3449.

Wood, H.B., May, G., Healy, L., Enver, T., and Morriss-Kay, G.M. (1997). CD34 expression patterns during early mouse development are related to modes of blood vessel formation and reveal additional sites of hematopoiesis. Blood *90*, 2300–2311.

Wulf, G.G., Jackson, K.A., and Goodell, M.A. (2001). Somatic stem cell plasticity: current evidence and emerging concepts. Exp. Hematol. *29*, 1361–1370.

Yoder, M.C., Hiatt, K., Dutt, P., Mukherjee, P., Bodine, D.M., and Orlic, D. (1997). Characterization of definitive lymphohematopoietic stem cells in the day 9 murine yolk sac. Immunity *7*, 335–344.

Note Added in Proof

We have recently found some rare GFP$^+$ clusters on the ventral wall of the dorsal aorta in transverse sections from E10 *Ly-6A GFP* transgenic embryos.

Chapter 6

Concluding remarks and future prospects

Concluding remarks

Although the experimental results have been presented and discussed in each chapter, in this chapter, I present here an overview of the main conclusions and discussions of my study.

In Chapter 2, in the *in vitro* studies examining a 14kb genomic fragment of the *Ly-6A* allele, I have shown that the sequences analogous to those in the *Ly-6E* allele are responsible for high-level γ-IFN induced expression in MEL cells. The deletion analyses performed in transfected MEL cells confirm that the two alleles, *Ly-6A* and *Ly-6E* are regulated very similarly on the level of transcription. Furthermore, I have shown that the 1 kb 3' distal *Ly-6A* fragment (containing HSS +8.7 and +8.9) directs high-level γ-IFN-induced expression from a heterologous promoter. It is highly homologous in DNA sequence to the *Ly-6E* 3' 1 kb distal fragment previously studied by Sinclair and colleagues (Sinclair et al., 1996). Taken together, these studies suggest that this fragment contains a potent enhancer that could be extremely useful upon further dissection, for the expression of exogenous genes specifically in hematopoietic stem cells.

From the *in vivo* studies, in which I examined hematopoietic cells of transgenic mice, the following conclusions can be made: the 14 kb genomic fragment which contains two 3' distal DNaseI hypersensitive sites (+8.7, +8.9 from transcriptional starting site) efficiently and faithfully drives the expression of inserted β-galactosidase marker gene in the expected hematopoietic lineages. Furthermore, the β-galactosidase marker gene driven by the *Ly-6A* cassette in transgenic adult bone marrow can be used in combination with the FDG substrate to enrich for HSCs by flow cytometric sorting (Chapter 3).

In direct comparisons of the *Ly-6ElacZ* and *Ly-6AlacZ* transgenic lines, no obvious differences were found between the two allelic expression cassettes as measured by transgene expression patterns in hematopoietic system or at the level of enrichment of functional HSCs obtained by FDG/Sca-1 flow cytometric sorting. Furthermore, no obvious differences were found between transgenic lines with different copy numbers of the *Ly-6AlacZ* transgene. These results suggest that the *Ly-6AlacZ* transgene is no better than the *Ly-6E lacZ* transgene for the HSC specific expression in the adult. These results are in direct contrast to the results of Spangrude (Spangrude and Brooks, 1993) in which the *Ly-6A* gene product is expressed in 100% of marrow repopulating cells while the *Ly-6E* gene product is expressed in only 25% marrow repopulating cells. Thus the incomplete expression of *Ly-6A* in all HSCs in our transgenic mice may be due to the lack of all the appropriate transcriptional control elements within the 14 kb *Ly-6A* cassette. Alternatively, β-galactosidase production may reach some physiological threshold in our transgenic mice, and thus, levels appear to be limited in the *Ly-6AlacZ* transgenic mice. Finally, our HSC enrichment procedure relies on the permeability of these cells for the FDG substrate. Limitation on the entry and/or retention of the substrate in the cells may affect efficiency.

To alleviate any possible problems of the β-galactosidase marker gene in HSC enrichment procedures, I performed *in vivo* studies using a *GFP* marker gene in the *Ly-6A* expression cassette (Chapter 4 and 5). I showed that all adult BM HSCs are GFP positive in such transgenic mice. Thus, the *Ly-6AGFP* transgenic model is better

and more appropriate than the *Ly-6AlacZ* transgenics for the study of HSC development. Within transgenic embryos, I also observed clear differences between *lacZ* and *GFP* reporter expression. In AGM tissue at 10-11 dpc, *GFP* is expressed in the AGM region: this includes the mesonephric tubules, Wolffian/Mullerian duct and most importantly, the dorsal aorta endothelium. Furthermore, GFP is also expressed in vitelline and umbilical vessels at this stage. A more limited expression pattern for the *lacZ* reporter was observed at this stage. While mesonephric tubes/Wolffian ducts expressed *lacZ*, expression was completely lacking around dorsal aorta (Chapter 3). FDG sorting for HSCs in the AGM was not successful in *Ly-6AlacZ* transgenic embryos, while the HSCs are exclusively within the GFP^+ compartment of AGM region, especially dorsal aorta sub-region of *Ly-6AGFP* transgenic embryos. Most likely, the expression difference between *Ly-6AlacZ* and *Ly-6AGFP* transgenes is related to the characteristics of the inserted marker gene itself: (1) The baterial *lacZ* gene may be constrained in its expression in certain cell types in the mice; (2) The long procedure of paraffin embedding at high temperature can reduce the X-gal staining signal on cross sections; (3) Permeabilization of cells for the FDG substrate can affect the sorting results. Nevertheless, the *Ly-6AGFP* transgene quite uniquely marks the HSCs in the tissue around dorsal-aorta subregion and localizes the 10-12 dpc functional HSCs to the endothelial layer of this vessel. These important results in *Ly-6AGFP* embryos have profound implications on our thinking concerning the origins of HSCs during ontogeny and suggest that intraembryonic HSCs arise most likely from hemogenic endothelial cells in the midgestation mouse embryo.

This thesis has not touched upon the topic of *Ly-6A/E* gene function *in vivo*. Since *Ly-6A/E* is a cell surface GPI linked glycoprotein it might be expected to have some role in signaling hematopoietic differentiation or HSC production/maintenance. The results of gene targeting experiments have demonstrated that *Ly-6A* mutant mice develop normally, are healthy and have normal numbers and percentages of hematopoietic cell lineages (Stanford et al., 1997). However since the *Ly-6A* gene is a member of the large highly homologous *Ly-6* multigene family, this result implies that the functional study *in vivo* of any single member of the *Ly-6* gene family could be difficult because of the suspected redundant/overlapping roles played by the members of this family. Recently, the first direct demonstration of physical linkage of several members of the *Ly-6* family on chromosome 15 has been reported (Patterson et al., 2000). It has been shown that the Sca-2 gene lies 35.4 kb downstream of the *Ly-A/E* (Sca-1) in the opposite transcriptional orientation and the *Ly-6M* gene (another novel identified, highly homologous to *Ly-6A/E* gene) localizes 13.4 kb downstream of Sca-1. Based on this physical map, the *in vivo* functional studies aimed at targeting this locus (including at least two genes) seems possible.

Interestingly, it has been found that the T lymphocytes from the *Ly-6A null* mice proliferate at a significantly higher rate in response to antigens and mitogens than wild-type T-cells. This implies that *Ly-6A* (Sca-1) plays some role in T cell biological function. Furthermore, during T cell differentiation and maturation, bone marrow derived $Sca\text{-}1^+$ prothymocytes seed the thymic cortex, lose Sca-1 expression and phenotypically obtain Sca-2 expression on double negative ($CD4^-/CD8^-$) immature T cells (Spangrude et al., 1988; Wu et al., 1991; Yeh et al., 1986). Later

Sca-1 is re-expressed on mature single positive ($CD4^+/CD8^-$, $CD4^-/CD8^+$) medullary thymocytes and peripheral T cells (Spangrude et al., 1988; Yeh et al., 1986). Taking these observations together, it appears that Sca-1 and Sca-2 both play role during T cell differentiation and maturation. It would be interesting to study single Sca-2 mutant mice and Sca-1/Sca-2 double knockout mice to further dissect the function of Sca-1 and Sca-2 during T cell development.

Future prospects

Since all HSCs at 11 dpc in the dorsal aorta express the *Ly-6AGFP* transgene, this transgenic model is an excellent marker system for studying the induction and expression of adult HSCs during mouse development. By using this model, temporal and spatial analysis of the expression patterns of the *Ly-6AGFP* and the function of these marked cell populations could be carefully and systematically performed. The successful tracing of emerging, expressing, migrating and colonizing HSCs during mouse embryonic development will provide insight into the generation of adult hematopoietic system. Although some experiments to test HSCs and progenitors at 10 and 11 dpc in embryonic tissues, such as AGM, subdissected aorta and urogenital ridges have been performed either by direct transplantation or by explant culture and then transplantation (Chapter 5), more data on these aspects should be collected in the slightly earlier stages like 8, 9 and 10 dpc. A careful fractionation of AGM cell populations based on surface phenotype prior to the onset of GFP expression and HSC function will be extremely useful in identifying the direct precursor cells to adult HSCs. The GFP marker can then be used to determine the molecular and cellular signals necessary for the onset of GFP expression and hence, HSC function.

Interestingly, preliminary data has shown that the cell cycle profiles of GFP^+ and GFP^- cells from dorsal aorta at 11 dpc are quite different (data not shown). It would be worthwhile to verify these results and to determine if cell cycle status is related to GFP^+ expression and HSC function. The experiments can start with sorting different cell cycle subsets within GFP^+ population of the dorsal aorta and/or bone marrow and transplantation into irradiated adult recipients.

In other related experiments, it would be interesting to use GFP^+ and/or GFP^- populations from dorsal aorta and/or urogenital ridges to test the stromal cell lines which others in our laboratory have isolated from aorta, urogenital ridges, fetal liver and BM. These kinds of experiments should lead to the uncovering of promising stromal cell lines that can maintain and/or proliferate HSCs. It would be very exciting if any GFP^- cells from AGM and/or subdissected aorta and urogenital ridges become GFP^+ and acquire HSC function after culture with certain stromal cell lines. This stromal cell and HSC precursor co-culture system can then be analysed for the specific cellular and molecular signals necessary to generate these stem cells.

The *Ly-6AGFP* transgenic embryos can also be used for characterizing the relationship between some transcription factors and the onset of HSC generation and function. *GATA-2* mutant, *AML-1* mutant /*CBFA2 lacZ* knockin transcription factor mouse models are presently available and being used in our laboratory. *CBFA2* is expressed in the ventral endothelium of the dorsal aorta, the ventral para-aortic mesenchyme and in the clusters of hemetopoietic cells associated with the ventral wall of the aorta (North et al., 1999). Furthermore, *CBFA2* is required for the

formation of intra-aortic hematopoietic clusters (North et al., 1999) and for the generation of HSCs in the AGM region (Cai et al., 2000). Taken together, these results suggest that *CBFA2* is expressed by the first HSCs as they arise in the mouse embryo. *GATA-2* transcription factor has also been found to affect definitive hematopoiesis, (causing lethality at day 10.5-11 in gestation) (Tsai et al., 1994). To determine if the expression of these transcription factors and their functions in HSCs and/or their precursors precede *Ly-6A* expression, crosses of *Ly-6AGFP* homozygous mice with *GATA-2* and *AML-1 (CBFA2)* mutant mice can be performed. Morphological and functional studies should uncover the relative positions of these two transcription factors and the *Ly-6A* gene in the gene expression and functional cascades related to HSC generation in the AGM.

Ongoing experiments will use the 14 kb *Ly-6A* expression cassette for the expression of Cre recombinase in the fate mapping of HSCs during ontogeny. Furthermore, genes thought to be related to HSC ontogeny can be conditionally knocked-out by using a *Ly-6A Cre-LoxP* system. The success of these experiments is dependent on the correct temporal and spatial expression pattern of Cre recombinase during development. For this reason careful examination and comparison with the pattern of *Ly-6AGFP* reporter gene expression results presented in this thesis are of outmost importance for these studies.

Finally, the Ly-6A expression cassette can be also used to drive expression of some oncogenes (i.e. *myc*, *bcl2*) in transgenic mice for the immortalization of definitive HSCs so as to better genetically characterize such cells. Previously, this strategy using another expression cassette has been successful for the immortalization of B cell progenitors (Strasser et al., 1990). However, improvements in this strategy will include the flanking of the oncogene with loxP recombination sites in the Ly-6A transgene construct. The presence of the loxP sites will allow for the removal of the immortalizing oncogene after cell lines are isolated from transgenic mice. Established lines characterized as being hematopoietic stem cell-like will be induced with Cre recombinase to remove the oncogene. Thus, the stem cells should return to an un-immortalized state, undergo normal differentiation and form a complete hematopoietic system in the irradiated mouse recipient.

References

Cai, Z., de Bruijn, M., Ma, X., Dortland, B., Luteijn, T., Downing, R.J. and Dzierzak, E. (2000) Haploinsufficiency of AML1 affects the temporal and spatial generation of hematopoietic stem cells in the mouse embryo. *Immunity*, **13**, 423-431.

North, T., Gu, T.L., Stacy, T., Wang, Q., Howard, L., Binder, M., Marin-Padilla, M. and Speck, N.A. (1999) Cbfa2 is required for the formation of intra-aortic hematopoietic clusters. *Development*, **126**, 2563-2575.

Patterson, J.M., Johnson, M.H., Zimonjic, D.B. and Graubert, T.A. (2000) Characterization of Ly-6M, a novel member of the Ly-6 family of hematopoietic proteins. *Blood*, **95**, 3125-3132.

Sinclair, A., Daly, B. and Dzierzak, E. (1996) The Ly-6E.1 (Sca-1) gene requires a 3' chromatin-dependent region for high-level gamma-interferon-induced hematopoietic cell expression. *Blood*, **87**, 2750-2761.

Spangrude, G.J., Aihara, Y., Weissman, I.L. and Klein, J. (1988) The stem cell antigens Sca-1 and Sca-2 subdivide thymic and peripheral T lymphocytes into unique subsets. *J Immunol*, **141**, 3697-3707.

Spangrude, G.J. and Brooks, D.M. (1993) Mouse strain variability in the expression of the hematopoietic stem cell antigen Ly-6A/E by bone marrow cells. *Blood*, **82**, 3327-3332.

Stanford, W.L., Haque, S., Alexander, R., Liu, X., Latour, A.M., Snodgrass, H.R., Koller, B.H. and Flood, P.M. (1997) Altered proliferative response by T lymphocytes of Ly-6A (Sca-1) null mice. *J Exp Med*, **186**, 705-717.

Strasser, A., Harris, A.W., Bath, M.L. and Cory, S. (1990) Novel primitive lymphoid tumours induced in transgenic mice by cooperation between myc and bcl-2. *Nature*, **348**, 331-333.

Tsai, F.Y., Keller, G., Kuo, F.C., Weiss, M., Chen, J., Rosenblatt, M., Alt, F.W. and Orkin, S.H. (1994) An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature*, **371**, 221-226.

Wu, L., Antica, M., Johnson, G.R., Scollay, R. and Shortman, K. (1991) Developmental potential of the earliest precursor cells from the adult mouse thymus. *J Exp Med*, **174**, 1617-1627.

Yeh, E.T., Reiser, H., Benacerraf, B. and Rock, K.L. (1986) The expression, function, and ontogeny of a novel T cell-activating protein, TAP, in the thymus. *J Immunol*, **137**, 1232-1238.

Summary

Mammalian hematopoiesis begins early during embryogenesis and proceeds throughout the entire life span of the animal to provide at least eight distinct lineages of mature blood cells that derive from pluripotential hematopoietic stem cells (HSCs). These rare cells, found in the bone marrow of adult mammals, are the basic of clinical transplantation therapies for leukemia and genetic diseases of the blood system. While HSCs are not morphologically distinguishable from the overwhelming numbers of other blood cells in the bone marrow, they have been characterized and enriched by a unique combination of markers expressed on their cell surface. In the mouse, the Sca-1 marker has been instrumental in such studies. This marker is a cell surface GPI linked glycoprotein that is found on bone marrow and fetal liver HSCs. Despite enrichment of HSCs by flow cytometric sorting for Sca-1 as well as other markers, the isolation of the necessary large numbers for genetic manipulation and transplantation is difficult due to the inability of these cell to proliferate in the adult microenvironment or *in vitro* culture systems. However, during embryonic development HSCs have been found to increase in number in an intraembryonic hematopoietically active region containing the dorsal aorta, gonads and mesonephros (AGM). Thus, the AGM region offers a unique opportunity to examine the earliest generation of definitive HSCs.

The AGM HSCs have been found to be positive for stem cell markers c-kit, CD34, Sca-1 and Mac and negative for mature lineage markers B220, CD4, CD8 and Gr-1. In order to better understand the transcription regulatory elements directing gene expression in HSCs so as to manipulate and express heterologous genes in these stem cells for the study of development and for possible therapeutic applications, identification and isolation of stem cell control elements of *Sca-1* (*Ly-6A/E*) gene was performed.

In this thesis studies two overlapping fragments (over 25 kb of the *Ly-6A* gene) from a phage library of the 129 mouse strain were cloned and selected by hybridization with *Ly-6E* cDNA probe. These fragments were mapped, subcloned and used in expression studies. Deletion analysis shows that the sequences analogous to those in the *Ly-6E* allele are responsible for high-level γ-IFN-induced expression *in vitro*. Furthermore, the 3' distal *Ly-6A* fragment containing two strong DNase I hypersensitive sites was proven to direct high-level γ-IFN-induced expression from a heterologous promoter, suggesting that it is a potent enhancer that could be useful for expression in hematopoietic stem cells. On the basis of *in vitro* studies, the 14 kb *Ly-6A* genomic fragment containing the 3' distal two strong DNase I hypersensitive sites was used as a expression cassette to drive heterologous report gene (lacZ and EGFP) for *in vivo* studies.

For the *Ly-6AlacZ* transgenic mice, the specific expression pattern of two lines (AZ1, 8 copies; AZ2, 20 copies) is similar. Furthermore, the expression pattern in the adult is also similar between *Ly-6AlacZ* and the previously generated *Ly-6ElacZ* lines. In general, *Ly-6AlacZ* transgene is expressed predominantly in the cells of the T lymphoid lineage and this transgene facilitates about 100-fold enrichment of bone marrow HSCs. Compared to the *Ly-6ElacZ* transgene, *Ly-6AlacZ* is expressed more widely in embryonic tissues (not only AGM and tail, but also limb bud) in

midgestation embryos. LacZ expression was not observed around the dorsal aorta at midgestation in *Ly-6ElacZ* or *Ly-6AlacZ* transgenic lines. Also, FDG sorting and enrichment of lacZ marked AGM HSCs has not been successful with either *Ly-6AlacZ* or *Ly-6ElacZ* lines. Thus, further studies were focused on *Ly-6AGFP* transgenic mice.

The *Ly-6A* GFP transgene is not only expressed in all functional repopulating HSCs in the adult bone marrow, but also is expressed in all of the HSCs in the midgestation AGM region, the first site of HSC emergence. While transgenic embryos express GFP in the aorta-mesenchyme and the gonad-mesonephros subregions, functional adult-repopulating HSCs are found only in the GFP^+ cells of the aorta-mesenchyme. Examination of whole and serially sectioned AGM regions shows that GFP expressing cells are localized to the wall of the dorsal aorta, strongly suggesting that HSC activity arises within the endothelial layer of this vessel. The results of these studies suggest that the 14 kb *Ly-6A* genomic fragment including the 3' DNaseI hypersensitive sites contains a potent enhancer for the exogenous gene (LacZ and GFP) expression in hematopoietic stem cells. Further examination of the GFP transgene expression patterns during mouse midgestation should be an extremely useful approach for fate mapping. The GFP^+, and even GFP^- populations from intraembryonic cells around dorsal aorta at midgestation stage are the important sources for HSCs differentiation and induction studies. Furthermore, the 14 kb *Ly-6A* expression cassette can be used as a starting point to further dissection of the transcriptional regulatory elements within the *Ly-6A* gene.

Samenvatting

De vorming van bloedcellen wordt hematopoiese genoemd. Het begint bij zoogdieren al vroeg in de embryogenese en gaat vervolgens gedurende het hele leven van het dier door. Het levert tenminste acht verschillende typen rijpe bloedcellen die allen afstammen van multipotente hematopoietische stamcellen (HSCs). Deze zeldzame cellen bevinden zich in het beenmerg van volwassen zoogdieren en vormen de basis voor transplantatie-therapiëën voor leukemie en genetische ziekten van het bloed. Omdat de HSCs morfologisch niet te onderscheiden zijn van de enorme aantallen andere bloedcellen in het beenmerg, zijn ze gekarakteriseerd en verrijkt door een unieke combinatie van kenmerken die aanwezig zijn op het celoppervlak. In de muis is Sca-1 zo'n kenmerk. Sca-1 is een glycoproteine op het celoppervlak dat verbonden is met GPI, en is gevonden op HSCs van het beenmerg en de foetale lever. Ondanks de verrijking van HSCs met behulp van een flow cytometer voor Sca-1, als ook andere kenmerken, is het verkrijgen van de noodzakelijke grote aantallen cellen voor genetische manipulatie en transplantatie moeilijk omdat deze cellen niet kunnen prolifereren in een volwassen omgeving of in *in vitro* kweeksystemen. Tijdens de embryonale ontwikkeling nemen HSCs echter wel in aantal toe in een intraembryonaal hematopoietisch aktief gebied, bestaande uit de dorsale aorta, de gonaden en de mesonephros (AGM). Het AGM gebied biedt dus een unieke mogelijkheid om de eerste vorming en proliferatie van definitieve HSCs te bestuderen.

Het is aangetoond dat de HSCs van de AGM positief zijn voor de stamcelkenmerken c-kit, CD34, Sca-1 en Mac en negatief voor kenmerken van rijpe bloedcellen als B220, CD4, CD8 en Gr-1. Om de genexpressie in stamcellen te kunnen manipuleren voor het bestuderen van de ontwikkeling en mogelijke therapeutische toepassingen is het noodzakelijk om meer te weten van de transcriptie regulerende elementen die verantwoordelijk zijn voor de genexpressie in HSCs. In dit kader zijn de stamcel controle elementen van het Sca-1 (*Ly-6A/E*) gen geïdentificeerd en geïsoleerd.

In dit proefschrift zijn twee overlappende fragmenten (van meer dan 25 kb van het *Ly-6A* gen) van een faag-bank van de muizenstam 129 gekloneerd en geselecteerd door middel van hybridisatie met een *Ly-6E* cDNA probe. Deze fragmenten zijn in kaart gebracht, gesubkloneerd en gebruikt in expressie-studies. Het blijkt dat delen van de fragmenten die analoog zijn aan die van het *Ly-6E* allel verantwoordelijk zijn voor het door γ-IFN geïnduceerde hoge expressie nivo *in vitro*. Bovendien bleek het 3' *Ly-6A* fragment, dat 2 sterke DNase I hypergevoelige gebieden bevat, een door γ-IFN geïnduceerde hoog expressie nivo van een andere promoter te regisseren. Dit suggereert dat het een krachtige "enhancer" is die bruikbaar is voor expressie in hematopoietische stamcellen. Daarom is het 14 kb *Ly-6A* genomisch fragment dat dit 3' fragment bevat gebruikt als een expressie cassette met de markeringsgenen lacZ en EGFP voor *in vivo* studies.

Er zijn 2 *Ly-6AlacZ* transgene muizenlijnen gegenereerd (AZ1/ 8 kopieen, AZ2/ 20 kopieen), waarvan het specifieke expressie patroon vergelijkbaar is. Bovendien is het expressie patroon in de volwassen muizen vergelijkbaar in de *Ly-6AlacZ* en in de eerder gegenereerde *Ly-6ElacZ* lijnen. *Ly-6AlacZ* komt overwegend

tot expressie tijdens de T-celdifferentiatie en met dit transgeen is een 100-voudige verrijking van beenmerg HSCs mogelijk. Tijdens de organogenese komt *Ly-6AlacZ* in meer embryonale weefsels tot expressie (niet alleen de AGM en staart, maar ook in de pootknop) dan *Ly-6ElacZ*. Op dit stadium is lacZ expressie echter niet waargenomen rond de dorsale aorta van de *Ly-6ElacZ* en de *Ly-6AlacZ* transgene lijnen. Tevens was het sorteren en verrijken van lacZ gemerkte AGM HSCs niet succesvol met noch de *Ly-6ElacZ* noch de *Ly-6AlacZ* lijnen. Daarom zijn verdere studies uitgevoerd met een *Ly-6AGFP* transgene muis.

Ly-6AGFP komt niet alleen tot expressie in alle functionele repopulerende HSCs van volwassen beenmerg, maar het komt ook tot expressie in alle HSCs van de AGM ten tijde van de organogenese. Dit gebied is het vroegste embryonale gebied waar HSCs voorkomen. Transgene embryo's vertonen GFP expressie in de gebieden van de aorta-mesenchym en van de gonade-mesonephros, maar functionele HSC die in staat zijn volwassen muizen te repopuleren, zijn alleen gevonden in de GFP^+ cellen van het gebied van de aorta-mesenchym. Onderzoek van totale en in serie gesneden AGM gebieden toont aan dat de cellen die GFP tot expressie brengen zijn gelokaliseerd in de wand van de dorsale aorta. Dit suggereert dat HSC aktiviteit onstaat in het endotheel van dit bloedvat. De resultaten van deze studies suggereren dat het 14 kb *Ly-6A* genomische fragment inclusief de 3' DNase I hypergevoelige gebieden een krachtige "enhancer" bevat voor het tot expressie brengen van een exogeen gen (lacZ en GFP) in hematopoietische stamcellen. Verder onderzoek van de GFP transgene expressie patronen tijdens de organogenese zal zeer bruikbaar zijn voor het in kaart brengen van het ontstaan van de HSC. De GFP^+ en zelfs de GFP^- celpopulaties van de embryonale dorsale aorta zijn belangrijke bronnen voor studies naar HSC differentiatie en inductie. Bovendien kan het 14 kb *Ly-6A* fragment gebruikt worden als startpunt in de verdere specificering van transcriptie regulerende elementen in het *Ly-6A* gen.

Curriculum Vitae

The author, Xiaoqian Ma was born on 26 May 1963 in Wuhan, the Capital City of Hubei Province, the People's Republic of China. When she was 14 years old, the Cultural Revolution was over and it became possible for anyone to have chance to study in university as long as he or she could pass the highly competitive annual national examination. This policy led to her decision to give up professional training for athlete and catch up normal studies from secondary school. In 1982, she enrolled in the Medical School of Southeast University in Nanjing, the Capital City of Jiangsu Province. In 1987, she finished her undergraduate studies and continued to pursue postgraduate studies at Morphological Research Centre (Histopathology and Embryology Department, under the supervision of Jianmin Xing) in the same school, obtained Master Degree of Medical Science and MD in 1990. Since then she worked as a resident in general internal medicine for two years. From 1992 she worked as a staff member at Morphological Research Centre in the same school.

In 1995, she got fellowship through the cultural exchange program between the Chinese and the Dutch governments and came to Rotterdam as a visiting scholar. She spent her fellowship in two different departments (Cell Biology-group of Dr. B. Scholte and Pharmacology-group of Prof. P. R. Saxena and Dr. H. Sharma, Erasmus University) and subsequently became a Ph.D. candidate in 1997 in Cell Biology Department of Erasmus University. Since late of 2001, she has been working as a Post Doctoral Fellow in the Pathology Department, ErasmusMC. Her research in Holland resulted in a number of publications, which are listed as follows.

Publications and Award

Zhongling Cai, Marella de Bruijn, **Xiaoqian Ma**, Bjorn Dortland, Tanya Luteijn, James R. Downing and Elaine Dzierzak. Haploinsufficiency of AML-1 affects the temporal and spatial generation of hematopoietic stem cells in the mouse embryo. Immunity 2000; 13: 423-431

Xiaoqian Ma, Kan-Wing Ling and Elaine Dzierzak. Cloning of the *Ly-6A* gene locus and identification of a 3' distal fragment responsible for high-level γ-interferon-induced expression in vitro. British Journal of Haematology 2001; 114: 724-730

Xiaoqian Ma, Marella de Bruijn, Catherine Robin, John Kong-A-San, Corne Snoijs and Elaine Dzierzak. Expression of the *Ly-6A (Sca-1) lacZ* transgene in mouse hematopoietic stem cells and embryos. British Journal of Haematology 2002; 116: 401-408

Marella de Bruijn*, **Xiaoqian Ma***, Catherine Robin*, Katrin Ottersbach, Maria-Jose Sanchez and Elaine Dzierzak. Hematopoietic stem cells localize to the endothelial cell layer in the midgestation mouse aorta. Immunity 2002; 16: 1-20

***These authors contributed equally to this work**

Xiaoqian Ma, Catherine Robin, Katrin Ottersnach, Elaine Dzierzak. The *Ly-6A (Sca-1) GFP* transgene is expressed in all adult mouse hematopoietic stem cells. Stem Cell 2002; 20: 514-521

Nicolas Coardon, Annette Schuh, Iman Hajar, **Xiaoqian Ma**, Helene Jouault, Elaine Dzierzak, Paul-Henri Romeo and Leila Maouche-Chretien. Ectopic expression of TAL-1 protein in *Ly-6E.1-htal-1* transgenic mice induces defects in B- and T-lymphoid differentiation. Blood 2002, 100 (2): 491-500

Catherine Robin, Katrin Ottersbach, Marella de Bruijn, **Xiaoqian Ma**, Karin van der Horn, Elaine Dzierzak. Development origins of Hematopoietic Stem Cells. Oncology Research 2002, 13: 315-321

Pankaj Bhalla, Haris S. Sharma, **Xiaoqian Ma**, Thierry Wurch, Petrus J. Pauwels and Pramod R. Saxena. Molecular cloning, pharmacological properties and tissue distribution of the porcine 5-$HT_{1\beta}$ receptor. British Journal of Pharmacology 2001; 133: 891-901

Xiaoqian Ma, Hetty van der Korput, Angelique Ziel-Van der made, Marcel Vermeij, Petra van Duijn, Paul Krimpenfort, Anton Berns, Theo van der Kwast, and Jan Trapman. Early onset of hyperplasia and intraepithelial neoplastic lesions in prostate-specific targeted *Pten* knockout mice. 2002 American Association for Cancer Research-AFLAC Travel Award.

Acknowledgments

First, I would like to thank Prof. Frank Grosveld and Prof. Elaine Dzierzak, for giving me the opportunity as a Ph.D. candidate. I enjoyed much from the academic freedom that you gave the people in the lab and the department.

I would like to express my gratitude to Dr. Arthur van der Kamp for your highly efficient paperwork for my specific case.

I thank all the members of my previous lab, Marian, Robert, Kam-Wing, Tanya, Bjorn, Marella, Catherine, Katrin, Karin, Claudia, Kirsty, Corne, special thanks to Corne Snoijs for excellent sorting skill, to Marian Peeters and Robbert J. Rottier for Dutch translating and checking the summary of this book, and to Catherine Robin for helping proof reading of some parts of this book.

I would also thank all the members of my present lab, Angelique, Eric Jan, Hetty, Marcel, Petra, Hanneke, Remko, Pascal, Martin, Karin, Binha, Eddy for all your cooperation and support during my staying. I would apologize for not making mention of everyone in my present department.

Special thanks to Professor Jan Trapman, my present supervisor, for giving me some room to finish my thesis and importantly, being open to discuss a lot on different aspects of doing science.

Thanks to all the people involved in the mouse work, especially Ton de Wit, John Kong-A-San, Danielle Zondervan. Thanks to photographers, people from kitchen, people in the ordering office, and Josselin de Jong for helping with microscope, Ton Verkerk and Sjozef van Baal for computer support.

Thanks to our secretaries Marike, Rita, Jasperina and Shannon for all the years of your help.

Thanks to Dubravka and Hanneke for being my paranimfen.

Thanks to Anja Veerman for going through the evaluation procedure and kind invitation of your nice birthday party.

I am greatly indebted to my former supervisors and colleagues in China, Prof. Jianming Zhang, Prof. Jianmin Xing, Prof. Jiwu Miao, Prof. Ning Yang and Hanyi Liu, for not only leading me to the medical field and scientific research but also your encouragement, support and understanding during my training and working in your departments.

I wish to express my sincere thanks to Prof. Huanming Yang who gave me the chance to learn some molecular techniques at Peking Union Medical University. Furthermore, you are the person whom I can always open to discuss all kinds of difficulties and problems.

I would like to give my thanks to my father, Junru Ma, and mother Xiu Xu, for your continual support, encouragement and love. I greatly appreciate my parents in

law, Houxuan Luo and Huaijin Yu for looking after my daughter when I started in Holland. It would not be possible to realize this book without your help.

Finally, I give special thanks to my husband, Chongde Luo, for your deep understanding, profound encouragement, and unlimited support in all aspects.

Stellingen

Behorende bij het proefschrift

***Ly-6A/E* (Sca-1) gene regulatory elements in hematopoietic stem cells in mouse**

---

1. The Ly-6A/E (Sca-1) antigen expression profiles vary between mouse strains. (Spangrude & Brooks, 1993, Blood 82, 3327-3332)

2. The *Ly-6A* and *Ly-6E* alleles are 98% homologous in the region containing the +8.7 and +8.9 HSS. (This thesis)

3. The 3' distal fragment of *Ly-6A/E (Sca-1)* gene is essential for high-level γ-interferon-induced expression *in vitro* and for tissue specific expression *in vivo* in hematopoietic system. (Sinclair et al., 1996, Blood 87, 2750-2761; Miles et al., 1997, Development 124, 537-547 and this thesis)

4. The 14 kb *Ly-6A* expression cassette does not contain all cis-acting control regions. (This thesis)

5. The unsuccessful enrichment of embryonic HSCs from the AGM region and the localization of embryonic HSCs in the AGM region of Ly-6A/lacZ transgenic mice might result from the properties of the lacZ marker itself. (This thesis)

6. Adult hematopoietic stem cells localize to the endothelial cell layer in the midgestation mouse aorta. (This thesis)

7. The first adult hematopoietic stem cells arise in the aorta-gonads-mesonephros (AGM) region. (Muller et al., 1994, Immunity 1, 291-301; Medvinsky & Dzierzak, 1996, Cell 86, 897-906)

8. Hematopoietic stem cell emergence occurs in close association with the major vasculature of the embryo. (Garcia-Porrero et al., 1995, Anat. Embryol. Berl. 192, 425-435; Shalaby et al., 1997, Cell 89, 981-990; Tavian et al., 1999, Development 126, 793-803; Wood et al., 1997, Blood 90, 2300-2311; de Bruijn et al., 2000, EMBO J 19, 2465-74)

9. The efficiency of the concurrent Cre-mediated inter-chromosomal recombination can be different in a same cell subset. (Vooijs et al., 2001, EMBO reports 2, 292-297)

10. Tumour suppressor PTEN regulates beta-catenin gene expression. (Sujata Persad et al., 2001, The Journal of Cell Biology 153, 1161-1174)

11. Oh, it is excellent to have a giant's strength, but it is tyrannous to use it like a giant. (Shakespeare)