1998

Cis-acting sequences and trans-acting factors regulating human globin gene expression.

Chiao-Wang Joe Sun
University of Alabama at Birmingham

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

CIS-ACTING SEQUENCES AND *TRANS*-ACTING FACTORS REGULATING HUMAN GLOBIN GENE EXPRESSION

by

CHIAO-WANG J. SUN

A DISSERTATION

Submitted to the graduate faculty of The University of Alabama at Birmingham
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

BIRMINGHAM, ALABAMA

1998

ABSTRACT OF DISSERTATION

GRADUATE SCHOOL, UNIVERSITY OF ALABAMA AT BIRMINGHAM

Degree Ph.D. Program Biochemistry

Name of Candidate Chiao-Wang J. Sun

Committee Chair Tim M. Townes

Title *Cis*-Acting Sequences and *Trans*-Acting Factors Regulating Human Globin Gene Expression

The human β-globin locus control region (LCR) is an essential element for β-like globin gene expression. The LCR not only acts as a powerful enhancer but also functions in organizing the entire β-globin locus into an open chromosomal. Hispanic β-thalassemia patients who have a deletion of DNase I hypersensitive sites (HS) 2-5 have a severe anemia due to lack of expression of the β-globin gene on the affected chromosome.

In work prior to this thesis, I cloned a transcription factor, LCR-Factor 1 (LCR-F1), that binds to the LCR. LCR-F1 was shown to stimulate transcription of globin genes specifically in erythroid cells (K562) in transient transfection assays; however, its *in vivo* function was unknown.

Gene knockout experiments in mice are a key to determining the biological function of genes. I have shown that Lcrf1 is necessary for mesoderm formation during mouse development (1). In order to demonstrate that embryonic lethality is due to the absence of LCR-F1, I produced transgenic mice that expressed different levels of human LCR-F1. Human Lcrf1 bacterial artificial chromosome (BAC) transgenes completely rescued Lcrf1 -/- animals. However, rescue with a human Lcrf1 cDNA was not complete because the murine stem cell virus-long terminal repeat (MSCV-LTR) was silenced in transgenic mice. To understand the embryonic lethality in Lcrf1 null animals, I determined the number of early hematopoietic colony-forming units-granulocyte, erythroid, macrophage, and megakaryocyte (CFU-GEMM) in the Lcrf1 -/-Tg fetal livers. This number was low compared with wild-type littermates. All the results were confirmed by FACS analysis of

cells from the same fetal liver. The data suggested that LCR-F1 may be involved in stem cell proliferation or differentiation.

Finally, I studied the *cis*-acting sequences in the human γ-globin promoter that are necessary for γ- to β-globin switching and found that the CACCC element at nucleotide position −140 was essential for γ- to β-globin switching during mouse development.

Taken together, the results of my studies demonstrate the following:

1. LCR-F1 regulates soluble signaling molecules that are required for mesoderm formation and for the efficient production of early hematopoietic progenitors (CFU-GEMM).

2. The CACCC box in human γ-globin gene promoter is essential for correct human γ- to β-globin gene switching *in vivo*. The CACCC box apparently binds a transcription factor that provides the γ-globin gene with a competitive advantage for interacting with the LCR in the fetal stage of development.

DEDICATION

I would like to dedicate this accomplishment to my family, my brother, Jiaw-Sen (Sunny) Suen and his wife, Ching-Hui Lai, and my ex-wife, Yu Zhou, for their support and encouragement. Especially for my dear mother, Nan-Hwa Lo, you taught me to be strong and to be what I am. My heart was broken when you passed away on January 02, 1997, and I will never recover again because part of me has gone with you. I will always follow your instructions and principles to be what I am. I love you so much and I want you to be proud of me. If life is a cycle, please let me be your son again.

ACKNOWLEDGMENTS

I thank my mentor, Dr. Tim Townes, for his great support, encouragement, and guidance in helping me to become a trained researcher in the field of molecular biology. I learned that a good scientist needs enthusiasm and self-motivation as well as knowledge.

I thank the members of my committee, Drs. Susan Ruppert, Peter Detloff, Casey Marrow, and Jeff Engler, for their continuous support through the difficult periods so that I could complete my training.

I thank all the people in Townes' laboratory, Tom Ryan, Kevin Pawlik, Vladimir Divoky, Zhiyong Liu, Evans Bailey, Wenyong Chen, Dominic Ciavatta, Susan Farmer, Clark Kelley, Howard Masuoka, Kumar Pandya, Sandra Pawlik, Jinxiang Ren, Chris Smith, Kelvin Whisenhunt, and Dakin Williams, for their help and friendship.

I extend my thanks to former Townes' laboratory members, John Caterina, David Donze, Jill Lahti, Steve McCune, Janine Askins, David Jones, Sandra Lo, Nan Martin, and Kathy Whitfield, for their help and friendship.

I thank all my Chinese friends in Birmingham for their friendship and kindness. I will always remember all the good times that we shared together.

Finally, I list some of phrases that are used on a daily basis in the Townes' laboratory:

1. He is a good citizen in the lab.
2. Has anyone seen Howard?
3. Nuts!
4. It's not going to work.
5. You got a "sec"?
6. Sorry, I just lost my concentration, can you say that again?

7. Let's assume everything works perfectly.
8. That is a great result!
9. We're golden!
10. You can do the experiment tonight!
11. You're kidding!
12. We just ought to have it!
13. Fantastic!
14. That's incredible!
15. It's interesting!

TABLE OF CONTENTS

Page

ABSTRACT ii

DEDICATION iv

ACKNOWLEDGMENTS v

LIST OF TABLES viii

LIST OF FIGURES ix

LIST OF ABBREVIATIONS xi

INTRODUCTION 1

Hemoglobin Gene Regulation 1
The Human Globin Locus 1
The Mouse Globin Locus 5
Transgenic Mouse Models 5
Locus Control Region (LCR) 6
Locus Control Region-Factor 1 (LCR-F1) 11
Development of Mouse Mesoderm Formation 15
Molecular and Cellular Biology of Hematopoiesis 18
Globin Gene Regulatory Proteins 22
Model for Globin Gene Switching 23
Studies of Molecular Mechanisms That Control Human γ-Globin Gene Switching During Development 26

EXPERIMENTAL METHODS 28

Mutagenesis of the Lcrf1 Gene in Mouse ES Cells 28
Generation of Chimeric Mice and Germ Line Transmission of the Lcrf1 Mutant Allele 29
Preparation of Slides and Histological Analysis of Embryos 30
In Situ Hybridization Analysis 30
RT-PCR Analysis 30
Glucose Phosphate Isomerase (GPI) Analysis 31
Embryoid Body *In Vitro* Cultures and Analysis 31
Murine Retrovirus Rescue Constructs 32
BAC DNA Purification 32
Generation of MSCV and BAC Transgenic Mice 32
Detection of Transgenes by PCR Analysis 33
Transgenic Examination by Southern Blot Analysis 33
Primers for Mega-Primer Mutations 34

TABLE OF CONTENTS (Continued)

Page

RESULTS 35

Hematopoietic Defects in Embryoid Body (EB) Derived from Lcrf1 -/- ES Cells 35
Generation of an Lcrf1 Mutant Allele in the Mouse Germ Line 37
Effect of the Lcrf1 Null Mutation Is Not Cell-Autonomous 38
Lcrf1 Homozygous Mutant Is Embryonic Lethal 43
No Mesoderm Formation Is Morphologically Detectable in Lcrf1 Null Embryos 44
Brachyury (T) Expression Is Absent in Homozygous Mutant Embryos 46
Human Lcrf1 Genomic BAC Clone Rescue 46
MSCV/Lcrf1 cDNA Partial Rescue 48
Erythroid Progenitor Assays 54
Human Globin Gene Switching 54

DISCUSSION 62

Lcrf1 Mutation Inhibits Mesoderm Formation 62
The Role of LCR-F1 in Globin Gene Regulation 64
Human Lcrf1 Genomic BAC and MSCV Rescue 64
Human Globin Gene Switching 65
Future Experiments 67

LIST OF REFERENCES 72

LIST OF TABLES

Table		Page
1	Geneotypes of offspring from +/- X +/- matting	43
2	Empty decidua from +/- X +/- matting	43

LIST OF FIGURES

Figure		Page
1	Human and murine α-globin and β-globin loci	3
2	Human and murine globin chains synthesized during development	4
3	Flow diagram of the steps involved in transgenic mice made by DNA injection	7
4	Summary of mutant constructs HS 2 (Kpn I-Puv II) β transgene expression	10
5	Schematic of LCR-F1 functional domains	13
6	LCR-F1 transactivates a heterologous promoter specifically in erythroid cells	14
7	Schematic of a simplified mouse mesoderm formation	16
8	Transcription factors involved in regulation of hematopoiesis	19
9	Schematic of a simplified model for globin gene switching	25
10	Analysis of embryoid bodies (EBs) produced by *in vitro* differentiation of Lcrf1 +/- and -/- ES cells	36
11	Targeted disruption of the LCR-F1 gene	39
12	GPI analysis of chimeras from injection of Lcrf1 -/- ES cells	42
13	LCR-F1 +/- and -/- 7.5-day-old embryos	45
14	LCR-F1 in situ hybridization of 6.5-day-old embryos	45
15	RT-PCR analysis of ES cells, 3.5-day-old blastocysts, and 8.0-day-old embryoid bodies (EBs)	47
16	LCR-F1 +/- and LCR-F1 -/- embryos were generated by in situ hybridization with an LCR-F1 probe	47
17	Comparison of the conserved peptide sequences between human and muouse Lcrf1 genes	49
18	Southern blot analysis of DNA from transgenic mice with mLcrf1 +/- and hLcrf1 mice	50
19	Partial rescue of LCR-F1 -/- embryos *in vivo*	52

LIST OF FIGURES (Continued)

Figure		Page
20	Fetal liver sections and blood smears from 16 d mLcrf1 +/- and mLcrf1 +/-, hLcrf1 Tg	53
21	LCR $\gamma\beta$ constructs used to define γ-globin gene promoter sequences that are involved in globin gene switching	56
22	Schematic diagram of the human Aγ-globin gene promoter and transcription factor binding sites	56
23	Primer extension analysis of RNA from 10.5-day-old yolk sac (YS), 15.5-day-old fetal liver (FL) and adult blood (BL) of LCR $\gamma\beta$ transgenic mice	57
24	Primer extension analysis of RNA from 10.5-day-old yolk sac (YS), 15.5-day-old fetal liver (FL) and adult blood (BL) of LCR $\gamma\beta$ transgenic mice containing mutations of transcription factor binding sites	60
25	Schematic of gene regulation in the Tet-Off and Tet-On system	68
26	Yolk sac λgt11 library clone probed with wild-type and mutant γ-globin CACCC box oligos	70

LIST OF ABBREVIATIONS

AGM	aorta-gonad-mesonephros
ALL	acute lymphoblastic leukemia
ALK3	Bmpr1 A receptor
AP-1	activator protein-1
BAC	bacteria artificial chromosome
BFU-E	burst-forming units-erythroid
BHLH	basic helix-loop-helix
BMP4	bone morphogenetic protein 4
bp	base pair
bZIP	basic leucine zipper
CFU-E	colony-forming units- erythroid
CFU-GM	colony-forming units-granulocyte, erythroid
CFU-GEMM	colony-forming units-granulocyte, erythroid, macrophage, megakaryocyte
CNC	Cap'N'Collar
dpc	days post coitus
EBs	embryoid bodies
EE	embryonic ectoderm
eed	embryonic ectoderm development gene
EKLF	erythroid Kruppel-like factor
ES	embryonic stem
FAK	focal adhesion kinase
FGF4	fibroblast growth factor 4 \
FGFR-1	fibroblast growth factor receptor-1

F6P	fructose-6-phosphate
GPI	glucose phosphate isomerase
G6P-DH	glucose-6-phosphate dehydrogenase
HNF3	hepatocyte nuclear factor 3
HNF4	hepatocyte nuclear factor 4
HS	hypersensitive
kb	kilo base
LCR	locus control region
LCR-F1	locus control region-factor 1
LTR	long terminal repeat
M	mesoderm
MSCV	murine stem cell virus
NF-E2	nuclear factor-erythroid 2
neo	neomycine
NADPH	β-nicotinamide adenine dinucleotide phosphate
NRF2	NF-E2-related factor 2
PCR	polymerase chain reaction
pGK	phosphoglycerkinase
Pc-G	polycomb group
PMS	phenazine methosulfate
rtTA	reverse tet-responsive transactivator
RT-PCR	reverse transcription-polymeras chain reaction
SCL	stem cell leukemia
tk	thymidine kinase
TGF-β-like factor	transforming growth factor-β like factor
TRE	tetracycline responsive element

VE	visceral endoderm
Ve	velvet coat

INTRODUCTION

Hemoglobin Gene Regulation

Studies of human hemoglobin gene expression provide an excellent model for investigating the mechanisms by which genes are regulated during development. Expression of the globin genes is regulated in three important ways:

1. Globin genes are expressed tissue-specifically, their expression confined to erythroid cells.

2. Globin genes are expressed during defined developmental stages.

3. Globin genes are expressed at high levels.

The Human Globin Locus

Hemoglobin was one of the first proteins to be purified to the point where its molecular weight and amino acid composition could be established accurately. The hemoglobin molecule has a total of four chains: two identical α chains with 141 amino acids and molecular weight 15,126 daltons and two identical β chains with 146 amino acids and molecular weight 15,867 daltons. The heme adds another 616 daltons to each chain, so the human hemoglobin tetramer is 64,450 daltons (2).

The role of hemoglobin is to bind molecular oxygen to its heme irons at the lungs and to deliver it to the tissues. The properties of hemoglobin in an oxygen-transporting system can be summarized as the following:

1. The carrier should have a high affinity for oxygen in the presence of a plentiful supply at the lungs and a lowered affinity for oxygen in the oxygen-poorer environment of the muscles.

2. The carrier should be able to transport carbon dioxide back to the lungs where it can be ejected as a waste product.

3. The carrier should release its oxygen more readily to working muscles, with its buildup of lactic acid and of carbonic acid from CO_2, than to resting muscle, which needs less oxygen.

The α and β families of hemoglobin genes are found on separate chromosomes in man, and the development of cloning methods has made it possible to map these regions of DNA to locate and to determine the base sequences of each individual genes. The human α-like globin locus is on chromosome 16, and the β-like globin locus is on chromosome 11. The α-like globin cluster contains three functional genes ($\zeta 2$, $\alpha 2$, and $\alpha 1$) and three pseudo genes in a 35-kb region on chromosome 16 (3). The β-like globin cluster contains five functional genes and one pseudo gene (ε, Gγ, Aγ, $\psi\beta$, δ, and β) which approximately span 65 kb on chromosome 11 (4, 5) (Fig. 1).

Human hemoglobin expression is confined to erythroid cells. Fig. 2 shows α– and β-globin polypeptides and their expression during different developmental stages; globin genes are tissue specific and developmental stage specific. At the β-like locus, ε gene expression begins at 3 weeks gestation in the yolk sac blood islets. At 8 weeks of gestation, the major site for erythropoiesis shifts from yolk sac to fetal liver. Expression of the embryonic ε gene is reduced while the expression of fetal Gγ and Aγ genes are increased during this developmental change. Globin gene expression switches again prior to birth, when the major site of erythropoiesis shifts from fetal liver to bone marrow and fetal γ expression is replaced with adult β and δ gene expression (6).

The α-globin genes also switch during development. The embryonic ζ gene is expressed at a high level at 2 weeks of gestation and then decreases. α-Globin is also expressed during the first 2 weeks, but expression increases until about 13 weeks and then remains constant throughout life (7, 8). A graphical representation of this information is shown in Fig. 2*A*. Each globin gene is expressed at a high level in erythroid cells at each

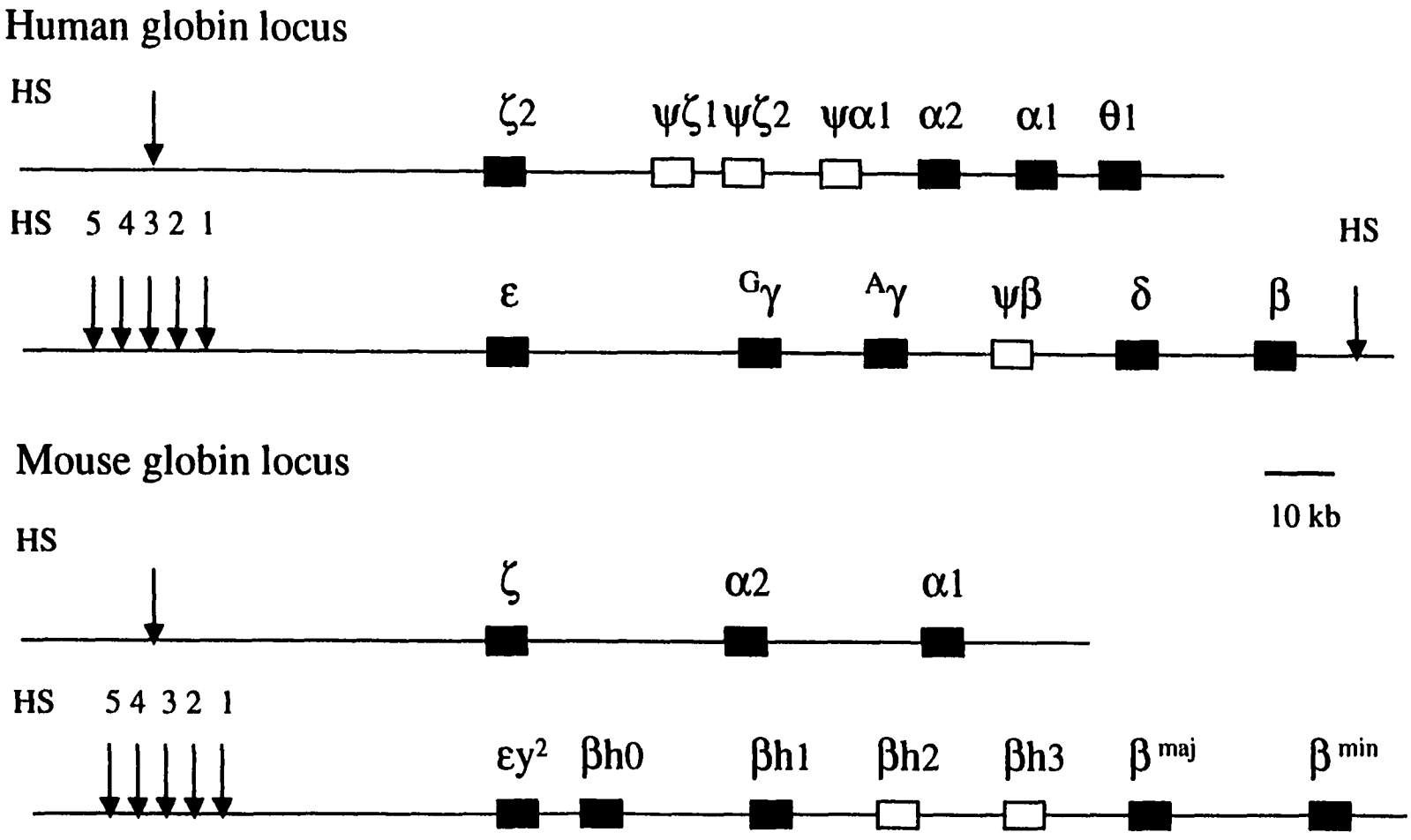


FIG. 1. **Human and murine α-globin and β-globin loci.** Individual genes of the human and mouse α- and β-globin Loci arranged 5' to 3' in the order in which they are expressed during development. Erythroid-specific hypersensitive site (HS) sites located 6-22 kb upstream of ε-globin and 22 kb downstream of β-globin are marked by arrows. Functional genes are representedby black boxes, while the pseudogenes are indicated in open boxes (Adapted from Ryan's dissertation [9]).

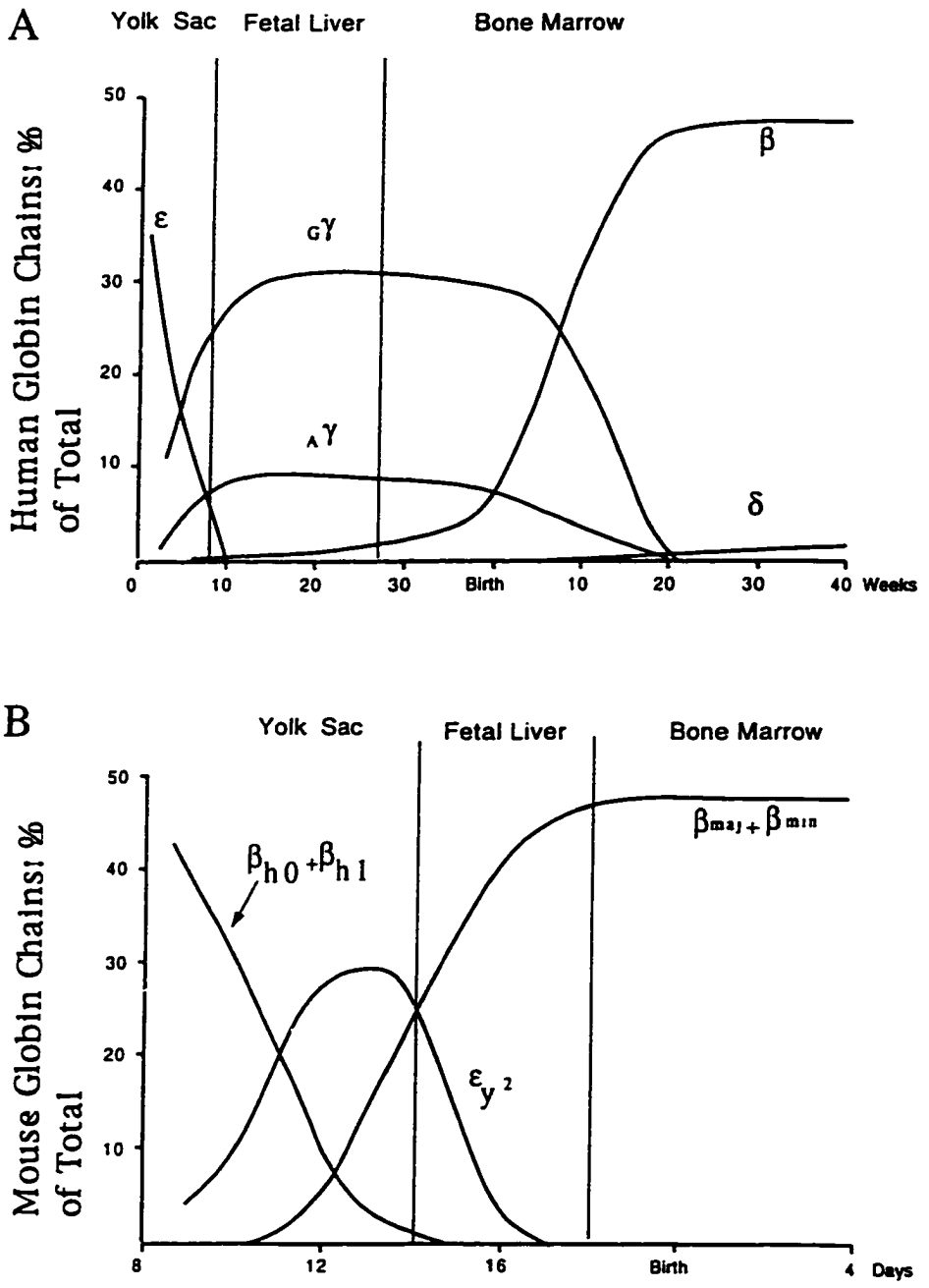


FIG. 2. **Human and murine globin chains synthesized during development.** The individual β-globin protein chains, expressed as a percentage of the total globin chains, for human *A* and mouse *B* are plotted against time. Two well-defined β-globin gene switches are evident in man, separating the embryonic, fetal, and adult stages of erythroid development. The embryonic to adult β-globin gene switch in the mouse occurs much earlier in the gestational life and is marked by quick rise and fall of expression of the εy^2-globin gene, the murine homologue to human fetal genes (Adapted from Bunn and Forget [6]; Ryan's dissertation [9] and Caterina's dissertation [10]).

developmental stage. Ninety percent of the protein in mature erythrocytes is hemoglobin. The study of globin switching provides an excellent model to understand gene regulation and expression during development.

The Mouse Globin Locus

The mouse β-like globin locus is composed of five functional genes (εy^2, βh0, βh1, β^{maj}, and β^{min}) and two pseudo genes (βh2 and βh3). Together the locus comprises a 65-kb region of DNA (11, 12). Unlike human globin gene switching, there is no clear switching between embryonic and fetal globin genes in the mouse β–globin locus. The βh0 and βh1 genes express early and switch off between days 12 and 14 of gestation. The mouse εy^2 gene expresses late and decreases between days 14 and 16 of gestation. The mouse adult globin genes, β^{maj} and β^{min}, are turned on when the fetal liver becomes the major site of erythropoiesis. β^{maj} and β^{min} gene expression persists when erythropoiesis shifts to bone marrow at birth.

The α-globin locus in mice is similar to humans. Both species have one embryonic ζ gene and two adult α1 and α2 genes. (3, 13). The mouse ζ gene is expressed during early embryonic yolk sac development until approximately 3 weeks of gestation when the α1 and α2 genes are turned on and persist through adulthood (Fig. 2*B*).

Transgenic Mouse Models

The development of transgenesis in the mouse in the early 1980s provided an opportunity to manipulate the genetic make-up of a higher vertebrate. To understand human globin gene regulation, transgenic mice provide the most stringent system *in vivo* for defining *cis*-acting elements required for the regulated expression of genes. Because transgenes can be designed with pre-determined *cis*-acting protein and enhancer sequences, the role of these sequences in tissue and developmental regulation can be determined. The mouse *trans*-acting factors regulate the human transgenes as they would their endogenous

genes. For example, transgenic mice with a locus control region (LCR) $\gamma\beta$ switching construct expressed human γ- and β-globin genes in a correct tissue- and temporal-specific manner; the human γ-transgenes are expressed at the highest level in 10.5-day-old yolk sac and the human β-transgenes are expressed at the highest level in 15.5-day-old fetal liver.

One major objective of this study was to define the DNA sequences responsible for human fetal to adult globin switching and to clone the factors that bind to these sequences. A series of deletions and site specific mutations were introduced into the γ-promoter region of the LCR $\gamma\beta$ switching constructs in transgenic mice. Mutations that alter normal hemoglobin switching have been identified.

A flow chart for the production of transgenic mice to define *cis*-elements involved in γ-gene switching is shown in Fig. 3. A total of approximately 400 injected eggs were transferred for each DNA construct. Each founder has one or multi-integration sites and the transgene can be one or many copies in a head-to-tail direction or other different combinations. The transgenic mice produced from these transferred eggs were mated to normal animals and embryos were removed at 10.5 or 15.5 days of development. Transgenes were identified by PCR and total RNA was analyzed by primer extension for human globin mRNA.

Locus Control Region (LCR)

The chromatin encompassing transcriptionaly active gene exists in an altered or "open" conformation that makes it more accessible to the cell's transcriptional apparatus or to exogenously added nuclease such as DNase I. The chromatin around tissue-specific genes is "open" in tissues that express the gene but is relatively insensitive to DNase I digestion in non-expressed cell types.

Tuan *et al.* (14) and Fraster *et al.* (15) mapped DNase I hypersensitive sites (HS) far upstream of the human β-globin gene on chromosome 11. Five DNase I (5' HS) 1-5 are located 50-65 kb upstream of the human β-globin gene, and one site (3' HS 1) is

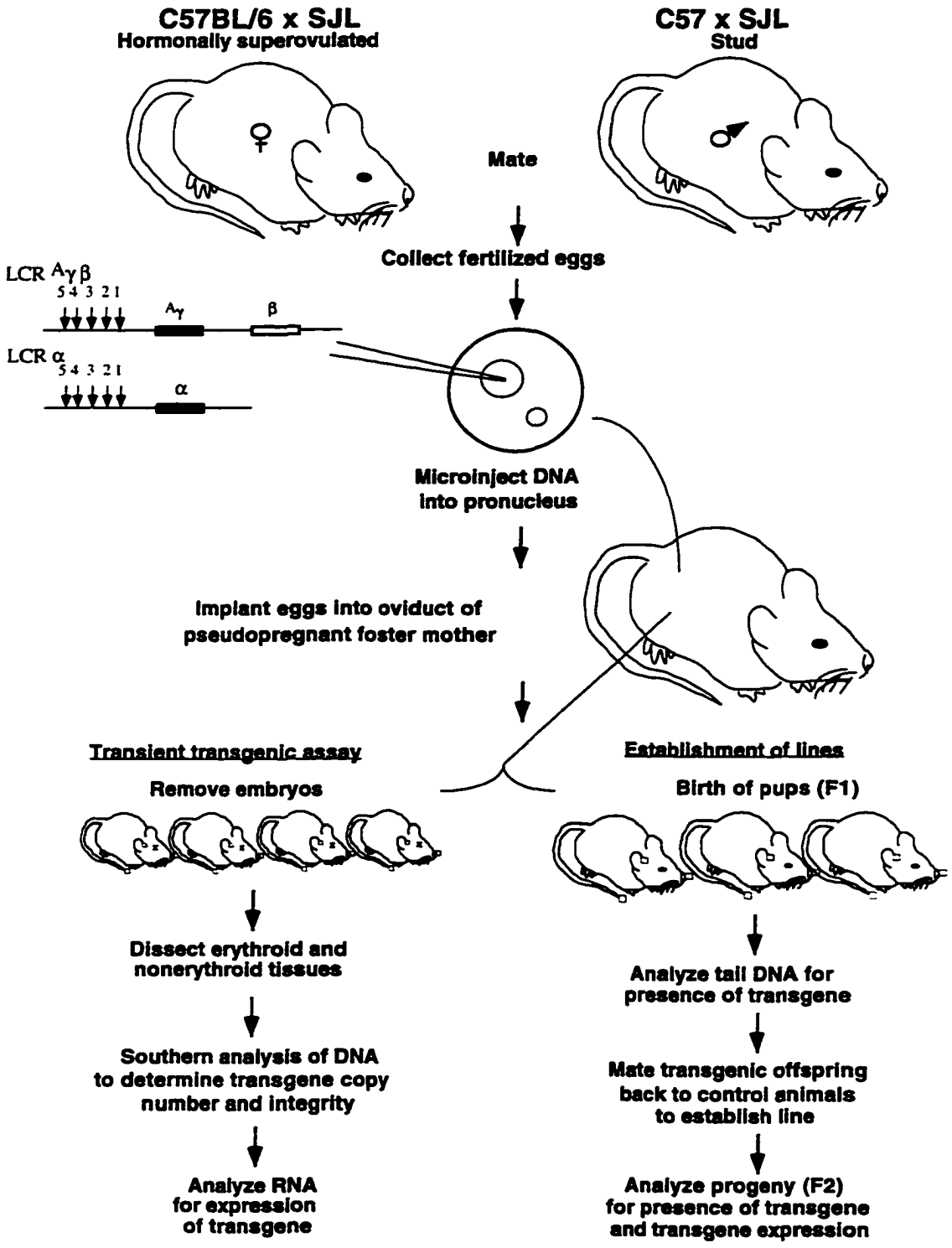


FIG. 3. **Flow diagram of the steps involved in transgenic mice made by DNA injection** (Adapted from Ryan's dissertation [9] and Caterina's dissertation [10]).

located 19 kb downstream of the human β-globin gene (Fig. 1). The 5' HS 1-5 are collectively referred to as the β-globin LCR. These hypersensitive sites are erythroid-specific. Unlike the DNase I hypersensitive sites that exist in the promoter proximal regions of globin genes only when they are being expressed at a particular developmental stage, these hypersensitive sites are developmentally stable; that is, they are present at all developmental stages.

Naturally occurring deletions that remove these DNase I hypersensitive sites result in severe β-thalassemia in patients even though these individuals retain an intact copy of the β-globin gene. Dutch and English γδβ-thalassemia patients have deletions upstream of β-globin that remove all of the 5' HS (16, 17). Although these deletions leave an intact copy of the β-globin gene, no β-globin gene expression can be detected. In these patients, the intact β-globin gene has lost its increased sensitivity to DNase I compared to the β-globin genes in normal erythroid cells. This suggests that the β-like globin chromosomal domain, which is usually open in erythroid cells, is now closed as a result of the deletion. The role of HS sequences is clearer in a Hispanic patient who had a deletion of only 5' HS sites 2-5 of the β-locus. The chromosome containing the deletion also contained a linked β^s-globin gene. No β^s- globin transcriptional activity was observed in this patient. Because all of the globin genes and a single 5' HS1 site were left intact in the mutant locus, this result demonstrates *in vivo* that 5' HS 2-5 are critical for globin gene expression.

Functional properties of the LCR were shown even more clearly when HS 1-5 (18, 19) or a single human β-globin gene linked to the LCR was tested in transgenic mice (15, 19-21). The constructs containing LCR β or HS 2 β were injected into fertilized eggs and transgene copy numbers and expression were analyzed by Southern blots in the resulting transgenic mice. All the animals with intact copies of the transgene expressed high levels of β-globin mRNA. Human β-globin transgenes without LCR sequences are not expressed during embryonic development in mice, but are turned on in the fetal liver. Human γ-globin transgenes linked to LCR sequences are expressed in the mouse yolk sac and

persist through adulthood. Therefore, developmental specificity is lost in these transgenes. However, when a β-globin gene was placed 3' to a γ-globin gene under control of a single LCR HS, the correct switching pattern was restored (22). The data suggests that the LCR acts as a powerful enhancer for globin transgene expression in erythroid cells; expression of the linked β-globin gene was enhanced 300-fold compared with a β-globin gene without the LCR. Additionally, the data suggest the LCR has domain-opening activity. β-Globin transgenes are usually expressed in only a fraction of the transgenic animals. However, when β-globin genes are linked to the LCR sequences, virtually all of the transgenic mice express β-globin. Therefore, the LCR is able to overcome site-of-integration position effects.

The mechanism(s) by which LCR sequences influence β-globin gene expression is unknown. Although the strongest effects are seen when all HS sites are linked together, individual sites can exert transcriptional effects and certain combinations of individual sites can also augment expression. The lack of β-globin gene expression from the Hispanic deletion chromosome indicates that 5' HS site 1 is not sufficient for LCR activity. The transgenic mice experiments showed that HS sites 2, 3, and 4 combined are sufficient for the full LCR function. These results demonstrated that the 5' HS sites 1 and 5 and 3' HS 1 are not essential for LCR activity. 5' HS sites 2, 3, and 4 each contain the LCR activity; however, when each individual HS is linked to a globin gene(s), the transgene expressed about 50 % of the full LCR's transcriptional enhancement capabilities (22). Careful comparison of the sequences between 5' HS 2, 3, and 4 show that they all contain the erythroid specific GATA-1 binding sites and the activator protein (AP)-1-like site(s) that may contribute to LCR activity. Ryan *et al.* (22) showed a 1.9-kb fragment containing the 5' HS 2 site enhances human β-globin gene expression 100-fold in transgenic mice and also conferred relative position independent expression (Fig. 4). Curtin *et al.* (20) and Fraser *et al.* (23) have obtained similar data. The 1.9-kb Kpn I-Pvu II 5' HS 2 was mapped carefully by DNase I footprinting with nuclear extracts from erythroid, K562 and

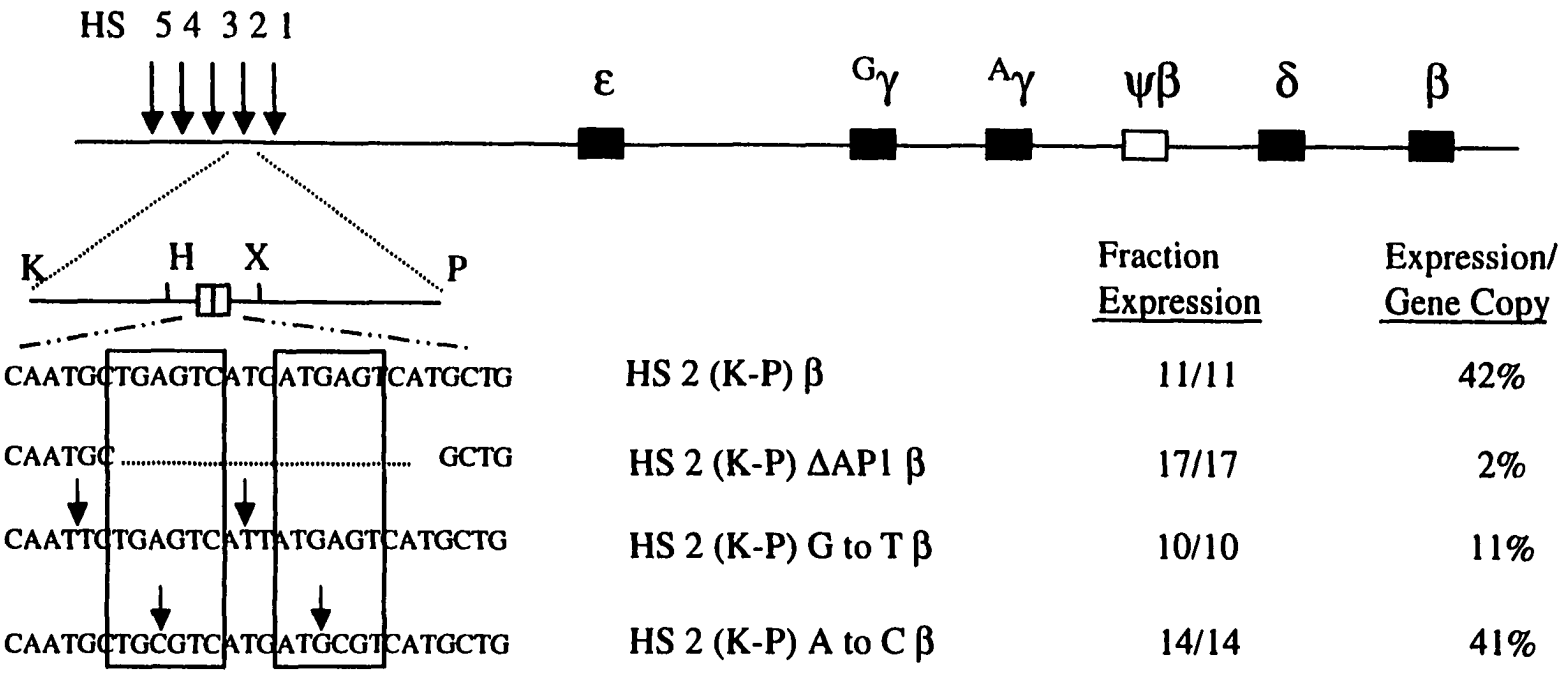


FIG. 4. **Summary of mutant hypersensitive sit HS 2 (Kpn I-Puv II) β transgene expression** (Adapted from Caterina *et al.* [30]).

MEL, and non-erythroid, HeLa and WEHI-3B, cells (24). A similar experiment was done by Talbot *et al.* (25). The enhancer core element of 5' HS 2 is made of two consensus AP-1 binding sites. Deletion of the duplicated AP-1 site results in dramatic reduction of 5' HS 2 enhancer activity (24, 26) (Fig. 4).

AP-1 is a heterodimeric complex of the two proto-oncogenes c-Jun and c-Fos. AP-1 has been implicated as a major *trans*-acting regulatory factor in a multitude of systems. Both c-Jun and c-Fos belong to the basic leucine zipper (bZIP) family of DNA binding proteins (27, 28). c-Jun and c-Fos form a heterodimer through the leucine zipper domains and bind an AP-1 site and activates transcription. Overexpression of c-Jun in K562 cells does not activate but rather inhibits globin gene expression (29). This result suggests that the Jun:Fos heterodimer does not bind to HS2 in erythroid cells. Although mutations in the AP-1 (for Jun:Fos heterodimer) consensus sequences in HS2, these mutations have no effect on HS 2 enhancer activity. In contrast, mutation at two bases upstream of the AP-1 consensus sequence reduced the 5' HS 2 enhancer activity four fold. This result suggests that a protein which binds to a site overlapping each AP-1 binding site is responsible for the enhancer activity directed through this element (30).

Three bZIP transcription factors, LCR-Factor 1 (LCR-F1) (28) or Nrf-1 (nuclear factor-erythroid 2 [NF-E2] related factor 1) (31), NF-E2 related factor 2 (Nrf2), (32) and nuclear factor-erythroid 2 (NF-E2) (33, 34), that bind to the duplicated AP-1 sites have been isolated. Of these factors, LCR-F1 has the highest transactivation activity in transient transfection assays in erythroid cells (30). Both NF-E2 and Nrf2 knockout mice have normal levels of β–globin expression. These results suggest these other factors may play a role in regulation of β–globin expression.

Locus Control Region-Factor 1 (LCR-F1)

Prior to the start of the work described in this thesis, I cloned LCR-F1 from a K562 λgt11 cDNA library (30). Five hundred thousand individual clones as a Southwestern blot

were screened with radioactively labeled multimerized copies of AP-1-like sites. I isolated and tested several clones for binding specificity with either wild-type AP-1-like or scrambled AP-1-like site DNA oligomers. One cDNA clone, which we designated LCR-F1, showed specific binding to the AP-1-like sites but did not bind to the mutant AP-1-like sites. After sequencing the cDNA, I named this transcription factor LCR-F1 (30). LCR-F1 contains a bZIP domain at the carboxyl terminal (Fig. 5). Interestingly, LCR-F1 contains a 30-amino acid domain that is 70 % homologous to a domain in the Drosophila melanogaster Cap N Collar (CNC) protein and 65 % homologous to a domain in the Caenorhabditis elegans SKN-1 protein. Both CNC and SKN-1 play important roles in development. CNC is expressed during Drosophila gastrulation for head formation. SKN-1 is involved in pharynx and body wall muscle formation in C. elegans. Therefore, LCR-F1 may play important roles in mouse development.

LCR-F1 is expressed in both erythroid and non-erythroid cells. LCR-F1 showed strong *trans*-activation activity for globin gene expression in transient transfection assays in K562 cells (30). LCR-F1 transactives an HS2/γ-globin reporter gene over 170-fold in erythroid cells (Fig. 6). Several CNC bZIP proteins that bind to the duplicated AP-1-like site have been isolated. They are LCR-F1 (30) or Nrf1 (31, 35), NF-E2 (33, 34), and Nrf2 (32). These proteins are members of a new family of regulatory factors that contain a 63-amino acid CNC domain overlapping the basic region. However, NF-E2 and NRF2 have no effect on regulation of globin gene development. Both Nfe2 and Nrf2 have been deleted in ES cells, and chimeras have been made from both mutant ES cells. The result showed that they are not necessary for β–globin expression *in vivo*, Nfe2 knockout mice have defects in platelet development, and Nrf2 knockout mice showed no phenotype. These results suggest that *in vitro* data do not always correlate with *in vivo* data due to the stringency in the biological system.

Caterina *et al.* (30) showed LCR-F1 *trans*-activates an HS 2/γ-globin reporter gene over 170-fold in transient transfection experiments, specifically in erythroid cells, and

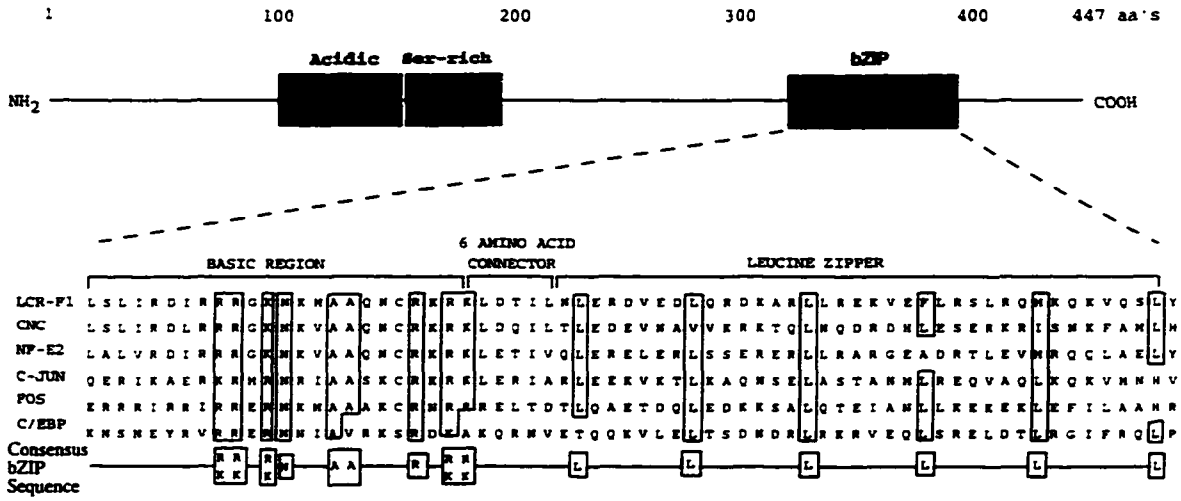


FIG. 5. **Schematic of LCR-F1 functional domains.** The *black boxes* represent acidic, serine-rich, and basic leucine zipper domains of the protein. Below the schematic, the bZIP domain of LCR-F1 is expanded and compared to 5 other bZIP proteins. *Open boxes* designate concensus bZIP amino acids. CNC is the Drosophila Cap N Collar protein. NF-E2 is another CNC family members. cJun, cFos, and C/EBP homologies in this region have been described previously (Adapted from Caterina *et al.* [30]).

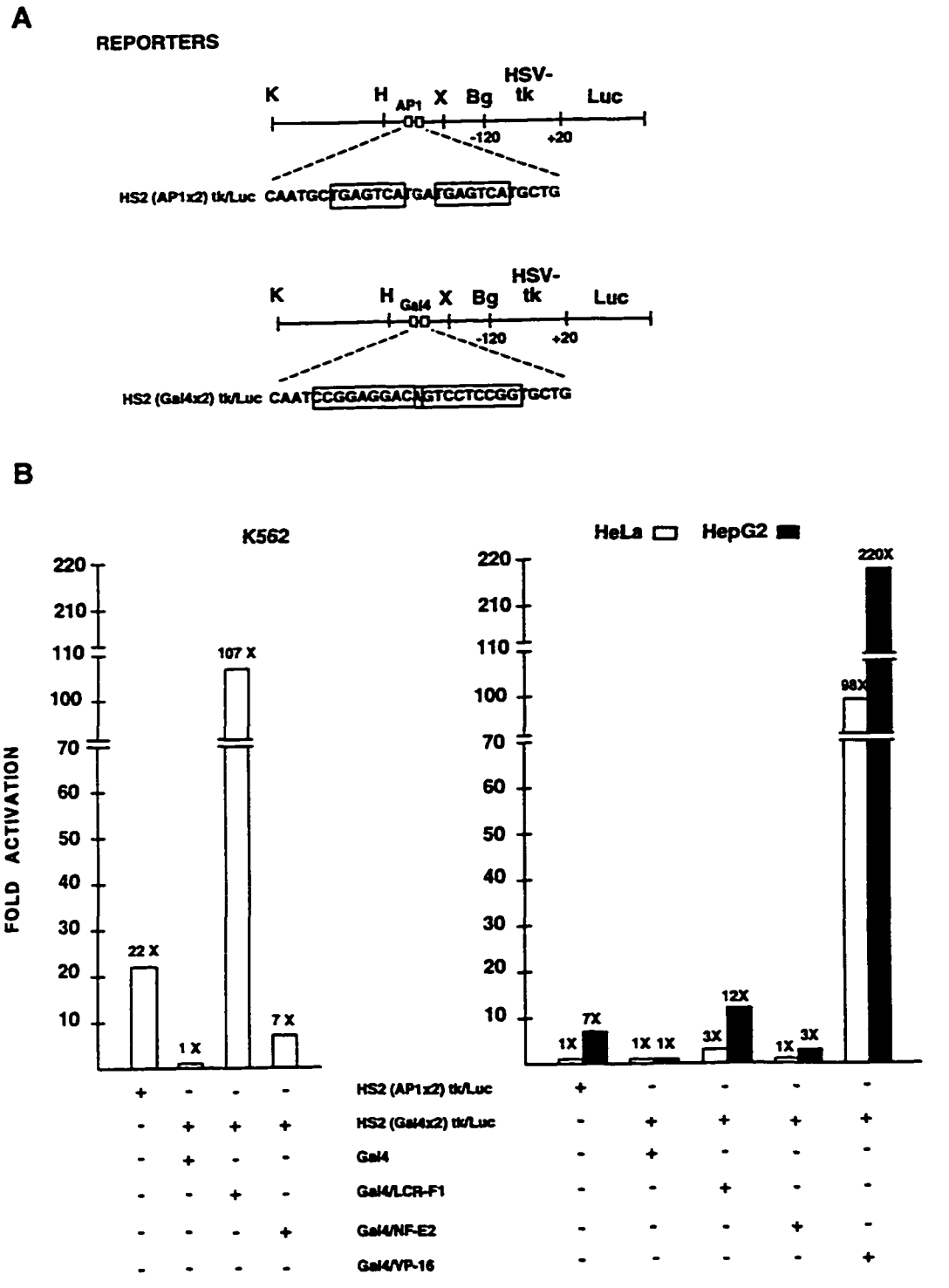


FIG. 6. **LCR-F1 transactivates a heterologous promoter specifically in erythroid cells.** *A*, reporter constructs used in transfections. *B*, results of transfection experiments. The constructs listed below the graph and an internal control plasmid was transfected into K562, Hela, and HepG2 cells, and luciferase activity was determined as shown (Adapted from Caterina *et al.* [30]).

suggested that LCR-F1 may be a critical factor involved in LCR-mediated human globin gene expression. However, this *in vitro* result does not prove that LCR-F1 does so *in vivo*. To explore the function of LCR-F1 in mouse, I knocked out the gene in ES cells, and chimeras were made from the mutant ES cells. The result showed that Lcrf1 knockout results in embryonic lethality in mice. In addition, I made transgenic mice containing human Lcrf1 gene that rescued the Lcrf1 -/- phenotype. The results will be described in more detail state where such as "in the first article."

Development of Mouse Mesoderm Formation

Gene knockout experiments in mice are a key to determining the biological function of genes. As described below, I have demonstrated that Lcrf1 null embryos die between 6.5 and 7.5 days of gestation. Fig. 7 illustrates the steps of early developmental at this stage. At the posterior of the 6.5 day post coitus (dpc) embryo at the border between embryonic epiblast (ectoderm), visceral endoderm and extraembryonic ectoderm, a small patch of columnar epithelial cells at the posterior rim of the cup-shaped epiblast delaminates and moves into the region between the ectoderm and endoderm at ~6.5-dpc (36). The structure resulting from this cellular migration is called the primitive streak, and the new layer of tissue is designated the mesoderm (Fig. 7). Mesoderm then becomes organized into populations with different fates: axial (prechordal plate and notochord), paraxial (somites), and lateral plate (splanchnopleure and somatopleure) mesoderm subsequent to the appearance of the node at the anterior of the primitive streak. A distinct subpopulation at the posterior of the streak gives rise to the extraembryonic mesoderm of the yolk sac and subsequently to primitive hematopoietic cells. The AGM (aorta-gonad-mesonephros) region derived later in development forms definitive hematopoietic cells (37-39).

Significant progress has been made over the past few years in the identification of genes required for mouse gastrulation and mesoderm formation (40, 41). However, indirect effects on epiblast (ectoderm) cell proliferation and survival have been difficult to

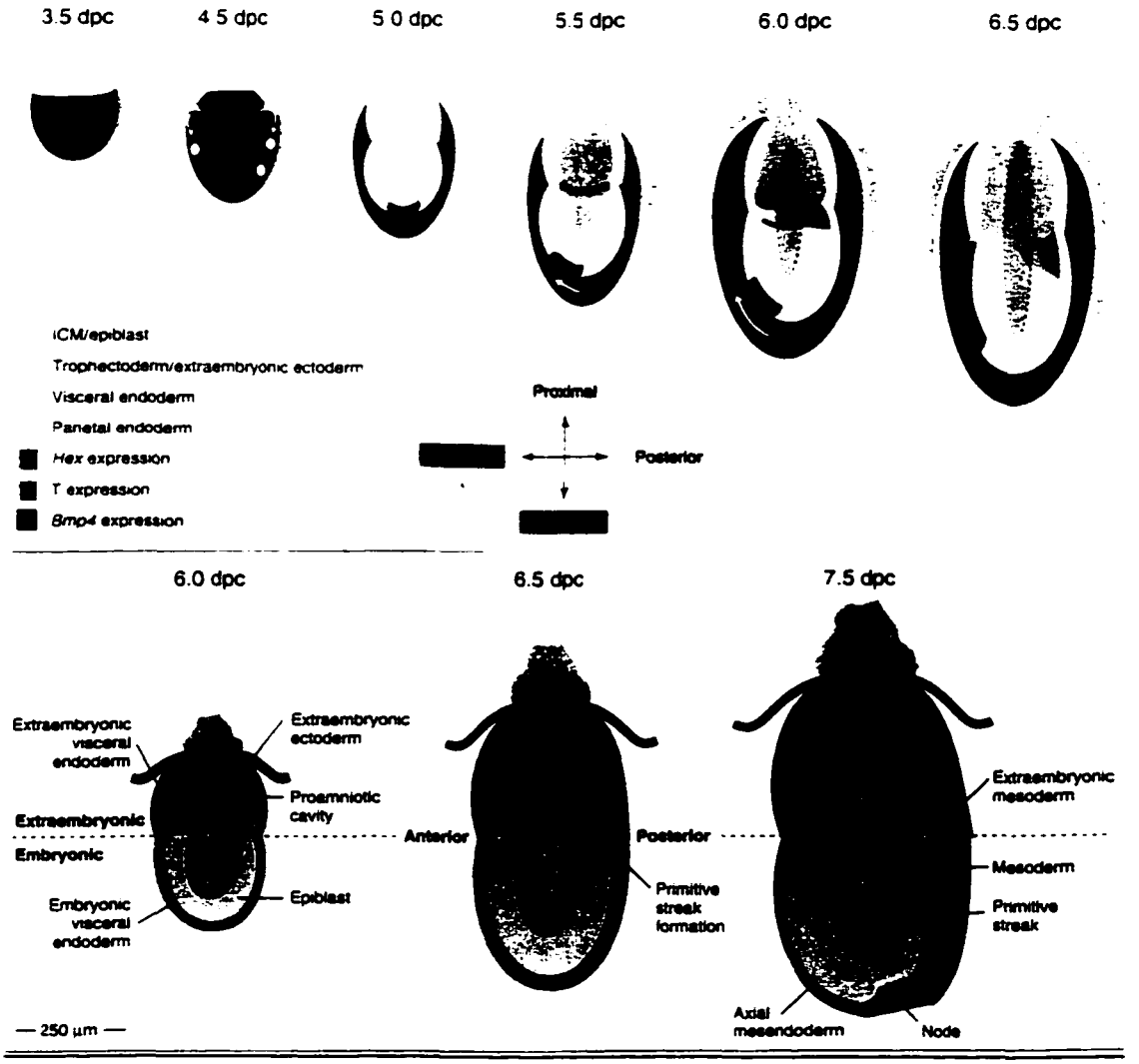


FIG. 7. **Schematic of a simplified mouse mesoderm formation.** Mouse gastrultion (3.5-7.5 days post coitum [dpc]) showing the egg cyclinder with the parietal yolk sac removed. The relationship of the epiblast to embryonic visceral endoderm, extraembryonic visceral endoderm, and extraembryonic ectoderm are shown. The primitive streak is first evident at 6.5-dpc, with the delamination of mesoderm at the embryonic-extraembryonic junction, and the site of streak formation defines the posterior margin of the embryo (Adapted from Beddington and Robertson [42]).

distinguish from direct effects on mesoderm induction and specification. Mutations in the gene encoding fibroblast growth factor 4 (FGF4) (43) and the fibroblast growth factor receptor-1 (FGFR-1) (44,45) appear to inhibit epiblast cell proliferation and subsequent organization of the egg cylinder. Deficient primitive ectoderm development also occurs for mice with homozygous mutation for the velvet coat (Ve) mutation (46). In contrast, mice with mutations in the β-catenine gene or at the Fu locus (47) die at gastrulation as a result of abnormal accumulation of ectoderm, and embryos with a null mutation in the fug1 gene have disorganized primitive ectoderm (48).

Normal endoderm function or the mesoderm pattern are inhibited by other mutations. Defect in embryonic endoderm morphogenesis and migration that affects mesoderm formation have been seen in mice with a β1 integrin gene deletion (49, 50). Mesoderm formation begins but is incomplete because of apparent effects on cell adhesion and motility in mice with mutations in the genes encoding protein tyrosine kinase focal adhesion kinase (FAK) (51, 52) or fibronectin (53).

Two transforming growth factor-β (TGF-β)-like factors (NODAL and BMP4) with mutation in coding region have been implicated in mesoderm induction. Normal primitive streak and node formation are inhibited when a null mutation exists in the nodal gene (54, 55). Consequently, mutant embryos die between 8.0- and 9.0-dpc with extensive overgrowth of the ectoderm and mesoderm cell patterning is abnormal. Mice with null mutations in Bmp4 (bone morphogenetic protein 4) (56) and ALK3 (Bmpr1A receptor) (57) also inhibit mesoderm formation. Most mouse embryos are arrested at 6.5-dpc development with these mutations. This phenotype suggests that BMPR1A tranduces signals important for mesoderm formation and that BMP4 is critical for mesoderm induction. However, Bmpr1A mutant embryos also affect epiblast cell proliferation and organization; therefore, it may have an indirect effect in mesoderm reduction.

Previously, no transcription factor has been reported to be involved directly with that mesoderm formation. A transcription factor, Brachyury (T), is expressed in primitive

mesoderm; however, in the T null mutation, early mesoderm formation and gastrulation occur normally (Fig. 7). Later development of posterior mesoderm and notochord are inhibited by this mutation possibly due to cell migration (58-60). Mutations in transcription factors hepatocyte nuclear factor 4 (HNF4) and 3 (HNF3), which expressed in embryos at 6.5-dpc, result in embryonic death later in development (61-63). HNF4 is required for completion of gastrulation. In the absence of HNF4, ectodermal cells die possibly due to the absence of a nutrient factor from endoderm (62). Formation of the node and notochord were inhibited in null mutations in the Hnf3 gene. In this case, mesoderm is formed but not patterned correctly. In the absence of the transcription factor MDM2 (64) or the putative transcription factor BRCA1 (65-68), mesoderm is not formed; however, these mesoderm deficiencies appear to result from cell cycle delays in the early egg cylinder. Development of specific subpopulations of later mesoderm is affected with mutations in transcription factor genes Lim1 (69), Otx2 (70-72), twist (73), and Notch1 (74). Finally, the primitive streak pattern is altered with mutations in the embryonic ectoderm development (eed) gene, which encodes a member of the Polycomb group (Pc-G) of transcriptional repressors (75). Formation of the streak is initiated, but most null mutants die at the mid-streak stage when anterior-posterior patterning is disrupted before segmentation.

Molecular and Cellular Biology of Hematopoiesis

Recently, gene knockout experiments have provided important insights in the area of differentiation and embryonic development. Several nuclear regulatory proteins SCL/Tal-1 (76), Rbtn2/LMO2 (77, 78), GATA-1 (79, 80), GATA-2 (81), E2A, and Pax-5 have been demonstrated to be involved in restricting cell lineages (Fig. 8). The process of lineage commitment has become increasingly clear over the past few years, and hematopoiesis is likely controlled through growth factors that permit cellular proliferation

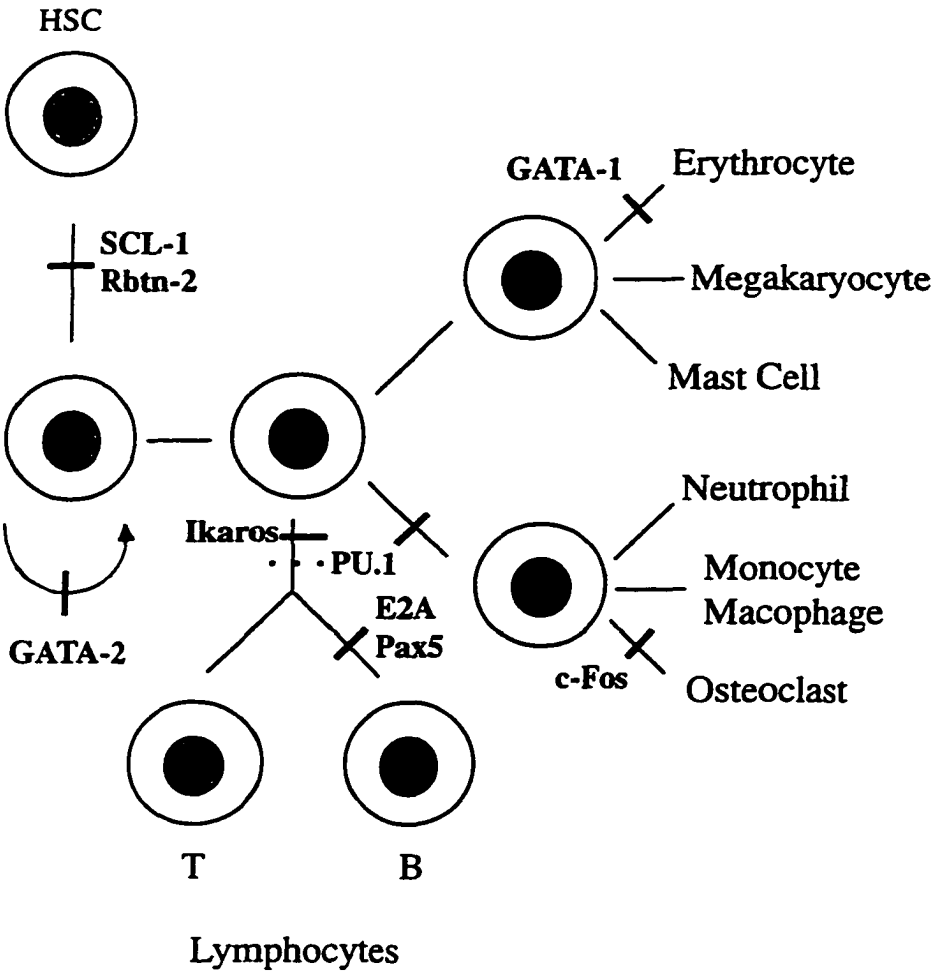


FIG. 8. **Transcription factors involved in the regulation of hematopoiesis.**

or differentiation and nuclear transcription factors that activate lineage specific genes and initiate, or arrest, cell cycling.

The stem cell leukemia (SCL) transcription factor for example is essential for hematopoietic cell differentiation. Lineage commitment and differentiation through a self-renewing population of multipotential stem cells produces hematopoietic lymphoid, erythroid, and myeloid cells of the blood. SCL -/- animals die due to the absence of all hematopoietic lineage cells (82, 83). However, GATA-1 is essential for erythroid cell differentiation (79, 80). Within the mammalian embryo, the site of hematopoiesis shifts from its initial position in the yolk sac blood island (primitive hematopoiesis) to the fetal liver and then to the bone marrow (definitive hematopoiesis). GATA-1 -/- animals die due to failure to form mature red blood cells, but other hematopoietic lineages are normal (79, 80).

SCL protein is a basic helix-loop-helix (bHLH) transcription factor which heterodimerizes with E2A gene family proteins. The SCL/Tal-1 gene was identified through its translocation in acute T-cell lymphoblastic leukemia (76). It is expressed specifically in hematopoietic cells, vascular endothelium, and developing brain. Embryos lacking SCL fail to develop past E9.5 because of the absence of yolk sac erythropoiesis (82, 83) and SCL -/- embryonic stem cells do not contribute to any hematopoietic lineage in chimeric mice (84). Recently, a study of transgenic rescue of hematopoietic defects of SCL -/- embryos and SCL -/- ES cells was published (85); the results show that SCL is essential for angiogenic remodeling of the yolk sac capillary network into complex vitelline vessels. These findings suggest that SCL functions in embryonic angiogenesis and argue for critical functions in both embryonic blood and vascular cells.

The Lim-domain nuclear protein rbtn2/Ttg/LMO2 has important connections with SCL/Tal-1. The genes encoding each protein may be associated with human T-cell acute lymphoblastic leukemia (ALL) (76). Although rbtn2/LMO2 is expressed in a variety of locations, including brain and spleen, it is not expressed in thymus; mRNA levels are

highest in hematopoietic tissues, including fetal liver, but neither gene is expressed in T lymphocytes (86). The homozygous rbtn2 null mutation leads to failure of yolk sac erythropoiesis and embryonic lethality at the same stage as the tal-1/SCL knockout mice (77). The similarity between SCL/Tal-1 and rbtn2/LMO2 knockout phenotypes suggests these proteins function in a common pathway of gene regulation (77, 78).

The DNA sequence motif, GATA, is present in important *cis*-regulatory elements of many erythroid-expressed genes. The zinc finger protein GATA-1 binds this sequence and, thus, poses as an important candidate for regulation of erythroid-specific gene expression and differentiation. Within hematopoietic cells, expression of GATA-1 is limited to the erythroid, megakaryocyte, eosinophil, and mast cell lineages and multipotential progenitors. GATA-1 knockout mice die in utero as a result of embryonic lethality (79, 80). In addition, GATA-1 -/- ES cells fail to contribute to mature red blood cells but are able to develop into other hematopoietic lineages as well as other tissues.

There are three related GATA factor family members. GATA-2 is involved in hematopoietic progenitor cell proliferation (81). GATA-3 expression is widespread during embryogenesis but restricted to T-lymphocytes and the nervous system in adults. GATA-4 is essential for heart tube formation and ventral morphogenesis (87, 88). GATA-2 expression is detected in erythroid and other hematopoietic progenitors and decreases as GATA-1 expression increases with erythroid differentiation (81). Over-expression of GATA-2 in chicken erythroid precursors promotes proliferation and blocks differentiation. Finally, in GATA-1 -/- ES cells, GATA-2 expression is 50-fold higher than in normal ES cells, presumably due to the absence of normal down-regulation from the high levels present in progenitors. Mice lacking GATA-2 exhibit a severe and early hematopoietic defect. They survive to E10-11, but are severely anemic due to a marked reduction in the number of embryonic red blood cells (81). Mice lacking GATA-3 also display hematopoietic abnormalities, including lethal embryonic hemorrhage, despite the presence of megakaryocytes and defective fetal liver hematopoiesis (89).

Globin Gene Regulatory Proteins

The CACCC nucleotide sequence is one of the *cis*-regulatory motifs critical for transcription of ubiquitous- and erythroid-expressed genes, including globin. Differential screening of an erythroleukemia cDNA library demonstrated an erythroid-specific CACCC-binding protein, designated erythroid Kruppel-like factor (EKLF), which is expressed at all stages of erythropoiesis in the developing mouse (90). EKLF binds preferentially to the β-globin CACCC promoter and activates globin gene switching from γ- to β-globin. Mice lacking EKLF die in utero at 15.5-dpc due to inefficient β-globin production (91, 92). Yolk sac erythropoiesis is unimpaired, and other genes whose promoters contain CACCC sites, including GATA-1, porphobilinogen deaminase, erythropoietin receptor, and carbonic anhydrase-1, are expressed at or near normal levels. Thus, *in vivo* EKLF is dispensable for erythroid commitment, primitive hematopoiesis, proliferation of erythroid precursors, and expression of most erythroid genes. Rather, this factor appears to be required principally for the expression of β-globin and perhaps a limited repertoire of other, as yet unidentified, target genes in erythroid cells.

Besides GATA and CACCC sequences, a third motif, the AP-1-like sites within HS 2 of the β-globin gene LCR, are important for globin gene expression *in vivo*. Drosophila CNC motif family members NF-E2, LCR-F1/Nrf1, and Nrf2, function at the AP-1-like sites as heterodimers.

To understand if these transcription factors regulated globin gene expression *in vivo*, these genes were knocked out in ES cells to produce chimeras. Mice lacking NF-E2 develop normally in utero; however, most animals die within the first week of life due to absence of circulating platelets. Megakaryocytes are present but show profound cytoplasmic abnormalities, including a dramatic reduction in granule numbers, disorganized membranes, and a failure to delimit platelet territories. Thus, NF-E2 is essential within the megakaryocyte lineage for regulated expression of genes that are required for cytoplasm maturation, granule formation, and platelet development. Mice lacking Nrf2 developed

normally, possibly due to a compensation by other CNC gene family members. LCR-F1 knockout mice die in utero at 7.5-dpc. These embryos failed to form a primitive streak and lack detectable mesoderm. We were also interested to know whether LCR-F1 regulates globin gene expression. Lcrf1 homozygous mutant ES cells were injected into wild-type blastocysts to produce chimeras. Blood from these high chimeras (base on coat color) were analyzed for GPIs (glucose phosphate isomerases). These results showed that the -/- ES cells contributed to all mesodermally derived tissues, including erythroid cells producing hemoglobin and suggested that LCR-F1 does not regulated globin gene expression.

The bZIP transcription factor Fos is one component of the AP-1 transcription factor complex. Mice carrying targeted disruptions of the c-fos gene display osteoporosis as a primary pathology and altered hematopoiesis as a secondary effect (93, 94). However, overexpression of c-Fos in transgenic mice leads to dysregulated bone growth and development of sarcomas, effects that can be traced to the osteoblast and chondroblast lineages. Osteoporosis in c-Fos null mice results from a failure of the appropriate hematopoietic progenitor to differentiate into functional osteoclasts; however, β-globin are normal.

Model for Globin Gene Switching

An understanding of the basic mechanisms involved in hemoglobin switching is necessary to develop new methods to increase fetal hemoglobin production in patients with sickle cell disease. Recently, experiments in transgenic mice have provided important insights into human hemoglobin switching (19, 22). The β-globin LCR has two important functions. First, these sequences "open" a chromosomal domain that extends over 200 kilobases, and then they function as a powerful enhancer to direct high level transcription of ε, γ, and β-globin genes specifically in erythroid cells. How are individual globin gene family members regulated during development? A model for human globin gene regulation

can be envisioned that incorporates the two important LCR functions and the concept of competition between various regulatory sequences (95) (Fig. 9).

The LCR sequences are activated in early erythroid cell precursors and presumably organize the entire β-globin locus into an open chromatin domain that is stable throughout development. Once the open domain is established, the LCR functions as a powerful enhancer to direct high levels of globin gene expression. In this model, correct temporal regulation results from competition of individual globin gene family members for interaction with LCR sequences (95). Promoter and proximal enhancer binding factors synthesized in yolk sac, fetal liver and bone marrow then influence these competitive interactions either positively or negatively and subsequently determine developmental specificity. For example, factors present in yolk sac derived erythroid cells may bind to the ε-globin gene promoter and stimulate interactions with proteins bound to LCR sequences (Fig. 9). These interactions would then lead to high level expression of the ε-globin gene in early embryos. When the major site of erythropoiesis shifts to the fetal liver, regulatory proteins synthesized in these cells could bind preferentially to γ-globin gene promoter and/or proximal enhancer sequences and provide the γ-globin gene with a selective advantage for interactions with the LCR-bound factors. Consequently, the γ-globin gene would be expressed at high levels at this stage of development. Similarly, when bone marrow becomes the major site of erythropoiesis at birth, factors that bind to β-globin promoter and/or proximal enhancers could stimulate preferential interactions between LCR and β-globin gene sequences and activate high levels of β-globin expression in adult erythroid tissue. Alternatively, repressors that inhibit interactions between specific globin gene family members and the LCR sequences could preferentially silence genes at different developmental stages.

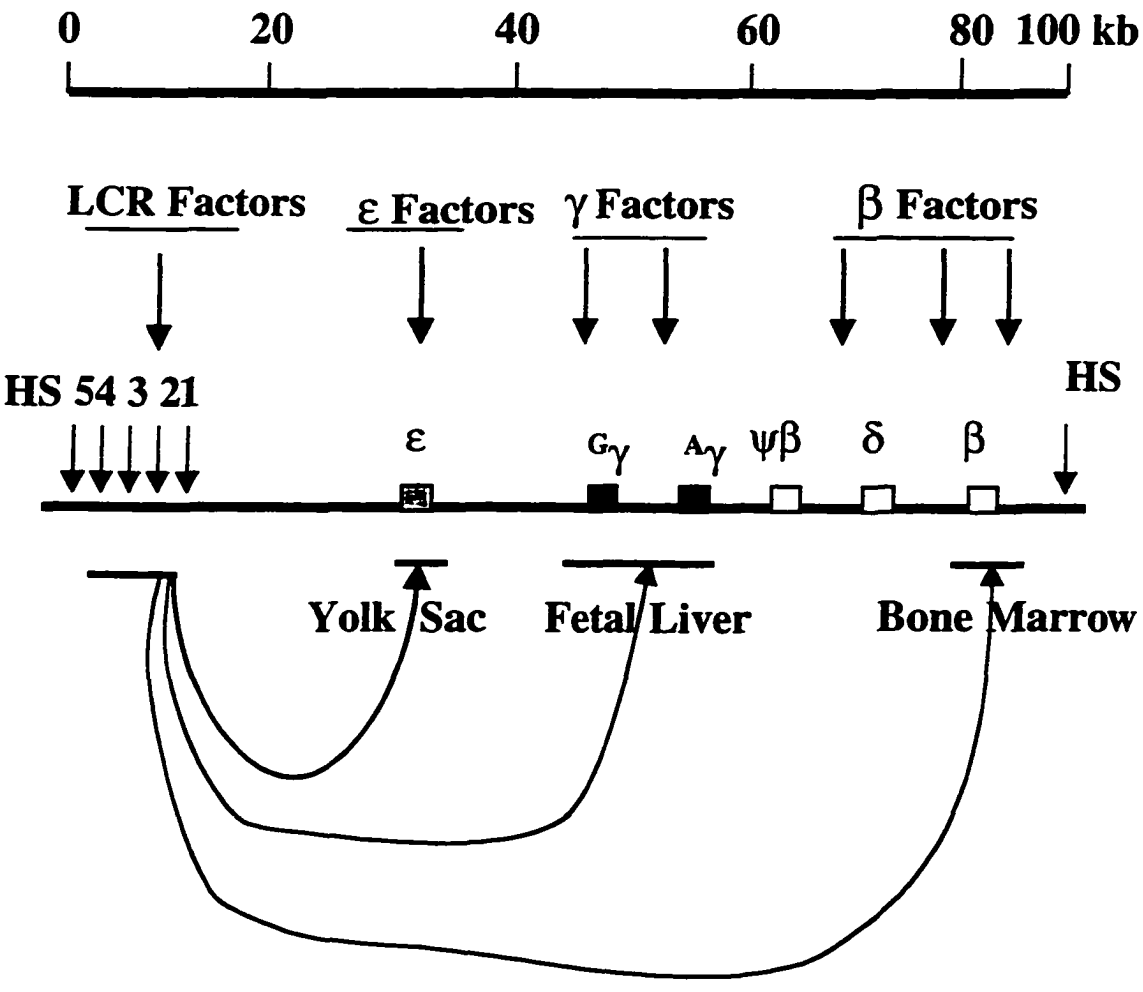


FIG. 9. **Schematic of a simplified model for globin gene switching.** Proteins bind to the LCR sequences in early erythroid cell precursors and organize the entire β-globin locus into an open chromatin domain that is stable throughout development. Once the open domain is established, the LCR functions as a powerful enhancer to direct high levels of globin gene expression. In this model, correct temporal regulation results from competition of individual globin gene family members for stable interactions with the LCR. Promoter and proximal enhancer binding factors synthesized in yolk sac, fetal Liver, and bone marrow influence these competitive interactions either positively or negatively and subsequently determine developmental specificity (Adapted from Townes *et al.* [95]).

Studies of Molecular Mechanisms That Control Human γ-Globin Gene Switching During Development

The goal of this project was to define the *cis*-acting sequences that control human γ-globin switching during development. Previously, Tom Ryan (unpublished data) demonstrated that human γ-globin gene promoter sequences located 72 to 202 base pair (bp) upstream of the transcription start site are essential for γ- to β-globin gene switching. I further defined these sequences to a 70-bp region between –202 and –130 and then defined the specific transcription factor binding site that is required for correct switching. My results demonstrate that the CACCC box at –140 is critical for human γ- to β-globin switching *in vivo*.

In transgenic mice, human γ-globin gene expression is restricted to embryonic life and is localized in a population of primitive nucleated erythroid cells derived from yolk sac blood islands. The β-globin gene is expressed in definitive erythropoiesis that originates in the fetal liver and shifts to the bone marrow at birth. If competition between ε-, γ- and β-globin genes for LCR sequences is an important aspect of temporal control, one might expect that deletion of γ-promoter sequences that removed factor binding sites will result in γ-globin loosing the competition advantage and will not be correctly regulated; although, yolk sac, fetal liver, and bone marrow regulatory factors are presumably present in appropriate concentrations. In order to define γ-globin sequences required for switching, we have established lines of transgenic mice that contain LCR sequences inserted upstream of both the human α- and γ-β-globin genes. In addition, mutations have been introduced within the γ-promoter region to interrupt correct switching.

Tom Ryan (unpublished data) recently demonstrated that sequences in the human γ-globin gene promoter between -202 and -72 are essential for correct switching. Mice containing an LCR (-202) γ-β transgene switch human globin genes correctly during development, but mice containing an LCR (-72) γ-β transgene do not switch correctly. In the latter mice, the adult β-globin gene is inappropriately expressed at a high level in early development. These results strongly suggest that γ-promoter sequences between -202 and

-72 are critical for correct globin gene switching. I undertook a deletion analysis of the γ-promoter in transgenic mice to determine which γ-promoter sequences were necessary for the competitive increase in β expression in mouse yolk sac.

This dissertation describes the deletion analysis of the γ-promoter and its effect on γ- to β-globin switching. I present results that show that a site located around nucleotide −140 of γ-globin promoter is critical for γ- to β-globin gene switching. The site encompasses a CACCC motif which has been predicted to be important for globin gene regulation. I show that the *cis*-element encompassing a CACCC site located at -140 relative to the γ-globin promoter is critical for γ- to β-globin gene regulation. This result confirms the hypothesis that CACCC site in the γ-globin promoter is critical for γ- to β-globin switching. This is the first *in vivo* demonstration of the importance of the γ-globin CACCC site. When the CACCC site is mutated, the developmental specificity of the human β-globin gene is lost and the β-globin gene is inappropriately activated in the yolk sac.

EXPERIMENTAL PROCEDURES

Mutagenesis of the Lcrf1 Gene in Mouse ES Cells

I screened a mouse 129/Sv λ genomic library from Stratagene with radioactively-labeled human Lcrf1 exon 6 as a probe. Several positive clones were isolated, and a 13.5-kb fragment containing the mouse Lcrf1 gene was analyzed by restriction endonuclease digestion. The Lcrf1 coding region was confirmed by DNA sequencing. A 5' homology ([a 5.0-kb BamHI-EcoRI]) fragment and a 3' homology ([a 1.7-kb *Hind*III-Sau3A]) fragment were used to construct a gene targeting vector (Fig. 11*A*, the EcoRI site was converted to BamHI and the Sau3A site was converted to SalI). The 5' homology sequence (5.0-kb BamHI fragment) and the 3' homology sequence (1.7-kb *HindIII*-SalI fragment) were subcloned into the pNTK targeting vector, a gift from Dr. Richard Mortensen, Harvard University, Boston, Massachusetts. The target vector was linearized with NotI and 25 μg of DNA were electroporated into $2x10^7$ embryonic stem (ES) cells in 1 ml ES cell media in a 0.4-cm gap cuvette at 400-450 V and 200-250 μf (BioRad Gene Pulser).

Two ES cell lines (D3 and R1) were used for this study. The D3 cells were from Dr. Tom Doetschman, University of Cincinnati, Ohio. The R1 cells were from Dr. Andras Nagy, Mt. Sinai Hospital, Toronto, Canada. All ES cells were grown on mouse primary embryonic fibroblast feeder layers during the process. G418 (GIBCO/BRL) at 300 μg/ml of active concentration and gancyclovir at 2-2.5 μM (a gift from Syntex) were applied to the media 24 h after electroporation. The selection media was changed every 48 h and surviving colonies were picked after 12-14 days of selection. All the colonies were grown in non-selective media in 24-well plates for 24-48 h. Colonies were expanded by trypsinization. Subsequently expanded cells were frozen for storage and a portion of cells

was used for DNA isolation. Genomic DNA from ES cells was digested with restriction endonuclease and homologous recombinants were identified by Southern blot analysis. The Lcrf1 gene targeting frequency in D3 ES cells was 1 in 76 and in R1 ES cells was 1 in 44. Homologous recombinants were checked for euploidy and absence of mycoplasma.

Generation of Chimeric Mice and Germ Line Transmission of the Lcrf1 Mutant Allele

Chimeras were generated by blastocyst injection as described in Bradley (96) and Hogan *et al.* (36). Three-and-a-half day blastocysts from C57BL/6 mice were collected and injected with Lcrf1 mutant ES cells. Pseudopregnant CD1 female recipient mice were surgically implanted with the injected blastocysts in the uterus to produce chimeras. Lcrf1 heterozygous D3 ES cells were converted into homozygous ES cells by increasing G418 concentration to 2.5 mg/ml in selection media. Chimeras for GPI assay were produced by injection of two lines of Lcrf1 -/- D3 ES cells. Germ line chimeras were generated from one clone of Lcrf1 +/- R1 cells. Mice carrying the Lcrf1 mutant allele were mated to C57BL Swiss outbred mice to produce Lcrf1 heterozygous mice. Southern blot analysis with a 3' probe was used to identify transgenics; progeny tail DNA was subsequently typed by PCR as shown in Fig. 1 and in the text. Lcrf1 heterozygous mice were mated to produce Lcrf1 -/- embryos with 129/Sv x C57BL Swiss background. Embryos from indicated ages were carefully dissected free of maternal tissues using fine forceps under sterile conditions, and polymerase chain reaction (PCR) was performed for genotype. All PCR reactions were done for 32 cycles as follows: 94 °C for 1 min, 59 °C for 1 min, and 68 °C for 1 min. The following oligonucleotides were used as primers: Lcrf1, 5'-CCACCCAGCACCCTCAAGAA-3' (forward) and 5'-GCAGCCGCCCAAACACCTC CT-3' (reverse) (product 385 bp) and neo, 5'-CGCCCGGTTCTTTTTGTC-3' (forward) and 5'-GCGGCCATTTTCCACCAT-3' (reverse) (product 479 bp).

Preparation of Slides and Histological Analysis of Embryos

Individual decidua at 6.5-dpc were fixed in 4 % paraformaldehyde at 4 ℃ overnight, dehydrated in several ascending methanol treatments and embedded in paraffin wax. They were then sectioned at 7-10 μm and stained with hematoxylin/eosin.

In Situ Hybridization Analysis

In situ hybridization was performed as described in Zhao and Hogan (97) and Zhao *et al.* (98). Briefly, decidua from Lcrf1 +/- mice cross mating were dissected out and sectioned. Sections were fixed in 4 % paraformaldehyde, rinsed in PBS, and dehydrated. Sections were rehydrated, permeabilized, washed, rinsed, and proteinase K treated before hybridizing overnight at 50 ℃ to α-35-S-labeled UTP of exon 6 Lcrf1 probe or mouse full-length Brachyury probe. After hybridization, slides were washed, dehydrated, then exposed to Kodak Bio-max film for 1 week.

RT-PCR Analysis

For RT-PCR, total RNA from pools of 60-70 wild-type 3.5-day-old blastocysts, from 100-mm plates of undifferentiated wild-type ES cells or from 100-mm plates of +/- or -/- embryoid bodies were amplified according to the manufacturer's protocol (Invitrogen). In brief, cDNA was synthesized from 3 μg of total RNA with MMLV RT in a 20 μl reaction. Amplification was performed with Taq Polymerase (Fisher) for 32 cycles as follows: 94 ℃ for 1 min, 58 ℃ for 1 min, and 68 ℃ for 2 min. The following oligonucleotides were used as primer: Lcrf1, 5'-CCAGAAGGAGCAGGATGTGGA-3' (forward) and 5'-GCAGCCGCCCAAACACCTCCT-3' (reverse) (product 1,472 bp) and hprt, 5'-GCTGGTGAAAAGGACCTCT-3' (forward) and 5'-CACAGGACTAGAACAC CTGC-3' (reverse) (product 249 bp). The PCR products were mixed at a 3:1 ratio (Lcrf1 to Hprt) and were run on a 1 % agarose gel.

Glucose Phosphate Isomerase (GPI) Analysis

GPI analysis was performed as described Nagy and Rossant (99). In brief, samples were lysed by three cycles of freeze/thaw and 5 µl of each sample was applied to Tris-glycine treated cellulose acetate plates. Samples were electrophoresed at 300 V for 90 min at room temperature. After the gel run, 10 milliliters of 1 % agarose gel mixture containing 10.8 mg magnesium acetate, 15 mg of fructose-6-phosphate (F6P), 2 mg β-nicotinamide adenine dinucleotide phosphate (NADPH), 0.36 mg of phenazine methosulfate (PMS), and 2 mg of methylthiazolium tetrazolium (DTT) at 55 °C were mixed with 10 units of glucose-6-phosphate dehydrogenase (G6P-DH) and immediately poured over the cellulose acetate plate. The plate was incubated for 10 min at 37 °C in the dark and then fixed in 1:3 acetic acid:glycerol for 10 min. Results were recorded by photography.

Embryoid Body In Vitro Cultures and Analysis

Homozygous Lcrf1 mutant ES cells were converted from heterozygous Lcrf1 ES cells by growth in media containing 2000 µg/ml G418 as described in Mortensen *et al.* (100) and Ausubel *et al.* (101). Homozygous mutant ES cells were confirmed by Southern blot analysis. Methylcellulose cultures were done as described Wiles (102). Both heterozygous and homozygous Lcrf1 mutant ES cells were cultured in methylcellulose in the presence of erythropoietin and 2000 µg/ml of G418 for 10 days. Embryoid bodies were stained with benzidine for hemoglobin detection in wet mounts as described Doetschman *et al.* (103), Pearse (104), and Stevens (105). Embryoid bodies were collected by centrifugation. *Northern blot analysis was performed as described in Ausubel et al.* (101) and Caterina *et al.* (30). Embryoid bodies were stained with Dip Quick (a stain for nucleated red blood cell precursors, Jorgensen Laboratories) for cytospin assay.

Murine Retrovirus Rescue Constructs

Recombinant murine stem cell virus (MSCV) retroviruses for expression of human Lcrf1 gene were constructed using a MSCV vector backbone. Exon 6 of human Lcrf1 cDNA was cloned downstream at the EcoRI site of a modified myeloproliferative sarcoma virus long terminal repeat (LTR) followed by a phosphoglycerkinase (PGK) promoter-puromycin N-acetyl transferase (pac) gene. The constructs were digested with KpnI, and the LTR-cDNA-pGK-pac-LTR fragment was separated from vector sequences on an agarose gel and was purified on DEAE paper for DNA injection.

BAC DNA Purification

Human genomic Lcrf1 BAC clones were isolated. In brief, a human Lcrf1 exon cDNA was radioactively-labeled as a probe to screen the BAC library on filters obtained from Genome System, Inc. The positive BAC DNA was isolated using alkaline lysis from 25 ml LB media as described by manufacturer's protocol (Genome Systems, Inc). DNA was dissolved in injection buffer (10 mM Tris-HCL [pH 7.5], 0.1 mM EDTA, 100 mM NaCl, 30 μM spermine, and 70 μM spermidine) and quantitated by spectrophotometer.

Generation of MSCV and BAC Transgenic Mice

DNA injection was done by Jinxiang Ren in Dr. Townes' injection facility at UAB. The linearized LTR-Lcrf1-pGK-pac-LTR and circular BAC DNA were injected at a concentration of 2 ng/μl and 1 ng/μl into fertilized oocytes from C57BL/6xSJL mice. Microinjections were performed using Leica micromanipulators attached to a Leica microscope. Oocytes that survived injection were transferred to both oviducts of pseudopregnant CD1 foster mothers. Transgenic mice were identified both by PCR and Southern blot analysis of genomic DNA prepared from tail biopsies as described below.

Detection of Transgenes by PCR Analysis

To identify transgenic animals the following sets of primers were used: 5' GACATTTCCAGCATAACAGAAGCAG and 5' CTTGCTGATGTGTTCACTTC (amplified a 900-bp fragment including intron 5) and neomycin resistance cassette primers 5' CGCCCGGTTCTTTTTGTC and 5' GCGGCCATTTTCCACCAT. Primers for both human and mouse Lcrf1 gene were: 5' CCACCCAGCACCCTCAAGAA and 5' GCAGCCGCCCAAACACCTCCT (amplified a 395-bp fragment). PCR assays were performed in a final volume of 20 µl consisting of 50 mM KCl, 10 mM Tris-HCl (pH8.9), 2.5 mM $MgCl_2$, 5 pmole each primer, 0.2 mM dNTPs, 1 µg genomic DNA, and 2 u of Taq DNA polymerase. PCR conditions were: 5 min at 100 °C; followed by 30 cycles of 30 sec at 94 °C, 30 sec at 60 °C, and 2 min at 68 °C. The reaction products were analyzed on a 1 % agarose gel. To determine the absence of the mouse Lcrf1 gene, PCR products with Lcrf1 primers were digested with EcoRI which cuts only the mouse Lcrf1 sequence but not human Lcrf1. The digested PCR products were separated on a 1.5 % agarose gel.

Transgenic Examination by Southern Blot Analysis

For transgene analysis, 10 µg of genomic DNA from tails were digested with BamHI and size fractionated on a 1 % agarose gel in 1x TBE. Gels were treated with 0.5 N NaOH and 1 M NaCl and neutralized with 1 M Tris-HCl pH7.4 and 1.5 M NaCl. DNA was blotted onto nitrocellular membranes and hybridized to radioactively-labeled 3' fragment DNA. Autoradiograms were developed using a PhosphoImager. Transgenic mice with 3 copies of transgenes in head-to-tail direction were identified by Southern blot analysis. Nine independent transgenic lines were established and were crossed with Lcrf1 heterozygous mice to produce Lcrf1 +/- transgenic strains, which were then mated to produce Lcrf1 -/- transgenic embryos. Progeny tail DNA was subsequently typed by PCR with neo primers and LTR leader sequence primers. Yolk sacs were removed from embryos and DNA from yolk sacs were genotyped by PCR.

Primers for Mega-Primer Mutations

Δ (159/72) (BamHI) primer, 5'-GCAGGGCCCCTTCCCCACACTATCTCAATGCAAA TATCTGTCTGAAACGGATCCAGTGAGGCCAGGGGC-3'.

Δ(200/160), 5'-GCAGGGCCC^GTCCCTGGCTAAACTCC-3'.

GATA (EcoRI) Distal primer, 5'-GCAGGGCCCCTTCCCCACAGAATTCCAATGCAA ATATCTGTCTGA-3'.

GATA (Bgl II) Proximal primer, 5'-AGCCAGGGACCGTTTCAGGAATTCATTTGCAT TGAGATAGTGTGG-3'.

^γCAC (EcoRI) mega primer, 5'-CAAGGCAAGGCTGGCCAACCCATGAATTCAGTTT AGCCAGGGACCG-3'.

OCT (EcoRI) primer, 5'GCAGGGCCCCTTCCCCACACTATCTCAGAATTCATATCTG TCTGAAACGGTCC-3'.

RESULTS

After cloning Lcrf1 for my master's project, I sought to determine its role in β-globin gene expression in the mouse. One way to study biological function of the gene in mouse is to remove the coding sequences by homologous recombination in ES cells. There are two ways to analyze the role of LCR-F1 in expression of β-globin. One is to study *in vitro* differentiated embryoid bodies (EBs) derived from -/- ES cells, and the other one is to study Lcrf1 null animals.

Hematopoietic Defects in Embryoid Bodies (EBs) Derived form Lcrf1 -/- ES Cells

Doetschman *et al.* (103) showed ES cells culture to form complex EB with endoderm, basal lamina, mesoderm, and ectoderm in 3 to 8 days and to continue to form visceral yolk sac, blood islands, and myocardium in 8 to 10 days. The hemoglobinized cells can be determined by benzidine (a stain for these erythroid proteins).

Embryoid bodies derived from Lcrf1 +/- and -/- ES cells were differentiated in methylcellulous cultures to investigate the effect of the Lcrf1 mutation on hematopoiesis and globin expression *in vitro*. After *in vitro* differentiation of -/- ES cells, red blood cell formation and hemoglobin synthesis were dramatically decreased (Fig. 10). Large numbers of hemoglobinized (benzidine-stained) cells were derived from +/- embryoid bodies; however, few hemoglobinized cells could be identified in -/- ES cell culture (Fig. 10*B,D*). When EBs were pelleted by centrifugation, a benzidine-stained layer of hemoglobin-producing cells was observed in +/- EBs but not in -/- EBs (Fig. 10*E*). Northern blot analysis of embryoid body RNA (Fig. 10*F*) confirmed that globin mRNA levels were significantly reduced in -/- EBs. In addition, nucleated red blood cell precursors were observed in +/- EBs (Fig. 10*G*) but not in -/- EBs (Fig. 10*H*).

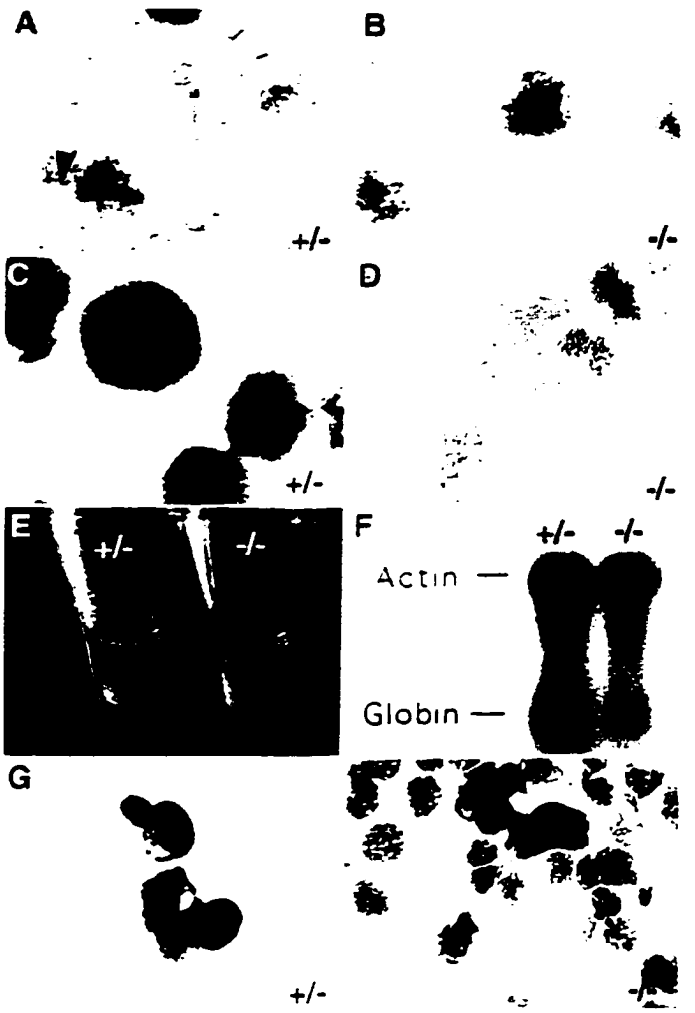


FIG. 7. **Schematic of a simplified mouse mesoderm formation.** Mouse gastrultion (3.5-7.5 days post coitum [dpc]) showing the egg cyclinder with the parietal yolk sac removed. The relationship of the epiblast to embryonic visceral endoderm, extraembryonic visceral endoderm, and extraembryonic ectoderm are shown. The primitive streak is first evident at 6.5-dpc, with the delamination of mesoderm at the embryonic-extraembryonic junction, and the site of streak formation defines the posterior margin of the embryo (Adapted from Beddington and Robertson [42]).

The possible interpretation of the *in vitro* data is that either the Lcrf1 mutation has an indirect effect via a pre-erythroid developmental block or a more direct effect on globin expression, as suggested by LCR-F1's ability to stimulate globin expression in transient transfection studies.

Generation of an Lcrf1 Mutant Allele in the Mouse Germ Line

Homologous recombination in ES cells can be used to delete the functional gene, and mutations can be engineered into any gene in the mouse. It has revolutionized the genetic analysis of cellular processes, particularly in the area of differentiation and embryonic development.

To assess the possible role of LCR-F1 in mammalian development, I made a knockout mutation in the mouse LCR-F1 gene. Gene targeting by homologous recombination allows the introduction of a specific mutation into any gene. The gene of interest is inactivated by interrupting its coding sequence with a positive selectable marker (pGK-neo) and a negative selectable marker (pGK-tk). To construct a knockout construct, I screened a mouse 129/Sv λ genomic library with human Lcrf1 exon 6 as a probe. A λ-bacterophage containing 13.5 kb of mouse Lcrf1 genomic DNA was isolated and mapped by restriction endonuclease digestions. I constructed a gene targeting construct by subcloning a 5' homology sequence [5.0-kb BamHI fragment (modified from BamHI-EcoRI)] and a 3' homology sequence [1.7-kb *Hind*III-SalI (modified from *Hind*III-Sau3A)] into the pNTK vector. The LCR-F1coding region was replaced with a 3.5-kb pGKneo cassette that was confirmed by sequencing (data not shown). This targeting construct is then transfected into cultured ES cells. ES cell lines are derived from the inner cell mass of a blastocyst-stage embryo. Homologous recombination occurs in a small number of the transfected cells, resulting in introduction of the mutation into the target gene. Homologous recombination removed the coding exons 3A, 3B, 4, 5, and 6 which contained all functionally important domains (transcriptional activation domains,

putative nuclear localization signal, CNC, and bZIP) (Fig. 11*A*). I screened all cells containing the correctly targeted allele by Southern blot analysis (data not shown). With assistance from Susan Farmer, I deleted the mouse Lcrf1 gene in both D3 and R1 ES cells and assayed all mutant ES cells for correct chromosome number (data not shown). The successfully targeted ES cells were injected into 3.5-day-old wild-type blastocysts from C57BL/6 mice. Upon reintroduction into a host blastocyst, they contribute to all tissues of a chimeric mouse, including germ cells. Only one +/- R1 cell line was passed though the germ line to produce Lcrf1 heterozygous mutant mice. Heterozygous mutant animals were mated to each other to generate homozygous Lcrf1 knockout animals.

Susan Farmer bred the chimeras to C57 Black Swiss females to produce heterzygous animals. She also analyzed the agouti offspring by Southern blot analysis or PCR analysis of tail DNA for heterozygous mutants (Fig. 11*B, C*). Tail DNA was digested with restriction endonuclease KpnI or BamHI, and the mutant allele generated a 39-kb KpnI or a 4.6-kb BamHI fragment when probed with 5' or 3' probes, respectively. To verify that no random integration occurred, a neomycin gene was radioactively labeled and used as a probe. This probe hybridized a 4.6-kb BamHI fragment. Heterozygous males (+/-) and females (+/-) were mated, and tail DNA of progeny were examined by PCR (Fig. 11*C*).

Effect of the Lcrf1 Null Mutation Is Not Cell-Autonomous

As described above, the Lcrf1 knockout resulted in an embryonic lethal phenotype in mice. We are also interested in knowing the role of LCR-F1 in regulating globin expression; therefore; we sought to dissect the different roles of LCR-F1. We made ES cells homozygous for the Lcrf1 mutation as described below and injected these cells into 3.5-dpc embryos (blastocysts) from C57BL/6 mice. We then transferred these injected blastocysts into the uteri of foster mothers. After approximately 17 days, chimeric mice were born. When the mice were approximately 8 weeks old, tissues from these chimeras

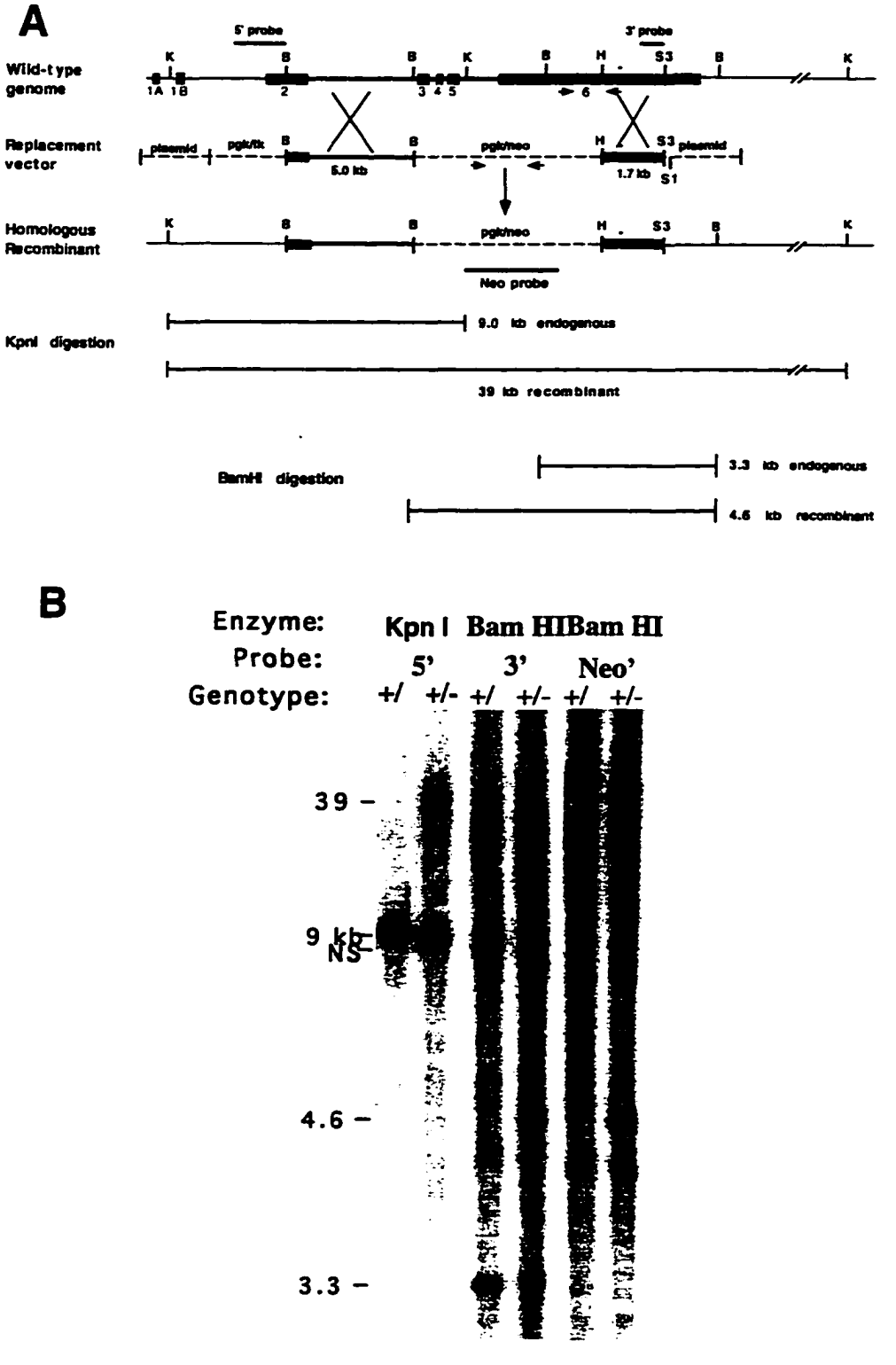


FIG. 11. **Targeted disruption of the LCR-F1 gene.** *A*, Knockout construct wild-type allele and targeted allele. *B*, Southern blot of wild-type and heterozygous mutant mice.

C

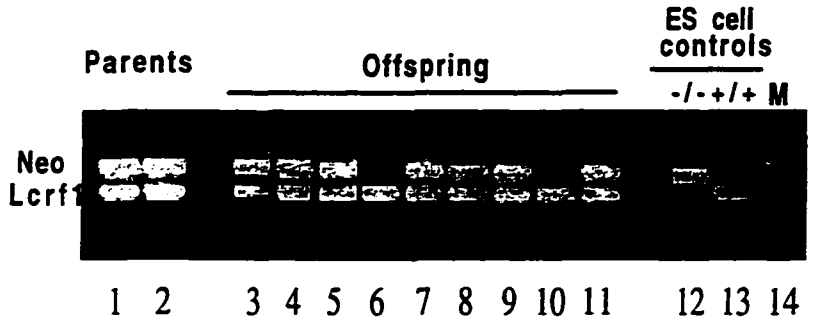


FIG. 11. (Continued.) *C*, PCR genotyping from LCR-F1 +/- parents and offspring. No homozygous mutant offspring are observed. These results suggest that the null mutation is embryonic lethal.

were analyzed for GPI. In this assay, tissues derived from ES cells and blastocysts are distinguished by differences in the charge of GPI isozymes. We wanted to determine whether Lcrf1 -/- ES cells could contribute to erythroid cells that made hemoglobin.

For GPI assay, Susan Farmer and I converted the heterozygous mutant D3 cells into homozygous mutant cells (double knockout) by increasing the G418 concentration in media. Lcrf1 -/- D3 cells were isolated from +/-D3 cells by culturing the cells in increasing concentrations of G418 (2.0 to 2.5 mg/ml). The more stringent selection conditions favor cells that contain two mutant alleles (pGK-neo) of the target genes as result of loss of heterozygosity. Southern blot analysis was performed to confirm the homozygousity (Fig. 12*A*) Two Lcrf1 -/- lines were injected to generate chimeras. Blood from five high-level chimeras (75 % to 90 % based on coat color) were analyzed by GPI analysis. The results indicated that Lcrf1 -/- cells contributed to all tissues examined, including blood that is derived from mesoderm (Fig. 12*B*). Animals with high blood cells from the -/- ES cells that were not anemic, and their hematological indices were not significantly different from those of non-chimeric (non-agouti) littermates (data not shown). These results demonstrate that Lcrf1 -/- ES cells differentiated into red blood cells that developed normally and produced hemoglobin. Analysis of chimeras also demonstrated that Lcrf1 -/- cells contributed to kidney, liver, lung, muscle, and spleen. These results indicate that the mutant phenotype can be rescued at gastrulation by +/+ cells derived from wild-type blastocysts; the effect of the Lcrf1 mutation is not cell-autonomous. Therefore, LCR-F1 must directly or indirectly regulate transcription of gene(s) encoding a secreted factor(s) or cell surface molecule(s) essential for gastrulation. This factor is secreted by Lcrf1 +/- cells and rescues Lcrf1 -/- cells, which then contribute to mesodermally derived tissues, including blood.

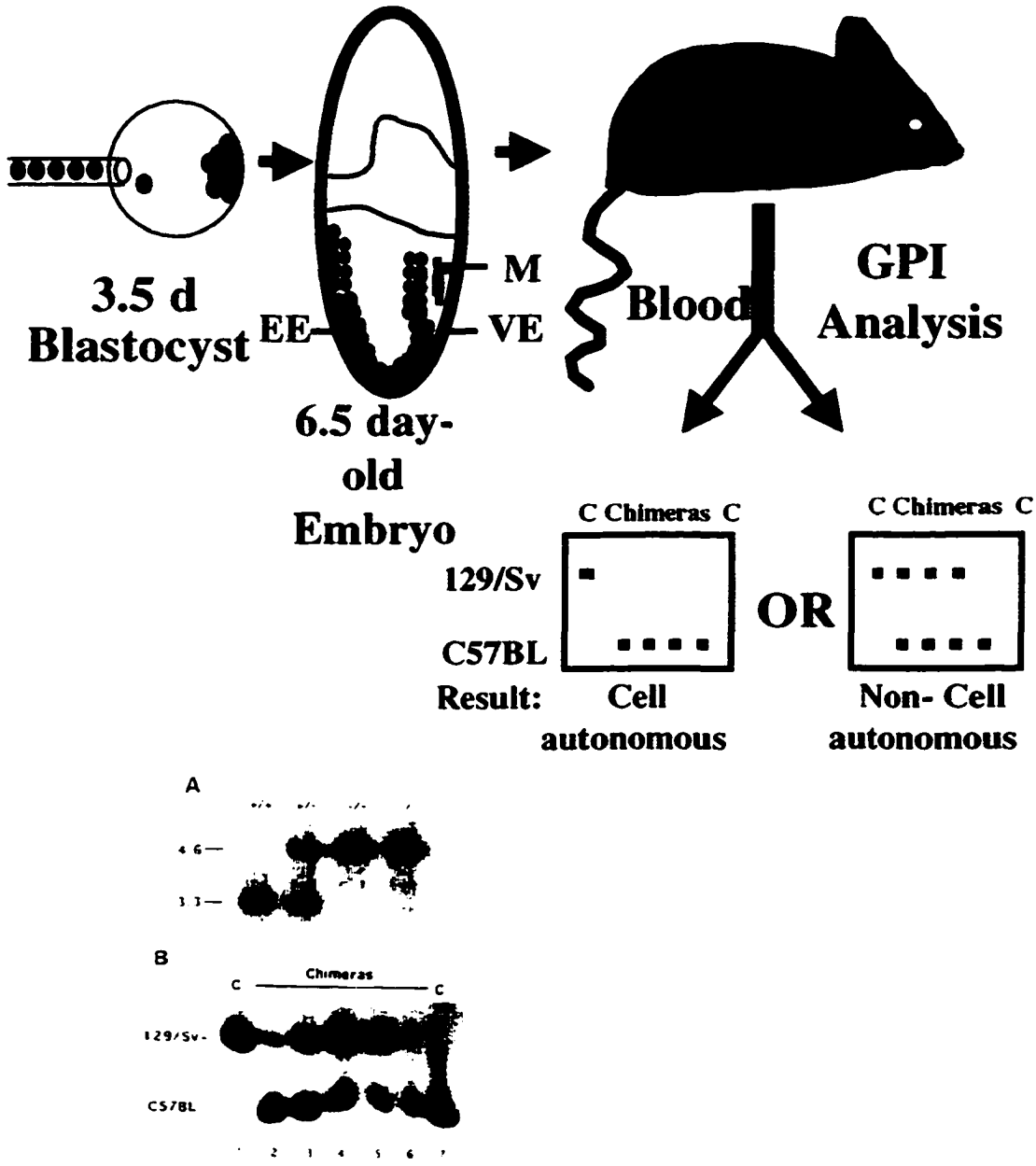


FIG. 12. **GPI analysis of chimera from injection of Lcrf1-/- ES cells.** Lcrf1-/- ES cells were derived from 129/Sv strain that were injected into C57/BL/6 blastocysts. Blood from the high percentage chimeras (based on coat color) were analyzed to determine whether the Lcrf1 -/- ES cells were able to contribute to the hemoglobin present in the blood. *A*, Southern blot analysis of DNA from Lcrf1+/+, +/- and -/- ES cells. *B*, GPI analysis of blood samples from chimeras produced with Lcrf1 -/- ES cells.

Lcrf1 Homozygous Mutant Is Embryonic Lethal

Susan Farmer found that the homozygous mutant embryos grow normally to the late egg cylinder stage at 6.5-dpc, but development was arrested before 7.5-dpc.

Embryos were obtained from timed matting between heterozygous (+/-) males and females. Embryos at 9.5- or 8.5-dpc were analyzed by PCR and no homozygous embryos were found (Table I); also, the percentage of empty decidua was high (19 %; Table II).

TABLE I

Genotypes of offspring from +/- X +/- matings

Age	Total	+/+ or +/-	-/-	% -/-
Adult	195	195	0	0
9.5-dpc	24	24	0	0
8.5-dpc	34	34	0	0
7.5-dpc	49	46	3	6
6.5-dpc	55	35	20	36

dpc = days post coitus

TABLE II

Empty decidua from +/- X +/- matings

Age	Total	# Empty	% Empty
9.5-dpc	31	6	19
8.5-dpc	62	12	19
7.5-dpc	137	27	20
6.5-dpc	196	15	8

dpc = days post coitus

Six percent of the embryos isolated at 7.5-dpc were -/- by PCR. This number is significantly below the 25 % rate expected from Mendelian segregation, suggesting that the majority of -/- embryos died before this stage. In addition, the percentage of empty descidua at 7.5-dpc was 20 %. Histological sections of these empty decidua from mating

between heterozygous demonstrated reabsorbing embryos. Fig. 13 shows a typical resorption site and a normal control. The results demonstrate that no homozygous mutant animals were obtained; therefore, the null mutation is embryonic lethal. PCR (Fig.11*C*) or *in situ* hybridization (Fig. 14) to genotype embryos was performed to determine the stage of embryonic death.

In situ hybridization of the histological sections from 6.5-dpc embryos was performed with a Lcrf1 exon 6 to further characterize the stage of embryonic lethality (Fig. 14). Homozygous mutant embryos were identified by the lack of Lcrf1 RNA. Thirty-six percent of the embryos lacked Lcrf1 RNA (Table I); this percentage is in the normal (25 %) range for Mendelian segregation. Only 8 % empty decidua were observed at 6.5-dpc (Table II), which is similar to the number observed in wild-type mating. These results demonstrate that the majority of Lcrf1 mutant embryos die between 6.5- and 7.5-dpc.

No Mesoderm Formation Is Morphologically Detectable in Lcrf1 Null Embryos

To determine the possible cause of death, histological sections of 6.5-dpc embryos from Lcrf1 heterozygous matting were examined. Normal layers of embryonic ectoderm, mesoderm, and endoderm were clearly distinguishable in Lcrf1 +/+ or +/- embryos (Fig. 14*A*, *B*). However, in Lcrf1 -/- embryos, no mesodermal layer (embryonic or extraembryonic) was observed, although embryonic ectoderm and visceral endoderm layers appeared normal. The egg cylinders of Lcrf1 null embryos were indistinguishable from wild type. However, mesoderm formation was absent, and no primitive streak was initiated. These results strongly suggest that LCR-F1 is essential for one of the earliest steps in mammalian mesoderm formation.

Susan Farmer analyzed RNA from undifferentiated ES cells and blastocysts to determine whether Lcrf1 expression is regulated developmentally in early mouse embryos. In wild-type, undifferentiated ES cells (Fig. 15, lane 1) and 3.5-day-old mouse blastocysts (Fig. 15, lane 2), Lcrf1 mRNA is absent by RT-PCR analysis. However, in 6.5-day- old

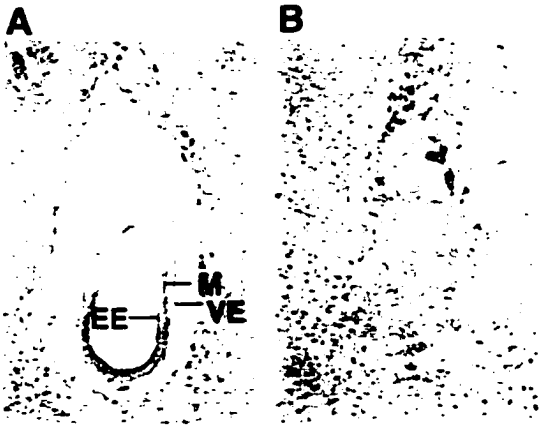


FIG. 13. **LCR-F1 +/- and -/- 7.5-day-old embryos.** *A*. Location of the three primary germ layers in normal embryos: (EE) embryonic ectoderm, (M) mesodermin the primitive streak, (VE) viseral endoderm. *B*. none of these three layers is observed in the resorbing embryo in *B*.

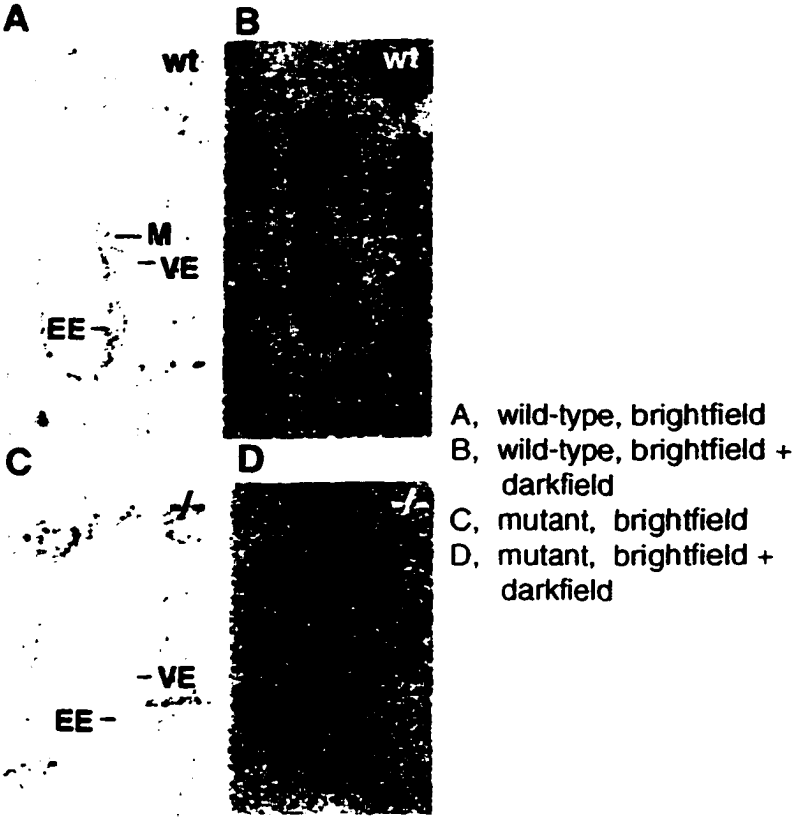


FIG. 14. **LCR-F1 in situ hybridization of 6.5-day-old embryos.** Expression of LCR-F1 is detected in all tissues of the wild-type embryo in *Panels A* and *B*. No LCR-F1 expression is detected in the mutant embryo in *Panels C* and *D*, and no mesodermal layer is observed.

embryos (Figs. 14*A*, *B*) and in 8.0-day-old embryoid bodies derived from differentiated, Lcrf1 +/- ES cells (Fig. 15, lane 3), Lcrf1 RNA is expressed at high levels. In 8.0 day-old embryoid bodies derived from differentiated, Lcrf1 -/- ES cells (Fig. 15, lane 4), no Lcrf1 mRNA is detected as expected. These results are consistent with a key role for LCR-F1 in early development. Lcrf1 expression is activated between 3.5 and 6.5 days of mouse development.

Brachyury (T) Expression Is Absent in Homozygous Mutant Embryos

Early markers of mesoderm formation may be the key to finding a master gene for mesoderm formation. LCR-F1 may be such a gene since it precedes T expression, which is the earliest mesodermal marker. A small number of mesodermal cells could be present in the Lcrf1 null embryos, although no morphologically distinguishable mesoderm is observed. Therefore, histological sections 6.5-day-old mutant embryos were analyzed for T by in situ hybridization. T is one of the earliest markers for mesoderm formation and is normally expressed at 6.5-dpc in the primitive streak of gastrulating embryos (58, 106, 107). It is considered to be a marker of mesoderm formation, and we wanted to ask whether LCR-F1 acted upstream of T in mesoderm formation. In the sagittal, and transverse sections shown in Fig. 16*A* and *B* respectively, T was detected in a normal pattern in Lcrf1-positive embryos. However, in Lcrf1 null embryos, no T expression was observed (Fig. 16*C*, *D*). These data confirm that Lcrf1 mutant embryos fail to form a primitive streak and mesoderm fails to form.

Human Lcrf1 Genomic BAC Clone Rescue

It is important to demonstrate that embryonic lethality phenotype is not due to the pGK-neo cassette affecting the neighbor genes. To rescue the lcrf1 -/- phenotype with either genomic DNA or cDNA, I sequenced the Lcrf1 genes from mouse and human. LCR-F1 amino acid sequences are highly conserved between mouse and human (97 %)

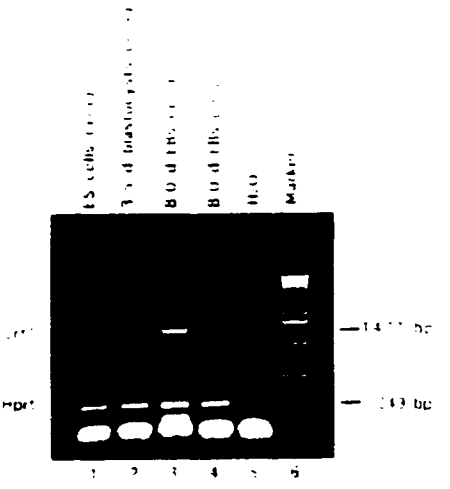


FIG. 15. **RT-PCR analysis of ES cells, 3.5-day-old blastocysts, and 8.0-day-old embryoid bodies (EBs).** LCR-F1 is not expressed in 3.5d blastocysts but is expressed in 6.5-day-old embryos and in 8.0-day-old EBs derived from differentiated LCR-F1 +/- ES cells. The results demonstrate that LCR-F1 is developmentally regulated.

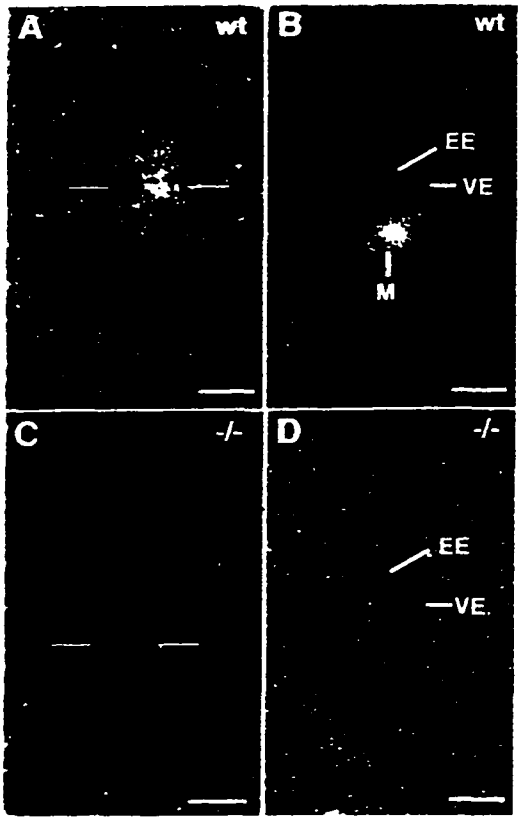


FIG. 16. **LCR-F1 +/- and LCR-F1 -/- embryos were genotyped by in situ hybridization with an LCR-F1 probe.** Adjacent sections were then hybridized with a Brachyury (T) probe. *A, B,* LCR-F1 wild-type embryos express T at the primitive streak. *C,D,* no T expression is detected in the LCR-F1 -/- embryos. *A, C,* sagittal sections, and *B, D,* transverse sections.

(Fig. 17), especially in the bZIP region (99.5 %) (Fig. 5). This sequence conservation suggests that human Lcrf1 might rescue embryonic lethality in Lcrf1 -/- mice.

To obtain a full-length human Lcrf1 genomic DNA, I screened a human genomic BAC library from Genome Systems, Inc. with a radioactively-labeled human Lcrf1 exon 6 probe. Two identical human genomic BAC clones containing Lcrf1 gene were isolated from the screen. Southern blot analysis was performed to compare the Lcrf1 gene in the BAC clone and in human genomic DNA. The Lcrf1 BAC clone DNA was digested with various restriction endonucleases, and the restriction fragments were separated by agarose pulse-field gel electrophoresis. Human genomic DNA was digested with the same restriction endonucleases as a control. DNA was denatured and transferred onto a nitrocellulose membrane, and probed with radioactively-labeled human Lcrf1 exon 6. The results of this analysis demonstrated the BAC clone contained an intact copy of the human Lcrf1 gene (data not shown).

The human BAC clone was injected into fertilized mouse eggs to generate transgenic mice. The injected eggs were transferred into foster mothers and tail DNA from pups was analyzed by PCR. Positive samples were confirmed by Southern blot analysis. Transgenic mice with human Lcrf1 BAC transgenes were mated to Lcrf1 +/- animals to generate transgenic that were heterozygous for the deletion and contained the transgene (Lcrf1 +/-, Tg). These transgenic animals were bred to Lcrf1 +/- animals and Lcrf1 -/-Tg pups were confirmed by PCR and Southern blot analysis (Fig. 18). Several Lcrf1 -/-Tg pups were born, and no abnormal phenotype was observed. These results demonstrate that the human Lcrf1 gene can substitute for the mouse Lcrf1 gene and suggest that Lcrf1 function is conserved in mammalian development.

MSCV/Lcrf1 cDNA Partial Rescue

The embryonic lethality of Lcrf1 homozygous mutants can be rescued by human Lcrf1 genome DNA. However, cDNA rescue shows direct evidence for Lcrf1 -/-

```
Human MEVNTSASEILYSAPPGDPLSTNYSLAPNTPINQNVSLHQASLGGCSQDFLLFSPEVESL 60
      MEVNTSASEILY APPGDPLS NYSLAPNTPINQNVSLHQASLGGCSQDF LFSPEVESL
Mouse MEVNTSASEILYNAPPGDPLSSNYSLAPNTPINQNVSLHQASLGGCSQDFSLFSPEVESL 356

Human PVASSSTLLPLAPSNSTSLNSTFGSTNLTGLFFPPQLNGTANDTAGPELPDPLGGLLDEA 120
      PVASSSTLLPLAPSNSTSLNSTFGSTNL G FFP QLNGTANDT GPELPDPLGGLLDEA
Mouse PVASSSTLLPLVPSNSTSLNSTFGSTNLAGPFFPSQLNGTANDTSGPELPDPLGGLLDEA 416

Human MLDEISLMDLAIEEGFNPVQASQLEEEFDSDSGLSLDSSHSPSSLSSSEGSSSSSSSSS 180
      MLDEISLMDLAIEEGFNPVQASQLEEEFDSDSGLSLDSSHSPSSLSSSEGSSSSSSSSS
Mouse MLDEISLMDLAIEEGFNPVQASQLEEEFDSDSGLSLDSSHSPSSLSSSEGSSSSSSSSS 476

Human SSSSASSSASSSFSEEGAVGYSSDSETLDLEEAEGAVGYQPEYSKFCRMSYQDPAQLSCL 240
        SSASSSASSSFSEEGAVGYSSDSETLDLEEAEGAVGYQPEYSKFCRMSYQDP QLSCL
Mouse ..SSASSSASSSFSEEGAVGYSSDSETLDLEEAEGAVGYQPEYSKFCRMSYQDPSQLSCL 534

Human PYLEHVGHNHTYNMAPSALDSADLPPPSALKKGSKEKQADFLDKQMSRDEHRARAMKIPF 300
      PYLEHVGHNHTYNMAPSALDSADLPPPS LKKGSKEKQADFLDKQMSRDEHRARAMKIPF
Mouse PYLEHVGHNHTYNMAPSALDSADLPPPSTLKKGSKEKQADFLDKQMSRDEHRARAMKIPF 594

Human TNDKIINLPVEEFNELLSKYQLSEAQLSLIRDIRRRGKNKMAAQNCRKRKLDTILNLERD 360
      TNDKIINLPVEEFNELLSKYQLSEAQLSLIRDIRRRGKNKMAAQNCRKRKLDTILNLERD
Mouse TNDKIINLPVEEFNELLSKYQLSEAQLSLIRDIRRRGKNKMAAQNCRKRKLDTILNLERD 654

Human VEDLQRDKARLLREKVEFLRSLRQMKQKVQSLYQEVFGRLRDENGRPYSPSQYALQYAGD 420
      VEDLQRDKARLLREKVEFLRSLRQMKQKVQSLYQEVFGRLRDE GRPYSPSQYALQYAGD
Mouse VEDLQRDKARLLREKVEFLRSLRQMKQKVQSLYQEVFGRLRDEHGRPYSPSQYALQYAGD 714

Human GSVLLIPRTMADQQARRQERKPKDRRK 447
      GSVLLIPRTMADQQARRQERKPKDRRK
Mouse GSVLLIPRTMADQQARRQERKPKDRRK 741
```

FIG. 17. **Comparison of the conserved peptide sequences between human and mouse Lcrf1 genes.**

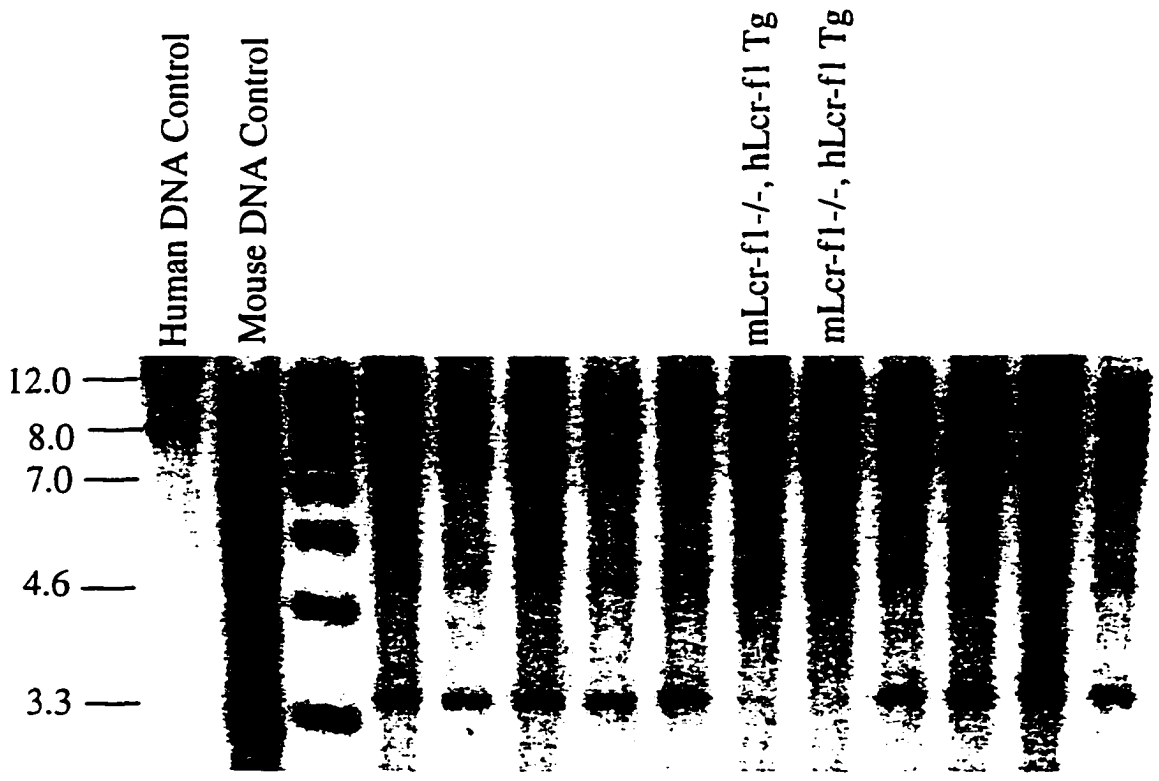


FIG. 18. **Southern blot analysis of DNA from transgenic mice with mLcr-f1 +/- and hLcr-f1 +/- mice.** DNA was digested with BamHI and hybridized with human Lcr-f1 cDNA. Positions of the endogenous 3.3-kb and 7.0-kb band and the homologous recombinant 4.6-kb bands are indicated on the left and are the same as the human Lcr-f1 endogenous 8.0-kb and 12.0-kb bands.

embryonic lethality. I generated transgenic mice that contain the human Lcrf1 cDNA inserted downstream of the MSCV LTR promoter. The construct also contained a yeast Gal 4 (1-94) tag at the 5' end of the cDNA. The construct was digested with Kpn I, which cuts at the identical position in both LTRs, and the LTR-(Gal4)Lcrf1-LTR fragment was separated from vector sequences by agarose gel electrophoresis. I identified transgenic mice with 3 copies of the transgene in a head-to-tail orientation; functional LTRs were reconstituted in these animals. Nine lines of animals carrying either human Lcrf1 cDNA or Gal4-Lcrf1 cDNA transgenes were generated. These animals were mated to Lcrf1 +/- animals to produce Lcrf1 +/- Tg animals that expressed the human transgenes. Finally, these animals were mated to produce Lcrf1 -/- and Lcrf1 -/- Tg animals. Over 200 pups were born and no Lcrf1 -/- pups were found; however, several Lcrf1 -/- Tg embryos were identified at 11-dpc. These Lcrf1 -/- Tg embryos were normal in morphology and alive in utero. After 500 embryo were analyzed, all the results indicated that homozygous mutant (Lcrf1 -/- Tg) embryos containing the human Gal4/Lcrf1 cDNA were arrested during gestation. Subsequently, some embryos developed a severe anemia (at approximately 16.5-dpc). These homozygous mutants were pale and smaller in comparison with their normal littermates (Fig. 19). Hematocrits of viable mutant embryos (Lcrf1-/-Tg) obtained at 16.0-dpc were 20-25 compared to 45-50 in wild-type and heterozygous littermates. In addition, the peripheral blood smears and fetal liver touch preparations from homozygous mutants at this stage showed a high percentage of nucleated red blood cells derived from the yolk sac (primitive erythroid cells) (Fig. 20). A dramatic decrease in erythroid precursors and mature red blood cells is observed in fetal livers of Lcrf1 -/-Tg embryos. Many vessels and sinuses in the liver are almost completely devoid of erythroid cells (Fig. 19). These results suggest that LCR-F1 is also essential for normal hematopoiesis in developing embryos.

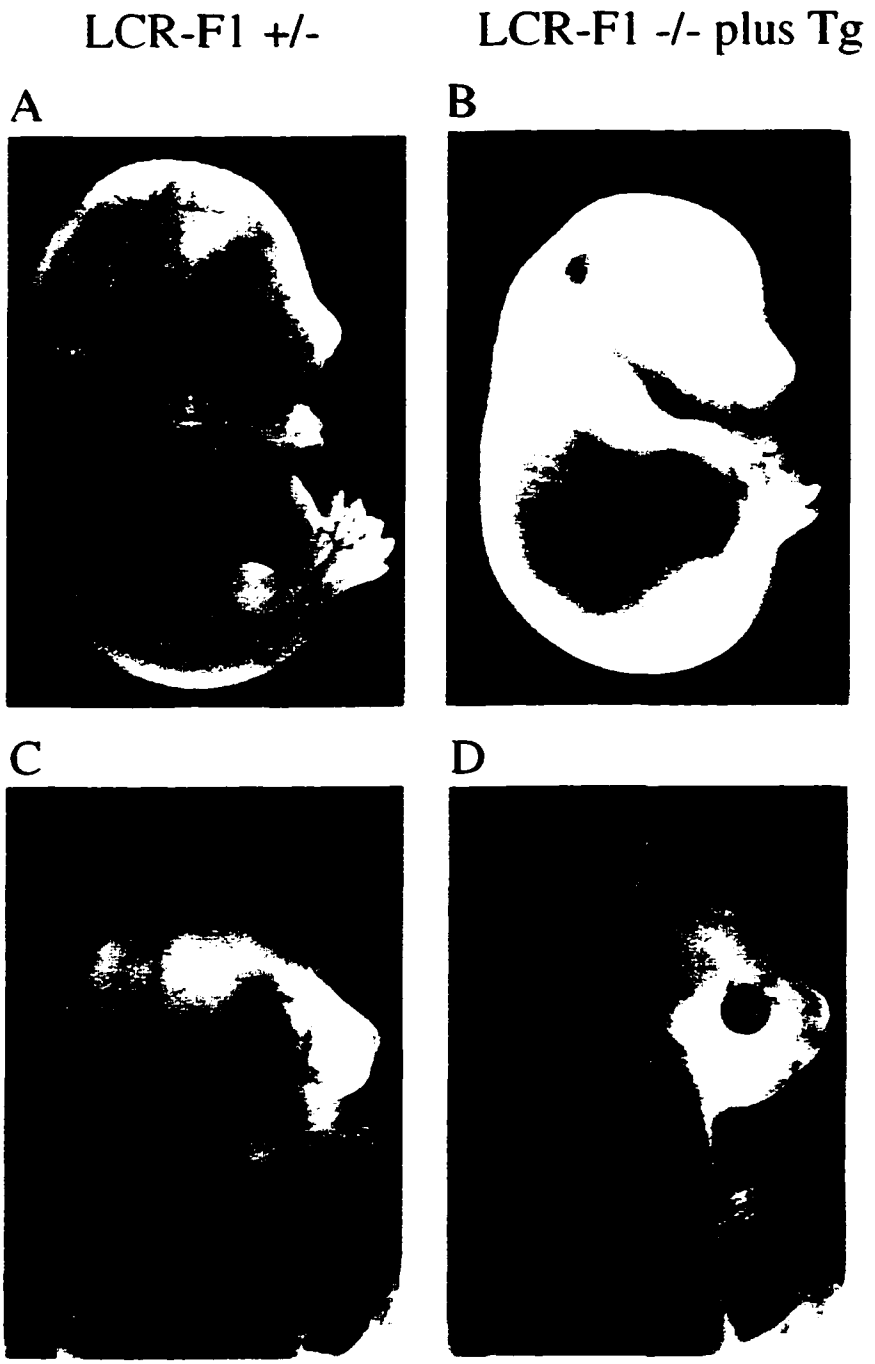


FIG. 19. **Partial rescue of LCR-F1 -/- embryos *in vivo*.** LCR-F1 +/- males were bred with animals carrying the MSCV/LCR-F1 transgene (Tg). Subsequently, *A* and *C*. LCR-F1+/- Tg progeny were bred with LCR-F1 +/- heterozygotes. *B* and *D*. LCR-F1 -/- Tg progeny survived the mesoderm block and then died of severe anemia at approximately 16.5 dpc. These results suggest an essential role for LCR-F1 after gastrulation.

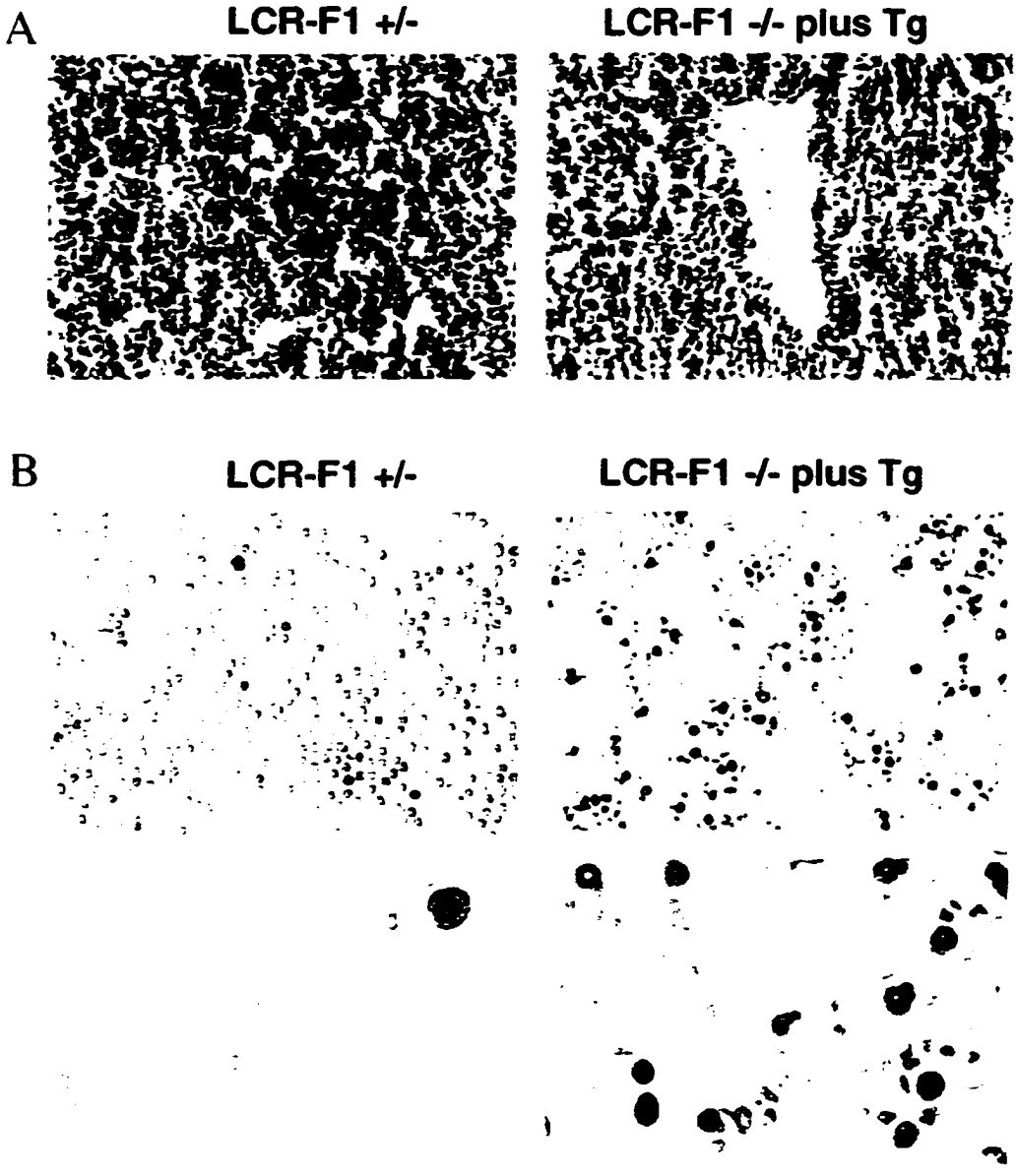


FIG. 20. **Fetal liver sections and blood smears from 16 d mLcrf1 +/- and m Lcrf1 +/-, hLcrf1 Tg.** The MSCV/LCR-F1 transgene (Tg) rescues Lcrf1 -/- mice from the mesodermal block at 6.5-dpc. However, development is subsequently arrested at approximately 16.0-dpc. This result demonstrates that LCR-F1 is essential for a second critical step in development. Lethality at 16.0-dpc is characterized by severe anemia. *A*, fetal liver sections above demonstrate a large reduction of hematopoietic progenitors in hepatic tissue and a decrease in the number of red blood cells in a hepatic sinus. *B*, blood smear of Lcrf1 -/- Tg mice contains numerous yolk sac derived (primitive), nucleated erythroid cells and a decreased number fetal liver derived (definitive) erythroid cells. The hematocrits of these animals are approximately 25. This result suggests a defect in definitive (fetal liver/ bone marrow) hematopoiesis.

Erythroid Progenitor Assays

In the partial rescue of mutant animals (Lcrf1 -/-) with human Lcrf1 transgene, the definitive red blood cells from the fetal livers are small, irregularly shaped, and pale compared to Lcrf1 +/- controls. To define the stage of erythroid differentiation that is affected by the Lcrf1 hypomorphic allele, I also quantitated erythroid progenitors (BFU-E and CFU-E) derived from 15.0-dpc fetal liver (Lcrf1 -/-Tg). The data demonstrates that the number of CFU-E and CFU-GEMM is decreased in Lcrf1 -/-Tg fetal liver cells compared with Lcrf1 +/- Tg controls. These results were confirmed by FACS analysis with the same fetal liver cells from Lcrf1-/-Tg mutant embryos. These data suggest that the number of hematopoietic stem cells is dramatically reduced in Lcrf1 -/-Tg mice.

We have previously demonstrated that Lcrf1-/- ES cells contribute to all cell lineages in adult animals derived from injection of these cells into wild-type host blastocysts. This result demonstrates that both the early defect (6.5-dpc) and the later defect (13.0 – 20.0-dpc) are non-cell-autonomous. Secreted factors or cell surface molecules from wild-type cells that rescue mutant cells at the later stage of development could be derived from hematopoiesis cells of stromal cells that support hematopoietic development. To distinguish between these possibilities, I transplanted Lcrf1 -/-Tg fetal liver cells into lethally irradiated C57/Bl6 mice. One mouse survived, and blood was collected 34 days after γ-irradiation. Blood smears were prepared, and hematocrits were measured. The hematocrit and blood smears of this animal is normal. Although more transplants need to be done, the results suggest that there is no intrinsic defect in hematopoietic cells. Apparently, wild-type host cells secrete factors that rescue Lcrf1 -/- Tg donor cells.

Hemoglobin Gene Regulation

The LCR not only functions in organizing the entire β-globin locus into an open chromosomal domain but also acts as a powerful enhancer. Individual globin gene family

members compete for interaction with the LCR and factors that influence this competition presumably determine developmental specificity. Our laboratory showed previously that linked human γ- and β-globin genes inserted down stream of LCR sequences (HS 1-5 γ–β) are correctly regulated in transgenic mice (19, 22). The human γ-globin gene was expressed in 10.5-day-old mouse yolk sac at the same time as the mouse βh1 gene. Human γ- to β-globin gene switching occurred at the time as mouse βh1- to mβ-globin gene switching. We predicted that deletion of 5' γ-promoter sequences, which normally activate γ-globin gene expression during development, would result in loss of γ competitive advantage and, therefore, to facilitate β-globin expression in embryonic yolk sacs of HS 1-5 γ-β transgenic mice. Tom Ryan and I made a series of 5' deletion constructs (-1400, -383, -202, -130, -72, and –54) of the human γ-globin promoter in a HS 1-5 γ–β cosmid (Fig. 21*A* and 22). We made transgenic mice lines with all these constructs. Mice that contained intact copies of transgenes were mated with non-transgenic animals, and embryos were removed at 10.5 days and 15.5 days of development. PCR was performed to determine which animals contained the transgene. RNA was extracted from positive embryos and from adult blood, and primer extensions were performed with radioactive-labeled primers to both mouse and human globin messages. Results from representative examples of each construct are shown in Fig. 23. Correct γ to β-globin gene switching was observed for LCR -1400 γ-β, LCR -383 γ-β and LCR -202 γ-β mice. Animals containing these constructs expressed human γ-globin mRNA early in development and human β-globin mRNA later in development when the endogenous mouse β-globin gene was activated. In contrast, the LCR -54 γ-β mice did not switch correctly (Fig. 23*A*). Little γ-globin mRNA was observed in yolk sac, fetal liver, or adult blood, and the human β-globin gene was expressed at high levels at all developmental stages. These results suggest that sequences between -202 and -54 of the human γ-globin promoter are required for correctly regulated expression of both γ- and β-globin genes. I further defined the sequences from –202 to –130, and a representative example is illustrated

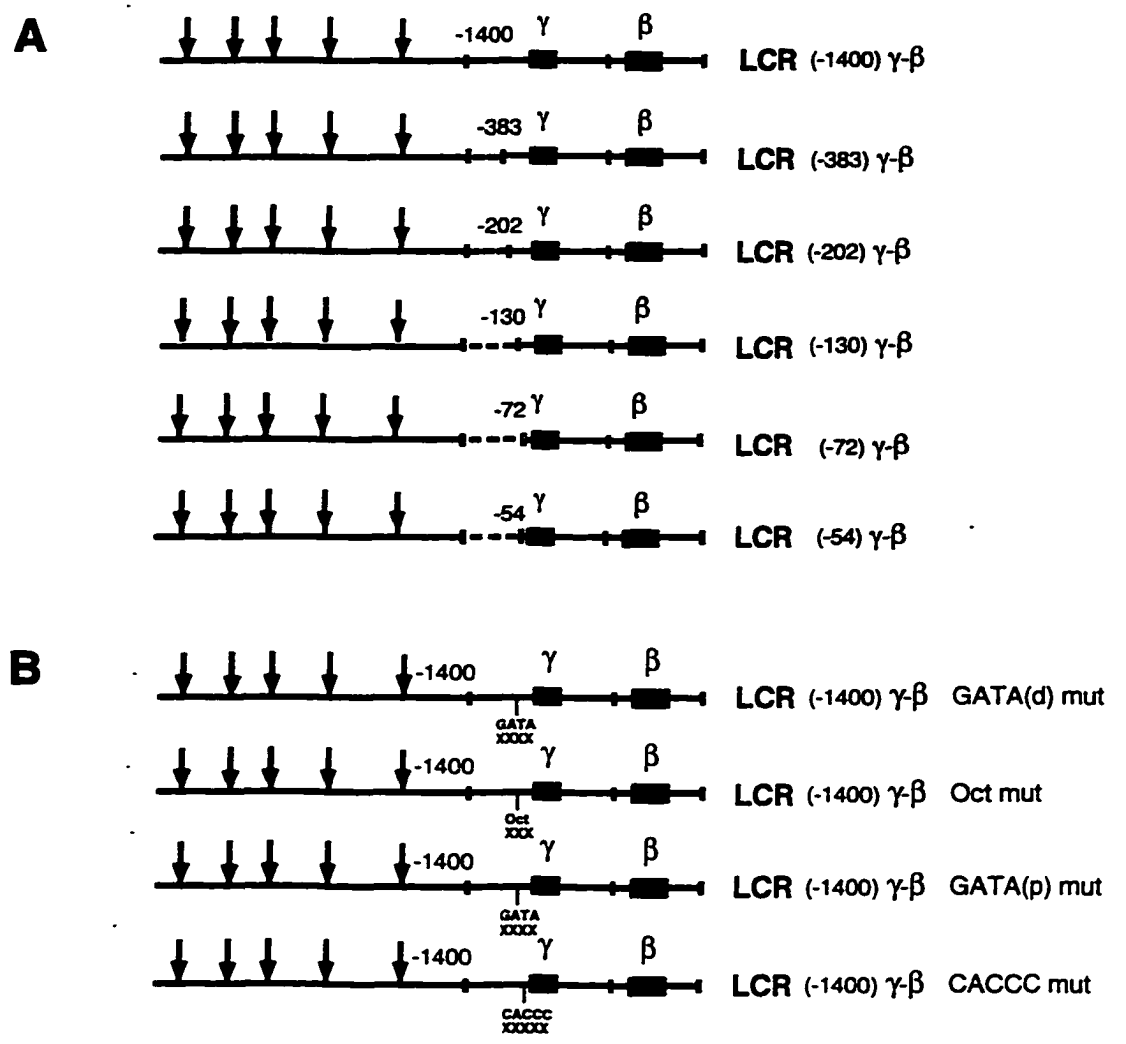


FIG. 21. **LCR γβ constructs used to define γ-globin gene promoter sequences that are involved in globin gene switching.**

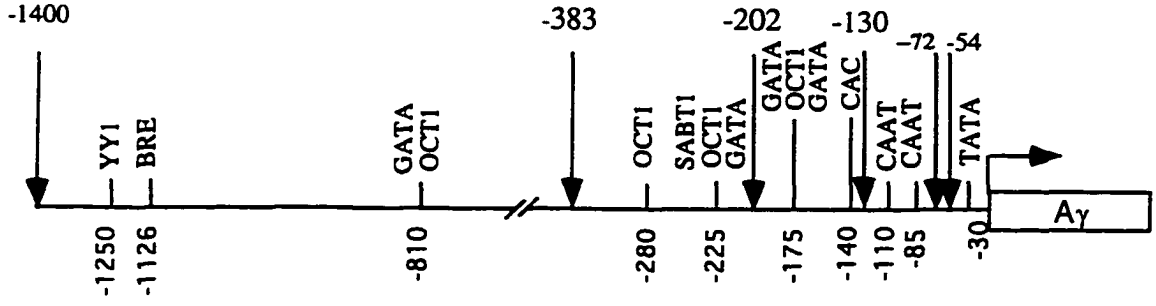
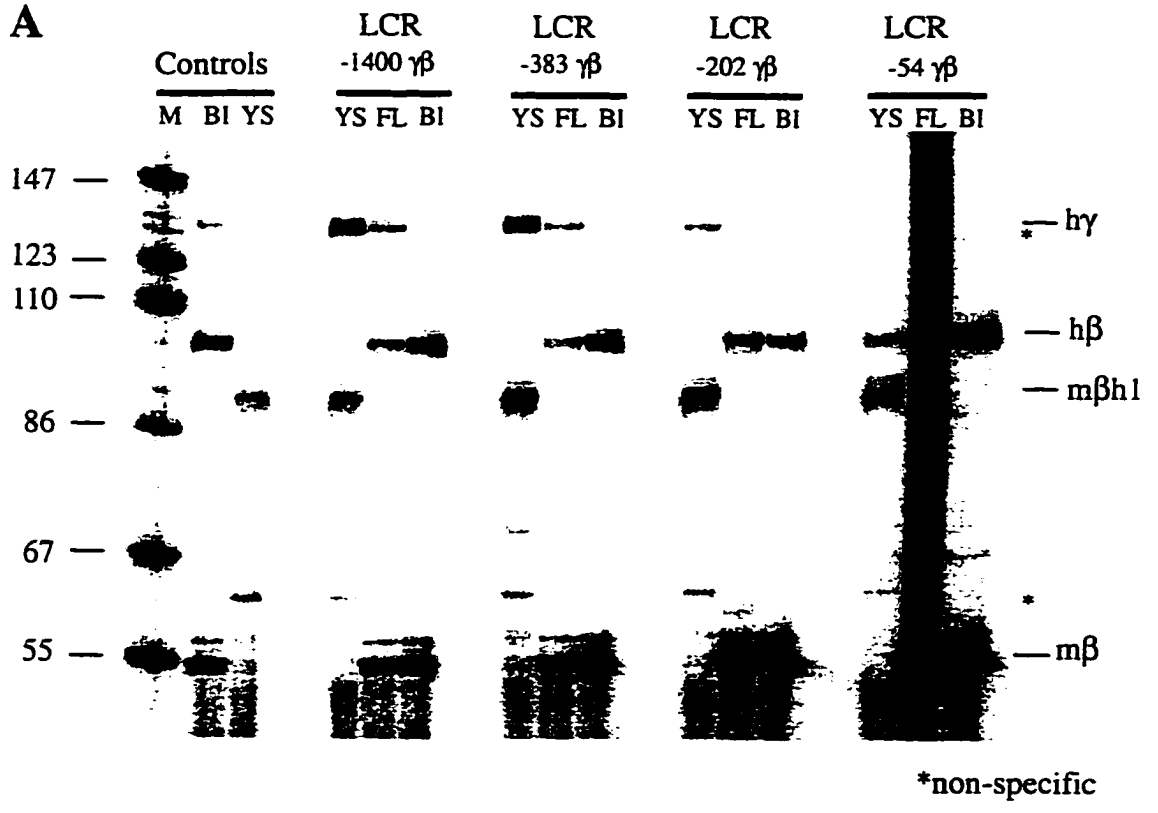
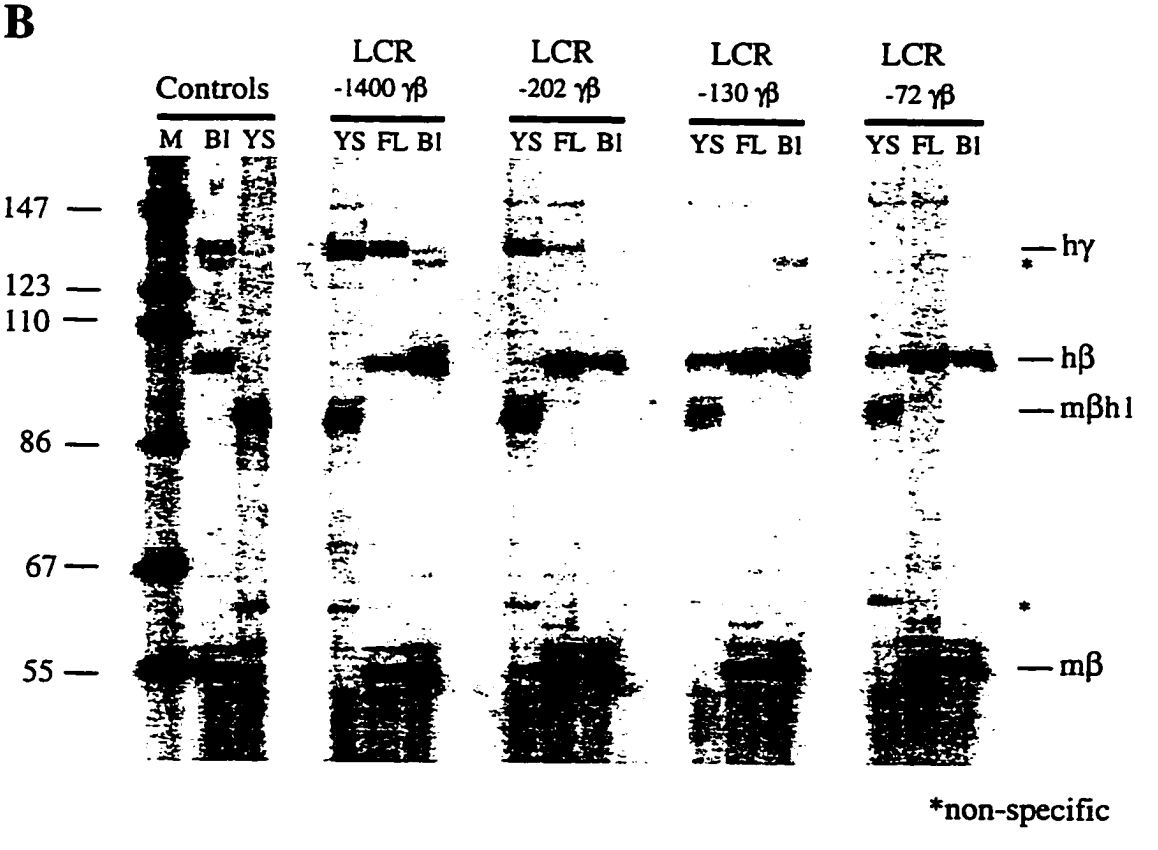


FIG. 22. **Schematic diagram of the human Aγ-globin gene promoter and transcription factor binding sites.**





in Fig. 23*B*. The β-globin gene is correctly regulated in the LCR -202 γ-β line included as a control. However, loss of developmental control is observed in the LCR -130 γ-β line; the β gene is expressed at a high level in the yolk sac. These results strongly suggest that sequences that bind γ-gene activating factors required for correct temporal control of γ- and β-globin genes are located between -202 and –130 (Fig. 23*B*).

Two GATA-1 sequences and an Oct-1 binding site are located around position –175, and the γCACCC box is located at –140. To assess the role of these sequences in γ- to β-globin gene switching, each sequence was scrambled by the substitution of 6 bp. Fig. 11 illustrates a primer extension assay of RNA from a representative animal containing each of these constructs. A minor band that migrates one bp shorter than correctly initiated human β-globin mRNA is observed in the distal GATA-1 mutant and Oct-1 mutant yolk sac samples. We are currently sequencing this band to determine whether it is an authentic β-globin transcript. No β-globin transcript is observed in the yolk sac sample of the proximal GATA-1 mutant. However, a high level of correctly initiated β-globin mRNA is observed in the yolk sac of the CACCC box mutant. Quantitative phosphorimager analysis demonstrates that the average hβ/hγ ratio for all 9 lines of CACCC box mutants is approximately 0.9. These results demonstrate that virtually all of the γ-globin gene's competitive advantage over β in the yolk sac is lost when the CACCC box is mutated (Fig. 24). We conclude that a positive regulatory factor, which is essential for the γ-globin gene to interact preferentially with the LCR at early developmental stages, is a CACCC box binding protein. This factor does not appear to be a basal transcription factor that is essential for γ-globin gene expression, but rather a regulatory protein that influences the delicate balance between the LCR and globin gene family members. Grosveld *et al.* have demonstrated elegantly that a dynamic equilibrium between γ- and β-globin gene (108) expression exists at the time of the fetal to adult switch. Expression of both genes in the same cell is the predicted result if equilibrium constants for LCR-γ and LCR-β interactions are equivalent when both fetal and adult regulatory factors are present. Our observation

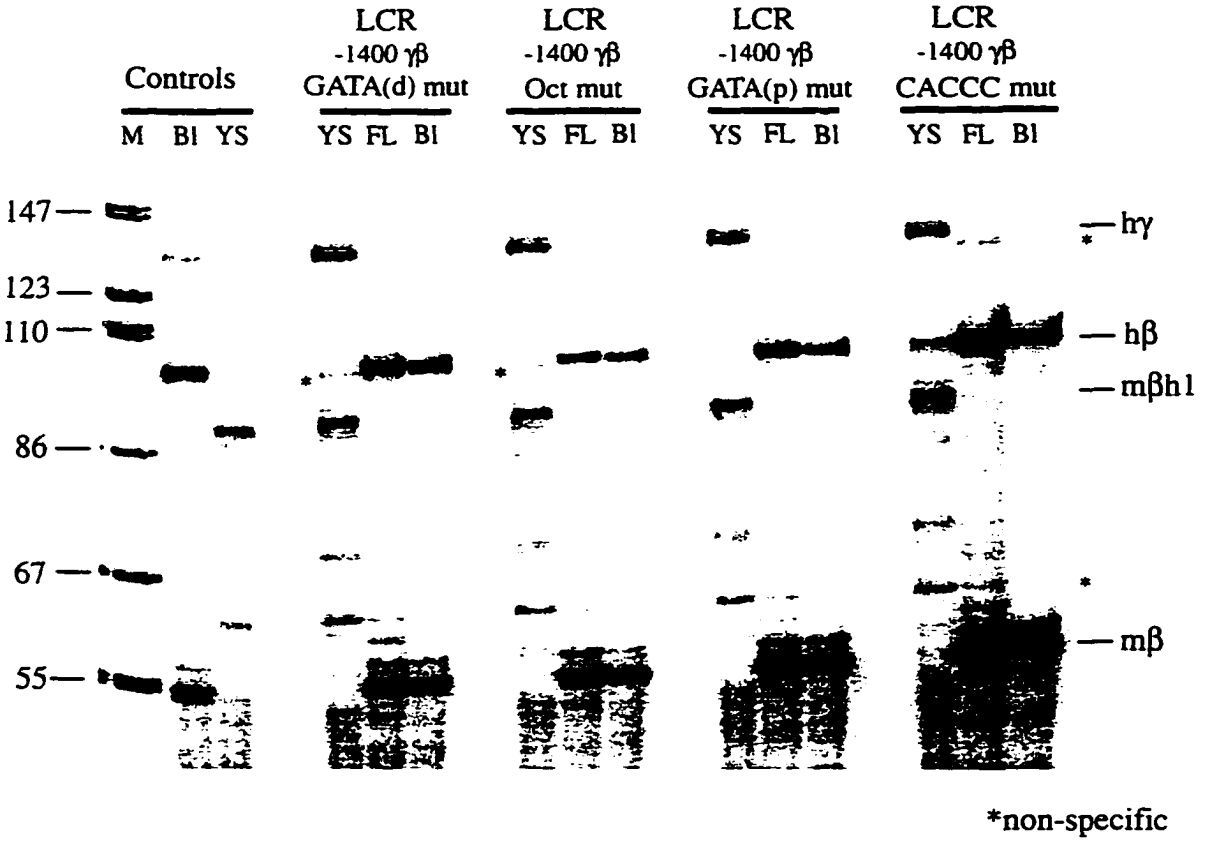


FIG. 24. **Primer extension analysis of RNA from 10.5-day-old yolk sac (YS), 15.5-day-old fetal liver (FL) and adult blood (BL) of LCR γβ transgenic mice containing mutations of transcription factor binding sites.** Correct globin gene switching is observed in the -1400 γβ GATA(d) mut, -1400 γβ Oct mut, and -1400 γβ GATA(p) mut mice. However, temporal specificity is lost in the -1400 γβ CACCC mut animals; β is expressed in the yolk sac. Quantitative, phosphorimager analysis demonstrates that the average hβ/hγ ratio for all 9 lines of CACCC box mutants is approximately 0.9. These results demonstrate that virtually all of the γ-globin gene's competitive advantage over β in the yolk sac is lost when the CACCC box is mutated. We conclude that a positive regulatory factor, which is essential for the γ-globin gene to interact preferentially with the LCR at early developmental stages, is a CACCC box binding protein.

that human γ- and β-globin genes are transcribed at roughly equivalent levels during a developmental stage that is normally restricted to embryonic/fetal globin gene expression demonstrates the importance of the CACCC box and CACCC box binding protein(s) for correct temporal regulation.

.

DISCUSSION

As described previously, I have performed the following experiments that define a novel *trans*-acting factor and a novel *cis*-acting sequence that are required for normal development and correct γ- to β-globin gene switching.

1. Gene targeting experiments demonstrated that LCR-F1 is essential for mesoderm formation.

2. Experiments involving Lcrf1 -/- ES cells injected into wild-type host blastocysts demonstrated that the -/- ES cells contributed to all cell lineages, suggesting that the mutation has a non-cell autonomous effect. I conclude that LCR-F1 controls the expression of a secreted factor that signals specific developmental pathways.

3. *In vitro* differentiation experiments of -/- ES cells demonstrated dramatic decrease in the level of β-globin mRNA. This result strongly suggests a role for LCR-F1 in globin gene expression.

4. Partial rescue experiments of the Lcrf1 -/- animals resulted in an anemia phenotype, which confirmed the role of LCR-F1 in globin gene expression.

5. I carried out γ-globin promoter mutation experiments in transgenic mice to define the *cis*-acting sequences important for γ- to β-globin switching. My results demonstrate a critical role for the CACCC element at –140 from transcription start site in γ- to β-globin switching.

Lcrf1 Mutation Inhibits Mesoderm Formation

Lcrf1 is the first example of a transcription factor that is essential for gastrulation of phenotypically normal late egg cylinder embryos. Mutations in other transcription factors either inhibit normal egg cylinder development or alter mesodermal patterning after

gastrulation has been initiated, as described above. Mutant embryos with homozygous Lcrf1 deletion develop normally to the late egg cylinder stage at ~6.5-dpc, but virtually all die before 7.5-dpc. The primitive streaks are absent, and the mesoderms are unable to be produced in the mutant embryos. These results demonstrate that LCR-F1 is essential for gastrulation in the mouse and suggest that mutation of this gene has a direct effect on mesoderm induction.

The nodal gene that encodes a TGF-β-like factor secreted by ectoderm cells at the egg cylinder stage is one possible target for LCR-F1 regulation in early development (109, 110). A distinct primitive streak and most mesoderm fail to form in homozygous nodal mutants (54, 55). However, when homozygous ES cells are injected into wild-type blastocysts, nodal -/- cells contribute to mesoderm (54). This is consistent with a role of NODAL protein as a secreted, signaling factor affecting mesodermal cell fate non-cell autonomously (54, 55, 111). As demonstrated in the results section, the non-cell-autonomous effect of the Lcrf1 mutation suggests that LCR-F1 regulates expression of genes encoding a secreted factor or a cell surface molecule capable of inducing mesoderm, and nodal is certainly a candidate because Lcrf1tm1uab mutation is non–cell-autonomous. However, some T is expressed in nodal mutants, and these embryos die later in development (8.0-9.0-dpc) than Lcrf1 -/-mutants, suggesting that other genes are regulated by LCR-F1. Another potential target is the Bmp4 gene. However, Lcrf1 -/- phenotype is distinct from the phenotypes of null mutation in these genes. Epiblast growth in the egg cylinder is retarded in Bmp4 and Bmpr1A null mutations, and Lcrf1tm1uab mutants appear normal. Alternatively, a novel factor that functions alone or conjunction with BMP4 and NODAL to induce mesoderm may be regulated by LCR-F1. Lcrf1 +/- and -/- embryos at 6.5-dpc could be used to identify novel factors by subtractive hybridization (112) or differential display (113) techniques. The identification of targets for LCR-F1 regulation may also be facilitated by isolating leucine zipper proteins that heterodimerize with LCR-F1. Relevant partners can be isolated by yeast two-hybrid

screens of 6.5-dpc cDNA libraries. Binding sites could be define with heterodimers, and, subsequently, target genes could be identified (114, 115).

The Role of LCR-F1 in Globin Gene Regulation

Why is LCR-F1 not required for globin gene expression? The simplest answer is that LCR-F1, NF-E2 p45, and NRF2 can compensate for each other's absence in mutant animals. All three related bZIP transcription factors are expressed in erythroid cells; however, no significant effect in globin gene expression is observed in any of the three null mutations *in vivo*. A globin phenotype may be observed in combinations of null mutations in these genes, but other CNC family members may also be involved in globin gene regulation. Factors that bind in LCR AP-1-like sites are critical for globin gene expression, but whether LCR-F1, NF-E2, NRF2, and/or other proteins function at these sites must be determined in further experiments.

In summary, a null mutation in the murine Lcrf1 gene is embryonic lethal without affecting globin expression or switching. Development is arrested at ~6.5-dpc. Remarkably, the Lcrf1*tm1uab* mutation does not appear to affect egg cylinder growth and organization. Mutant embryos fail to form a demonstrable primitive streak and are unable to produce mesoderm. The Lcrf1 mutation is not cell-autonomous. When homozygous mutant ES cells are injected into wild-type blastocysts to produce chimeras, mutant ES cells contribute to all cell lineages examined. All of the results described above demonstrate that LCR-F1 is essential for gastrulation in the mouse and suggest that this transcription factor controls expression of secreted factors or cell surface molecules that are critical for the earliest events in mesoderm formation.

Human Lcrf1 BAC and MSCV Rescue

Embryos lacking the Lcrf1 gene die at 7.0-dpc due to the absence of mesoderm. Embryonic lethality was rescued completely by the human Lcrf1 genome BAC transgene.

Homozygous mutant mice with the BAC transgene (Lcrf1-/-Tg) developed normally and survived to adulthood. Blood samples were collected from these animals, and the hematocrits and the peripheral blood smears were normal. In contrast, the human Lcrf1 cDNA transgene driven by MSCV-LTR was shut down in host cells and the rescue was never completed. All the Lcrf1-/-Tg mutants died during gestation or shortly after birth. There were no living Lcrf1-/-Tg pups among 286 pups examined. However, some of homozygous mutant embryos developed a severe anemia at approximately 16.5-dpc. The hematocrits reading of the homozygous mutants (Lcrf1-/-Tg) were reduced to 50 % compared with wild-type and heterozygous littermates. Lcrf1 -/-Tg embryos were pale compared with controls, and many of the vessels that were normally visible due to circulating red blood cells were invisible. Peripheral blood smears and fetal liver touch stains from 16.0-dpc Lcrf1-/-Tg embryos contain numerous nucleated red blood cells that were derived from the yolk sac (primitive erythroid cells). Fetal liver sections of Lcrf1-/-Tg embryos showed a dramatic decrease in erythroid precursors and mature red cells compared with littermate controls. Mature definitive erythroid cells, which are derived from the fetal liver, are small (microcytic), irregularly shaped (poikilocytic), and pale (hypochromic) compared with Lcrf1 +/- controls.

Because of the blood phenotype in some of partial rescued embryos, I would like to know that if Lcrf1 knockout affects erythroid precursors in definitive erythroid tissues. I compared the hematopoietic progenitor cells that were derived from the fetal livers of mutant and the wild-type littermates by *in vitro* differentiation assays. I quantitated erythroid progenitors (CFU-E) to define the stage of erythroid differentiation that was affected by the partial rescue. The results indicated that the numbers of CFU-E were decreased two fold. The numbers of colony-forming units-granulocyte, erythroid, macrophage, megakaryocyte (CFU-GEMM) were decreased four fold. The decrease in early hematopoietic progenitors was confirmed by FACS analysis. The number of Sca+, c-kit+, Lin- fetal liver cells was decreased eight fold. These results strongly suggest that

the Lcrf1 hypomorphic allele inhibits normal hematopoietic stem cell formation and/or proliferation.

Human Globin Gene Switching

The competition model predicts that deletion of sequences which normally activate γ-globin gene expression during development will stimulate the competition model predicts that deletion of sequences which normally activate γ-globin gene expression during development will stimulate the β-globin gene in embryos of HS 1-5 γ–β transgenic mice. The goal of these studies was to define the molecular mechanisms that control human hemoglobin switching during development. In order to define sequences regulated for switching, human γ- and β-globin genes must be introduced into a functional assay system in which their expression is correctly regulated. We have demonstrated correct tissue- and temporal-specific expression of human γ- and β-globin genes in transgenic mice. When large DNA fragments containing these genes and the LCR are injected into fertilized mouse eggs, the human γ-globin gene is expressed at a high level in early embryos. Between 12 and 14 days of development, a complete switch occurs; the human γ-globin gene is turned off, and the human β-globin gene is transcribed at a high level. This switch occurs at the same time as the switch in endogenous mouse globin gene expression. Interestingly, correct switching occurs only when both γ- and β-globin genes are linked on the same DNA fragment. These results suggested a competition model for hemoglobin genes switching in which each globin gene competes for interaction with LCR sequences located 50-kb upstream of the human β-globin locus (95). These interactions are presumably controlled by temporal-specific factors that bind within and/or surrounding the human γ- and β-globin genes and determine which gene will be enhanced by the powerful LCR. If competition between γ- and β-globin genes for LCR sequences is an important aspect of temporal control, one might expect that deletion of γ-promoter sequences that removed factor binding sites would result in γ-globin loss, and the competition advantage and would

not be correctly regulated. Although yolk sac, fetal liver, and bone marrow regulatory factors are presumably present in appropriate concentrations. By introducing site-directed mutations into the human Aγ-promoter region, the CACCC box at -140 was found to be critical for correct γ- to β-globin gene switching. Interestingly, Jane *et al.* (116) recently tested an HS 2 -54 γ-β construct in K562 cells. These cells presumably mimic the human fetal stage of development because they express human γ-globin but not β-globin mRNA. In this experiment correct regulation was observed with the HS 2 -54 γ-β construct; γ was expressed at a low level and β expression was not detected. Subsequently, a HS 2 -35 γ-β construct was tested. In this case γ expression was absent and β-globin gene expression was observed. The authors concluded that sequences required for correct γ- to β-globin gene switching are located from -54 to -35 in the γ-globin promoter. This is clearly not the case in our experiments, which suggest that critical regulatory sequence is located at –140 CACCC box. The reasons for this discrepancy are not clear, but the inclusion of all five HS sites in our constructs may be important for correctly regulated γ- and β-globin expression. Also, transgenic mice may provide a more stringent system than K562 cells for defining sequences required for γ- to β-globin gene switching.

Future Experiments

The role of LCR-F1 could be defined more easily if a complete block of Lcrf1 expression could be achieved after birth. Therefore, I also produced a conditional knockout of the gene. I produced over 20 lines of transgenic animals that contain the human Lcrf1 cDNA driven by the tetracycline responsive element (TRE) and a reverse tet-responsive transactivator (rtTA) clone or tet-responsive transactivator (tTA) as a control (Fig. 25). These animals were bred with Lcrf1 knockout animals to produce Lcrf1-/- Tg (rtTA) mice that are maintained on doxycycline. For developmental studies, Lcrf1-/- Tg (rtTA) homozygous animals will be mated and females will be treated with doxycycline in their drinking water. At 1-day intervals between 6.5-dpc and birth, doxycycline will be

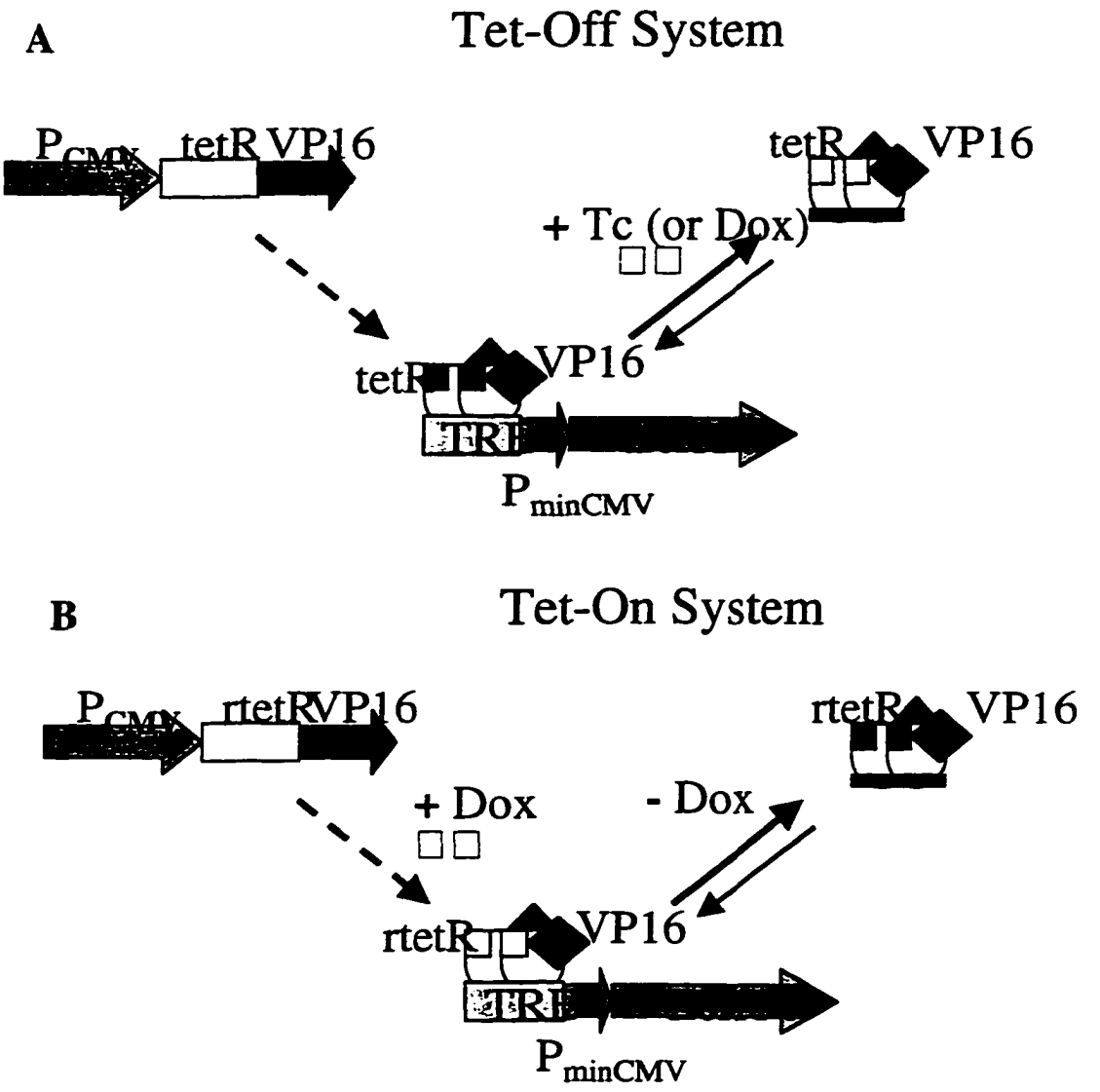


FIG. **25. Schematic of gene regulation in the Tet-Off and Tet-On systems.** *A*, transcription turned on by remove of Tc or Dox. The tet-responsive transcriptional activator (rTA) is a fusion of the wild-type Tet repressor (TetR) to the VP16 activation domain (AD) of herpes simplex virus. The TRE is located upstream of the minimal immediate early promoter of cytomegalovirus (PminCMV), which is silent in the absence of activation from the TRE. rTA binds the TRE and activates transcription of Lcrf1 in the absence of tetracycline (Tc) or Tc derivates such as doxycline (Dox). *B*, transcription turned on by addition of Dox. The reverse Tet repressor (rTetR) was created by four amino acid changes that reverse the protein's response to Tc and Tc derivates. Thus, the reverse tet-responsive transcriptional activator (rtTA) binds TRE and activates transcription in presence of Dox (Adapeted from CLONETECH Protocol #PT3001-1).

removed to inhibit Lcrf1 expression. My goal is to produce embryos that synthesize the minimum amount of LCR-F1 to overcome the mesoderm block so that a null phenotype for Lcrf1 can be defined after gastrulation. Both primitive (yolk sac) and definitive (fetal liver and bone marrow) hematopoiesis will be examined in these animals. These experiments should provide insights into LCR-F1 function at multiple stages of development.

To isolate the transcription factor(s) that bind to the γ CACCC box, I constructed a λgt11 cDNA expression library from C57/SJL 10.0-dpc yolk sac and screened one million clones with radioactively-labeled γ CACCC sequence as a probe (Fig. 26). I have isolated 16 positive clones that bind to the wild-type CACCC box but not to the mutant CACCC box probe. I sequenced all of these clones and analyzed their expression pattern by Northern blot hybridization. Many of the first set of clones I obtained were the ubiquitous transcription factors in the SP1 family, which have a zinc finger DNA binding domain. Although SP1 may be involved in globin gene switching, I predict that further screens of my yolk sac cDNA library will result in the isolation of yolk sac-specific factors that bind to the γ-globin CACCC box and regulate human globin gene switching. I will continue to screen the library until I find erythroid-specific, embryonic cDNA clones.

In addition, I have also searched the human fetal liver and spleen EST databases for erythroid specific cDNA clones containing EKLF-like zinc finger motifs. Several cDNAs that are present only in human erythroid tissues have been isolated. These cDNAs encode proteins that are candidates for γ CACCC box binding factors which regulate γ-globin switching. When these clones are isolated, I will analyze the clones in multi-tissue Northern blots and *in situ* hybridization to determine the tissue and developmental specificity of each clone. I will subclone the cDNA into expression vectors and cotransfected MEL cells with reporter plasmids containing HS2 γ/Luc-β/CAT. Authentic γ-globin gene regulatory factors should preferentially activate the γ-globin gene. The clones that preferentially activate the γ-globin gene in MEL cells will be over-expressed in adult erythroid tissue of transgenic mice containing LCR γ-β constructs. These

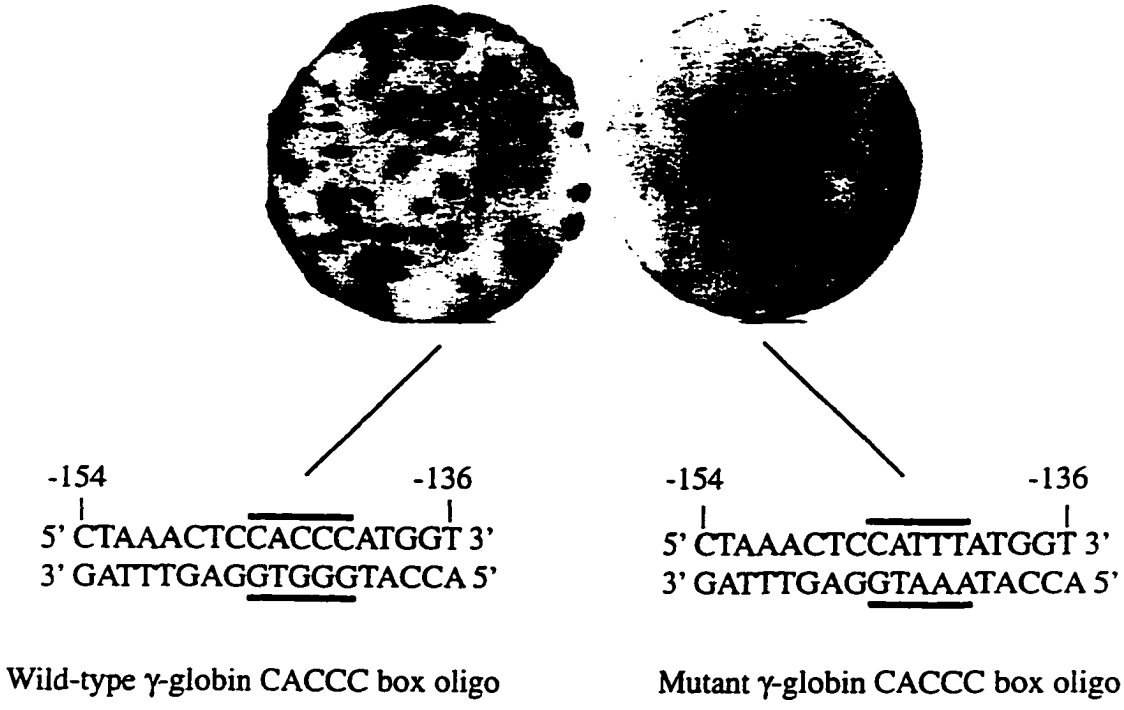


FIG. 26. **Yolk sac λgt11 library clone probed with wild-type and mutant γ-globin CACCC box oligos.** One of 15 yolk sac λgt11 library clones expressing a protein which binds specifically to the γ-globin gene CACCC box. Equal numbers of phage derived from a single clone were plated on two plates, and plaques were blotted onto nitrocellulose. After renaturation of proteins, the filters were probed with multime rized, double-stranded oligos containing wild-type or mutant γ-globin gene CACCC box sequences. The oligos were labeled by nick-translation.

experiments will determine whether the clones reactivate human γ-globin expression in adult erythroid cells. If they do reactivate γ-globin genes expression, I plan to introduce LCR β/γ-factor/β constructs into hematopoietic stem cells of our sickle mice in an attempt to reactivate the human γ-globin gene and correct the disease. These studies should provide a foundation for similar therapeutic strategies in human sickle cell patients.

REFFERENCES

1. Farmer, S. C., Sun, C. W., Winnier, G. E., Hogan, B. L., and Townes, T. M. (1997) *Genes Dev.* **11,** 786-798

2. Dickerson, R. E., and Geis, I. (1983) *Hemoglobin,* The Benjamin/Cummings Publishing Co., Inc., Menlo Park, CA 94025

3. Lauer, J., Shen, C. K., and Maniatis, T. (1980) *Cell* **20,** 119-130

4. Fritsch, E. F., Lawn, R. M., and Maniatis, T. (1980) *Cell* **19,** 959-972

5. Collins, F., and Weissman, S. (1984) *Progr. in Nucleic Acid Res. and Mol. Biol.* **31,** 315-463

6. Bunn, H. F., and Forget, B. G. (1986) *Hemoglobin: Molecular, Genetic, and Clinical Aspects.*, W.B. Saunders, Philadelphia, PA

7. Orkin, S. H., and Goff, S. C. (1981) *Cell* **24,** 345-351

8. Peschle, C., Mavilio, F., Migliaccio, G., Migliaccio, A. R., Russo, G., Mastroberardino, G., and Marinucci, M. (1985) *Prog. Clin. Biol. Res.* **191,** 383-396

9. Ryan T. (1990) Dissertation The University of Alabama at Birmingham.

10. Caterina, J. (1993) Dissertation The University of Alabama at Birmingham.

11. Jahn, C. L., Hutchison, C. A. D., Phillips, S. J., Weaver, S., Haigwood, N. L., Voliva, C. F., and Edgell, M. H. (1980) *Cell* **21,** 159-168

12. Leder, P., Hansen, J. N., Konkel, D., Leder, A., Nishioka, Y., and Talkington, C. (1980) *Science* **209,** 1336-1342

13. Leder, A., Weir, L., and Leder, P. (1985) *Mol. Cell. Biol.* **5,** 1025-1033

14. Tuan, D., and London, I. M. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 2718-2722

15. Fraser, P., Hurst, J., Collis, P., and Grosveld, F. (1990) *Nucleic Acids Res.* **18,** 3503-3508

16. Curtin, P., Pirastu, M., Kan, Y. W., Gobert-Jones, J. A., Stephens, A. D., and Lehmann, H. (1985) *J. Clin. Invest.* **76,** 1554-1558

17. Driscoll, M. C., Dobkin, C. S., and Alter, B. P. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 7470-7474

18. Grosveld, F., van, A. G., Greaves, D. R., and Kollias, G. (1987) *Cell* **51,** 975-985

19. Ryan, T. M., Behringer, R. R., Townes, T. M., Palmiter, R. D., and Brinster, R. L. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 37-41

20. Curtin, P. T., Liu, D., Liu, W., Chang, J. C., and Kan, Y. W. (1989) *Proc. Natl. Acad. Sci., USA* **86,** 7082-7086

21. Collis, P., Antoniou, M., and Grosveld, F. (1990) *EMBO J.* **9,** 233-240

22. Ryan, T. M., Behringer, R. R., Martin, N. C., Townes, T. M., Palmiter, R. D., and Brinster, R. L. (1989) *Proc. Natl. Acad. Sci. USA* **3,** 314-323

23. Fraser, P., Pruzina, S., Antoniou, M., and Grosveld, F. (1993) *Genes & Dev.* **7,** 106-113

24. Caterina, J. J., Ryan, T. M., Pawlik, K. M., Palmiter, R. D., Brinster, R. L., Behringer, R. R., and Townes, T. M. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 1626-1630

25. Talbot, D., Philipsen, S., Fraser, P., and Grosveld, F. (1990) *EMBO J* . **9,** 2169-2177

26. Talbot, D., and Grosveld, F. (1991) *EMBO J.* **10,** 1391-1398

27. Landschulz, W. H., Johnson, P. F., Adashi, E. Y., Graves, B. J., and McKnight, S. L. (1988) *Genes Dev.* **2,** 786-800

28. Gentz, R., Rauscher, F. J., III, Abate, C., and Curran, T. (1989) *Science* **243,** 1695-1699

29. Moi, P., and Kan, Y. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 9000-9004

30. Caterina, J. J., Donze, D., Sun, C.-W., Ciavatta, D. J., and Townes, T. M. (1994) *Neuleic Acids Res.* **22,** 2383-2391

31. Chan, J. Y., Han, X. L., and Kan, Y. W. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 11371-11375

32. Moi, P., Chan, K., Asunis, I., Cao, A., and Kan, Y. W. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 9926-9930

33. Andrews, N. C., Erdjument-Bromage, H., Davidson, M. B., Tempst, P., and Orkin, S. H. (1993) *Nature* **362,** 722-728

34. Andrews, N. C., Kotkow, K. J., Ney, P. A., Erdjument-Bromage, H., Tempst, P., and Orkin, S. H. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 11488-11492

35. Chan, J., Han, X., and Kan, Y. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 11366-11370

36. Hogan, B., Beddington, R., Costantini, F., and Lacy, E. (1994) *Manipulating the mouse embryo, a laboratory manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

37. Cumano, A., Dieterlen-Lievre, F., and Godin, I. (1996) *Cell* **86,** 907-916

38. Medvinsky, A. L., Samoylina, N. L., Muller, A. M., and Dzierzak, E. A. (1993) *Nature* **364,** 64-67

39. Medvinsky, A., and Dzierzak, E. (1996) *Cell* **86,** 897-906

40. Copp, A. J. (1995) *Trends Genet*. **11,** 87-93

41. St-Jacques, B., and McMahon, A. P. (1996) *Curr. Opin. Genet. Dev*. **6,** 439-444

42. Beddington, R. S. and Robertson, E. J. (1998) *Trends in Genetics* **14,** 277-284

43. Feldman, B., Poueymirou, W., Papaioannou, V. E., DeChiara, T. M., and Goldfarb, M. (1995) *Science* **267,** 246-249

44. Deng, C. X., Wynshaw-Boris, A., Shen, M. M., Daugherty, C., Ornitz, D. M., and Leder, P. (1994) *Genes Dev*. **8,** 3045-3057

45. Yamaguchi, T. P., Harpal, K., Henkemeyer, M., and Rossant, J. (1994) *Genes Dev.* **8,** 3032-3044

46. Rossant, J., and Vijh, K. M. (1981) *J. Embryol. Exp. Morphol.* **66,** 43-55

47. Jacobs-Cohen, R. J., Spiegelman, M., Cookingham, J. C., and Bennett, D. (1984) *Genet. Res.* **43,** 43-50

48. DeGregori, J., Russ, A., von Melchner, H., Rayburn, H., Priyaranjan, P., Jenkins, N. A., Copeland, N. G., and Ruley, H. E. (1994) *Genes Dev.* **8,** 265-276

49. Fassler, R., and Meyer, M. (1995) *Genes Dev.* **9,** 1896-1908

50. Stephens, L. E., Sutherland, A. E., Klimanskaya, I. V., Andrieux, A., Meneses, J., Pedersen, R. A., and Damsky, C. H. (1995) *Genes Dev.* **9,** 1883-1895

51. Furuta, Y., Ilic, D., Kanazawa, S., Takeda, N., Yamamoto, T., and Aizawa, S. (1995) *Oncogene* **11,** 1989-1995

52. Ilic, D., Furuta, Y., Kanazawa, S., Takeda, N., Sobue, K., Nakatsuji, N., Nomura, S., Fujimoto, J., Okada, M., and Yamamoto, T. (1995) *Nature* **377,** 539-544

53. George, E. L., Georges-Labouesse, E. N., Patel-King, R. S., Rayburn, H., and Hynes, R. O. (1993) *Development* **119,** 1079-1091

54. Conlon, F. L., Barth, K. S., and Robertson, E. J. (1991) *Development* **111,** 969-981

55. Collignon, J., Varlet, I., and Robertson, E. J. (1996) *Nature* **381,** 155-158

56. Winnier, G., Blessing, M., Labosky, P. A., and Hogan, B. L. (1995) *Genes Dev.* **9,** 2105-2116

57. Mishina, Y., Suzuki, A., Ueno, N., and Behringer, R. R. (1995) *Genes Dev.* **9,** 3027-3037

58. Beddington, R. S., Rashbass, P., and Wilson, V. (1992) *Dev. Suppl.* **1,** 157-165

59. Herrmann, B. G. (1992) *Ciba Found. Symp.* **165,** 78-86

60. Herrmann, B. G., and Kispert, A. (1994) *Trends Genet .* **10,** 280-286

61. Ang, S. L., and Rossant, J. (1994) *Cell* **78,** 561-574

62. Chen, W. S., Manova, K., Weinstein, D. C., Duncan, S. A., Plump, A. S., Prezioso, V. R., Bachvarova, R. F., and Darnell, J. E., Jr. (1994) *Genes Dev.* **8,** 2466-2477

63. Weinstein, D. C., Ruiz i Altaba, A., Chen, W. S., Hoodless, P., Prezioso, V. R., Jessell, T. M., and Darnell, J. E., Jr. (1994) *Cell* **78,** 575-588

64. Jones, S. N., Roe, A. E., Donehower, L. A., and Bradley, A. (1995) *Nature* **378,** 206-208

65. Gowen, L. C., Johnson, B. L., Latour, A. M., Sulik, K. K., and Koller, B. H. (1996) *Nat. Genet.* **12,** 191-194

66. Hakem, R., de la Pompa, J. L., Sirard, C., Mo, R., Woo, M., Hakem, A., Wakeham, A., Potter, J., Reitmair, A., Billia, F., Firpo, E., Hui, C. C., Roberts, J., Rossant, J., and Mak, T. W. (1996) *Cell* **85,** 1009-1023

67. Jensen, R. A., Thompson, M. E., Jetton, T. L., Szabo, C. I., van der Meer, R., Helou, B., Tronick, S. R., Page, D. L., King, M. C., and Holt, J. T. (1996) *Nat. Genet.* **12,** 303-308

68. Liu, C. Y., Flesken-Nikitin, A., Li, S., Zeng, Y., and Lee, W. H. (1996) *Genes Dev.* **10,** 1835-1843

69. Shawlot, W., and Behringer, R. R. (1995) *Nature* **374,** 425-430

70. Acampora, D., Mazan, S., Lallemand, Y., Avantaggiato, V., Maury, M., Simeone, A., and Brulet, P. (1995) *Development* **121,** 3279-3290

71. Matsuo, I., Kuratani, S., Kimura, C., Takeda, N., and Aizawa, S. (1995) *Genes Dev.* **9,** 2646-2658

72. Ang, S. L., Jin, O., Rhinn, M., Daigle, N., Stevenson, L., and Rossant, J. (1996) *Development* **122,** 243-252

73. Chen, Z. F., and Behringer, R. R. (1995) *Genes Dev.* **9,** 686-699

74. Swiatek, P. J., Lindsell, C. E., del Amo, F. F., Weinmaster, G., and Gridley, T. (1994) *Genes Dev.* **8,** 709-719

75. Faust, C., Schumacher, A., Holdener, B., and Magnuson, T. (1995) *Development* **121,** 273-285

76. Begley, C. G., Aplan, P. D., Denning, S. M., Haynes, B. F., Waldmann, T. A., and Kirsch, I. R. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 10128-10132

77. Valge-Archer, V. E., Osada, H., Warren, A. J., Forster, A., Li, J., Baer, R., and Rabbitts, T. H. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 8617-8621

78. Wadman, I., Li, J., Bash, R. O., Forster, A., Osada, H., Rabbitts, T. H., and Baer, R. (1994) *EMBO J.* **13,** 4831-4839

79. Pevny, L., Simon, M. C., Robertson, E., Klein, W. H., Tsai, S. F., D'Agati, V., Orkin, S. H., and Costantini, F. (1991) *Nature* **349,** 257-260

80. Simon, M. C., Pevny, L., Wiles, M. V., Keller, G., Costantini, F., and Orkin, S. H. (1992) *Nat. Genet.* **1,** 92-98

81. Tsai, F. Y., Keller, G., Kuo, F. C., Weiss, M., Chen, J., Rosenblatt, M., Alt, F. W., and Orkin, S. H. (1994) *Nature* **371,** 221-226

82. Robb, L., Lyons, I., Li, R., Hartley, L., Kontgen, F., Harvey, R. P., Metcalf, D., and Begley, C. G. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 7075-7079

83. Shivdasani, R. A., Mayer, E. L., and Orkin, S. H. (1995) *Nature* **373,** 432-434

84. Porcher C, S. W., Rockwell K, Fujiwara Y, Alt FW, Orkin SH. (1996) *Cell* **86,** 47-57

85. Visvader, J. E., Fujiwara, Y., and Orkin, S. H. (1998) *Genes Dev.* **12,** 473-479

86. Foroni, L., Boehm, T., White, L., Forster, A., Sherrington, P., Liao, X. B., Brannan, C. I., Jenkins, N. A., Copeland, N. G., and Rabbitts, T. H. (1992) *J. Mol. Biol.* **226,** 747-761

87. Kuo, C. T., Morrisey, E. E., Anandappa, R., Sigrist, K., Lu, M. M., Parmacek, M. S., Soudais, C., and Leiden, J. M. (1997) *Genes Dev.* **11,** 1048-1060

88. Molkentin, J. D., Lin, Q., Duncan, S. A., and Olson, E. N. (1997) *Genes Dev.* **11,** 1061-1072

89. Pandolfi, P. P., Roth, M. E., Karis, A., Leonard, M. W., Dzierzak, E., Grosveld, F. G., Engel, J. D., and Lindenbaum, M. H. (1995) *Nat. Genet.* **11,** 40-44

90. Miller, I. J., and Bieker, J. J. (1993) *Mol. Cell. Biol.* **13,** 2776-2786

91. Nuez, B., Michalovich, D., Bygrave, A., Ploemacher, R., and Grosveld, F. (1995) *Nature* **375,** 316-318

92. Perkins, A. C., Sharpe, A. H., and Orkin, S. H. (1995) *Nature* **375,** 318-322

93. Wang, Z. Q., Ovitt, C., Grigoriadis, A. E., Mohle-Steinlein, U., Ruther, U., and Wagner, E. F. (1992) *Nature* **360,** 741-745

94. Johnson, R. S., Spiegelman, B. M., and Papaioannou, V. (1992) *Cell* **71,** 577-586

95. Townes, T. and Behringer., R. (1990) *Trends Genet.* **6,** 219-223

96. Bradley, A. (1987) *Production and analysis of chimeric mice. In teratocarcinomas and embryonic stem cells: a practical approach*, 1st Ed., IRL Press, Oxford, UK , 113-152

97. Zhao, G. Q., Deng, K., Labosky, P. A., Liaw, L., and Hogan, B. L. (1996) *Genes Dev.* **10,** 1657-1669

98. Zhao, G. Q., and Hogan, B. L. (1996) *Mech. Dev.* **57,** 159-168

99. Nagy, A., and Rossant, J. (1993) *Production of completely ES cell-derived fetuses. In gene targeting, a practical approach*, 1st Ed., IRL Press, New York, NY. , 107-146

100. Mortensen, R. M., Conner, D. A., Chao, S., Geisterfer-Lowrance, A. A., and Seidman, J. G. (1992) *Mol. Cell. Biol.* **12,** 2391-2395

101. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., Struhl, K., Albright, L. M., Coen, D., Varki, M. A., and Janssen, K. (1994) *Current protocols in molecular biology*, John Wiley and Sons, New York, NY.

102. Wiles, M. V. (1993) *Methods Enzymol.* **225,** 900-918

103. Doetschman, T. C., Eistetter, H., Katz, M., Schmidt, W., and Kemler, R. (1985) *J. Embryol. Exp. Morphol.* **87,** 27-45

104. Pearse, A. G. E. (1985) *Pigments and pigment precursors*, Churchill Livingstone, New York, NY.

105. Stevens, A. (1990) *Pigments and minerals. In theory and practice of histoclogical techniques*, 3rd Ed., Churchill Livingstone, Avon. UK. , 245-268

106. Wilkinson, D. G., Bhatt, S., and Herrmann, B. G. (1990) *Nature* **343,** 657-659

107. Kispert, A., and Herrmann, B. G. (1994) *Dev. Biol.* **161,** 179-193

108. Wijgerde, M., Grosveld, F., and Fraser, P. (1995) *Nature* **377,** 209-213

109. Zhou, X., Sasaki, H., Lowe, L., Hogan, B. L., and Kuehn, M. R. (1993) *Nature* **361,** 543-547

110. Conlon, F. L., Lyons, K. M., Takaesu, N., Barth, K. S., Kispert, A., Herrmann, B., and Robertson, E. J. (1994) *Development* **120,** 1919-1928

111. Jones, C. M., Kuehn, M. R., Hogan, B. L., Smith, J. C., and Wright, C. V. (1995) *Development* **121,** 3651-3662

112. Sive, H. L., and St John, T. (1988) *Nucleic Acids Res.* **16,** 10937

113. Liang, P., and Pardee, A. B. (1992) *Science* **257,** 967-971

114. Kinzler, K. W., and Vogelstein, B. (1989) *Nucleic Acids Res.* **17,** 3645-3653

115. Wright, W. E., and Funk, W. D. (1993) *Trends Biochem. Sci.* **18,** 77-80

116. Jane, S. M., Ney, P. A., Vanin, E. F., Gumucio, D. L., and Nienhuis, A. W. (1992) *EMBO J.* **11,** 2961-2969

GRADUATE SCHOOL
UNIVERSITY OF ALABAMA AT BIRMINGHAM
DISSERTATION APPROVAL FORM
DOCTOR OF PHILOSOPHY

Name of Candidate Chiao-Wang J. Sun

Major Subject Biochemistry

Title of Dissertation Cis-Acting Sequences and Trans-Acting Factors Regulating Human Globin Gene Expression

I certify that I have read this document and examined the student regarding its content. In my opinion, this dissertation conforms to acceptable standards of scholarly presentation and is adequate in scope and quality, and the attainments of this student are such that he may be recommended for the degree of Doctor of Philosophy.

Dissertation Committee:

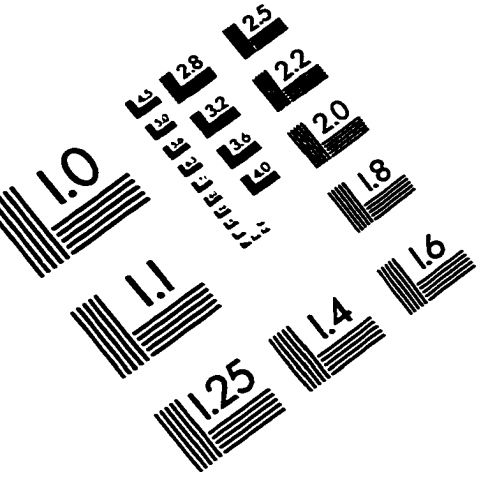
Name		Signature
Tim M. Townes	**, Chair**	Tim Townes
Jeffrey A. Engler		Jeffrey A. Engler
Peter J. Detloff		
Susan M.Ruppert		Susan M. Ruppert
Casey D. Morrow		

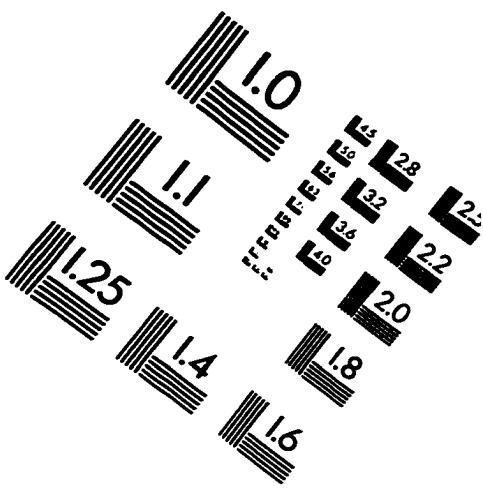
Director of Graduate Program Jamila I. Horabin

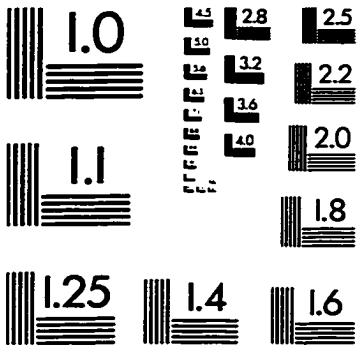
Dean, UAB Graduate School

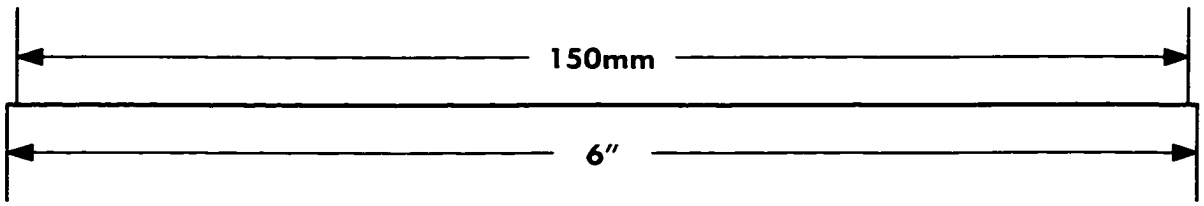
Date 12/8/98

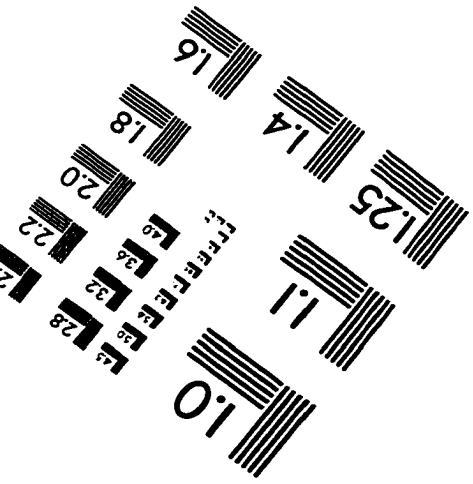
IMAGE EVALUATION
TEST TARGET (QA-3)











APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989