To the Graduate Council:

I am submitting herewith a dissertation written by Hsi-Hsien Lin entitled " Identification and Characterization of Genes Involved in Hematopoiesis Using *Cmyb* Mutant Mice as a Model System." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Biomedical Sciences.

Michael L. Mucenski, Major Professor

We have read this dissertation
and recommend its acceptance:

Accepted for the Council:

Associate Vice Chancellor and
Dean of the Graduate School

Identification and Characterization of Genes Involved in Hematopoiesis Using *Cmyb* Mutant Mice as a Model System

A Dissertation

Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Hsi-Hsien Lin

December, 1997

DEDICATION

To the memory of my grandmother, Mrs. Lin, Chang Zing (1920 - 1997)

ACKNOWLEDGEMENTS

I would like to express my appreciation to my major advisor, Dr. Michael L. Mucenski, for his guidance, advice, and assistance during the course of my research in his laboratory. The advice and suggestions provided by the members of my committee, Drs. Michael L. Mucenski, Dabney K. Johnson, Lisa J. Stubbs, Joyce Merryman, and Philip N. Bochsler, are greatly appreciated.

I am most grateful to the members of Mucenski laboratory, Drs. Peter R. Hoyt, Theresa A. Banks, and Mark Shannon as well as Bill Lee, Marilyn K. Kerley, and John Geisler for their assistance, helpful discussion and friendship.

I would like to thank Drs. M. Handel and E. Michaud for the use of cDNA libraries, Drs. J. Ihle and T. Bender for hematopoietic cell lines, IL-2, and IL-3, Drs. Eugene Rinchik and Lisa Stubbs for the interspecific backcross mapping panel, and Drs. Steve Kennel and Siamon Gordon for the use of monoclonal antibodies.

I would like to thank my family for their support and encourgement. Special thanks to my wonderful wife, Gin-Wen, for her always being there for me.

This research was supported by the Office of Health and Environmental Research, United States Department of Energy, under Contract No. DE-AC05-96OR22464 with Lockheed Martin Energy Research Corp.

ABSTRACT

Of key interest in the present study is to identify genes that may be involved in hematopoiesis. Here, a mutant mouse model system, in which the fetal hematopoietic development was impaired resulting from a mutated *Cmyb* proto-oncogene *in vivo*, is utilized. The expression levels of four genes, including *Gata1*, *Epor*, *Scl*, and *Nfe-2,* were found to be downregulated in the livers of *Cmyb* mutant fetuses in comparison to the phenotypically normal littermates. To identify novel genes that may be involved in hematopoiesis, the differential display technique was conducted to compare the differences in mRNA expression patterns in *Cmyb* wild type, heterozygous, and homozygous mutant fetal livers. Differentially expressed cDNA fragments were isolated and analyzed. A total of eight putative novel genes that are differentially expressed were identified. One of these genes, designated DD7A5-7, is expressed exclusively in cells of the myeloid lineage and was characterized as a member of a receptor family consisting of seven transmembrane segments. DD7A5-7 maps to mouse chromosome 17 and was shown to be the murine homolog of human *EMR1* gene. DD7A5-7 was therefore renamed *Emr1*. The amino acid sequence encoded by *Emr1* was found to be identical to that encoded by the *F4/80* gene, whose gene product is expressed specifically on the cell surface of mouse macrophages. The biological function of the Emr1/F4/80 protein was evaluated using a gene targeting approach. *Emr1* null mutant mice were generated, and no histological abnormality was observed. These results indicate that the Emr1/F4/80 protein is not required for mouse embryonic or hematopoietic development.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Prologue

Our understanding of mammalian hematopoiesis has benefitted greatly from the study of mouse embryonic and fetal hematopoietic development. Murine embryonic and fetal hematopoietic development is characterized by rapid and dramatic changes in the sites of erythropoiesis, erythroblast morphology, and hemoglobin synthesis (1). The mouse hematopoietic system is derived from the mesodermal germ layer, which is formed in primitive-streak-stage embryos at 6.5 days postcoitum (dpc). Blood islands first appear in the extraembryonic mesoderm at 7.5 dpc and hematopoietic progenitors are first detected in the visceral yolk sac mesoderm of embryos at 8 dpc (2, 3). By mid-8 dpc, nucleated, primitive red blood cells are visible in the vasculature of the yolk sac. These primitive red blood cells enter the primitive circulation at approximately 8.5 dpc when the embryonic and extraembryonic vessels are linked.

The yolk sac remains as the primary hematopoietic organ until approximately 12 dpc. Beyond this stage of development, yolk sac hematopoiesis declines and the fetal liver becomes the predominant hematopoietic organ for the remainder of fetal life. Erythropoiesis finally shifts to bone marrow around the time of birth. Yolk sac-derived erythrocytes, termed primitive erythrocytes, are nucleated, while those produced later in development, termed definitive erythrocytes, are non-nucleated. The changing patterns of globin proteins synthesized during development, termed globin switching, have been demonstrated in all vertebrate species (1). In mice, three embryonic globin tetramers, designated EI (x2y2), EII (α2y2), and EIII (α2z2) are found (4). The globin proteins x, y ,and z are the products of the genes ξ, βH1, and εy^2, respectively (5). In addition to these embryonically expressed genes, there are two very similar adult β-globin cluster genes, β major (β^{maj}) and β minor (β^{min}), and two adult α-globin cluster

genes (α1 and α2). Primitive erythrocytes arising from the yolk sac contain largely embryonic hemoglobins (EI, EII, and EIII) and a small quantity of adult hemoglobins (6). Definitive erythrocytes formed later in the fetal liver, spleen, and bone marrow contain only adult hemoglobin ($\alpha_2\beta_2^{maj}$, $\alpha_2\beta_2^{min}$) (7).

Differentiation of hematopoietic cells involves the progressive restriction of developmental potential as pluripotent hematopoietic stem cells give rise to multipotential progenitors that subsequently mature along single lineages. Functional hematopoietic assays have been developed to determine the lineage potential of hematopoietic progenitors and stem cells. These assays have demonstrated that the most committed progenitors (colony forming units - culture, CFU-C), including colony forming units - erythroid (CFU-E), colony forming units - granulocyte and macrophage (CFU-GM), and colony forming units - megakaryocyte (CFU-Meg), are near the final stages of cell expansion and lineage differentiation. The cells of CFU-C can lead to mature blood cell production in 5-14 days in *in vitro* culture. More primitive progenitors (colony forming units - spleen, CFU-S), some of which can self-renew, generate many committed progenitors and lead to mature hematopoietic cells after 7-14 days in *in vitro* culture (8, 9).

Although the mechanisms responsible for decision-making in hematopoietic development are not well defined, it is likely that they involve two distinct, but interrelated pathways, one involving nuclear regulatory proteins that directly control and coordinate gene expression, and a second involving signal transduction from growth factors through their cognate surface receptors (10). Recent progress in genetic analysis using mouse mutants generated by gene targeting in embryonic stem (ES) cells and transgenic mice containing marker genes have provided valuable information on hematopoietic development in mice (11-14). Data from these mouse models and from functional assays have suggested that there are two types of stem

cell activity in the developing mouse embryo, primitive and definitive (2). Although the origins of the two activities in the mouse have not been unequivocally demonstrated, it was known that the primitive hematopoietic stem cells function differently from the definitive hematopoietic stem cells. For example, the hematopoietic stem cells of the adult bone marrow require many weeks for full multilineage engraftment, while embryonic hematopoiesis is much more rapid, with the primitive erythrocytes appearing only days after mesoderm formation (2). With precise identification of the molecular events that lead to the hematopoietic defects, the genetically engineered mutant mouse models can be used to pinpoint the mechanism whereby hematopoietic stem cells develop and differentiate.

One of the mutant mouse models used for studying hematopoietic development is the *Cmyb* homozygous mutant (-/-) mice (12). The *Cmyb* proto-oncogene is the cellular homolog of the *v-myb* oncogene that was initially identified as the oncogenic agent in two replication-defective avian acute leukemia viruses, avian myeloblastosis virus (AMV) and E26 leukemia virus (15). Both viruses are able to induce leukemia *in vivo* and transform cells of erythroid (E26) and/or myeloid (E26, AMV) lineages *in vitro* (16, 17). The precise structures of the *v-myb* oncogenes of AMV and E26 virus have been determined by molecular cloning and nucleotide sequencing (18-20). The *v-myb* gene of AMV encodes a 45 KDa product which lacks 71 and 199 amino acids from the amino- and carboxy-terminal ends of the Cmyb gene product, respectively. The *v-myb* gene of E26 is expressed as part of a 135 KDa gag-Myb-Ets fusion protein, in which the v-Myb portion lacks 80 and 278 amino acids from the amino- and carboxy-terminal ends of the Cmyb protein, respectively. Taken together, these data indicate that the oncogenic activation of the *Cmyb* gene by virus transduction is accomplished by extensive truncation of the Cmyb protein. Diseases involving cells of the hematopoietic lineages such as B lymphomas and plasmacytoid

lymphosarcomas are also associated with the expression of an altered Myb gene product (21-23).

The Myb gene product is a 75 kDa nuclear protein that contains three functional domains : a DNA-binding domain, a transactivating domain, and a negative regulatory domain (22, 23). The DNA binding domain, located near the amino terminus of the protein, consists of three imperfect 50-52 amino acid repeats that are designated R1, R2, and R3. The specific target for DNA binding is the consensus nucleotide sequence PyAAC(G/T)G (24). Deletion and linker insertion analysis of *Cmyb*, as well as *v-myb*, have revealed a 50-residue trans-activation domain located downstream of the DNA-binding domain. In addition, a negative regulatory domain has been defined, which is carboxy-terminal to the activator region (25-27). A potential α-helical 'leucine zipper'-like structure, which is assumed to mediate protein-protein interaction, has been localized to the negative regulatory domain of Myb protein. A similar structure has been identified in other DNA-binding proteins including *Cmyc*, *Cfos*, and *Cjun* (24). Additionally, a major phosphorylation site has been mapped near the amino terminus of the protein. Phosphorylation of this site by casein kinase II down-modulates the binding of immunoenriched Myb to a Myb response element (MRE) *in vitro*. Loss of this region in most oncogenically activated Myb proteins suggests a possible regulatory role of this phosphorylation site.

The *Cmyb* gene is highly expressed in immature cells of the myeloid, erythroid and lymphoid lineages and the level of expression decreases markedly as these cells differentiate (28, 29). Expression of *Cmyb* proto-oncogene is also found in acute myelogenous leukemia, chronic myelogenous leukemia, acute lymphocytic leukemia, F9 teratocarcinoma stem cell line, breast and colon carcinoma cells and chicken embryo fibroblasts (CEF) (15). Down-regulation of *Cmyb* proto-oncogene during the differentiation of a myeloid tumor cell line suggested a role in either initiation of growth

arrest or differentiation (28). Antisense oligonucleotides to *Cmyb* have been used to inhibit endogenous Myb protein expression, resulting in the inhibition of growth of various human myeloid leukemia cell lines (30-33). Constitutive expression of exogenously introduced *Cmyb* cDNA blocks chemical- or erythropoietin-induced terminal differentiation in mouse erythroleukemia (MEL) cell lines (34-36). However, an alternatively spliced *Cmyb* mRNA encoding an amino- and carboxy-terminally truncated Myb protein which retains the DNA-binding domain and the nuclear localization signal but lacks the activation domain has been shown to induce MEL cell differentiation (37). These studies suggest that *Cmyb* gene product is involved in the cellular proliferation and/or differentiation of hematopoietic cells.

To study the biological function of *Cmyb* gene *in vivo*, the *Cmyb* proto-oncogene has been disrupted by targeted mutagenesis in murine embryonic stem (ES) cells (12). Germline chimeric animals were generated. Heterozygous (+/-) progeny from germline chimeric mice appear to be phenotypically normal. However, homozygous mutant (-/-) mice from heterozygous matings died at approximately 15.5 dpc due to anemia. The *Cmyb* -/- mice were phenotypically indistinguishable from wild type (+/+) and heterozygous (+/-) littermates up to day 13.5 of gestation, when most of the major phases of organogenesis have been completed. However, by day 15 of gestation, the *Cmyb* -/- fetuses are very pale in color, though grossly normal in form and size. Histological analysis showed that yolk sac erythropoiesis is unaffected, but fetal hepatic erythropoiesis is reduced significantly in *Cmyb* -/- fetuses. In addition, β major globin gene expression is highly reduced in the *Cmyb* -/- fetal livers. Further studies using *in vitro* methocellulose assays, which determine the progenitor potential of various hematopoietic lineages, show that erythropoiesis and myelopoiesis in *Cmyb* -/- fetuses are dramatically reduced in comparison to phenotypically normal

littermates (38). It is concluded that the *Cmyb* gene is required for normal murine fetal hepatic hematopoiesis (12).

Since the *Cmyb* gene product contains DNA-binding and transcriptional activation domains, and since Myb consensus DNA-binding sites are found in the promoter regions of hematopoietic and non-hematopoietic genes (39-41), Cmyb is strongly implicated as a transcriptional regulator. Indeed, several genes have been identified that are transcriptionally regulated by Myb. The *Mim1* gene, whose gene product is found in the granules of normal promyelocytes, was identified by using differential hybridization screening in cells transformed by a temperature sensitive mutant of the oncogene (39). Other examples of Myb modulated genes include *Gata1* (42), *Cd34* (43), and *Cd4* (44). These observations suggest that the *Cmyb* proto-ongene may play an essential role in hematopoietic development by regulating the expression of other genes.

Given that the Myb protein can transactivate various target genes and that fetal hepatic hematopoiesis is severely affected in the *Cmyb* -/- fetuses, *Cmyb* -/- mice might be a useful model system for the identification of additional Myb target genes or genes that are responsible for the anemic phenotype. Furthermore, the disturbed fetal hematopoiesis in *Cmyb* -/- fetuses also provides us an excellent opportunity to identify genes that may be involved in this important developmental process.

Identification of genes whose expression levels are altered by the inactivation of *Cmyb* gene *in vivo* was therefore conducted in the present study. Numerous gene cloning techniques, including subtractive hybridization (45, 46) and differential hybridization (39), have been developed to identify genes whose expression patterns were altered due to the physiological changes, such as overexpression of oncogenes and/or down-regulation of tumor suppressor genes in cancer cells. The subtractive hybridization technique uses a cDNA library from cells of interest and

hybridizes the cDNAs with RNA from the cells whose genes were not being selected. Unhybridized cDNAs that represent unique mRNAs of the cells of interest are analyzed further, while the cDNA/RNA hybrids are removed (45). Similarly, the differential hybridization technique compares the differences in hybridization intensity between two normalized cDNA libraries and the unique cDNAs are isolated for further analysis. Although potentially powerful, these techniques are laborious and technically difficult.

The recent invention of the differential display technique has simplified the identification of differentially expressed genes (47). The method is based upon comparisons of mRNAs expressed in two or more cell populations, by running their reverse transcribed PCR products on sequencing gels in adjacent lanes. The differentially expressed cDNA fragments can be easily detected, eluted out of the gels, and amplified by PCR to use as probes in Northern blots for verification. By using four 3' oligo(dT) primers ($T_{12}MA$, $T_{12}MC$, $T_{12}MG$, and $T_{12}MT$) in the reverse transcription reactions, the differential display technique divides the expected 10,000-15,000 different mRNAs into four groups. The 5' arbitrary decamer primers used in this technique hybridize with complementary sequences located at varying distances from the 3'-end of each cDNA strand. The carefully chosen primers and PCR reaction conditions of differential display analysis usually result in cDNA fragments of 100-600 base pairs, which are easily detected on DNA sequencing gels. Comparisons of the differentially displayed RT-PCR products of different samples in parallel offers an advantageous selection strategy for isolating differentially expressed genes. Since its introduction, differential display has been successfully used to identify a number of genes, including a candidate tumor suppressor gene (48) and a transcriptional target of the p53 protein (49). Differential display is thus the method of choice in this study to unravel the molecular events resulting from the inactivation of *Cmyb* gene *in vivo*.

The disruption of genes in mouse ES cells (gene targeting) has been widely used in recent years to study the biological functions of specific genes *in vivo* (50, 51). Gene targeting in ES cells was made possible by the isolation of pluripotent ES cells from mouse blastocysts (52, 53), and by the ability of the cellular recombination machinery to mediate homologous recombination (54-56). Pluripotent ES cells are derived from the inner cell mass (ICM) of mouse blastocysts and retain their ability to contribute to all mouse tissues after they are reintroduced into a blastocyst (52, 53). The pluripotency of ES cells results in the generation of germ-line chimeric mice, which can transmit the genome of ES cells to the offspring. For example, when pluripotent ES cells that were derived from a mouse homozygous for the black coat color allele were introduced into a blastocyst that was isolated from a mouse homozygous for the albino allele, the resulting chimeric mice would be composed of cells of both genotypes throughout the body, including germ cells. The chimeric mice thus would be able to transmit the black coat color allele of the original ES cells through germ line to generate black offspring when they were mated with albino mice (57).

The ability of cells to mediate homologous recombination between endogenous chromosomal DNAs and exogenously introduced DNAs suggested that one can easily manipulate the genome of cells. Therefore, desired modification of the DNA structure of specific genes, such as an insertion, deletion, or substitution is possible. Additionally, selectable exogenous genes, such as the neomycin resistant gene (Neo^r), can be introduced into cells, enhancing the ability to select for targeted cells. The application of gene targeting techniques in ES cells provides the means to study the functions of specific genes *in vivo* and to generate mice of any desired genotype to serve as potential animal models for a variety of human diseases (11, 51). It is

therefore of interest in our study to adapt this technique to study the biological function(s) of the differentially expressed genes that are identified.

The present study was designed to use *Cmyb* -/- mice as a model system in an attempt to identify genes that may be involved in hematopoiesis, and to characterize the function of the identified gene(s). The first issue to be addressed is to verify whether *Cmyb* -/- mice could be used as a model system for the purpose of this study, and to develop the differential display technique. Chapter 1 and 2 describe the validity of using *Cmyb* -/- mice as a model system and the differential display technique for the identification of novel genes that may be involved in hematopoiesis. Chapter 3 pertains to the second issue of the present study, that is to characterize the function of the novel differentially expressed gene(s). In order to address this question, a null homozygous mutant mouse strain with an inactivated *Emr1* gene was generated by gene targeting in ES cells. Finally, a new screening strategy for selecting targeted ES cell clones, developed in the course of this study, is presented in Chapter 4.

Chapter 1 Alteration of gene expression of known and putative novel genes in *Cmyb* wild-type, heterozygous, and homozygous mutant fetal livers

Abstract

The expression level of 25 genes previously shown to be expressed in the mouse fetal liver at 14.5 days postcoitum (dpc) was determined in *Cmyb* wild-type (+/+), heterozygous (+/-), and homozygous mutant (-/-) mice by the reverse transcriptase-mediated polymerase chain reaction (RT-PCR) and Northern blot analysis. These studies showed that four of these genes, *Gata1*, *Epor*, *Scl*, and *Nf-e2* were expressed at much lower levels in the livers of *Cmyb* -/- fetuses in comparison to the *Cmyb* +/+ and +/- fetuses at 14.5 dpc. The differential expression patterns of these hematopoietic genes suggest that *Cmyb* -/- mice could be used as a model system for the identification of genes that may be involved in hematopoiesis. To test this hypothesis, differential display was successfully carried out utilizing total liver RNA from *Cmyb* +/+, +/-, and -/- fetuses at 14.5 dpc. Eight putative novel genes were identified. The chromosomal location and the expression pattern of each differentially expressed gene were determined by chromosomal mapping and Northern blot analysis, respectively. While five of these eight putative novel genes were shown to be expressed in all hematopoietic cell lineages tested, the other three genes had a more restricted pattern of expression. These results demonstrate the validity of using the unique *Cmyb* -/- fetuses and differential display to identify novel genes involved in hematopoiesis.

Introduction

The Myb proto-oncogene gene product is a 75 kilodalton (kDa) nuclear protein that contains three functional domains : a DNA-binding domain, a trans-activation domain, and a negative regulatory domain (15). Three imperfect tandom repeats (R1, R2, R3) of 51-52 amino acids that are highly conserved in various Myb-related proteins are located in the DNA-binding domain, which is in the amino-terminus of the Myb protein (22). A concensus nucleotide sequence, PyAAC(G/T)G, has been identified as the specific Myb DNA-binding target site (24). Through the deletion and linker insertion analyses, the ability of Myb protein to trans-activate gene expression has been defined to a trans-activation domain of 50 amino acid residues located immediately adjacent to the DNA-binding domain (25, 27). In addition, a negative regulatory domain has also been identified, which is located 3' to the trans-activation domain (26). The expression of several genes, including *Mim1* (39), *Cd34* (43), and *Gata1* (42), has been found to be modulated by Myb protein. Collectively, these data indicate that Myb protein is a sequence-specific transcription factor that modulates the expression of its target genes.

The *Cmyb* proto-oncogene is highly expressed in immature cells of the myeloid, erythroid, and lymphoid lineages and the levels of expression decrease markedly as these cells differentiate (28, 29). Gene expression has also been detected in acute myelogenous leukemia, chronic myelogenous leukemia, lymphocytic leukemia, breast and colon carcinoma cells (15). Endogenous Myb protein expression and the growth of various human myeloid leukemia cell lines are inhibited by antisense oilgonucleotide to the *Cmyb* proto-oncogene (30-33). Two avian acute leukemia viruses, AMV (avian myeloblastosis virus) and E26, which contain a truncated version of the *Cmyb* proto-oncogene, designated the *v-myb* oncogene, can induce leukemia *in vivo* and

transform cells of erythroid and/or myeloid lineages *in vitro* (16, 17). These studies suggest that Myb protein is involved in the cellular proliferation and/or differentiation of hematopoietic cells.

The direct evidence of the involvement of Myb protein in hematopoietic development *in vivo* came from the targeted mutagenesis study of the *Cmyb* proto-oncogene. Mouse embryos that are homozygous for the mutated *Cmyb* proto-oncogene display a severe anemia phenotype and die at approximately 15.5 dpc (12). Molecular and histological analyses revealed that yolk sac erythropoiesis was not affected by the inactivation of the *Cmyb* gene *in vivo*, however, fetal liver erythropoiesis was severely impaired. Moreover, using *in vitro* methocellulose assays, erythropoiesis and myelopoiesis in *Cmyb* -/- fetuses were shown to be dramatically reduced in comparison to phenotypically normal littermates (38).

Mouse embryonic and fetal hematopoietic developments are characterized by rapid and dramatic changes in the site of erythropoiesis, erythroblast morphology, and hemoglobin synthesis (1). The hematopoietic system is established early during embryogenesis from the extra embryonic mesoderm of the yolk sac (9). The yolk sac remains the primary hematopoietic organ until approximately day 12 of gestation. Beyond this stage of development, yolk sac hematopoiesis declines and the fetal liver becomes the predominant hematopoietic organ for the remainder of fetal life. Erythropoiesis finally shifts to the bone marrow around the time of birth. Yolk sac-derived erythrocytes, termed primitive erythrocytes, are nucleated, while those produced later in development, termed definitive erythrocytes, are non-nucleated. Differentiation of hematopoietic cells involves progressive restriction of developmental potential as pluripotent hematopoietic stem cells give rise to multipotential progenitors that subsequently mature along lineage-specific pathways. Although mechanisms responsible for decision-making in hematopoietic development are not well defined, it

is likely that they involve nuclear regulatory proteins that directly control and/or coordinate gene expression and signal transduction pathways transmitted from hematopoietic growth factors through their cognate surface receptors (10).

Identification of genes for the nuclear regulatory proteins and growth factors and/or their receptors that are involved in hematopoiesis will help to elucidate the complex mechanism(s) whereby hematopoietic stem cells proliferate and differentiate into mature blood cells of various lineages. In this study, we describe the continued analysis of *Cmyb* -/- fetuses and the attempt to identify novel genes that may be involved in hematopoiesis using these fetuses as a model system.

Materials and methods

Cell lines

Hematopoietic cell lines of various cell lineages were maintained as described (96-99). Briefly, lymphoid cell lines including NFS 70 (pre-B cell), NFS 112 (B cell), EL4, DA2, DA25, and CTLL (T cell) were cultured in RPMI 1640 medium (Life Technologies) supplemented with 10 % fetal bovine serum (FBS). Exogenous cytokine, IL-2 (20 ng/ml), was added in the medium for CTLL cells. Myeloid cell lines, including NFS 60, NFS 78, WEHI-3, DA1, DA33, and FDCP-1, were cultured in RPMI 1640 medium with 10% FBS and 20 U/ml IL-3. The fetal liver cell line, FL8-Myb, was maintained in the same medium as the myeloid cell lines. The erythroid cell line, MEL-C19, was cultured in DMEM plus 10% FBS.

RNA isolation

The livers of mouse fetuses at 13.5 dpc and 14.5 dpc were dissected, frozen in liquid nitrogen, and stored at -80°C until needed. Total RNA was isolated from frozen

fetal livers according to the methodology of Chomczynski and Sacchi (58). Briefly, individual frozen tissues were homogenized in 500 µl of guanidium thiocyanate RNA isolation solution (4M guanidium thiocyanate, 25mM sodium citrate, pH 7.0, 0.5% sarcosyl, 0.1M 2-mercaptoethanol) by passing each through a 18-gauge needle several times. The homogenate was then sequentially mixed with 50 µl of 2M sodium acetate (pH 4.0), 500 µl of water-saturated phenol, and 200 µl of chloroform-isoamyl alcohol (49:1), followed by 15 minutes of incubation on ice. Homogenate suspension was separated by centrifugation at 10,000g for 15 minutes at 4°C. The upper aqueous phase containing total RNA was collected and mixed with an equal volume of isopropanol and stored at -20°C for at least 1 hour. RNA was precipitated by centrifugation at 14,000 g for 15 minutes at 4°C, air dried, and resuspended in ribonuclease (RNase) free H_2O. Total RNA samples used in the expression studies for differentially expressed genes were isolated from various hematopoietic cell lines using the same methodology.

Reverse transcriptase mediated - polymerase chain reaction (RT-PCR) analysis

Total RNA was subjected to a RNase-free deoxyribonuclease (DNase) digestion step prior to cDNA synthesis by reverse transcriptase. To obtain DNA-free RNA samples, 5 µg of total RNA was treated with 5 units of RNase-free DNase (Stratagene) in 10 µl of 1X DNase reaction buffer (50 mM Tris-HCl, pH7.5, 200 mM NaCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$) at 37 °C for 50 minutes, followed by phenol-chloroform extraction. DNA-free RNA samples were recovered by ethanol precipitation.

For the first strand cDNA synthesis, DNA-free RNA was reverse transcribed using Moloney murine leukemia virus reverse transcriptase (M-MuLV RTase) and

random hexamer primers in conditions recommended by the manufacturer (GeneAmp® RNA PCR Kit, Perkin Elmer Cetus). Briefly, 0.5 µg of total RNA was mixed with 2.5 µM random hexamer, RNase inhibitor (1U/µl), and M-MuLV RTase (2.5 U/µl) in a total 20 µl of 1X PCR buffer II (50 mM KCl, 10 mM Tris-HCl, pH 8.3) containing 5 mM $MgCl_2$ and 1 mM of dNTPs. The reaction mixture was incubated at room temperature for 10 minutes, followed by 20 minutes of incubation at 42°C and stopped by heating at 99°C for 5 minutes.

For cDNA amplification, one tenth (2 µl) of the reverse transcribed products was amplified by PCR using AmpliTaq DNA polymerase (Perkin Elmer Cetus) in a total volume of 20 µl. PCR amplification was carried out using a DNA Thermal Cycler 480 (Perkin Elmer Cetus). PCR conditions were as follows : hot start at 95°C for 5 minutes, followed by 40 cycles of denaturation for 30 seconds at 94°C, annealing for 30 seconds at optimal temperature for individual primer sets (see Table 1), and extension for 2 minutes at 72°C. After the final cycle, an extension step was carried out at 72°C for 10 minutes. PCR products were analyzed using 1.2% agarose gels.

Northern blot hybridization analysis

Total RNA (10 µg) was denatured in 10 mM sodium phosphate buffer (pH 6.8) containing 54% formamide and 7.0% formaldehyde at 65°C for 10 minutes. Following denaturation, RNA was separated by electrophoresis through a 1% agarose gel containing 7.0% formaldehyde in 10 mM sodium phosphate buffer (pH 6.8). Separated RNA in gel was then transferred onto nylon filters (Duralon-UV, Stratagene) using standard blotting techniques (59). For prehybridization, nylon filters were incubated in 20 ml of prehybridization buffer (4XSSC, 10% dextran sulfate, 40% formamide, and 100 µg/ml denatured salmon sperm DNA) at 42°C for at least 3 hours.

Gene-specific cDNA fragments were labelled with α-^{32}P-dCTP (3000 Ci/mmol, Amersham) using random 9-mers (Prime-IT® II Kit, Stratagene) as primers. ^{32}P-labelled cDNA probes, at specific activity $\geq 10^8$ dpm/μg, were denatured in 4N NaOH for 5 minutes, neutralized by mixing with 70 mM Tris-HCl, pH 7.5 and 0.2N HCl, and added directly into the prehybridization buffer. Hybridization was carried out at 42°C for at least 12 hours. Hybridized filters were washed twice in 2XSSC, 0.1%SDS at room temperature for 15 minutes, followed by the wash in 0.2XSSC, 0.1%SDS at room temperature for 20 minutes and at 50-55°C for 20 minutes. Washed filters were exposed to X-ray films (Kodak) and kept at -80°C for 1-2 days before development.

Differential display analysis

DNA-free total RNA from the livers of *Cmyb* +/+, +/-, and -/- fetuses at day 14.5 dpc was used. Differential display (47) was carried out using procedures recommended by the manufacturer (RNAmap™ kit, GeneHunter). Four reverse transcription reactions for each sample were performed, each containing one of four degenerate oligo(dT) primers (T12MN). Briefly, 0.2 μg of DNA-free total RNA was added to a total 20 μl of reaction mixture containing 1X RT buffer (25 mM Tris-HCl, pH8.3, 37.6 mM KCl, 3 mM $MgCl_2$, and 5 mM DTT), 1 μM of oligo(dT) primer, 20 units of RNase inhibitor, and 20 μM of dNTPs. The entire reaction mixture was incubated at 65°C for 5 minutes, transferred to 37°C for 60 minutes, then heat inactivated at 95°C for 5 minutes. M-MuLV RTase (100 units) was added to each tube after 10 minutes of incubation at 37°C.

At the end of the reaction, one tenth (2 μl) of the reverse transcribed products was amplified in PCR in a total 20 μl of reaction mixture for each primer set. The reaction mixture contained 1X PCR buffer (10 mM Tris-HCl, pH 8.4, 50 mM KCl, 1.5 mM

$MgCl_2$, and 0.001% gelatin), 2 μM dNTPs, 1 μM of oligo(dT) primer, 0.2 μM of AP-primer (arbitrary decamer), 1 μl of ^{35}S-dATP (1,000 Ci/mmol, Amersham), and 0.5 units of AmpliTaq DNA polymerase (Perkin Elmer Cetus). The PCR conditions for differential display were : hot start at 95°C for 5 minutes, followed by 40 cycles of denaturation at 94°C for 30 seconds, annealing at 40°C for 2 minutes, and extension at 72°C for 30 seconds, and the reaction was extended at 72°C for 10 minutes. In all, 4 oligo(dT) primers and 5 AP-primers were used.

^{35}S-labeled PCR products were then separated on a 6% polyacrylamide DNA sequencing gel. After exposure to X-ray films , differential banding patterns were detected. Differentially expressed cDNA fragments were excised from the gel, boiled in 100 μl of water for 5 minutes, ethanol precipitated in the presence of 50 μg glycogen and resuspended in 10 μl of TE buffer. Recovered cDNA fragments (4 μl) were reamplified by PCR using the same oligonucleotide primer set and the reaction conditions described above. Reamplified cDNA fragments were electrophoresed through 1% agarose gels, isolated, and ligated into the pCRII vector (pCR™ vector, Invitrogen).

The cloned cDNA inserts were isolated from the vector plasmids by appropriate restriction endonuclease digestion and used as probes in Northern blot analysis to confirm whether they are really differentially expressed in mice of different genotypes. The nucleotide sequence of each cDNA fragment was determined by dideoxy chain termination DNA sequencing. These DNA sequences were then analyzed using the University of Wisconsin Genetics Computing Group (GCG) software.

Chromosomal mapping analysis

To determine the chromosomal positions of the 8 differentially expressed genes, the segregation of sequences detected by the differential display-isolated cDNA probes was followed in 119 progeny of a *Mus musculus X Mus spretus* interspecific backcross [(C3Hf/RI-*Mgf*$^{sl\text{-}ENURg}$/+ x *Mus spretus*) X C3Hf/RI] (60). The polymorphic DNA fragments of variant sizes detected by these cDNA probes in *M. musculus* DNA and *M. spretus* DNA were followed in the interspecific backcross panel. Mapping data were stored and map positions, with standard errors, were calculated using standard statistical methods with the aid of the Map Manager data analysis program (61, 62).

Results

Specific hematopoietic genes are differentially expressed in *Cmyb* mutant fetuses

To further explore the fetal hematopoietic phenotype of *Cmyb* -/- mouse fetuses, we first examined the expression levels of genes previously shown to be expressed in various hematopoietic cell lineages. Fetal livers of *Cmyb* +/+, +/-, and -/- animals at 13.5 dpc and 14.5 dpc were analyzed because the liver is the major hematopoietic center at this stage of mouse fetal development (2). More importantly, the liver appears to be the only organ affected in *Cmyb* homozygous mutant fetuses (12).

Table 1 lists the specific hematopoietic genes that were characterized by the RT-PCR analyses. These genes were chosen because they are genes that are expressed early in hematopoiesis, such as *Cd45* (63), and *Scl* (64), and/or genes known to be involved in hematopoietic development, such as *Gata1* (65), and *Nf-e2*

Table 1. Oligonucleotide primers used for RT-PCR analysis.

Gene	Forward Primer (5'--->3')	Reverse Primer (5'--->3')	Tm (°C)
IL-1α	AAGGAGAGCCGGGTGACAGTAT	GATGCCTGTCAAGCTCAGAGGA	50
IL-1β	TAGGCTCATCTGGGATCCTCTC	TACCAGTTGGGGAACTCTGCAG	50
IL-1R	CTGTTGGTGAGGAATGTGGCTG	AGCCGTGAGGATGATAAAGCCC	55
IL-3	TTCTTGCCAGCTCTACCACCAG	ACTTTAGGTGCTCTGCCTGCTG	50
IL-3R	TTCCTTTGGGCTCTTCTATCGC	CCAGGAGCAAGGTGAAGATGAG	50
IL-4	TCTGTAGGGCTTCCAAGGTGCT	GATGCTCTTTAGGCTTTCCAGG	50
IL-6	AACCACGGCCTTCCCTACTTCA	GCATAACGCACTAGGTTTGCCG	50
CSF-1	GTCAGAACACTGTAGCCACATG	CTGTCAACGGCCTGTCTGTTAT	60
CSF-1R	GCGATGTGTGAGCAATGGCAGT	AGACCGTTTTGCGTAAGACCTG	60
GCSF	CCAACTTTGCCACCACCATCTG	GGAGCAGCAGCAGGAATCAATA	62
GCSFR	CCCCTCAAACCTATCCTGCCTC	TCCAGGCAGAGATGAGCGAATG	60
EPO	ACGTCCCACCCTGCTGCTTTTA	CGTGTACAGCTTCAGTTTCCCC	64
EPOR	TTCTCCTCGCTATCACCGCATC	CCTCAAACTCGCTCTCTGGGCT	64
c-kit	CAACAGCAATGGCCTCACGAGT	GTGGTACACCTTTGCTCTGCTC	64

Table 1. (continued)

Gene	Forward Primer (5'--->3')	Reverse Primer (5'--->3')	Tm (°C)
SL(steel factor)	CGCTGCCTTTCCTTATGAAGAAGA	CGGGACCTAATGTTGAAGAGAGCA	64
GATA-1	ATGCCTGTAATCCCAGCACT	TCATGGTGGTAGCTGGTAGC	55
GATA-3	CCTTAAAACTCTTGGCGTCC	AGACACATGTCATCCCTGAA	50
CD45	CCTGAGTCTGCATCTAAACCCC	TGCTTGGCCAGTATTCTGCGCA	64
NF-E2	CATCTACTCCCCATGTCCCA	CTCGTCCCGACTCCCTGCCT	60
SCL	ATTGCACACACGGGATTCTG	CATACAGTACGACACTGACG	64
c-vav	GTATAACGTGGAGGTCAAGC	GGAGGACTGGAGAAAATCAG	50
cdc-2	GGAGAAGGTACTTACGGTGTGG	CGCTCTGGCAAGGCCGAAATCA	62
c-myc	AGAGCAAGAAGATGAGGAAG	GAATGGACAGGATGTAGGCG	55
(βH1) globin	AGTCCCCATGGAGTCAAAGA	CTCAAGGAGACCTTTGCTCA	55
(βmajor) globin	CTGACAGATGCTCTCTTGGG	CACAACCCCAGAAACAGACA	55
HPRT	CACAGGACTAGAACACCTGC	GCTGGTGAAAAGGACCTCT	60

Tm : annealing temperature used in RT-PCR analysis.

(66). Hypoxanthine guanine phosphoribosyl transferase (*Hprt*) gene expression was used in the analysis as a positive control. The nucleotide sequence and the annealing temperature for each primer set are presented in Table 1.

Of the 25 genes analyzed, only *Gata1* was found to be expressed at a lower level in *Cmyb* -/- fetal livers in comparison to phenotypically normal littermates (Fig. 1). Gata1 is a tissue-specific DNA-binding zinc-finger protein that binds to the WGATAR or 'GATA' motif found in the regulatory elements of several genes expressed specifically in erythroid cells (67, 68). *Gata1* is expressed not only in erythroid cells (69, 70) but also in megakaryocytes and mast cells (71) and is believed to play an important role in gene regulation during the development of these hematopoietic cell lineages.

Since the PCR reaction is highly sensitive and the RT-PCR analysis used in this study was not quantitative, results from these analyses may not be informative as to the comparison of gene expression among *Cmyb* +/+, +/-, and -/- fetuses. To circumvent this potential problem, Northern blot hybridization analysis was also performed. 10 μg of total RNA from *Cmyb* +/+, +/-, and -/- fetal livers at 13.5 dpc and 14.5 dpc was analyzed. A chick tubulin cDNA probe that is highly homologous to the mouse tubulin gene was used for a RNA loading control (72). Results from this analysis revealed that in addition to *Gata1*, three other genes, *Epor*, *Scl*, and *Nf-e2,* were also expressed at much lower levels in *Cmyb* -/- fetal livers in comparison to phenotypically normal littermates (Fig. 2).

Erythropoietin (Epo) is an important cytokine that regulates the development of erythroid progenitor cells. Specific responsiveness to Epo appears in committed erythroid progenitor cells at a stage somewhere between the burst forming unit - erythroid (BFU-E) and the colony forming unit - erythroid (CFU-E) during erythroid differentiation (73). Therefore, the expression of the receptor for Epo (EpoR) on the

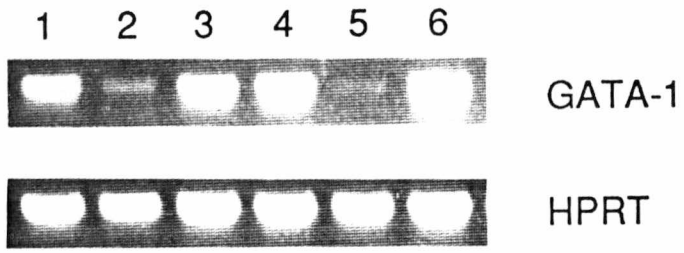


Fig. 1 RT-PCR analysis of hematopoietic gene expression in *Cmyb* wild-type (+/+), heterozygous (+/-), and homozygous mutant (-/-) fetal livers. Total liver RNA isolated from +/- (lanes 1 and 4), -/- (lanes 2 and 5), and +/+ (lanes 3 and 6) fetuses at 13.5 (lanes 1-3) and 14.5 (lanes 4-6) dpc was analyzed by RT-PCR using primers specific for *Gata1* (upper panel) and *Hprt* (lower panel). RT-PCR products were separated in 1% agarose gels.

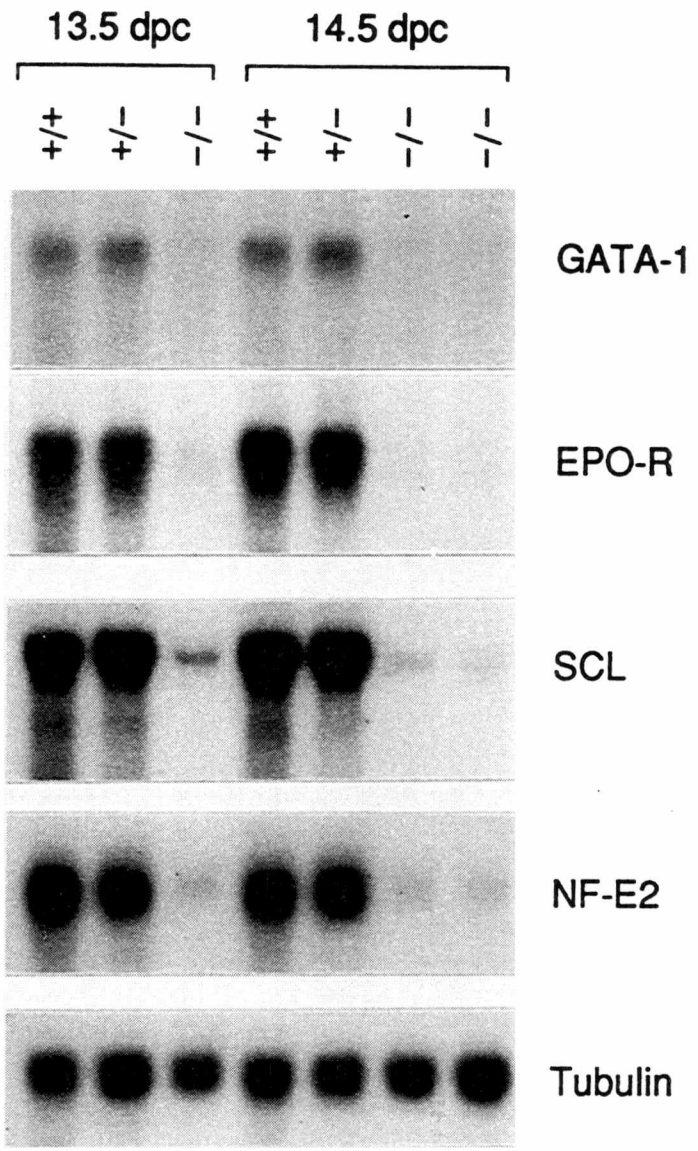


Fig. 2 Differential expression patterns of hematopoietic genes in *Cmyb* +/+, +/-, and -/- fetal livers. Northern blot analysis was performed to examine the expression patterns of hematopoietic genes, including *Gata1*, *Epor*, *Scl*, and *Nf-e2,* using total liver RNA (10 μg) isolated from fetuses at 13.5 and 14.5 dpc. The filter was stripped and rehybridized using a tubulin cDNA as a RNA loading control.

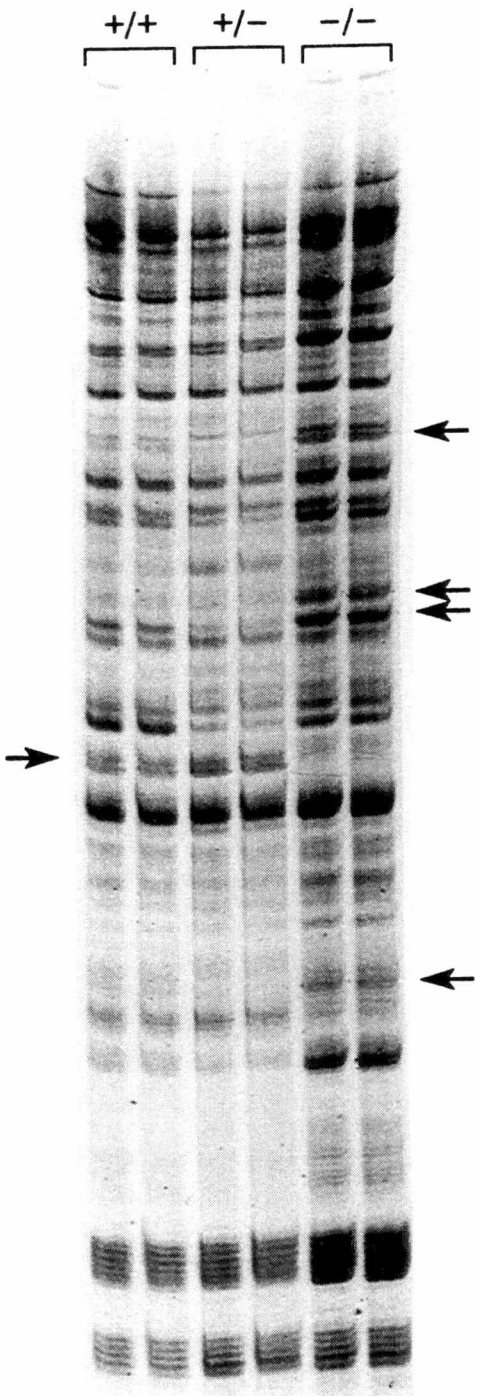
cell membrane of the progenitor cells is critical during the erythroid differentiation process. The *Scl* gene, which is involved in a chromosomal translocation associated with a stem cell leukemia, is a member of the helix-loop-helix (HLH) DNA-binding protein family (74, 75). *Scl* is expressed in fetal livers, bone marrows, and various hematopoietic cell lines. The *Nf-e2* gene is an erythroid-specific member of the family of basic-region-leucine zipper (b-zip) transcription factors (66). *Nf-e2* recognition sites are found in the α- and β-globin locus control region (LCR) as well as in the promoter regions of the porphobilinogen deaminase and ferrochelatase genes, which encode enzymes for heme biosynthesis (76, 77). Expression studies showed that *Nf-e2* is expressed in erythroid cell lines, hematopoietic progenitor lines, megakaryocytes, and mast cell lines (66).

Taken together, RT-PCR and Northern blot analyses revealed that four specific hematopoietic genes are down-regulated in *Cmyb* -/- mouse fetuses. The fact that most of these genes are expressed in erythroid cells correlated well with the anemia phenotype seen in *Cmyb* -/- fetuses. These data indicated that the normal gene expression patterns during fetal liver hematopoiesis have been perturbed in these embryos. The *Cmyb* -/- fetuses thus represent an excellent model system for the identification of genes that may be involved in fetal liver hematopoiesis.

Identification of putative novel genes differentially expressed in *Cmyb* homozygous mutant mice by differential display

Differential display is one of the most widely used techniques to identify differential gene expression (47, 78). In an attempt to identify from *Cmyb* -/- fetal livers differentially expressed genes that may represent novel genes involved in hematopoiesis, differential display was carried out. Total RNA isolated from the fetal livers of *Cmyb* +/+, +/-, and -/- fetuses at 14.5 dpc was analyzed. Fig. 3 shows the

Fig. 3 Differential display gel electrophoresis. Total liver RNA from *Cmyb* +/+, +/-, and -/- fetuses at 14.5 dpc was used in differential display analysis using a oligo(dT) primer and a 5'-end arbitrary decamer to amplify a subset of RNA species. Amplified cDNA fragments were separated in a 6% DNA sequencing gel and the differential banding patterns among samples were revealed (arrows).



differential banding patterns seen in one of the typical differential display sequencing gels. Using a specific oligo(dT) primer and an arbitrary decamer in RT-PCR analysis, differential gene expression patterns can be detected simultaneously among different samples. Therefore, cDNA fragments that were expressed at higher levels in *Cmyb* +/+ and +/- fetal livers as well as those expressed at higher levels in *Cmyb* -/- fetal livers can be isolated and analyzed.

From the utilization of four oligo(dT) primers and five arbitrary decamers, 45 partial cDNA fragments that showed differential expression patterns were identified (Table 2). These partial cDNA fragments were eluted out of the sequencing gels and reamplified by PCR in the presence of the same primer set. To date, 38 cDNA fragments have been used as probes in Northern blot analysis to confirm whether they are really differentially expressed. Fig. 4 shows four examples of Northern blot analysis for screening differentially expressed cDNA fragments. Total RNA from fetal livers of 13.5 dpc and 14.5 dpc animals were used in the analysis. Here, two cDNA probes, DD3-3 and DD5-5, detected stronger gene expression in *Cmyb* +/+ and +/- RNA samples, while the other two cDNA probes, DD2-2 and DD7A5-7, detected higher expression levels in *Cmyb* -/- RNA samples.

Using the same screening method, twenty-two of the 38 cDNA fragments analyzed (57.9%) were found to be differentially expressed (Table 2). Of the 22 differentially expressed cDNAs, five were expressed at higher levels in fetal livers of *Cmyb* +/+ and +/- embryos, while the remaining seventeen were expressed at highest levels in *Cmyb* -/- fetal livers (data not shown). Twelve differentially expressed cDNA fragments were partially sequenced, and eight were initially designated as novel based upon a lack of identity with sequences found in the GenBank/EMBL database (data not shown).

Table 2. Summary of differential display experiments

Oligo-dT primer used : $T_{12}MA$, $T_{12}MC$, $T_{12}MG$, $T_{12}MT$.	
5'-end primers used :	AP-1(5'-AGCCAGCGAA-3') AP-2 (5'-GACCGCTTGT-3') AP-3 (5'-AGGTGACCGT-3') AP-4 (5'-GGTACTCCAC-3') AP-5 (5'-GTTGCGATCC-3')
Number of cDNA fragments subcloned : 45	
Number of cDNA probes used in Northern Blot Analysis : 38	
Number of real differentially expressed cDNA fragments: 22	
Number of differentially expressed cDNA sequenced : 12	
Number of putative novel genes : 8	
Known genes identified :	phospholipase C - alpha (GenBank accession no. M73329) prothrombin/thrombin (GenBank accession no. X52308) cysteine proteinase inhibitor (MS1) (GenBank accession no. M92417)

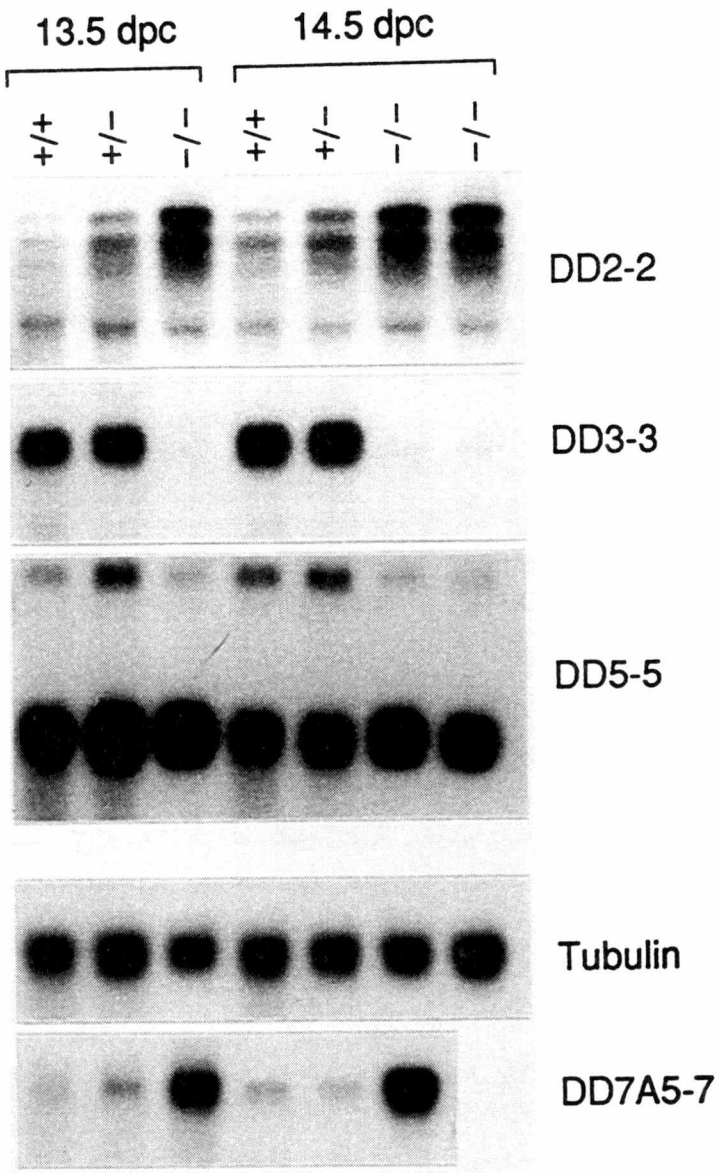


Fig. 4 Differential expression patterns of cDNAs identified by differential display. Northern blot analysis was performed to verify the differential expression patterns using total liver RNA (10 μg) isolated from *Cmyb* +/+, +/-, and -/- fetuses at 13.5 and 14.5 dpc. The cDNA fragments isolated from the differential display gels (DD2-2, DD3-3, DD5-5, and DD7A5-7) were used as probes. The filter was stripped and rehybridized using a tubulin cDNA as a RNA loading control.

The sequences of the other four partial cDNAs were identical to those for phospholipase C-α, which was identified twice in the analysis, prothrombin/thrombin, and a cysteine proteinase inhibitor (Table 2). Expression of these known genes had been previously reported in hematopoietic cells (79-81).

Hematopoietic cell lineage-specific expression patterns of differentially expressed genes

To evaluate whether these eight putative novel genes are possibly involved in hematopoiesis, we examined the expression patterns of these genes in various hematopoietic cell lines using Northern blot analysis. While five of the differentially expressed genes (DD2-2, DD3-2, DD5-5, DD8G1-1, and DD8G1-5) showed ubiquitous expression in all cell lines tested, the remaining three genes (DD3-3, DD5-3, and DD7A5-7) appeared to have intriguing cell lineage-specific expression patterns. As shown in Fig. 5, the DD7A5-7 gene transcript (3.3 kb in size) is detected exclusively in myeloid cell lines but not in erythroid and lymphoid cell lines tested. The DD5-3 gene, whose transcript is approximately 5.0 kb in length, is expressed predominantly in myeloid cell lines and differentiated erythroid cell lines. No DD5-3 gene expression is detected in lymphoid cell lines tested. In addition, since two different sizes (~5.0 kb, ~3.0 kb) of DD5-3 RNA transcript were evident in myeloid cell lines but only one transcript (~5.0 kb) is in erythroid cell lines, it appears that the DD5-3 gene transcript is alternatively spliced in myeloid cells. The DD3-3 gene is predominantly expressed in T lymphoid cell lines with apparent alternatively spliced transcripts. At least three different sizes (~1.6 kb, ~2.0 kb, ~2.4 kb) of DD3-3 RNA transcript were identified. Moreover, the expression levels of the three alternatively spliced transcripts are dramatically different among the cell lines. This indicates that the regulation of DD3-3 gene expression is controlled differently in these T cell lines. No DD3-3 gene

Fig. 5 Hematopoietic cell lineage-restricted expression patterns of differentially expressed genes. Total RNA (10 μg) isolated from erythroid (MEL), myeloid (FDC-P1, DA-33, DA-1, WEHI-3, NFS-78, NFS-60), T-cell (CTLL-2, EL-4, DA-25, DA-2), B-cell (NFS-112, NFS-70) lines, and a fetal liver cell line (FL8-myb) was analyzed by Northern blot hybridization using DD3-3, DD5-3, and DD7A5-7 cDNA probes. The same filter was stripped and rehybridized with a tubulin cDNA probe as a RNA loading control.

DD7A5-7
DD5-3
DD3-3
Tubulin
Lymphoid
B Cell
T Cell
Myeloid
Erythroid
NFS 70
NFS 112
DA-2
DA-25
EL-4
CTLL
NFS 60
NFS 78
WEHI-3
DA-1
DA-33
FDCP-1
MEL undiff.
MEL 1 day diff.
MEL 2 day diff.
MEL 3 day diff.
MEL 4 day diff.
MEL 5 day diff.
FL8-Myb

transcripts were detected in other hematopoietic cell lines tested except for the weak expression identified in one of the myeloid cell lines, NFS 78. Overall, the Northern blot analysis in this study demonstrated that all the differentially expressed putative novel genes are expressed in hematopoietic cells, either ubiquitously or specifically. In addition, alternatively spliced transcripts are evident for most of the hematopoietic cell-lineage specific genes, indicating highly regulated gene expression mechanisms for these genes in the hematopoietic cell lines.

Chromosomal mapping of the differentially expressed genes

In order to determine the chromosomal location of the eight putative novel genes, interspecific backcross analysis was conducted (82). The rationale for this work was to determine whether these genes mapped to the chromosomal regions that are associated with the heritable hematopoietc disorders. The interspecific backcross was conducted using DNA samples isolated from 119 progeny of an interspecific backcross between *Mus musculus* and *Mus spretus* and has been previously mapped with various existing genetic markers (60). Specific polymorphisms can be detected in the 119 DNA samples digested with specific restriction enzymes using various specific DNA probes. Comparison of the distribution patterns for the polymorphisms detected by a uncharacterized new DNA probe with the ones that were generated by known genetic markers reveals the approximate chromosomal position for the new gene. Although the chromosomal locations of the eight differentially expressed genes were dispersed over the mouse genome (Table 3), no heritable hematopoietic or other diseases in mouse or human were found to be associated with these map positions after using Online Medelian Inheritance in Man (OMIM) database.

Table 3. Summary of Northern blot and chromosomal mapping analysis of the differentially expressed genes

Gene	Gene expresssion		Chromosomal location
	Fetal liver[a]	Cell lines[b]	
DD2-2	mutant	E+B+T+M	12
DD3-2	mutant	E+B+T+M	2
DD3-3	wild-type	T	6
DD5-3	mutant	E+M	9
DD5-5	wild-type	E+B+T+M	17
DD7A5-5	mutant	M	17
DD8G1-1	wild-type	E+B+T+M	16
DD8G1-5	wild-type	ND	3

a : gene expression is higher in *Cmyb* homozygous mutant (mutant) fetal liver or in wild-type fetal liver as determined by Northern blot analysis.
b : gene expression is detected in erythroid (E) , B lymphoid (B), T lymphoid (T), or myeloid (M) cell lines. ND : not determined.

Discussion

Using RT-PCR and Northern blot analysis, we have demonstrated that four hematopoietic-specific genes, *Gata1*, *Epor*, *Scl*, and *Nf-e2*, were expressed at much lower levels in *Cmyb* -/- fetal livers than in their phenotypically normal littermates (Fig. 1). These four genes encode a hematopoietic cell-specific growth factor receptor and/or sequence-specific DNA-binding transcription factors. These proteins belong to two categories of gene products that are believed to be important in the regulation of hematopoietic cellular proliferation and/or differentiation (10). In fact, targeted mutagenesis experiments have revealed the critical roles for these genes in normal hematopoietic development *in vivo* (65, 73, 83-86). For example, Pevny *et al.* showed that targeted embryonic stem (ES) cells carrying a mutated X-linked *Gata1* gene failed to contribute to the mature red blood cell populations in chimeric mice (87). This result indicated Gata1 is a necessary transcription factor during normal *in vivo* erythroid lineage development. *In vitro* colony assay studies performed by the same group further demonstrated that the development of erythroid progenitors in yolk sacs was not affected by the inactivation of the *Gata1* gene; however, the cells develop only to an early point in terminal differentiation, the proerythroblast stage, leading to a developmental arrest and finally death. Mice homozygous for a mutated *Epor* gene died at approximately 13.5 days of gestation due to the failure of definitive fetal liver erythropoiesis (73). *Epor* mutant embryos contained normal numbers of primitive erythrocytes in yolk sacs; however, they were reduced in size and the proliferation potential of the cells was severely retarded. On the other hand, virtually no definitive erythrocytes beyond the late progenitor (CFU-E) stage were detected *in vivo* despite the presence of committed erthyoid BFU-E and CFU-E progenitors in *Epor* homozygous mutant fetal livers. Thus, it is believed that Epor is crucial for the

proliferation and survival of CFU-E progenitors and their irreversible terminal differentiation *in vivo* (86).

The decreased expression levels of these specific hematopoietic genes correlated very well with the phenotype seen in *Cmyb* homozygous mutant fetuses, in which the definitive fetal liver erythropoiesis and myelopoiesis are severely impaired (12). More interestingly, the phenotype described for *Epor* mutant embryos is very similar to the one found in *Cmyb* mutant fetuses. Therefore, it is tempting to suggest a possible connection between *Cmyb* proto-oncogene and these specific hematopoietic genes. For example, it is possible that Myb protein, as a transcription factor, can modulate the expression of other hematopoietic genes, *e.g.*, *Epor.* Accordingly, decreased expression of the *Epor* gene resulting from the inactivation of *Cmyb* proto-oncogene *in vivo* may lead to the anemia phenotype. Using transient expression assays in the ME26 virus infected FDC-P2 myeloid cell line, Aurigemma *et al.* (41) demonstrated that the 135 kDa gag-Myb-Ets fusion protein encoded by the ME26 recombinant mouse retrovirus can transactivate the *Gata1* promoter. Furthermore, they showed that ME26 virus can cooperate with the Gata1 protein to enhance the expression of the *Epor* gene. Similarly, Chiba and co-workers (88) demonstrated that the Gata1 protein can transactivate the *Epor* gene in a mouse erythroleukemia cell line, SKT6, using a series of transient expression assays, gel mobility shift assays, and DNase I footprinting experiments.

Data from these *in vitro* experiments support the hypothesis presented; however, we can not rule out other possible explanations for the reduced expression of hematopoietic genes in *Cmyb* -/- fetuses. For example, the fetal livers of *Cmyb* -/- animals are hypocellular and have a reduced number of hematopoietic progenitor cells. Therefore, it is also possible that the lower levels of gene expression are merely due to reduced number of cells that normally expressed the specific

hematopoietic genes. Unfortunately, the real mechanism(s) for the reduced expression levels of the hematopoietic genes in *Cmyb* -/- fetuses can not be deduced from the studies presented here. Nevertheless, data from our studies definitely indicates that normal gene expression patterns are severely affected in *Cmyb* -/- fetuses. The *Cmyb* mutant fetuses may thus be an unique system for the identification of novel genes that may be involved in hematopoiesis.

To test this hypothesis, differential display analysis was carried out. Differential display is a PCR-based alternative method to substractive and differential hybridization techniques (47). By comparing radioisotope-labelled PCR products in parallel, this technique can identify differentially expressed genes (e.g., candidate tumor suppressor genes vs. oncogenes) among cell or tissue samples of the same origin but with different phenotypes (89). This novel method was thereby chosen in our study to attempt to identify differentially expressed genes in *Cmyb* +/+, +/-, and -/- fetuses (Fig. 3). The fact that 22 out of 38 purified cDNA fragments (57.9%) were really differentially expressed (Table 2) indicates the usefulness of this methodology in our model system. Furthermore, 8 out of 12 sequenced differentially expressed cDNA fragments showed no homology to sequences reported in GenBank/EMBL database, suggesting that a large number of novel genes may be identified using this methodology. To date, only four oligo(dT) primers and five degenerate decamers were used in the differential display analysis. Since a total of 80 possible combination of primer sets may be used (47), differential display represents a powerful methodology for identifying a large number of novel genes involved in hematopoiesis using this system. Indeed, among the eight differentially expressed putative novel genes, a recent search in GenBank / EMBL database using BLAST program showed that while DD2-2 gene is highly homologous to Rho-associated protein kinase (GenBank accession number, U58513, 187/195, 96%), DD5-5 gene is most similar to

cytochrome C oxidase subunit VII (GenBank accession number, X80899, 77/81, 95%), and DD7A5-7 is identical to mouse *Emr1* and *F4/80* cDNA (100%)(see Chapter 2), the other five genes remain unknown as no significant sequence matches were found in the database (data not shown).

To determine whether the putative novel genes are indeed involved in hematopoiesis, definite screening strategies must be designed. Northern blot analysis using total fetal liver RNA from 13.5 and 14.5 dpc *Cmyb* +/+, +/-, and -/- fetuses was used to verify that the genes identified were indeed differentially expressed (Fig. 4). Hematopoietic cell lineage expression pattern as well as the chromosomal mapping of each gene were then used to prioritize the order in which the genes were analyzed. While all eight differentially expressed putative novel genes were shown to be expressed in hematopoietic cell lines, three of them (DD3-3, DD5-3, and DD7A5-7) were found to have intriguing cell lineage-specific expression patterns (Fig. 5), suggesting possible function(s) in the hematopoietic cells. It is therefore of interest to study the hematopoietic cell lineage-specific genes. However, it should be made clear that the ubiquitously expressed genes may also play a role in hematopoiesis. Additional screening strategies may be applied to the system in the future. For instance, determination of gene expression in hematopoietic cells of different developmental stages may reveal a temporal expression pattern of the gene and may be suggestive of a possible role for the gene in the differentiation of hematopoietic cells. Comparison of the gene expression among normal and transformed and/or diseased hematopoietic cells may also help to elucidate the importance of the gene in the occurrence of hematopoietic diseases. The chromosomal mapping study failed to reveal any association between the novel genes and known disease loci. Nevertheless, this remains a good strategy for identifying important hematopoietic genes in the future.

Overall, in this study we provide evidence for perturbed hematopoietic gene expression patterns in *Cmyb* -/- fetuses and prove that differntial display can be used to identify putative novel genes that may be involved in hematopoiesis. Isolation and characterization of full-length cDNAs of the differentially expressed genes identified herein will be the next step toward the unraveling of the role(s) of the genes in hematopoiesis. Likewise, continued application of differential display in the *Cmyb* -/- mouse model system will also be beneficial to our understanding of hematopoiesis.

Chapter 2 Identification and characterization of a seven transmembrane hormone receptor using differential display

Abstract

Targeted mutagenesis analysis has shown that the *Cmyb* proto-oncogene, which encodes a sequence-specific DNA binding protein, is required for normal murine fetal liver erythropoiesis and myelopoiesis. To identify novel genes involved in hematopoiesis, differential display analysis was conducted using total liver RNA isolated from 14.5 days post coitus *Cmyb* wild-type, heterozygous, and homozygous mutant littermates. Using four oligo-dT 3' primers and five arbitrary decamers as 5' primers, twenty-two differentially expressed genes have been identified. Eight putatively novel genes were identified out of twelve cDNAs that were sequenced. One gene, initially designated DD7A5-7, is primarily expressed in cells of the myeloid lineage. The full length DD7A5-7 cDNA is 3,239 nucleotides in length, encoding a putative protein of 931 amino acids. The protein is a member of a family of hormone receptors containing seven transmembrane segments. The receptor also contains seven epidermal growth factor-like (*Egf*-like) motifs at the amino terminal of the predicted protein. The gene is alternatively spliced, resulting in the deletion of one or more copies of the *Egf*-like motif. DD7A5-7 maps to murine chromosome 17 and is the putative homologue of *EMR1,* a recently described *Egf*-like module containing, mucin-like hormone receptor with seven transmembrane segments in humans. Our results indicate that the *Cmyb* mutant fetuses represent a unique resource for identifying genes involved in hematopoiesis.

Introduction

Hematopoiesis is a very complex biological process in which stem cells not only self renew, but also give rise to committed progenitor cells that proliferate and differentiate into the various hematopoietic lineages. Hematopoiesis in the mouse initiates in the mesoderm of the yolk sac, with subsequent stem cell colonization of the fetal liver, spleen, and bone marrow (2, 90). A large number of genes, including growth factors and their receptors as well as transcription factors, have been shown to be involved in the development of various hematopoietic lineages. However, many of the genes involved in the proliferation and/or differentiation of specific cellular lineages are presently unknown (91, 92).

The *Cmyb* proto-oncogene plays a critical role in the proliferation and/or differentiation of multiple hematopoietic lineages. This proto-oncogene encodes a sequence-specific DNA-binding protein that is primarily expressed in immature hematopoietic cells (24, 93). The Myb protein functions as a transcription factor that modulates the expression of numerous genes, including *Mim1* (39), *Cd4* (44), *Cd34* (42), and *Gata1* (43). Targeted mutagenesis studies have shown that *Cmyb* mutant fetuses die at 15.5 days postcoitum (dpc) due to anemia (12). Histological analysis indicates that primitive yolk sac erythropoiesis is unaffected in mutant animals while definitive fetal liver erythropoiesis is severely impaired. The fetal livers of mutant animals are hypocellular, having a reduced number of hematopoietic progenitors in comparison to phenotypically normal littermates. *In vitro* hematopoietic assays confirmed these results as the number of both erythroid and myeloid progenitors were dramatically reduced in mutant fetal livers from 12.5 to 15.5 dpc (38). These results indicate that the normal pattern of gene expression during fetal liver hematopoiesis has been perturbed in *Cmyb* homozygous mutant fetuses.

The comparison of gene expression between *Cmyb* wild-type, heterozygous, and homozygous mutant fetal livers should lead to the identification of novel genes involved in hematopoiesis. To test this hypothesis, differential display analysis was conducted using total liver RNA isolated from *Cmyb* wild-type, heterozygous, and homozygous mutant fetuses at 14.5 dpc. Differential display is a PCR-based methodology that can be used to detect altered gene expression in eukaryotic cells and to isolate partial cDNAs of the affected genes (47, 94). This methodology has been successfully used to identify a number of genes, including a candidate tumor suppressor gene (48) and a transcriptional target of the p53 protein (49). In this paper, we describe the identification and characterization of a seven hormone receptor that is predominantly expressed in cells of the myeloid lineage. Sequence analysis and chromosomal mapping studies strongly suggest that this gene is the murine homologue of a human gene, *EMR1*, a member of the family of hormone receptors with seven transmembrane segments (95).

Materials and methods

RNA analysis

Total RNA was isolated from cell lines and frozen tissues using the acid guanidium thiocyanate-phenol-chloroform method (Chomczynski and Sacchi, 1987). Total RNA (10 μg) was electrophoresed through formaldehyde/agarose gels and transferred onto nylon filters (Duralon-UV, Stratagene). Northern blot analysis was carried out using standard methodologies (59). Filters were stripped of radiolabeled probes by washing in 0.2X SSC/1% SDS for 15 minutes at 90°C and rehybridized with a tubulin cDNA probe to compare the amount of RNA loaded in each lane.

RT-PCR analysis to identify alternatively spliced transcripts was performed as previously described (38). PCR amplification was performed using a DNA Thermal Cycler 480 (Perkin Elmer) and the following conditions: denaturation for 30 seconds at 94°C, annealing for 30 seconds at 64°C, and extension for 2 minutes at 72°C for 40 cycles. PCR products were analyzed using 1.2% agarose gels. The primers used in this analysis included: 5'-end forward primer (O-170, 5' GTACGATGTGGGGCTTTTGG CT 3'), 5'-end reverse primer (O-167, 5' CCAAGATCCCTGCCCTGCAC 3'), 3'-end forward primer (O-200, 5' GCATCTAGCAATGGACAGCTG 3'), 3'-end reverse primer #1 (O-201, 5' AGGACTGGAAGCCCATAGCCAA 3'), and 3'-end reverse primer #2 (O-202, 5' CCGTTGAAATAGGACGTGCTC 3'). Three different sizes of 5'-end RT-PCR products were identified, purified, and cloned into the pCRII vector following the manufacturer's instructions (Invitrogen). The DNA sequence of each cDNA was determined using the dideoxy chain termination method.

Differential display

DNA-free total RNA was prepared from the livers of *Cmyb* wild-type, heterozygous, and homozygous mutant fetuses at 14.5 dpc. Differential display was performed as previously described using the RNAmap™ Kit A (GenHunter) (47, 94). The oligo-dT primers used in this study were $T_{12}MA$, $T_{12}MC$, $T_{12}MG$, and $T_{12}MT$. The five arbitrary decamers used were: AP-1 (5' AGCCAGCGAA 3'), AP-2 (5' GACCGCTTGT 3'), AP-3 (5' AGGTGACCGT 3'), AP-4 (5' GGTACTCCAC 3'), and AP-5 (5' GTTGCGATCC 3'). PCR amplification conditions used were: 95°C for 5 minutes followed by 40 cycles at 94°C for 30 seconds, 40°C for 2 minutes, and 72°C for 30 seconds. [α-^{35}S]-labeled PCR amplified cDNA fragments were separated on a 6% polyacrylamide DNA sequencing gel, excised, processed, reamplified, and then subcloned into the pCRII vector (Invitrogen) using standard molecular cloning

techniques. Partial cDNAs were isolated from the vector plasmid by appropriate restriction endonuclease digestion and used as probes in Northern blot analysis.

Cell lines

Total RNA was isolated from eleven cell lines representing several distinct hematopoietic lineages and stages of differentiation. Cell lines were maintained as previously described (96-99).

Isolation and characterization of cDNA clones

Two positive cDNA clones (1.6 and 2.2 kb in length) were identified using the DD7A5-7 cDNA probe in an adult mouse testis cDNA library that was constructed using the l-ZAP vector (Stratagene). A murine thymoma RACE cDNA library, constructed using the Marathon RACE kit (Clontech), was used to identify the 5'-end of the full-length DD7A5-7 cDNA. A 1.4kb DNA fragment was generated by PCR amplification using the AP-2 primer and one gene-specific primer (O-139, 5' GAGACACTTGTAGCTGCA 3'). Southern blot analysis using an internal gene-specific oligonucleotide probe (O-140, 5' CCAAAGTGAGCAGATACA 3') was used to verify the RACE product's authenticity.

DNA and protein sequence analysis

DNA sequencing was performed using the dideoxy chain termination method and the Sequenase™ Version 2.0 DNA Sequencing Kit (United States Biochemical). Homology searches of both nucleotide and putative amino acid sequences were performed by FASTA and TFASTA programs of the Genetics Computer Group (GCG) sequence analysis package by searching the GenBank/EMBL database.

Chromosome mapping analysis

To determine the chromosomal position of the *Emr1* gene, the segregation of sequences detected by a 1.2 kb *Eco*RI - *Xho*I cDNA probe was followed in 119 progeny of a *Mus musculus X Mus spretus* interspecific backcross [(C3Hf/RI-*Mgf*$^{sl-ENURg}$/+ x *Mus spretus*) X C3Hf/RI]. The cDNA probe hybridized to fragments of 15.0, 9.8, 3.2, and 2.0 kb in *Bam*HI-digested *M. musculus* DNA and fragments of 7.7, 5.2, 3.6, 3.2, 2.0, and 1.5 kb in *Bam*HI-digested *M. spretus* DNA. The four variant *M. spretus* fragments were followed in the interspecific backcross panel. Probes and variant fragments used to map neighboring genes *Rfx2, C3, Vav, Nfya,* and *Xdh* in the same cross, and the segregation patterns for these genes were recently described (60). Mapping data were stored and map positions, with standard errors, were calculated using standard statistical methods with the aid of the Map Manager data analysis program (61, 62).

Results

Putative novel cDNAs identified by differential display

Differential display was carried out to identify novel genes involved in hematopoiesis, using total RNA isolated from the fetal livers of 14.5 dpc *Cmyb* wild-type, heterozygous, and homozygous mutant fetuses. From the utilization of four oligo-dT primers and five arbitrary decamers, forty-five partial cDNAs were identified that showed differential patterns of expression. To date, thirty-eight have been used as probes in Northern blot analysis of total liver RNA isolated from *Cmyb* wild-type, heterozygous, and homozygous mutant fetuses at 13.5 and 14.5 dpc. Twenty-two of

the thirty-eight cDNAs analyzed (57.9%) were found to be differentially expressed (38). Of the twenty-two differentially expressed cDNAs, five were expressed at higher levels in *Cmyb* wild-type and heterozygous fetal livers in comparison to homozygous mutant livers while the remaining seventeen were expressed at highest levels in *Cmyb* homozygous mutant fetal livers (data not shown). Twelve differentially expressed cDNAs were partially sequenced and eight were initially designated as novel based upon a lack of identity with sequences found in the GenBank/EMBL database (data not shown). The sequence of the other four partial cDNAs were identical to those for phospholipase c-alpha (80), which was identified twice, prothrombin/thrombin (79), and a cysteine proteinase inhibitor (81). Expression of these known genes had previously been reported in hematopoietic cells.

DD7A5-7 is expressed primarily in cells of the myeloid lineage

To determine whether the eight novel genes identified were expressed in hematopoietic cells, we examined their expression patterns by Northern blot analysis using total RNA isolated from a variety of hematopoietic cell lines. All were shown to be expressed in various hematopoietic cell lineages. Interestingly, the 3.3 kb transcript of gene DD7A5-7 was detected in all myeloid cell lines analyzed (DA-1, DA-33, FDC-P1, NFS 60, NFS 78, WEHI-3) but not in the erythroid (C19 MEL) or lymphoid (CTLL-2, DA-2, DA-25, EL-4, NFS 70, NFS 112) cell lines (data not shown). To determine the temporal and spatial pattern of expression of DD7A5-7, Northern blot analysis showed that DD7A5-7 was expressed at highest levels in the 14.5 dpc fetal liver and in adult bone marrow. Intermediate levels of expression were detected at 14.5 dpc in the head and body as well as in adult brain, spleen, and thymus. Lowest levels of expression were seen in the heart and thymus of 18.5 dpc fetuses and in the adult testis (data not shown).

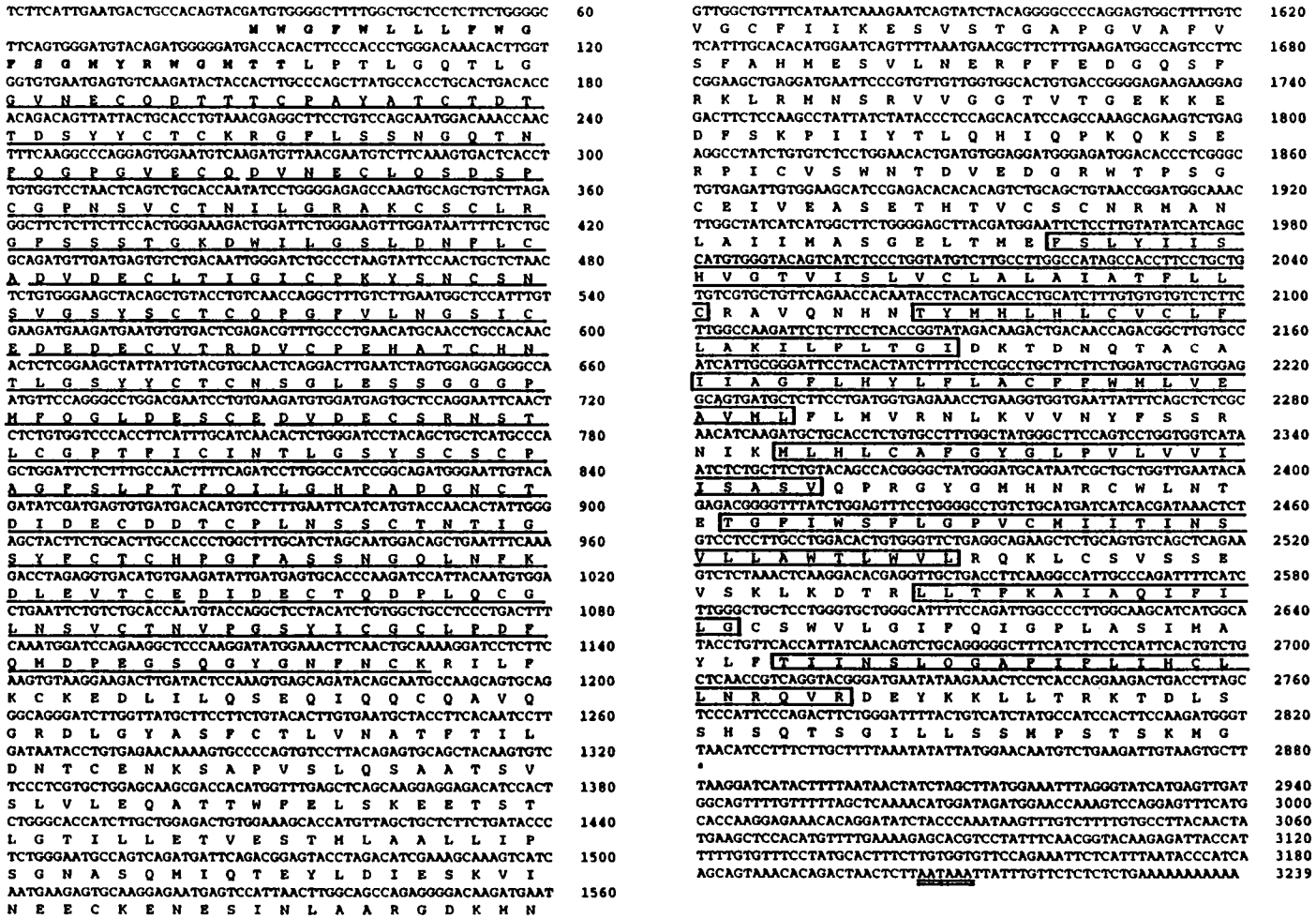
Cloning of the DD7A5-7 cDNA

Two overlapping cDNAs, of 1.6 kb and 2.2 kb in length, were isolated from an adult mouse testis cDNA library. Restriction endonuclease digestion patterns and nucleotide sequence analysis revealed that each clone contained a poly(A) tail that defined the 3'-end of the gene. To obtain a full-length cDNA, the 5'-end of the gene was amplified by PCR using a murine thymoma RACE cDNA library. A 5'-end RACE product of approximately 1.4 kb was amplified, subcloned into the pCR II vector, and sequenced. The complete nucleotide sequence of the DD7A5-7 cDNA is 3,239 bp in length, including 27 bp of 5'-end untranslated sequence, 2,796 bp of coding sequence, and 416 bp of 3'-end untranslated sequence (Fig. 6). The region including the first ATG contains a favorable Kozak translational initiation signal (100). The open reading frame encodes a putative protein of 931 amino acids with an estimated molecular weight of 100 kDa.

Chromosomal mapping of the DD7A5-7 gene

To determine the chromosomal location of the DD7A5-7 gene in mice, the segregation pattern of sequences detected by a 1.2 kb EcoRI-XhoI cDNA probe was followed in a *Mus musculus X Mus spretus* interspecific backcross (IB) (82). Four variant *Bam*HI *Mus spretus* DNA fragments were identified. All of the variant fragments co-segregated, and were linked to markers on the distal region of murine chromosome 17. DD7A5-7 is closely linked to *Vav, C3, and Rfx2* genes (Fig. 7). Map positions of these and other neighboring genes were determined previously in this same IB system (60). IB analysis indicated that DD7A5-7, *Vav*, *C3*, and *Rfx2* are located together, approximately 12.6 centimorgans (cM) (+/- 3.0 cM) distal of *Nfya*

Fig 6. The cDNA sequence and its deduced amino acid sequence of DD7A5-7. The open reading frame starts at nucelotide 28 and ends at the termination codon at nucelotide 2823. The amino acid sequence is shown immediately below the nucelotide sequence. Amino acids of the putative leader sequence are highlighted in bold lettering, the seven *Egf*-like motifs are underlined, the seven transmembrane domains are enclosed in boxes, and the polyadenylation signal is underlined twice.



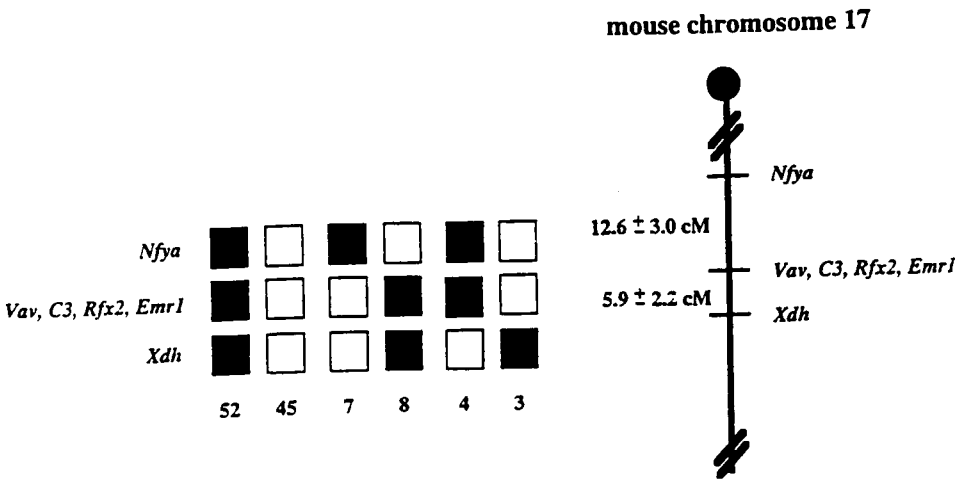


Fig 7. Genetic mapping of the *Emr1* gene. The linkage of *Emr1* to *Nfya* and *Xdh* genes on mouse chromosome 17 was determined by Southern blot hybridization of *Bam*HI-digested DNA from 119 interspecific backcross progeny. Black boxes represent the inheritance of a C3Hf allele for a particular set of loci, while white boxes represent the *M. spretus* allele. Gene linkage, order, and intergenic distances with associated standard errors were determined using Map Manager data analysis software.

and 5.9 cM (+/- 2.2 cM) proximal of *Xdh* (61). The *Vav, C3,* and *Rfx2* region of murine chromosome 17, as determined by physical and comparative mapping, is known to be syntenic to human chromosome 19p13.3 (60). Recently *EMR1,* a hormone receptor containing *Egf*-like motifs and a seven transmembrane segment was mapped to the same region of human chromosome 19p13.3 by both somatic cell hybrid and *in situ* hybridization analyses (95). In addition to these mapping results, comparison of the nucleotide and amino acid sequences as well as protein structure between these two genes strongly suggests that the DD7A5-7 gene encodes the murine homologue of *EMR1* and will thus be designated *Emr1.*

Comparison of EMR1 and Emr1

Initial comparison of the nucleotide and amino acid sequences of *Emr1* did not show striking homology to any reported in the GenBank/EMBL database. However, moderate homology was noted between the N-terminal third of Emr1 and proteins such as murine and human fibrillin-1 (101), fibulin-2 (102), and human latent transforming growth factor-β binding protein (103), which contain epidermal growth factor-like (*Egf*-like) motifs. A total of seven *Egf*-like domains, characterized by the presence of six cysteine residues within each motif, were identified in the predicted Emr1 protein, starting from amino acid position 32 to 367 without interruption. A consensus sequence for post-translational β-hydroxylation of aspartic acid or asparagine (CX(D,N)XXXX(F,Y)XCXC) was found in six of the seven *Egf*-like motifs. This is similar to what has been reported for a number of proteins containing *Egf*-like motifs (104). When the C-terminal third of the Emr1 protein (amino acids 645-931) was compared to sequences in the GeneBank/EMBL database, significant homology was found with proteins containing seven transmembrane domains, including the secretin and calcitonin receptors (105). A hydropathy plot of the predicted Emr1

protein obtained by using the Kyte-Doolittle algorithm confirmed the presence of a seven transmembrane domain (data not shown) (106).

When the cDNA sequence of *EMR1* gene was available in the database, comparison of the human *EMR1* and the murine *Emr1* genes revealed a 74.3% and a 69.0% identity at the nucleotide and amino acid levels, respectively (data not shown). The corresponding *Egf*-like motifs share 63.9% amino acid identity. The Emr1 gene product also contains an additional *Egf*-like motif in comparison to its human homologue. An alignment of *Egf*-like motifs indicates that the fifth Emr1 *Egf*-like repeat is absent in EMR1 (Fig. 8A). A consensus sequence for β-hydroxylation of aspartic acid or asparagine was found in every *Egf*-like motif of the EMR1 and Emr1 proteins, with the exception of the second *Egf*-like motif in Emr1. The middle third of each protein is serine/threonine-rich with 27 serine and 25 threonines in the human protein compared to 27 serines and 22 threonines in the murine protein. Within this region, the two proteins share 61.2% amino acid identity (data not shown). Serine and threonine residues are potential sites of O-glycosylation, indicating the potential for high levels of glycosylation in both proteins, which may function to inhibit the formation of secondary structure (95). The proteins are most highly conserved at the amino acid level in the carboxy-terminal region which contains the seven transmembrane domain (84.0% identity). The alignment of the amino acid sequences in this region is shown in Figure 8B. Most of the conserved residues found in the transmembrane region of seven transmembrane hormone receptors are conserved in the Emr1 protein (Fig. 8B) (95).

Emr1 mRNA is alternatively spliced

To confirm the accuracy of the nucleotide sequence of RACE-generated products, RT-PCR analysis of total liver RNA isolated from a *Cmyb* wild-type fetus at

Fig 8. Comparison of predicted amino acid sequences encoded by *EMR1* and *Emr1*. (A) The *Egf*-like domains of *EMR1* (top) and *Emr1* (bottom) were aligned using GCG software, revealing a 63.9% amino acid identity between the two domains. Conserved amino acids are represented by dashes in *Emr1* and the conserved cysteine residues are underlined. (B) The seven transmembrane domains of *EMR1* (top) and *Emr1* (bottom) were similarly aligned, revealing a 84% amino acid identity in this region. The amino acids conserved in the seven transmembrane region between these two genes and other family members are underlined.

(A)

```
EGF-like 1   KGNN CRDSTL  CPAYAT CTNTVDSYY  CTC KQ GF LSSNGQNHFKDPGVR    CK
egf-like 1   GV-E -Q-T-T  ------ --D-T----  --- -R -- ------TN-QG---E    -Q

EGF-like 2   DIDE CSQSPQP CGPNSS CKNLSGRYK  CSC LD GF SSPTGNDWVPGKPGNFS  CT
egf-like 2   -VN- -L--DS- -----V -T-IL--A-  --- -R -- --S--K--IL-SLD--L  -A

EGF-like 3   DINE CLTSRV  CPEHSD CVNSMGSYS  CSC QV GF ISRNST              CE
egf-like 3   -VD- ---IGI  --KY-N -S--V----  -T- -P -- VLNG-I              --

EGF-like 4   DVNE CADPRA  CPEHAT CNNTVGNYS  CFC NP GF ESSSGHLSCQGLKAS    CE
egf-like 4   -ED- -VTRDV  ------ -H--L-S-Y  -T- -S -L ---G-GPMF---DE-    --

egf-like 5   DVDE CSRNSTL CGPTFI CINTLGSYS  CSC PA GF SLPTFQILGHPADGN    CT

EGF-like 5   DIDE CTEM    CPINST CTNTPGSYF  CTC HP GF APSSGQLNFTDQGVE    CR
egf-like 6   ---- -DDT    --L--S ----I----  --- -- -- -S-N-----K-LE-T    -E

EGF-like 6   DIDE CRQDPST CGPNSI CTNALGSYS  CGC IV GF HPNPEGSQKDGNFS     CQ
egf-like 7   ---- -T---LQ --L--V ---VP---I  --- LP D- QMD-----GY---N     -K
```

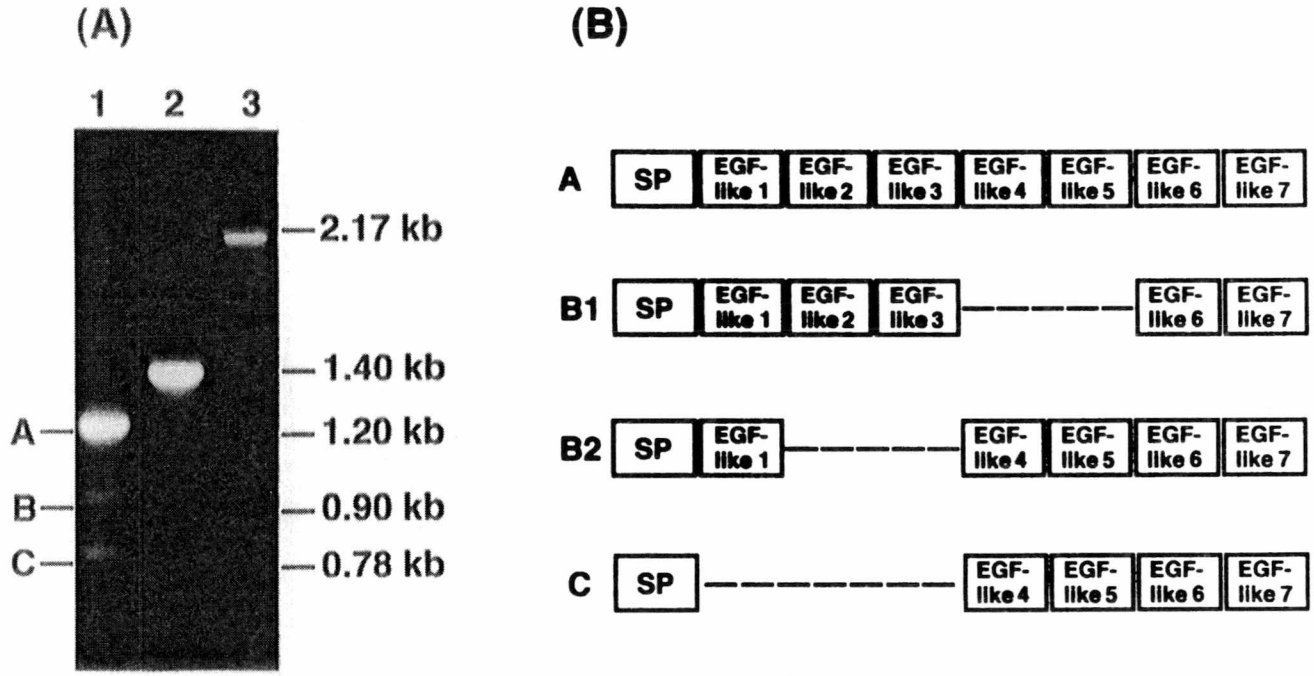
(B)

```
TM 1   FSLYIISHVGIIISLVCLVLAIATFLLC
tm 1   ----------TV------A---------

TM 2   TYLHLHLCVCLLLAKTLFLAGI
tm 2   --M--------F---I---T--

TM 3   IIAGFLHYLFLACFFWMLVEAVIL
tm 3   ----------------------M-

TM 4   MLHICAFGYGLPMLVVVISASV
tm 4   ---L--------V---I-----

TM 5   TGFIWSFLGPVCTVIVINSLLLTWTLWIL
tm 5   ------------MI-T---V--A----V-

TM 6   LLTFKAFAQLFILGCSWVLG
tm 6   ------I--I----------

TM 7   TIINSLQGAFIFLIHCLLNGQVR
tm 7   -------------------R---
```

14.5 dpc was conducted. Three alternatively spliced transcripts were detected when primers O-167 and O-170 were used (Fig. 9A). Primer O-170 is located at the 5'-end of the full-length cDNA while primer O-167 is located downstream of the seventh *Egf*-like motif. The three DNA fragments were cloned into the pCR II vector and sequenced. Fragment A contains *Egf*-like motifs 1 through 7, thus encoding the 5'-end of the full-length message. Fragments B1 and B2 contain only five *Egf*-like motifs, with the fourth and fifth motifs being absent in B1 and the second and third *Egf*-like motifs being absent in B2. Fragment C contains only *Egf*-like motifs 4 through 7. RT-PCR assays were also conducted to analyze the 3'-end of the *Emr1* transcript, using oligonucleotide primers O-200, O-201, and O-202. Only single fragments of the appropriate size were amplified using these oligonucleotide primers, indicating that this region of the transcript was not alternatively spliced (Fig. 9A, lanes 2 and 3). Overall, these results indicate that there are several different isoforms of the Emr1 protein that vary in the number and/or combination of *Egf*-like motifs (Fig. 9B).

Northern blot analysis indicates that *Emr1* is predominantly expressed in murine myeloid cell lines. Similar results were reported for *EMR1* although basal levels of expression were also detected in all tumor cell lines and normal hematopoietic cells examined by RT-PCR analysis (95). Overall, *Emr1* and *EMR1* share a high degree of similarity in nucleotide and amino acid sequences as well as protein structure and pattern of expression. However, differences in the number of *Egf*-like motifs as well as the presence or absence of β-hydroxylation domains for aspartic acid or asparagine residues were noted between the two proteins. In addition, *Emr1* is alternatively spliced, resulting in at least four protein isoforms that differ in the number and/or combination of *Egf*-like motifs. This suggests that the each *Egf*-like motif may be encoded by an individual exon and that *EMR1* may also be alternatively spliced in a similar manner.

Fig 9. Alternative splicing of the *Emr1* gene. Oligonucleotide primers used in this analysis are described in the Materials and Methods (A) Lane 1 shows the RT-PCR analysis at the 5'-end of the gene. Three fragments (A, B, and C) were identified. Fragment A contains all seven *Egf*-like motifs, while two different nucleotide sequences were identified in fragment B. B1 is an alternatively spliced message which is missing *Egf*-like motifs 4 and 5, while B2 is missing *Egf*-like motifs 2 and 3. Fragment C is also alternatively spliced and is missing *Egf*-like motifs 1-3. Lanes 2 and 3 represent results using oligonucleotide primers from the serine/threonine-rich extracellular region and the same extracellular region and the transmembrane region, respectively. Only the expected sized transcripts were detected using these oligonucleotide primers. (B) Diagram of the alternatively spliced transcripts detected. SP represents the signal peptide sequence and each *Egf*-like motif is represented individually.


(A)
1
2
3
2.17 kb
1.40 kb
A
1.20 kb
B
0.90 kb
C
0.78 kb
(B)
A
SP
EGF-like 1
EGF-like 2
EGF-like 3
EGF-like 4
EGF-like 5
EGF-like 6
EGF-like 7
B1
SP
EGF-like 1
EGF-like 2
EGF-like 3
EGF-like 6
EGF-like 7
B2
SP
EGF-like 1
EGF-like 4
EGF-like 5
EGF-like 6
EGF-like 7
C
SP
EGF-like 4
EGF-like 5
EGF-like 6
EGF-like 7

Discussion

Differential display has been widely used in recent years to identify differentially expressed genes in eukaryotic cells. To identify novel genes involved in hematopoiesis, this methodology was used to analyze the alteration of gene expression in the liver of *Cmyb* wild-type, heterozygous, and homozygous mutant fetuses at 14.5 dpc. This report describes the identification and characterization of a receptor, *Emr1*, which is predominantly expressed in cells of the myeloid lineage. The cDNA and amino acid sequences of *Emr1* indicate the gene encodes a seven transmembrane hormone receptor containing seven *Egf*-like motifs in the amino-terminal region of the extracellular domain. RT-PCR analysis indicates that the gene is alternatively spliced, resulting in a minimum of four protein isoforms with different numbers and/or combination of *Egf*-like motifs. These results suggest that each motif is encoded by an individual exon, similar to the gene structure found in the human genes Factor X (107) and Protein C (108). The different isoforms may interact with different ligands or have different ligand-binding affinities. It is also possible that one or more conserved *Egf*-like motif found in all isoforms may be responsible for protein-protein interactions, as seen in the interaction between the *Drosophila* Notch protein and either the Delta or Serrate proteins (109). An additional similarity between *Emr1* and other proteins containing *Egf*-like motifs which are involved in Ca^{++}-dependent protein-protein interactions is the presence of a consensus sequence for the β-hydroxylation of aspartate or asparagine residues at defined positions within the *Egf*-like domain. Overall, these results suggest that *Emr1* may play a critical role in calcium-dependent protein-protein interactions.

All members of G-protein-coupled receptor families have hydrophobic regions that cross the plasma membrane seven times. The superfamily of G-protein-coupled

receptors functions by transmitting signals across the plasma membrane of cells after ligand binding. Through the activated G-proteins, various second messages, including the activation of adenylate cyclase, trigger a cascade of intracellular signaling which results in various changes in cellular physiology. Recently, a new family of receptors has been identified based upon structural similarities. This family includes receptors that bind secretin, calcitonin, and glucagon (105). The predicted amino acid sequence of *Emr1* indicates that it is the newest member of this family. The predicted protein structure and restricted pattern of expression suggest that *Emr1* may function as a receptor on myeloid cells that binds to a specific secreted ligand or interacts with other cell surface proteins through its extracellular domain. *Emr1* may therefore be involved in the proliferation and/or differentiation of myeloid cells through these interactions.

Recently, McKnight *et al.* reported the cloning of a murine macrophage-restricted cell surface glycoprotein, recognized by the monoclonal antibody F4/80, which shares homology with the seven transmembrane hormone receptor family (110). Comparison of 3,239 nucleotides between the *Emr1* and F4/80 cDNAs revealed only two nucleotide differences in the 3' untranslated region, indicating that both cDNAs encode the same protein. The F4/80 monoclonal antibody recognizes a 160 kDa cell surface glycoprotein on the surface of most murine macrophage populations (111). It has been suggested that F4/80 may be involved in cell adhesion interactions within specific microenvironments (110). The presence of both *Egf*-like motifs and a serine/threonine-rich mucin-like domain in the extracellular region of the Emr1 (F4/80) protein further strengthen this hypothesis. The purification of ligand(s) for *Emr1* and the generation of null mutants through the use of targeted mutagenesis of embryonic stem cells will definitely help to elucidate the function of this protein. Targeted mutagenesis studies are currently ongoing in our laboratory.

Chromosomal mapping studies indicate that *Emr1* is located on murine chromosome 17 and is closely linked with the *Vav, C3,* and *Rfx2* genes. This region of murine chromosome 17 has been shown to be syntenic to human chromosome 19p13.3 (60). The fact that *EMR1* has been localized to chromosome 19p13.3 by both fluorescence *in situ* hybridization and radiation hybrid analyses (95) is consistent with our results, indicating that we have identified the murine homologue of this gene.

Chapter 3 Generation and characterization of *Emr1 / F4/80* homozygous mutant mice

Abstract

The Emr1/F4/80 protein is a 160 kilodalton (kDa) glycoprotein that is specifically expressed on the cell surface of mouse macrophages. The recent cloning of the *Emr1/F4/80* cDNA has revealed an unusual protein structure for the Emr1/F4/80 protein, which contains features of two protein superfamilies, the epidermal growth factor (EGF) family and the seven transmembrane (7TM) receptor family. However, little is known about the function of the Emr1/F4/80 protein. In this study, we used gene targeting to mutate the *Emr1* gene in mouse embryonic stem (ES) cells and generate *Emr1* homozygous mutant (-/-) mice to define the biological function(s) of the F4/80 protein. A gene targeting vector was constructed to replace the coding sequence of the first exon of the *Emr1* gene with a drug-resistance gene cassette. Five targeted ES cell clones were successfully generated. Eight chimeric male mice derived from one targeted ES cell clone transmitted the mutated gene allele to the offspring and generated the *Emr1* heterozygous (+/-) mice. *Emr1* homozygous mutant (-/-) mice were born at the expected Mendelian ratio from the mating of *Emr1* +/- animals. *Emr1* -/- mice were healthy, fertile, and had normal hematopoietic cell development. The macrophages of the *Emr1* -/- mice developed normally and expressed appropriate macrophage-specific cell surface proteins. These data suggest that the F4/80 protein is not required for normal mouse embryonic and fetal development and may not be essential for maintaining a normal natural immune response. Further analysis will be necessary to determine the immunological role of this cell surface receptor.

Introduction

In a previous study, we have identified a gene for a seven transmembrane (7TM) receptor, designated *Emr1*, through the differential display analysis of *Cmyb* wild-type (+/+), heterozygous (+/-), and homozygous mutant (-/-) fetuses (112). *Emr1* maps to mouse chromosome 17 and was shown to be the mouse homolog of the human *EMR1* gene. The predicted Emr1 gene product consists of 931 amino acids that are identical to that of a recently cloned mouse macrophage (Mϕ)-specific F4/80 cDNA, which encodes a 160 kilodalton (kDa) cell surface glycoprotein (112). Initially identified as the antigen (Ag) for the F4/80 monoclonal antibody (MAb) (111), the F4/80 glycoprotein is now considered one of the most highly specific cell surface markers for murine Mϕ due to its restricted expression on most mouse Mϕ populations (3, 113).

The expression of the F4/80 protein has been found to be tightly regulated according to the physiological status of the cells. For example, F4/80 is expressed at lower levels on activated Mϕ isolated from Bacille Calmette-Guérin (BCG) infected animals in comparison to unstimulated resting Mϕ (114). Similarly, F4/80 expression is down-regulated on Mϕ in response to interferon-γ (115). The precursor of tissue Mϕ, the blood monocyte, is also known to express fewer F4/80 molecules than its mature counterparts (116). In addition, heterogeneities of F4/80 protein expression are noted in tissues and cells. For example, the epidermal Langerhans cells are F4/80 expressing cells, while their derivative interdigitating dendritic cells are not (117). In lymph nodes and spleens, the F4/80 protein is detected only on Mϕ that are located outside the T-cell areas (117). Collectively, these data suggested that the F4/80

protein may be involved in cell adhesion and/or migration in certain tissues. However, little is known about the real function of the F4/80 protein.

The analysis of amino acid sequences has revealed an intriguing protein hybrid structure for F4/80 (110, 112). Three different protein motifs are identified in the F4/80 protein. They include seven epidermal growth factor (EGF)-like domains at the extracellular amino-terminus, a mucin-like domain in the central portion, and a 7TM domain at the carboxyl-terminus of the protein. Furthermore, the reverse transcriptase-mediated polymerase chain reaction (RT-PCR) analysis conducted in our laboratory indicated that different F4/80 protein isoforms may exist based on the identification of the alternatively spliced RNA transcripts that differ in the number and/or combination of the EGF-like domains. To date, only two other proteins, the human EMR1 protein and the human CD97 leukocyte antigen, are found to possess the same protein hybrid structure and a 7TM protein subfamily, termed EGF-TM7, has been designated for these three proteins (116). A ligand for CD97 has been identified recently (118). Using an adhesion assay, it was shown that CD97 plays an important role in cell-cell interactions by binding to another cell surface protein on heterogeneous cells. Whether the F4/80 protein has a similar function remains to be determined. Nevertheless, a dual ligand interaction has been proposed for the EGF-TM7 proteins based on their protein hybrid structure, which consists of two protein superfamilies, the EGF-family, and the 7TM receptor family (116).

The EGF-like domain is characterized by the formation of three disulfide bonds by the six cysteine residues located at the consensus positions within each domain (119). To date, a large number of EGF-like domain containing proteins, such as Notch, fibrillin-1, fibulin-2, and lamilins have been identified (101, 102, 109). A function for the EGF-like domain in protein-protein interactions has been demonstrated between the *Drosophila* Notch protein and its counterreceptors, the Delta and Serrate proteins

(109). In addition, proteins with EGF-like domains often contain a conserved amino acid sequence for aspartate/asparagine β-hydroxylation within the domain. This post-translational modification was believed to play a role for the EGF-like domain in Ca^{2+}-binding and the stabilization of the interactions mediated by the EGF-like domains (120). In F4/80, six of the seven EGF-like domains are found to contain the consensus Ca^{2+}-binding sequence (112). This would suggest a possible role for the F4/80 protein in Ca^{2+}-dependent protein-protein interactions between cells and /or cells-to-extracellular matrix (ECM).

All 7TM proteins identified to date are found to be coupled with heterotrimeric G-proteins inside the cell. These G-protein coupled receptors can transmit extracellular signals as diverse as hormones, growth factors, neurotransmitters, and sensory stimuli such as photons and odorants (121, 122). The G-protein coupled receptors make up a large 7TM protein superfamily with hundreds of members. Moreover, on the basis of the characteristics of the ligands they bind and the corresponding consensus amino acid sequences in their 7TM domains, 7TM proteins can be further divided into different subfamilies (123). The F4/80 protein is found to be most similar to the members of the peptide hormone receptor subfamily (95, 112). This would suggest that F4/80 might serve as a Mϕ-specific hormone receptor that transmits extracellular signal(s) through the G-protein mediated pathway upon ligand binding.

Macrophages serve as professional phagocytes and professional antigen presenting cells in the innate as well as the adaptive immune systems to help defend against pathogens. Some of their important functions are carried out by the cell surface proteins, such as the complement receptors, CR1 and CR3, for the binding and ingestion of opsonized targets (124), and the mannose receptor for the phagocytosis of saccharide-coated particles (125). It is therefore of interest to study

the possible role(s) the F4/80 protein may play in Mϕ biology. To delineate the biological function of the F4/80 protein, we took the *in vivo* gene inactivation approach utilizing targeted mutagenesis techniques in embryonic stem (ES) cells. Here, the generation and characterization of the *Emr1*/F4/80 -/-mice will be discussed.

Materials and methods

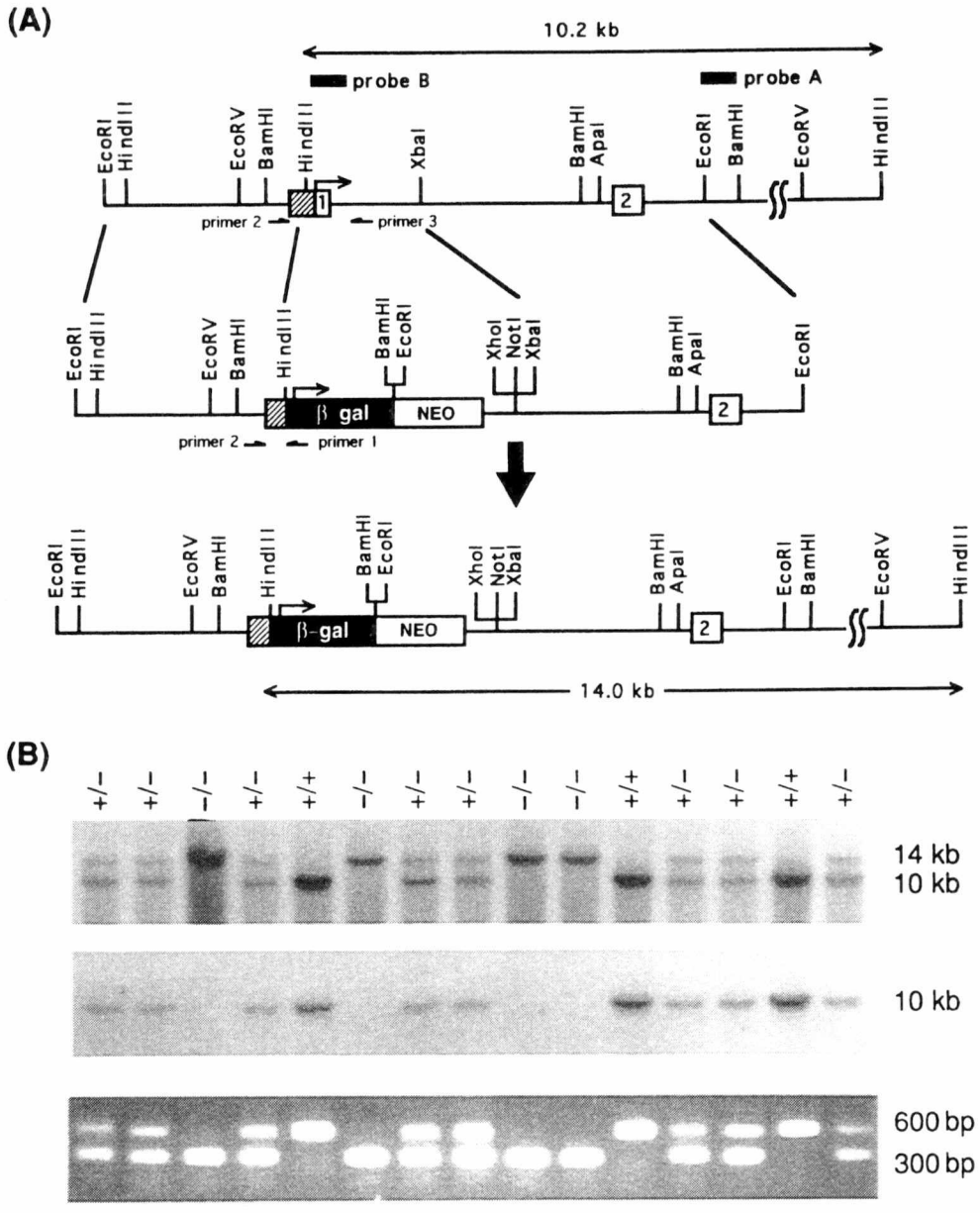
Trgeting construct

A positive cosmid clone containing exons 1 and 2 of the *Emr1* gene was identified from a 129/Sv mouse cosmid library using a 540 bp *Eco*R I-*Xho* I cDNA fragment as a probe. To make the targeting construct, a 1.8 kb *Eco*R I-*Hind* III fragment containing the 5'-UTR of the *Emr1* gene was subcloned into the *Not* I and *Hind* III sites of the pSAβgal-Neo vector (126) using standard molecular cloning techniques. The splice acceptor sequence located between *Not* I and *Hind* III sites was therefore deleted in the resulting construct. A 2.8 kb *Xba* I-*Eco*R I fragment containing the exon 2 of the *Emr1* gene was then engineered and subsequently inserted into the previous construct through the *Xba* I and *Kpn* I sites. The completed construct, pEmrβgal, resulting in the deletion of the exon 1 coding sequence of the *Emr1* gene, is shown in Fig. 10A.

Generation of the *Emr1* -/- mice

The targeting construct, pEmrβgal, was linearized with *Kpn* I and electroporated into CJ7 mouse ES cells (127) as described previously (12). Briefly, 25 μg of the linearized targeting construct was added to a total of 1 X 10^7 ES cells resuspended in 0.8 ml of phosphate buffered saline (PBS). Electroporations were carried out using a BioRad Gene Pulser at settings of 0.24 Kvolts and 500 μF. Electroporated ES cells

Fig. 10 Gene targeting of the mouse *Emr1* gene. (A) Partial restriction maps of the mouse *Emr1* gene, the targeting construct, and the resulting targeted allele. Selected restriction sites are shown. The β-galactosidase and pGK-Neo gene cassettes are shown as black and white boxes, respectively. White boxes with numbers indicate exons. The 5'-UTR (untranslated region) is represented by a stippled box. DNA probes A and B used for Southern blot analysis of genomic DNA are shown. The expected *Hind* III fragments of the wild-type and the targeted alleles detected by probes A and B are also shown. Arrows with numbers indicate the positions and orientations of the primers used for the PCR genotyping reaction. The map is not drawn to scale. (B) Southern blot and PCR genotyping. Tail DNA from 2 litters of 4 week-old offspring derived from the mating of the *Emr1* +/- animals was digested with *Hind* III and hybridized with probe A (upper panel) and probe B (middle panel). The same genomic DNA was analyzed by the PCR genotyping reaction using primers 1, 2, and 3, and was visualized by EtBr staining (lower panel). The sizes of the DNA fragments of interest are indicated. The genotype of individual animals is shown.


(A)
10.2 kb
probe B
probe A
EcoRI
HindIII
EcoRV
BamHI
HindIII
XbaI
BamHI
ApaI
EcoRI
BamHI
EcoRV
HindIII
primer 2
primer 3
1
2
EcoRI
HindIII
EcoRV
BamHI
HindIII
BamHI
EcoRI
XhoI
NotI
XbaI
BamHI
ApaI
EcoRI
β gal
NEO
primer 2
primer 1
2
EcoRI
HindIII
EcoRV
BamHI
HindIII
BamHI
EcoRI
XhoI
NotI
XbaI
BamHI
ApaI
EcoRI
BamHI
EcoRV
HindIII
β-gal
NEO
2
14.0 kb
(B)
+/−
+/−
−/−
+/−
+/+
−/−
+/−
+/−
−/−
−/−
+/+
+/−
+/−
+/+
+/−
14 kb
10 kb
10 kb
600 bp
300 bp

were cultured in the presence of X-ray irradiated mouse embryonic fibroblast (MEFs) feeder cells and selected in standard ES medium containing 250 μg/ml G418 (Life Technologies, USA) for 10-14 days. Individual drug resistant ES colonies were picked and placed into individual wells of 24-well plates containing X-ray irradiated MEFs. Approximate 10 μg of *Hin*d III digested genomic DNA isolated from the resulting ES cell clones was analyzed by Southern blot hybridization using a ^{32}P-labelled 0.2 kb *Eco*R I-*Bam*H I fragment (probe A in Fig. 10A). Targeted ES cell clones were identified by the homologous recombination event at the *Emr1* locus characterized by the presence of a 14 kb *Hin*d III fragment detected by probe A.

Targeted ES cells were microinjected into C57BL/6J blastocysts as described (12). Briefly, blastocysts were collected from the uteri of superovulated females and incubated in M16 medium. Approximately 15 targeted ES cells were injected into the blastocoel cavity of each blastocyst. Injected blastocysts were implanted into the uteri of pseudopregnant females. The resulting chimeric progeny were recognized by the presence of agouti coat color in black background at approximately 10 days of age. Chimeric male animals were mated to C57BL/6J females to produce heterozygous mice, which were characterized by agouti coat color and Southern blot analysis of tail DNA as described above or by PCR analysis. The PCR genotyping reaction was performed using approximately 200 ng of tail DNA in a total 50 μl of reaction mixture containing 1X PCR buffer (20 mM Tris-HCl, pH8.4, 50 mM KCl), 1.2 mM $MgCl_2$, 100 μM of dNTPs, 2.5 U of Taq DNA polymerase (5 U/μl, Life Technologies), a primer specific for the β-galactosidase gene (primer 1 : 5'-TCCGCCATGTCACAGATC AT-3', 5ng) and two *Emr1* gene-specific primers (primer 2 : 5'-GGAAGGAAATGGAG AGAAAG-3', 5 ng, and primer 3 : 5'-GAAGATCTACCCTGGTGAAT-3', 5 ng) (Fig. 10A). The reaction conditions for the PCR genotyping were : hot start at 95°C for 5 minutes, followed by a PCR touchdown program of three sequential 5 cycles of

denaturation at 94°C for 30 seconds, annealing at 65°C, 62°C, and 59°C for 5 cycles each for 30 seconds, and extension at 72°C for 90 seconds, and a final 25 cycles of the same amplification but at 55°C annealing temperature. Finally, the PCR products were extended at 72°C for 10 minutes, and stored at 4°C indefinitely. The amplified PCR products were separated in 1% agarose gels. Primer 2 and 3 should amplify a 600 bp fragment from the wild-type allele, while primer 1 and 2 would produce a 300 bp fragment from the targeted allele (Fig. 10).

Emr1 -/- mice were generated from the breeding of *Emr1* +/- animals. For genotyping, tail DNA from 4-week old progeny was analyzed by Southern blot hybridization using probes A and B or by PCR as dscribed above. DNA probe B, generated by PCR with primer 5'-GCAAGCTTTTAAAGGGAAGGC-3' and primer 5'-GAAGATCTACCCTGGTGAAT-3', was a 0.27 kb DNA fragment located in the intron 1 region that is deleted in the targeted allele (probe B, Fig. 10A).

Northern blot hybridization and RT-PCR analysis

Total RNA isolated from thymi, spleens, and livers of 5-week old *Emr1* +/+, +/-, and -/- mice was analyzed by Northern blot hybridization using an *Emr1* full-length cDNA probe and by RT-PCR. Gene-specific primers (5'-AGGACTGGAAGCCCATAG CCAA-3', and 5'-GCATCTAGCAATGGACAGCTG-3') used in the RT-PCR analysis are from the 7TM region and the mucin-like domain of the Emr1 protein, respectively (112). Primers for the hypoxanthine-guanine phosphoribosyl transferase (*Hprt*) gene were included in the reaction as an internal control (see Chapter 1). The reaction conditions for Northern blot and RT-PCR analyses are as previously described (Chapter 1).

Whole mount X-gal staining of mouse embryos

Mouse embryos (9.5-11.5 dpc) from the intercross of the *Emr1* +/-animals were isolated and fixed at 4°C for 30-60 minutes in 1X PBS containing 2% formaldehyde, 0.2% glutaraldehyde, 0.02% NP-40, and 0.01% deoxycholate. Genomic DNA isolated from the dissected yolk sacs was used in the PCR genotyping reaction to determine the genotype of individual embryos. The fixed embryos were washed in 1X PBS for 20-30 minutes twice, then stained in freshly made staining solution (1 mg/ml X-gal, 5 mM potassium ferricyanide, 5 mM potassium ferrocyanide, and 2 mM magnesium chloride in 1X PBS) at 30°C for 1-2 days depending on the staining intensity. Finally, stained embryos were fixed in 4% formaldehyde in 1X PBS at 4°C overnight.

Peritoneal cell preparation

Resident peritoneal cells (RPC) and thioglycollate elicited peritoneal cells (TPC) were obtained from 8-10 week-old *Emr1* +/+, +/-, and -/- mice. RPC was harvested from untreated animals by peritoneal lavage using 5 ml of cold sterile 1X PBS. Similarly, TPC was harvested from animals that were injected intraperitoneally (i.p.) with 3.0 ml of 3.0% (w/v) sterile Brewer thioglycollate medium (Difco) 4 days eariler. The cells were collected by centrifugation at 1,000 rpm (250xg) for 10 minutes, and then resuspended in S-PBS^{s} (see Flow Cytometry) for flow cytometric analysis. Cells were counted in a hematocytometer.

Antibodies

Antibodies used for the immunohistochemical analysis were kindly provided by Professor Siamon Gordon (Oxford University, UK). Rat anti-mouse polyclonal and monoclonal F4/80 antibodies are specific for the F4/80 protein (110, 111), MAb 3D6 recognizes sialoadhesion (Sn) specifically (128), MAb 2F8 is specific for Mϕ

scavanger receptor (MSR) (129), MAb 5C6 is used to detect the type 3 complement receptor (C3R) (130), and MAb FA-11 is specific for macrosialin (MS) (131). MAbs used for flow cytometric analysis were phycoerythrin (PE)-conjugated anti-mouse CD4 Ab (Caltag, 0.5 mg/ml), fluorescein isothiocyanate (FITC)-conjugated anti-mouse CD8 Ab (Caltag, 0.5 mg/ml), FITC-conjugated anti-mouse Thy-1.2 Ab (Caltag, 1 mg/ml), PE-conjugated anti-mouse B220 Ab (PharMingen), FITC-conjugated anti-mouse Mac-1 Ab (PharMingen), PE-conjugated anti-mouse F4/80 Ab (Serotec), and PE- or FITC-conjugated rat IgG_{2b}(K) Ab (Pharmingen) for isotype control. Anti-mouse FcγII/IIIR Ab was used as Fc block in the analysis to block non-specific binding of Abs.

Flow cytometry

To determine the number and the distribution of B and T lymphocytes in the *Emr1* +/+, +/- and -/- mice, flow cytometric analysis was conducted as described (132). Briefly, spleens and thymi of 6-8 week old animals were removed, and single cell suspensions were prepared by repeatedly passing small pieces of tissues through a 27-gauge hypodermic needle in S-PBS^s (calcium- and magnesium-free 1X PBS containing 0.1% BSA and 0.1% sodium azide). Cells were washed with S-PBS^s at 4°C, and centrifuged at 1,000 rpm (250xg) for 10 minutes. The number of cells per milliliter in the cell suspension was counted using a hematocytometer.

For each sample, 10^6 cells were incubated with 0.1 μg of purified anti-mouse FcγII/IIIR antibody (PharMingen) on ice for 10 minutes to block the non-specific binding of antibodies to the Fc receptor, then optimal amounts of FITC-conjugated and/or PE-conjugated anti-mouse IgG antibodies were added and incubated on ice for 30 minutes. Following the incubation, samples were washed with 2.0 ml of S-PBS^s twice to remove unbound antibodies. Thereafter, samples were resuspended in 1.0 ml of S-PBS^s. For B cell identification, anti-mouse B220 Ab was used. For T cell identification,

anti-mouse CD4, CD8, and Thy-1.2 Abs were used. PE- or FITC-conjugated rat $IgG_{2b}(k)$ was used as isotype control in each analysis. For the detection of F4/80 expression on Mϕ, RPC and TPC isolated from 6-week old *Emr1* +/+, +/-, -/- mice were used. For each sample, 10^6 cells were blocked with anti-mouse FcγII/IIIR antibody, washed, and incubated with FITC-conjugated Mac-1 Ab and PE-conjugated F4/80 Ab on ice for 30 minutes. Thereafter, samples were washed with 2.0 ml of $S\text{-}PBS^s$ twice, and resuspended in 1.0 ml of $S\text{-}PBS^s$. Flow cytometric analysis was performed on a $FACStar^{plus}$ (Becton Dickinson) flow cytometer and for each sample, 10,000 events (approximate 300 cells/second) were collected for data analysis. Data collected from the flow cytometer were analyzed using CELLQuest™ software.

Immunohistochemical analysis

Immunohistochemical staining of tissue sections of *Emr1* +/+, +/-, and -/- mice was carried out according to Martinez-Pomares *et al.* (133). Tissue and organs were collected and frozen in O.C.T. compound (Miles Inc., Elkhart, IN, USA)) cooled in isopentane (Sigma) over dry ice. Tissue sections were cut at 5 μm, air dried, and stored at -20°C until used. Before staining, frozen sections were thawed to room temperature and fixed in 2% paraformaldehyde in Hepes-buffered saline for 10 minutes at 4°C. In some cases, tissues were collected and fixed in PLP fixative (0.02M sodium periodate, 0.1M Lysine, and 0.25% paraformaldehyde) at 4°C for 24 hours. Fixed tissues were then dehydrated in ethanol and embedded in low melting point wax. Wax-embedded tissues were cut and processed according to standard protocols.

All the staining procedures were carried out at room temperature unless specifically stated. Fixed tissue sections were washed in 0.1% Triton X-100 in PBS for 15 minutes. A solution containing 0.01M glucose, 1mM sodium azide, and 40U

glucose oxidase (Sigma) in 0.1M phosphate buffer was then used to inactivate the endogenous peroxidase activity by incubation at 37°C for 15 minutes. After another wash for 15 minutes, 5% (vol/vol) normal goat or rabbit serum in PBS (blocking buffer) was reacted to the sections for 30 minutes to block any non-specific binding. Specific primary antibodies diluted to optimal concentrations in blocking buffer were used to bind to tissue sections for 60 minutes. Antibody binding was detected by staining sections with 1% (vol/vol) biotinylated rabbit anti-rat IgG or goat anti-rabbit IgG in blocking buffer for 30 minutes, followed by avidin-biotin-peroxidase complex and DAB staining solution (0.5 mg/ml diaminobenzidine, 0.024% H_2O_2, 10 mM Imidazole in PBS, pH7.4). The staining reaction was stopped by several washes in PBS and sections were counterstained using 0.1% crystal violet. For Mϕ staining in liver granuloma, 8-10 week old animals were injected i.p. with inactivated *Corynebacterium parvum* (*C. parvum*) (a gift from Prof. Siamon Gordon, Oxford University, UK) at 500 µg / 0.2 ml PBS / mouse, 7-10 days before sacrifice. Tissues and organs from the treated animals were processed and stained as described above.

Results

Emr1 gene targeting

A mouse *Emr1* genomic DNA fragment was identified from a 129/Sv cosmid DNA library and its physical map was determined by restriction enzyme digestion and Southern blot hybridization using various *Emr1* cDNA fragments as probes. This genomic DNA fragment is approximate 35 kb in length and contains exons 1 and 2 of the mouse *Emr1* gene. A partial restriction map of this fragment is shown in Fig. 10A.

An *Emr1* targeting construct was prepared by replacing a ~1.0kb *Hind* III-*Xba* I fragment containing the coding sequence of the first exon and the 5'-end of the first intron with a promoterless β-gal/pGK-Neo cassette (126). This region was chosen because the first exon of *Emr1* gene encodes the first ten amino acids of the signal peptide of the Emr1 protein (112), which is important for exporting membrane proteins to the extracellular surface (134). Deletion of these amino acids would render the translated protein with incorrect cellular location and therefore abrogate its normal function as a transmembrane receptor even if there is any leaky transcription and translation. The use of this targeting construct to replace the wild-type allele through homologous recombination is therefore predicted to create a null mutation. Furthermore, since the promoterless β-gal/pGK-Neo cassette was inserted into the *Hind* III site of the 5'-UTR, the endogenous *Emr1* gene promoter should direct the expression of the β-galactosidase gene to allow for the detection of *Emr1* gene expression *in situ* in transgenic animals. The homologous regions in both arms of the targeting construct are a 1.8 kb fragment 5' to the first exon and a 2.8 kb fragment containing the exon 2 and the 3'-end of intron 1 (Fig. 10A).

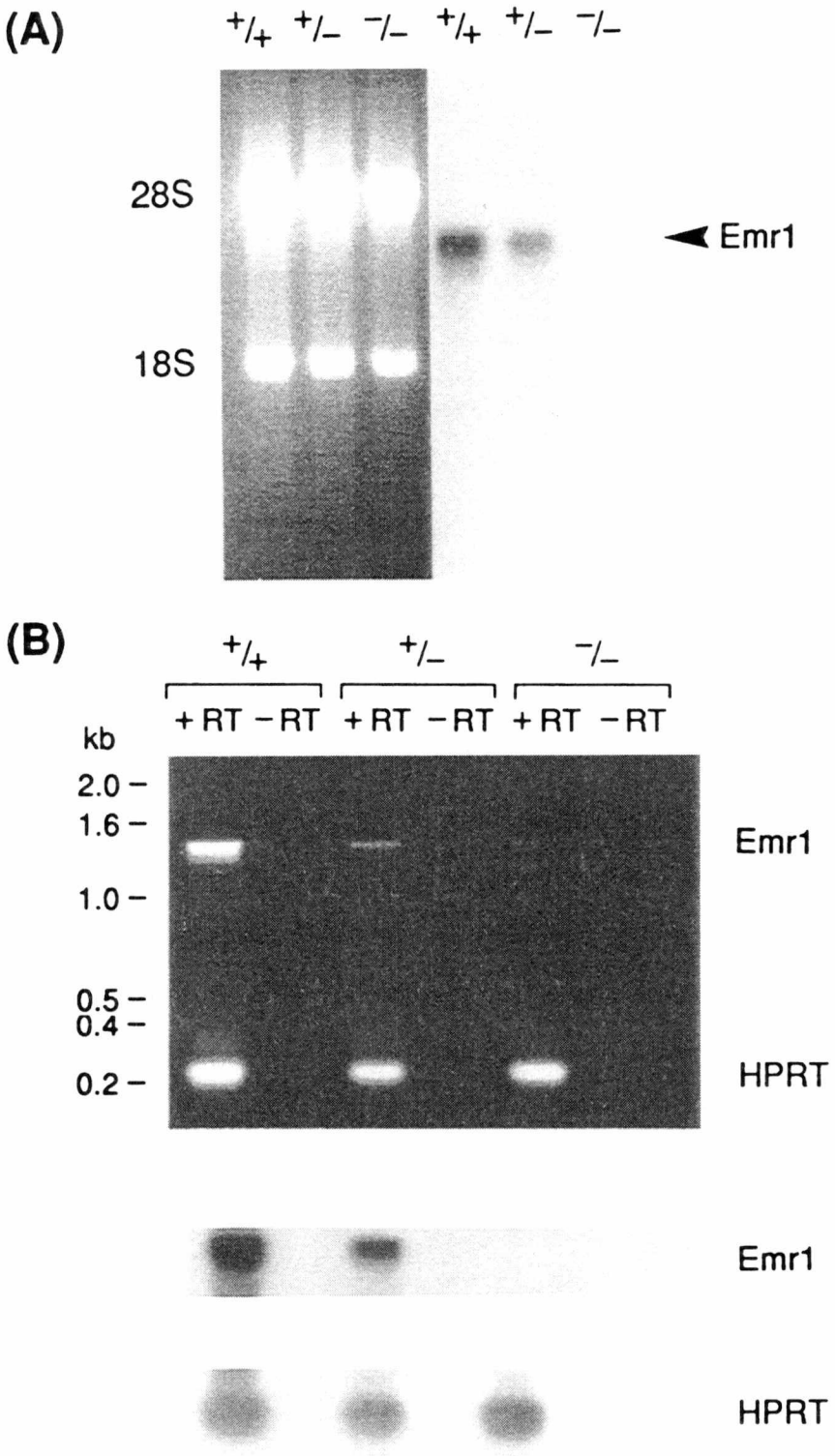
The targeting vector was linearized and electroporated into CJ7 ES cells. Following drug selection, 5 targeted ES cell clones were identified from 576 G418-resistant colonies by Southern blot analysis (data not shown). Targeted ES cells were microinjected into blastocysts, implanted into the uteri of pseudo-pregnant female mice, and chimeric animals were generated. Eight chimeric male mice derived from one targeted ES cell clone transmitted the mutation through the germ line, resulting in *Emr1* +/- animals. *Emr1* +/- animals were bred to produce -/- mice, which were identified by genotyping of tail DNA using PCR analyses and Southern blot hybridization with probes A and B (Fig. 10B). Probe A is a DNA fragment external to the homology region of the targeting construct and detects *Hind* III fragments of 10 kb

and 14 kb specific to the wild-type allele and the targeted allele, respectively. In contrast, probe B is derived from the intron 1 region that was deleted in the targeted allele and therefore only detect the wild-type alleles. The presence of a 14 kb *Hind* III fragment detected by probe A and the absence of probe B hybridized DNA fragments in *Emr1* -/- DNA samples indicated that homologous recombination has indeed occurred at the *Emr1* locus. This result was further verified by PCR genotyping analyses in which only the 300 bp fragment specific to the *Emr1* targeted allele but not the wild-type allele-specific 600 bp fragment was amplified in *Emr1* -/- DNA samples (Fig. 10B).

To determine whether the DNA alteration at the *Emr1* locus affects gene transcription, Northern blot hybridization and RT-PCR analyses were carried out using total thymus RNA isolated from 5-week old *Emr1* +/+, +/-, and -/- animals. As shown in Fig. 11, the expression level of *Emr1* transcript in *Emr1* +/- animals was reduced compared to that detected in +/+ animals, probably due to the gene disruption in one of the two *Emr1* alleles. More importantly, no *Emr1* transcript was identified in *Emr1* -/- animals by either method. The same results were observed with RNA samples isolated from spleens and livers (data not shown). The complete absence of *Emr1* transcript in *Emr1* -/- mice indicates that we have generated a *Emr1* null mutation. Taken together, results from the analyses of DNA and RNA samples of *Emr1* -/- mice show that the genomic DNA structure at the *Emr1* locus has been altered and the *Emr1* gene transcription was disrupted.

The expression of the β-galactosidase gene was also examined by Northern blot analysis using a full-length lacZ cDNA as a probe. No positive signal was detected in RNA isolated from thymus, spleen, and liver of the *Emr1* +/-, and -/- mice, even after prolonged exposure to X-ray film (data not shown). In addition, *Emr1* +/+, +/-, and -/- mouse embryos at 9.5, 10.5, and 11.5 dpc were collected, fixed, and stained for

Fig. 11 Northern blot hybridization (A) and RT-PCR analysis (B) of total RNA from thymus of *Emr1* +/+, +/-, and -/- animals. (A) Total RNA (10 μg) separated in a formaldehyde agarose gel was visualized by EtBr staining (left panel), transferred to a nylon membrane, and hybridized with an *Emr1* cDNA probe (right panel). The *Emr1* transcript as well as 18S, and 28S rRNAs are indicated. (B) RT-PCR analysis was carried out using the primers specific for the *Emr1* and *Hprt* cDNAs. The amplified cDNA fragments were transferred to a nylon membrane, and hybridized with specific *Emr1* and *Hprt* cDNA probes. RT-PCR products were shown in EtBr staining gel (upper panel) and in autoradiography films (middle and lower panels). The RT-PCR reaction was conducted with (+RT) or without (-RT) the reverse transcriptase for each sample to ensure no DNA contamination. Sizes (kb) of 1kb DNA ladder are shown.


(A)
+/+ +/− −/− +/+ +/− −/−
28S
18S
Emr1
(B)
+/+ +/− −/−
+RT −RT +RT −RT +RT −RT
kb
2.0
1.6
1.0
0.5
0.4
0.2
Emr1
HPRT
Emr1
HPRT

β-galactosidase activity using X-gal staining solution as described (135). No specific staining patterns were observed in the *Emr1* +/-, and -/- embryos (data not shown), while non-specific background coloring was noted after prolonged staining in all embryos tested, including the *Emr1* +/+ embryos (data not shown). These results indicate that the β-galactosidase gene is not expressed in the *Emr1* mutant mice.

Characterization of *Emr1* -/- mice

Emr1 -/- mice were born at the expected Mendelian ratio. As shown in Table 4, out of 700 progeny produced from the breeding *of Emr1* +/- animals to date, 168 (24.0%) were +/+, 358 (51.1%) were +/-, and 174 (24.9%) were -/- mice. This result clearly indicates that the Emr1/F4/80 protein is not required for normal mouse embryonic and neonatal developments. *Emr1* -/- mice appear healthy, have normal bodyweight, and are fertile. Histological examination of various organs and tissues, including bone marrow, liver, lung, kidney, heart, spleen, thymus, and lymph nodes from *Emr1* -/- animals failed to show any gross abnormalities (Table 5). Biochemical and microscopic analyses of peripheral blood samples carried out at 2 months of age were also found to be normal. The life span of the *Emr1* -/- mice is indifferent from that of the +/+ and +/- mice, with the oldest -/- mice living up to 1 year of age to date in a conventional mouse facility.

Since the Emr1/F4/80 protein is expressed specifically on Mϕ, which plays an important role in the innate and adaptive immune systems, we next examined the development of Mϕ as well as B and T lymphocytes. The number of resident peritoneal cells (RPC) isolated from *Emr1* -/- mice was comparable to that isolated from *Emr1* +/+ and +/- mice (Table 6). The same was true for thioglycollate elicited peritoneal cells (TPC), as both *Emr1* +/+ and -/- mice were found to produce similar

Table 4. Genotype of offspring derived from *Emr1* heterozygous parents

Emr1 +/- X *Emr1* +/-

wild-type (+/+)	*Emr1* +/-	*Emr1* -/-
168 / 700 (24.0%)	358 / 700 (51.1%)	174 / 700 (24.9%)

Table 5. Summary of tested properties of Mϕ.

Genotypes	+/+	*Emr1* +/-	*Emr1* -/-
Histology[a]	N	N	N
Hematopoietic cell compartments[b]	N	N	N
Formation of peritoneal exudate[c]	N	N	N
Cell surface markers expression[d]	N	N	N
Blood histology[e]	N	N	N
Formation of granuloma	N	N	N

N : Normal
a : Assayed by microscopic examination of tissue sections.
b : Assayed by FACS analysis of spleen, thymus, and resident peritoneal cells with lymphoid markers, CD4, CD8, B220, and Thy-1; Mϕ markers, Mac-1 and F4/80.
c : Assayed by cell counts and morphology.
d : Assayed by immunohistochemical analysis of tissue sections using Mϕ-specific Mabs, F4/80, 3D6, 5C6, 2F8, and FA-11.
e : Assayed by cell counts and morphology.

Table 6. Number of resident peritoneal cells (RPC) and thioglycollate elicited peritoneal cells (TPC) isolated from *Emr1* +/+, +/-, and -/- mice[a]

Genotypes	+/+	+/-	-/-
RPC[b]	8.35±0.54	8.07±1.07	8.08±1.48
TPC[c]	3.10±0.26	ND	2.96±0.15

a. RPC and TPC were isolated by peritoneal lavage and counted by hemocytometer. Data represents mean S.D. of cell numbers from three animals. ND: Not done.
b. Total cell number x 10^{-6}.
c. Total cell number x 10^{-7}.

numbers of TPC. These data not only indicate that the lack of the F4/80 protein did not affect the proliferation and/or differentiation of Mϕ cells, but also demonstrates that F4/80$^-$ monocytes/Mϕ were able to respond to exogenous stimili and migrate to the peritoneal cavity as well as the wild-type cells, suggesting the F4/80 protein may not be involved in the process of Mϕ mobility.

Flow cytometric analysis of RPC and TPC of *Emr1* +/+, and -/- mice showed no F4/80 Ab staining on the *Emr1* -/- cells (Fig. 12). These data confirm that the mutation at the *Emr1* locus disrupted the synthesis of the F4/80 protein. In addition, flow cytometric analysis of cell surface marker expression on *Emr1* -/- splenic, and thymic cell populations showed that the number and distribution of B and T lymphocytes was normal in comparison to *Emr1* +/+, and +/- animals. As shown in a representative analysis in Fig. 13, both *Emr1* +/+, and -/- mice have normal expression of CD4 and CD8 on thymic T cells as well as on splenic cells. The expression of B220 and Thy-1.2 on splenic cells was also comparable in *Emr1* -/- and +/+ mice. These findings suggest that although Mϕ participate in the innate and adaptive immune responses and are believed to be involved in the development of lymphoid cells (136), the inactivation of the *Emr1* gene *in vivo* does not have effect on the development of mature lymphoid cells from hematopoietic precursor cells.

Immunohistochemical analysis of tissue Mϕ populations of the *Emr1* -/- mice

To investigate whether the characteristics of tissue Mϕ populations of the *Emr1* -/- animals were affected by the inactivation of the *Emr1* gene, the expression profile of various Mϕ-specific cell surface proteins was examined using immuno-histochemical analysis. In naive mice, F4/80 is normally expressed in blood

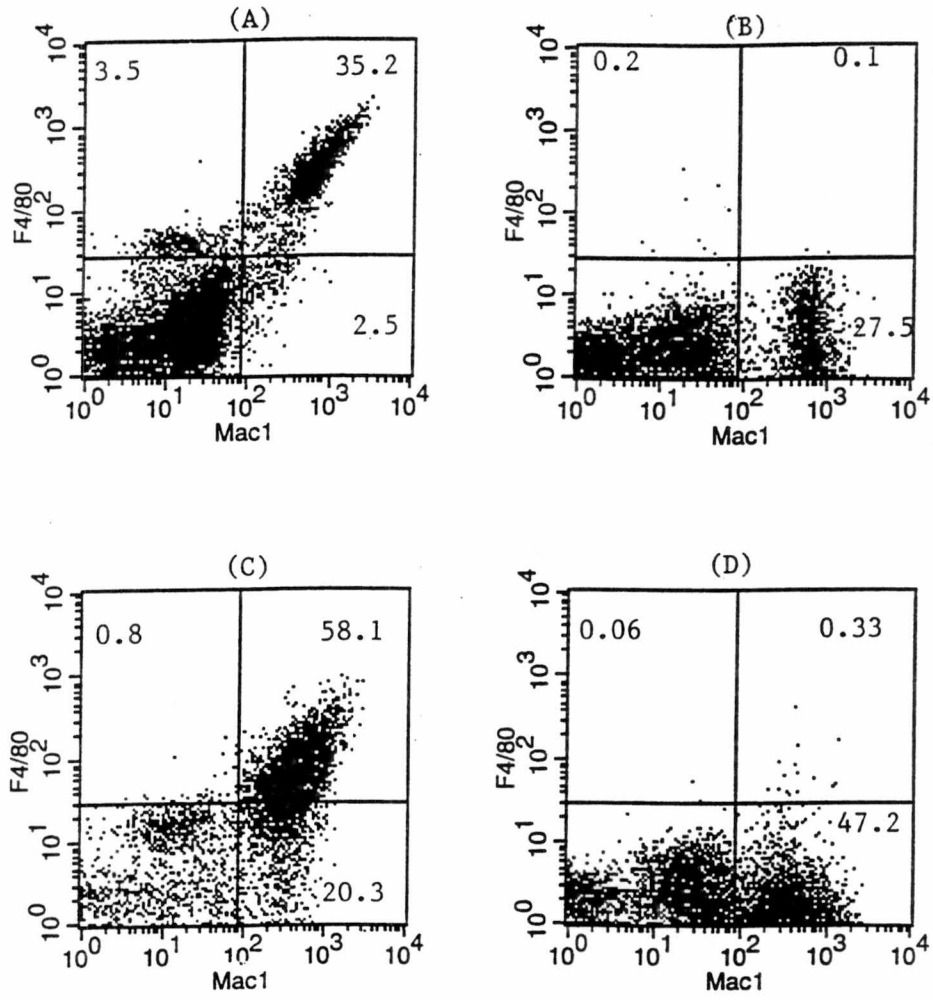


Fig. 12 Flow cytometric analysis of resident peritoneal cells (A, B) and thioglycollate elicited peritoneal cells (C, D) of *Emr1* +/+ (A, C), and -/- (B, D) mice. Cells were incubated with PE-conjugated anti-mouse F4/80 Ab and FITC-conjugated anti-mouse Mac-1 Ab, washed, and analyzed using FACStar^{plus} (Becton Dickinson). Numbers in each profile indicate the percentage of events.

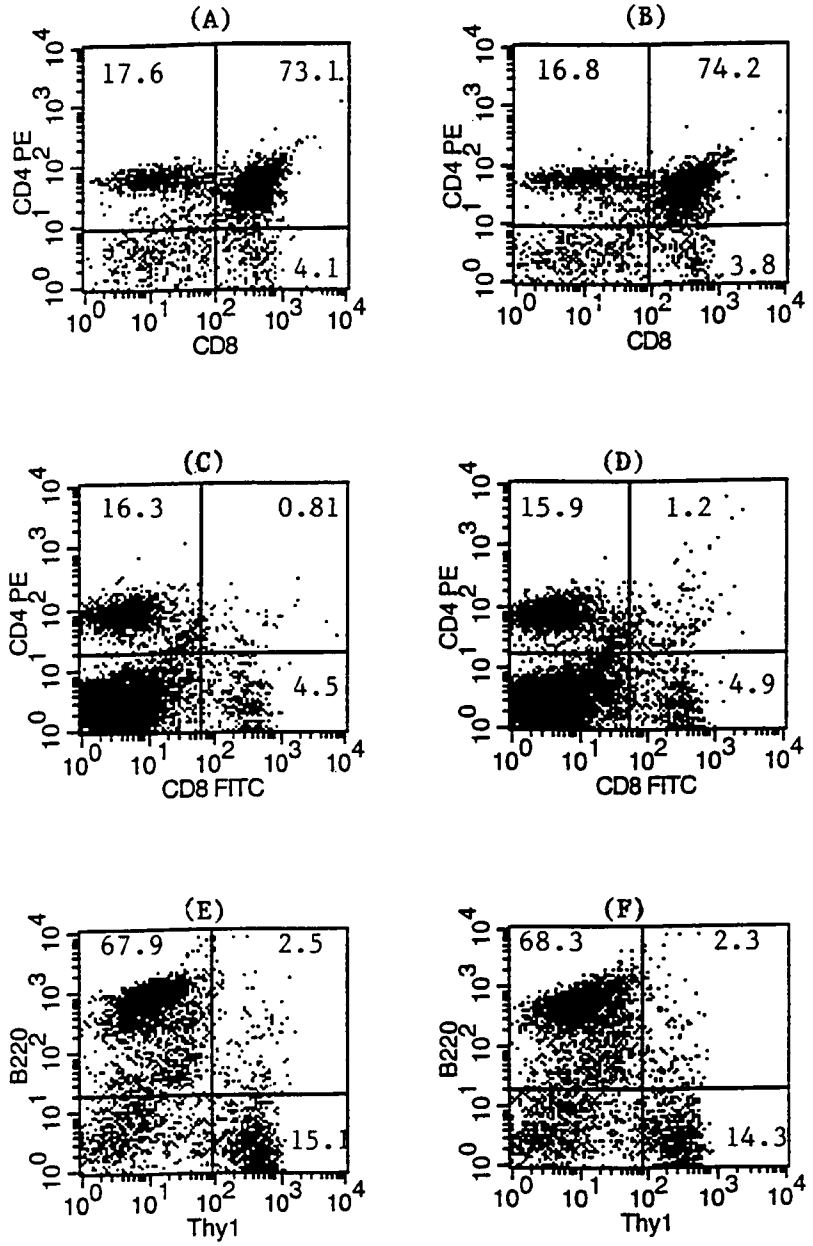


Fig. 13 Flow cytometric analysis of thymic (A, B) and splenic (C, D, E, and F) cell populations of *Emr1* +/+ (A, C, E), and -/- (B, D, F) mice. Cells were incubated with PE-conjugated anti-mouse CD4 Ab and FITC-conjugated anti-mouse CD8 Ab (A, B, C, D), or PE-conjugated anti-mouse B220 Ab and FITC-conjugated anti-mouse Thy-1.2 Ab (E, F), washed, and analyzed using FACStar^{plus} (Becton Dickinson). Numbers in each profile indicate the percentage of events.

monocytes, and most of the mature Mϕ populations in almost every tissue, including bone marrow (stromal Mϕ), liver (Kupffer cell), red pulp of spleen , lymph node, epidermis (Langerhans cell), etc. (113). Here, both polyclonal and monoclonal anti-F4/80 antibodies detected the typical F4/80 expression patterns in the *Emr1* +/+ tissues. On the contrary, no F4/80 protein expression was identified in the *Emr1* -/- mice. For instance, F4/80 expression was detected on Mϕ in the red pulp of spleens and on Kupffer cells of the livers of the *Emr1* +/+ animals (Fig. 14A, C) but was undetectable on Mϕ of the same tissues of the *Emr1* -/- mice (Fig. 14B, D). The complete absence of the F4/80 protein once again confirmed that a *Emr1/F4/80* null mutation has been generated.

In contrast to the F4/80 protein expression, MAb 3D6 staining profiles were identical in the *Emr1* -/- as well as the *Emr1* +/+ tissues, *e.g.*, in spleen and liver (Fig. 15). Mab 3D6 specifically recognized sialoadhesion, a 185 kDa Mϕ-restricted adhesion molecule that mediates sialic acid-dependent binding to cells (128). Sialoadhesion is normally expressed on bone marrow stromal Mϕs, and Mϕs in secondary lymphoid organs, including spleen, lymph nodes, liver, and Peyer's patches (137). In *Emr1* +/+ and -/- spleens, 3D6 staining was easily detected in the red pulp areas, where F4/80 is normally expressed ($F4/80^+$), however, strong staining was also detected in the inner region of the marginal zone, the marginal metaphils, where F4/80 is normally not expressed ($F4/80^-$) (Fig. 15C, D) (138). This result indicates that the distribution of different Mϕ subpopulations ($F4/80^+3D6^+$, and $F4/80^-3D6^+$) is not affected in the *Emr1* -/- mice.

MAbs 2F8, 5C6, and FA/11 that specifically recognize the Mϕ scavanger receptor (129), the type 3 complement receptor (130), and macrosialin (131), respectively, were also used in this study and the same staining patterns for each

Fig. 14 MAb F4/80 staining in the spleens (A, B) and livers (C, D) of the *Emr1* +/+ (A, C) and -/- (B, D) animals. $F4/80^+$ Mϕ (brown) were found in the red pulp of spleen of the *Emr1* +/+ (A) but not of the *Emr1* -/- (B) mice. Similarly, $F4/80^+$ Kupffer cells (brown) were observed in the liver of the *Emr1* +/+ (C) but not of the *Emr1* -/- (D) mice. Magnification X200 (A, B), X400 (C, D).

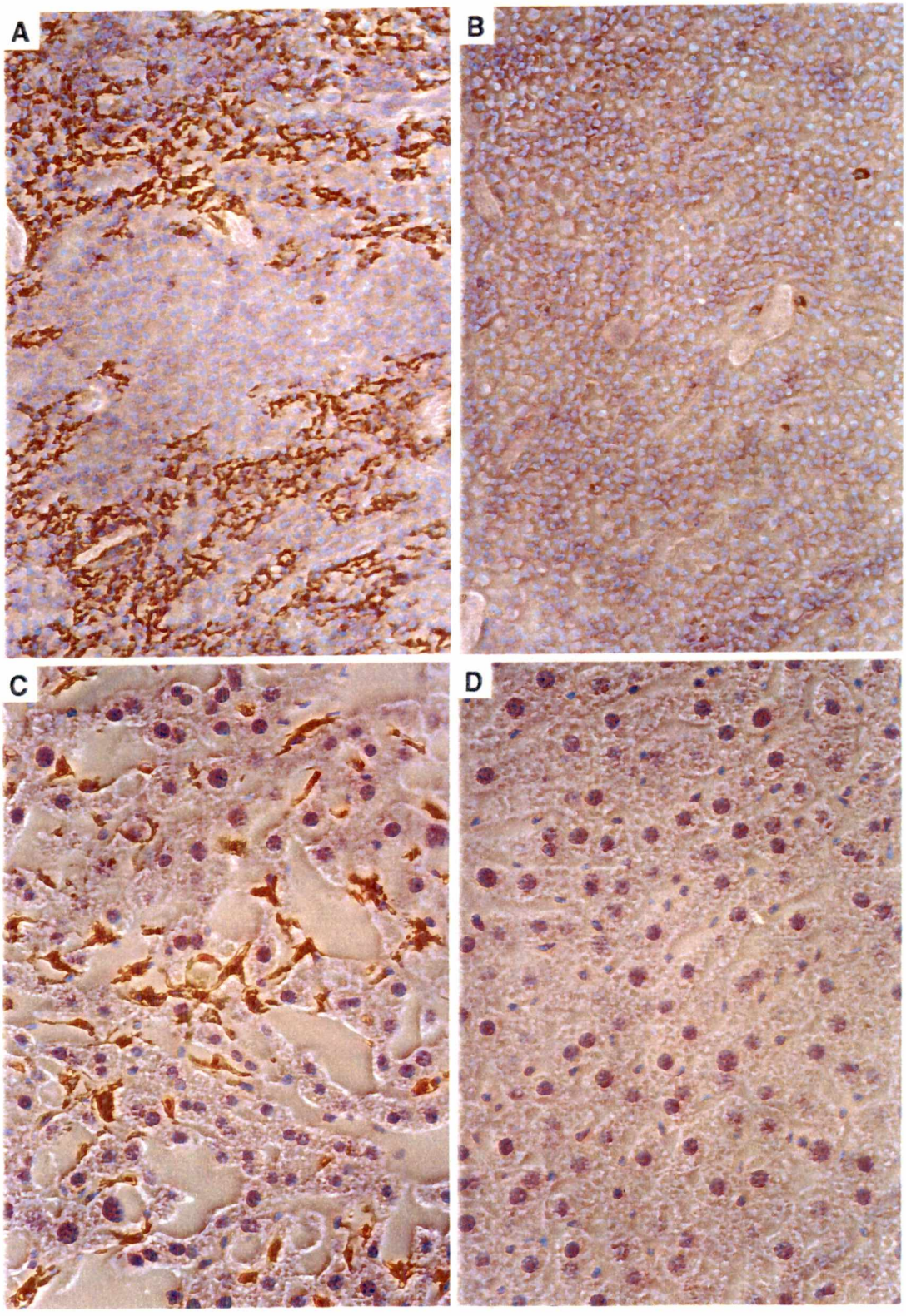

A
B
C
D

Fig. 15 MAb 3D6 staining in the livers (A, B) and spleens (C, D) of the *Emr1* +/+ (A, C) and -/- (B, D) animals. MAb 3D6 specifically recognizes sialoadhesion, a Mϕ-restricted sialic acid-dependent adhesion protein. $3D6^+$ Kupffer cells (brown) were found in the livers of both *Emr1* +/+ (A) and -/- (B) mice. $3D6^+$ Mϕ (brown) were also found in the red pulp and marginal zone areas of spleens of the *Emr1* +/+ (C) and -/- (D) mice. The strongest 3D6 staining is on Mϕ that are in the inner region of the marginal zone (marginal metallophil), forming a tight ring (arrow) around the white pulp. Magnification X400 (A, B), X200 (C, D).

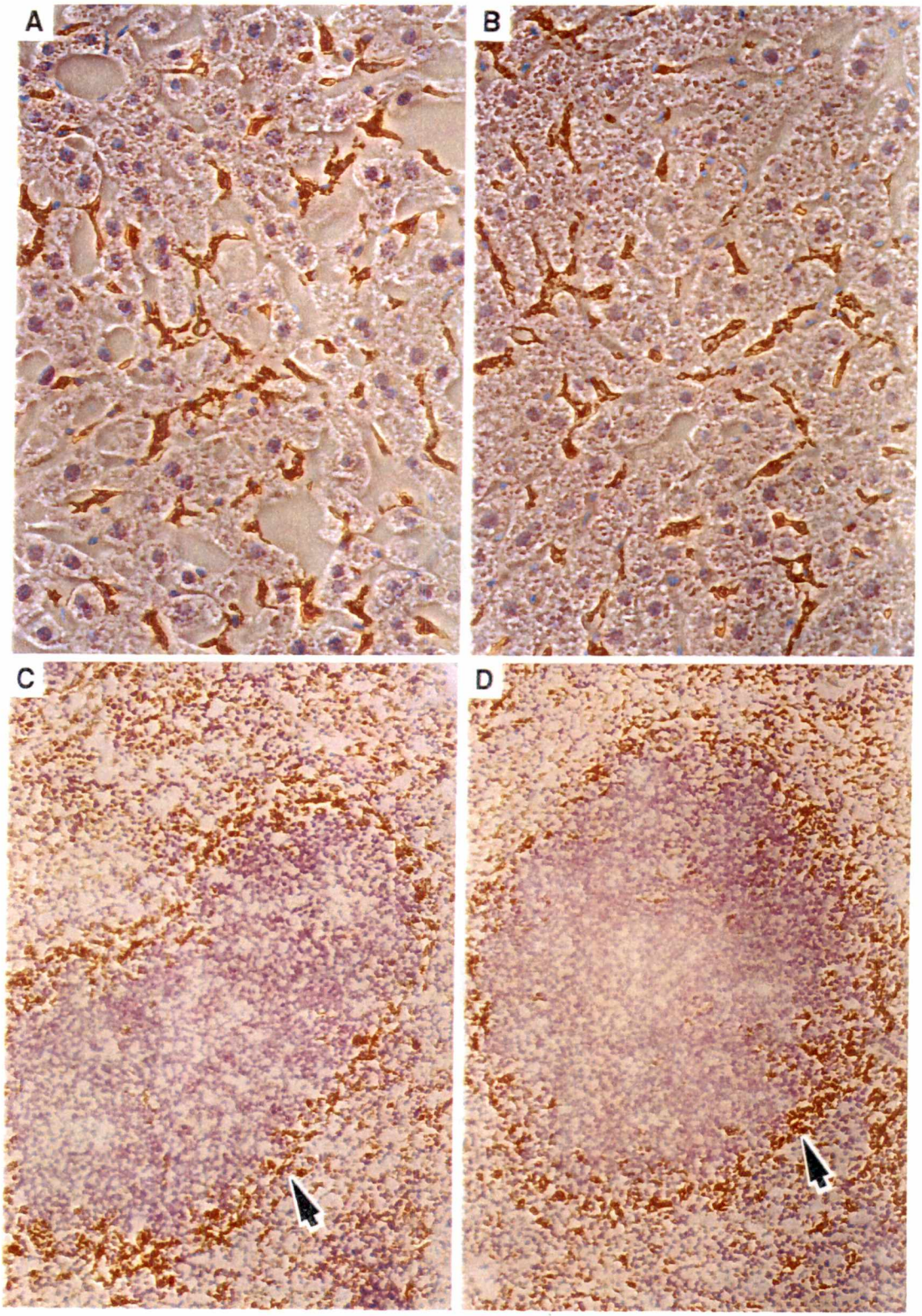

A
B
C
D

Fig. 16 MAb FA-11 staining in the livers of the *Emr1* +/+ (A) and -/- (B) animals, and MAb 5C6 staining in the spleens of the *Emr1* +/+ (C) and -/- (D) animals. MAb FA-11 is specific for macrosialin, while MAb 5C6 specifically recognizes C3R. FA-11^{+} Kupffer cells (brown) were identified in the livers of both *Emr1* +/+ (A) and -/- (B) mice. 5C6^{+} Mϕ (brown) were found in the red pulp, marginal zone of spleens of the *Emr1* +/+ (C) and -/- (D) mice. A few 5C6^{+} Mϕ can be identified inside the white pulp area as well. Magnification X400 (A, B), X200 (C, D).

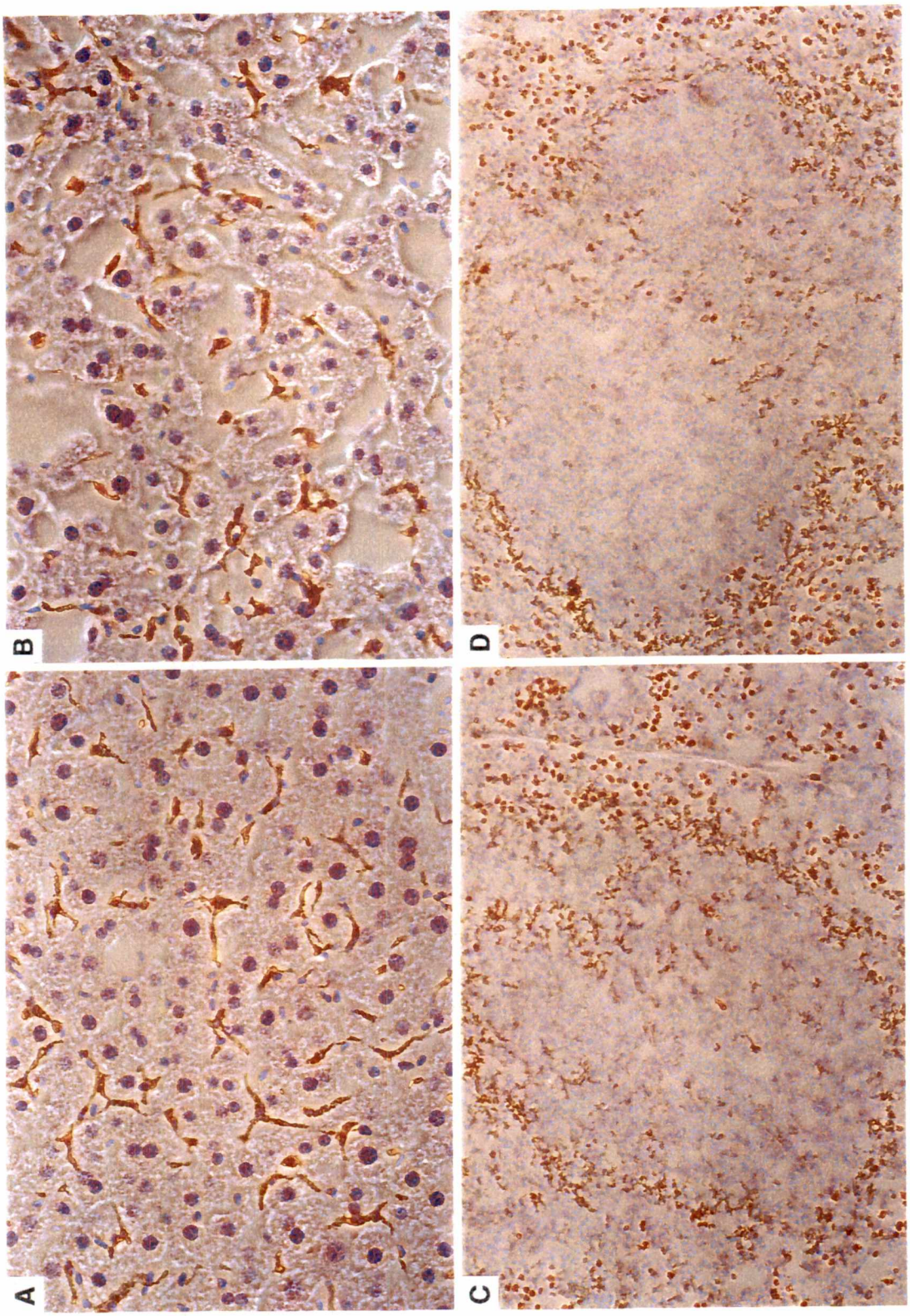

B
D
A
C

Table 7. Summary of immunohistochemical analysis

Tissue	Cell type	Reactivity to antibodies									
		F4/80		3D6		5C6		2F8		FA-11	
		W	M	W	M	W	M	W	M	W	M
Spleen	Red pulp Mϕ	++	-	++	++	+	+	+	+	+	+
	White pulp	-	-	+/-	+/-	+	+	+/-	+/-	++	++
	Outer marginal zone	+/-	-	+	+	+	+	+++	+++	++	++
	Marginal metallophil	-	-	+++	+++	+	+	+	+	+	+
Thymus		+	-	+/-	+/-	-	-	++	++	++	++
Lung	Alveolar Mϕ	+	-	+	+	+/-	+/-	+++	+++	+	+
Heart		+/-	-	++	++	+/-	+/-	++	++	+	+
Intestine		++	-	+	+	+	+	++	++	+	+
Liver	Kupffer cells	++	-	+	+	+/-	+/-	+	+	++	++
Liver(*C.parvum*)	Activated Mϕ	+++	-	++	++	++	++	+++	+++	++	++

W : *Emr1* wild-type mice

M : *Emr1* homozygous mutant mice

MAb were observed in *Emr1*+/+ and -/- tissues (Fig. 16, Table 7, and data not shown). In addition, injection of mice i.p. with inactivated *C.pavum* was found to induce granuloma formation equally in the *Emr1*+/+ and -/- livers. Moreover, in concert with the immunostaining patterns described above for the normal tissue Mϕ, the activated Mϕ in the *Emr1* +/+ and -/- granulomas also express the same levels of Mϕ-specific cell surface marker proteins (Fig. 17). These results show that the *Emr1* -/- tissue Mϕs express all the tested cell surface marker proteins, except F4/80, at comparable levels, and respond to the exogenous stimili in a normal manner.

Discussion

Using the gene targeting technique in ES cells, we have generated a mouse strain in which the *Emr1/F4/80* gene was mutated and the synthesis of the F4/80 protein was disrupted. Several lines of evidence show that the mouse strain generated in this study is a null mutant. First, analysis at the genomic DNA level using PCR and Southern blot hybridization demonstrated that in the *Emr1* -/- mice the DNA structure/organization at the *Emr1* locus was altered (Fig. 10B), evidently resulting from the homologous recombination between the *Emr1* wild-type allele and the targeted construct in ES cells. Secondly, analysis of total RNA isolated from the *Emr1* -/- mice using RT-PCR and Northern blot hybridization failed to show any *Emr1* RNA expression (Fig. 11). Finally, the F4/80 protein expression, analyzed by the flow cytometry and the immunohistochemical staining, was found to be completely absent in the *Emr1* -/- mice (Fig. 12, 14). The *Emr1* -/- mice therefore provide us an excellent animal model for studying the biological function of the F4/80 glycoprotein.

Since the F4/80 protein is a specific cell surface marker for mouse Mϕ, it was of great interest to study its expression patterns *in vivo* in order to gain more insight into

Fig. 17 Immunohistochemical staining of liver sections of animals injected with inactivated *C.parvum* using MAb 2F8 (A,B) and MAb 5C6 (C,D). MAb 2F8 is specific for the Mϕ scavanger receptor. Granuloma was formed mostly by activated Mϕ that were recruited to liver to fight off *C.parvum*. Predominant $2F8^{+}$ Mϕ were evident in the liver granuloma of *Emr1* +/+ (A) and -/- (B) mice. $5C6^{+}$ Mϕ were also identified in *Emr1* +/+ (C) and -/- (D) liver granuloma. Magnification X200 (A, B), X400 (C, D).

A
B
C
D

the ontogeny and the distribution of Mϕ in mouse tissues. In an attempt to recapitulate the expression of the F4/80 protein *in vivo* using X-gal staining, a promoterless β-galactosidase gene was inserted into the targeting construct right behind the 5'-UTR of the *Emr1* gene (Fig. 10A). The rationale for this strategy is to use the endogenous *Emr1* gene promoter to direct the expression of the β-galactosidase gene, therefore mimicking the F4/80 protein expression by the β-galactosidase activity. Unfortunately, neither RNA nor protein expression of the β-galactosidase gene was identified in the *Emr1* +/- and -/- mice (data not shown), even though DNA sequencing analysis of the *Emr1* -/- DNA samples showed that the DNA organization at the *Emr1* locus is exactly as shown in Fig. 10A, and predicted the Emr1-directed expression of the β-galactosidase gene (data not shown).

This unexpected result indicates that either the β-galactosidase gene transcript is not expressed at all or the β-galactosidase RNA transcript is extremely unstable in this system. There are several possible explanations for this result : (1) An enhancer element requiring for the optimal promoter activity may have been deleted in the targeted allele, thereby adversely affecting the β-galactosidase gene expression. An approximately 1.0 kb fragment containing the coding sequence of the exon 1 and 5'-end of the intron 1 was deleted in the targeted allele. Whether this DNA fragment has the enhancer activity remains to be determined. (2) The inserted exogenous DNA in the targeted allele may cause a change in chromosomal DNA structure, which in turn inhibits the activity of the endogenous *Emr1* gene promoter. The real cause for the inability to express β-galactosidase gene in the mutant mice can not be elucidated here; however, attention should be given to *in vitro* and/or *in vivo* experiments of the *Emr1* gene promoter activity, such as the use of the *Emr1* gene promoter for Mϕ-specific expression studies in Mϕ cell lines or transgenic mice models.

The findings that the *Emr1* -/- mice were born at the expected mendelian ratio, and that the mutant mice appeared healthy, showed no gross abnormalities in various organs and tissues, and were fertile in a conventional mouse facility indicate not only that the F4/80 protein is not required for normal mouse embryonic and neonatal developments but also suggest that the F4/80 protein is not necessary for maintaining a normal immune response. This was further supported by the finding that the *Emr1* -/- mice have normal number of B and T lymphocytes as determined by flow cytometry (Fig. 13). In addition, the typical expression patterns of several Mϕ-specific cell surface markers (Sn, CR3, MSR, and MS) in various tissues of the *Emr1* -/- mice detected by immunohistochemical staining (Fig. 15-17) and the normal number of RPC and TPC found in the *Emr1* -/- mice also indicate that the F4/80 protein is not involved in the cell survival / proliferation process or in the regulation of expression of other cell surface proteins.

Other attempts to reveal phenotypes for the *Emr1* -/- mice have proven unsuccessful to date (data not shown). The lack of novel phenotypes of the *Emr1* -/- mice may be interpreted by several ways : (1) The F4/80 protein is an orphan receptor with no biological function(s); (2) The function of the F4/80 protein was complemented by other cell surface receptor(s) on Mϕ; (3) The assays used in this study were not sensitive enough to detect subtle phenotypic changes in the *Emr1* -/- Mϕ or the F4/80 receptor is involved in specific biological functions that have yet to be analyzed. In light of the unusual protein structure, the Mϕ-specific expression patterns, and the highly conserved amino acid sequence of the F4/80 protein between species, we would argue that the F4/80 protein should have certain biological function(s). The recent cloning and mapping of EGF-TM7 genes, including human *EMR1* (95), human *CD97* (139), and mouse *Emr1/F4/80* (110, 112), are suggestive of a gene cluster of a novel protein family, whose members not only have

a similar protein structure but also are linked on the same chromosome. Both human *EMR1* and *CD97* have been mapped to human chromosome 19p13 region (95, 139), whereas mouse *Emr1/F4/80* is mapped to the mouse chromosome 17 region that is syntenic to human chromosome 19p13.3 (112). Indeed, analysis of the gene structure of human *CD97* gene has suggested an exon shuffling mechanism for the generation of the EGF-TM7 receptor family (140). It is therefore reasonable to suggest that other members of the EGF-TM7 protein subfamily may exist and if it is the case, the new EGF-TM7 protein member(s) and the F4/80 protein may have overlapping function(s), which may compensate for the loss of the F4/80 protein in the *Emr1* -/- mice. On the other hand, numerous functional tests have been developed for Mϕ (132); however, only a limited number of them were conducted in this study. It is very possible that a specific function for the F4/80 protein has not yet been identified due to the inappropriate or insensitive assay systems employed here. Further studies, especially *in vivo* immune-reponsive assays, such as the reactions of mice to *Listeria monocytogenes* infection and/or virus infction, are needed to explore the possible function(s) of the F4/80 protein.

Regardless of the lack of novel phenotypes, the *Emr1* -/- mice do represent a useful resource for studying F4/80-related biological questions : (1) As mentioned above, additional EGF-TM7 protein members may exist, it is therefore of interest to identify the putative unknown members. It has been demonstrated that degenerate primers deduced from highly conserved regions of members of structurally and/or functionally-related protein families can be used to amplify and identify new protein members by RT-PCR (141). Given that EGF-TM7 proteins are highly homologous at the 7TM domain and that no *Emr1/F4/80* RNA transcript is detected in the *Emr1* -/- mice, the application of RT-PCR analysis using RNA sample isolated from the *Emr1* -/- mice and degenerate primers deduced from the conserved regions among the EGF-TM7

proteins may allow one to amplify homologous RNA species and isolate new EGF-TM7 genes. (2) Although the human *EMR1* gene and the mouse *Emr1/F4/80* gene have similar RNA expression patterns in hematopoietic cell lines (95, 110, 112), it is not known whether they also have similar protein expression patterns in tissues. To date, there are no known antibodies that specifically recognize the human EMR1 protein. Furthermore, it is not known whether the anti-mouse F4/80 MAb can cross-react to the human EMR1 protein. A highly-specific Ab against the human EMR1 protein therefore is needed for studying the cellular distribution of the protein in human tissues. However, due to the high degree of conservation of the amino acid sequence of the EMR1 proteins between species, poor immunogenicity is expected to be a problematic obstacle for this objective. To circumvent similar problems with poor immunogenicity in mice in an attempt to generate specific blocking MAb against rat P-selectin, Walter *et al.* (142) recently reported the identification of a novel MAb that blocks the adhesion function of rat as well as mouse P-selectin by immunizing P-selectin knock-out mice. Since no F4/80 protein is being made in the *Emr1* -/- mice, the *Emr1* -/- mice would be an excellent host for the production of highly-specific Abs for the human EMR1 protein.

Chapter 4 Identification of Targeted Embryonic Stem Cells by Slot Blot Hybridization Analysis

Abstract

An alternative method for identifying targeted embryonic stem cells is presented. Two specific DNA probes, one from the genomic region adjacent to the homology target of the targeting vector (external probe) and the other from the genomic region deleted after homologous recombination (internal probe), are used in slot blot hybridization analysis. By comparing the hybridization intensities of DNA samples detected by the external DNA probe to those detected by the internal DNA probe, the number of wild-type alleles of the gene of interest can be determined. This information is then used to identify putative targeted cells. The sensitive slot blot hybridization method offers a rapid, inexpensive alternative to identify homologous recombination events in embryonic stem cells.

Introduction

The disruption of genes in mouse embryonic stem (ES) cells (gene targeting) has been widely used in recent years to study the biological functions of specific genes *in vivo* (50, 51, 143). A gene targeting construct is often made by replacing one or more exons of the gene of interest with a selectable marker (*e.g.*, neomycin phosphotransferase [NEO] or hypoxanthine-guanine phosphoribosyl transferase [HPRT] gene cassette). Linearized targeting vectors are electroporated into pluripotent ES cells and drug-resistant colonies are subsequently selected and analyzed. The frequency of homologous recombination and the efficiency of identifying targeted ES

cells are therefore critical in obtaining correctly targeted ES cell clones. Several methods have been introduced to improve the frequency of homologous recombination in ES cells, including the use of isogenic DNA in the targeting construct, positive-negative drug selection strategies, and increasing the size of the homologous regions in both arms of the targeting construct (143). However, little effort has been made to enhance the efficiency of identifying targeted ES cells. Two techniques, the polymerase chain reaction (PCR) and Southern blot analysis, are commonly used to screen for targeting events. Although these techniques are effective, both are expensive, time-consuming, and labor-intensive.

To simplify the screening process, an alternative method for identifying targeted ES cells is presented. This new method takes advantage of the sensitivity of slot blot hybridization analysis and the fact that targeted ES cells contain only one copy of the wild-type allele and one copy of the engineered mutant allele for the gene of interest. The slot blot hybridization analysis, which can detect different copy numbers of specific DNA sequences, enables us to discriminate the targeted ES cells from the non-targeted ES cells.

Materials and methods

Electroporation and selection of ES cells

25 μg of a linearized gene targeting construct was added to a total of 1 X 10^7 CJ7 ES cells (127) resuspended in 0.8 ml of phosphate buffered saline (PBS). Electroporations were carried out using a BioRad gene pulser at settings of 0.24 Kvolts and 500 μF. Electroporated ES cells were then cultured and selected in

standard ES cell medium containing 250 μg/ml G418 (Life Technologies, Grand Island, NY, USA) as previously described (12).

Genomic DNA isolation

ES cells in 24-well petri dishes were washed twice with PBS and incubated with 0.5 ml of 0.05% trypsin-EDTA (Life Technologies, Grand Island, NY, USA) at 37°C for 3-5 minutes in a 5% CO_2 incubator. One ml of culture medium containing 10% fetal bovine serum (FBS) was added to each well and a single suspension of cells was generated by repeated pipetting. Cells were pelleted by centrifugation at 5,000 rpm for 5 minutes at room temperature and resuspended in 400 μl of cell lysis buffer (100 mM NaCl, 10 mM Tris-HCl, pH 8.0, 25 mM EDTA, pH 8.0, and 0.5% SDS) with proteinase K (10 μg/ml) and incubated at 50°C overnight. After phenol/chloroform extraction, two volumes of 100% ethanol were added to the aqueous layer to precipitate the genomic DNA. The DNA was pelleted by centrifugation at 14,000 rpm for 5 minutes and resuspended in 200 μl of TE buffer (10 mM Tris-HCl, pH 7.6, 1 mM EDTA, pH 8.0).

Slot blot hybridization analysis

To denature the genomic DNA, 10 μl of each DNA sample was mixed with 40 μl of TE and an equal volume (50 μl) of 2X denaturation buffer (0.5N NaOH, 1M NaCl), and incubated in a 55°C water bath for 30 minutes. After cooling at room temperature for 10 minutes, the denatured DNA samples were applied into the wells of the PR600 slot blot apparatus (Hoefer, San Francisco, CA, USA). After a 10 minute incubation at room temperature, a gentle vacuum was applied to facilitate the binding of the DNA to the Duralon-UV™ nylon membrane (Stratagene, La Jolla, CA, USA). After disassembling the apparatus, the nylon membrane was wetted with 1X neutralization

buffer (1M NaCl, 0.5M Tris-HCl, pH 7.0) for 10 minutes at room temperature, then baked in a vacuum oven at 80°C for 2 hours.

The prehybridization and hybridization of the slot-blotted membrane were carried out as described (112). The DNA probes were generated by PCR and represented specific internal and external DNA fragments in relation to the *Emr1/F4/80* gene targeting construct. The hybridized membrane was washed twice at room temperature in 2XSSC/0.1%SDS for 20 minutes, followed by a wash in 0.2XSSC/0.1%SDS at room temperature for 20 minutes and at 55°C for 15 minutes. The membrane was exposed to X-ray film (Eastman Kodak, New Haven, CT, USA) at -80°C for 24 hours. The hybridization intensities of DNA samples were analyzed using a Fujix MacBAS1000 phosphor autoradiography imaging system (Fuji Medical Systems, Stamford, CT, USA) with the Advanced Macintosh™ Image Analysis software program.

Southern blot hybridization analysis

Genomic DNA (10 µg) was digested with *Hin*d III at 37°C overnight, electrophoresed through a 1% agarose gel, and transferred to a nylon membrane using the standard Southern blotting method (59). Prehybridization and hybridization were carried out in 4XSSC/10% dextran sulfate/40% formamide at 42°C. DNA probes were labeled using random oligonucleotides as primers in conditions recommended by the manufacturer (Prime-It®II, Stratagene, La Jolla, CA, USA). Membranes were washed and processed as described in slot blot analysis.

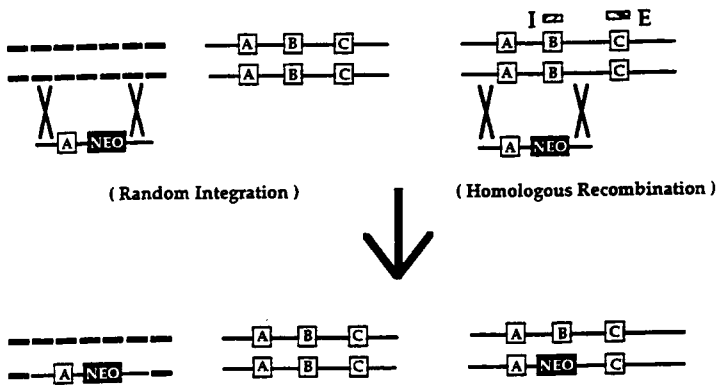
Results and discussion

Following electroporation and drug selection, two types of G418-resistant ES cells are typically identified. Type A ES cells have a randomly integrated targeting vector and two wild-type alleles of the gene of interest (Fig. 18A, left). Type B ES cells, in which homologous recombination has occurred, have one targeted and one wild-type allele of the gene of interest (Fig. 18A, right). Identification of the number of wild-type alleles of the gene of interest in individual G418-resistant colonies will therefore distinguish random integrations (Type A) from targeting events (Type B). Using specific external and internal DNA probes, it is possible to determine the number of the wild-type alleles of the gene of interest (Fig. 18B). Here, an external probe is a specific DNA fragment adjacent to the homologous region of the targeting construct, while an internal probe is from within the genomic region that is deleted after homologous recombination. More specifically, putative targeted ES cell clones could be identified by comparing the hybridization intensity of the external DNA probe to the internal DNA probe in individual samples. For example, equivalent hybridization intensities detected by both internal and external DNA probes would indicate a random integration event (Type A). Conversely, a two-fold increase in the hybridization intensity of the external probe compared to the internal probe would suggest a homologous recombination event (Type B) (Fig. 18).

This approach was tested in an ongoing gene targeting experiment using slot blot hybridization analysis due to the sensitivity of the method. The gene of interest in this study is *Emr1/F4/80*, a gene that encodes the F4/80 glycoprotein that is expressed on the cell surface of most murine macrophage populations (110-112). A gene targeting construct was engineered in which the 3'-end of exon 1 and the 5'-end of intron 1 of *Emr1/F4/80* gene had been deleted and replaced with a promoterless

Fig. 18 Diagram illustrating the gene of interest, the gene targeting construct and the outcome of drug selection of electroporated ES cells. (A) Exons (A, B, and C) and the selectable marker gene (Neo) are represented by white and black boxes, respectively. The bold dashed lines represent the chromosome into which the gene targeting construct is randomly integrated. The specific external (E) and internal (I) DNA probes are represented by hatched boxes. The chromosomal alterations resulting from a random integration or a homologous recombination event are illustrated. (B) Number of copies of the external and internal DNA probes detected in a random integration (Type A) and a homologous recombination (Type B) event.

(A)



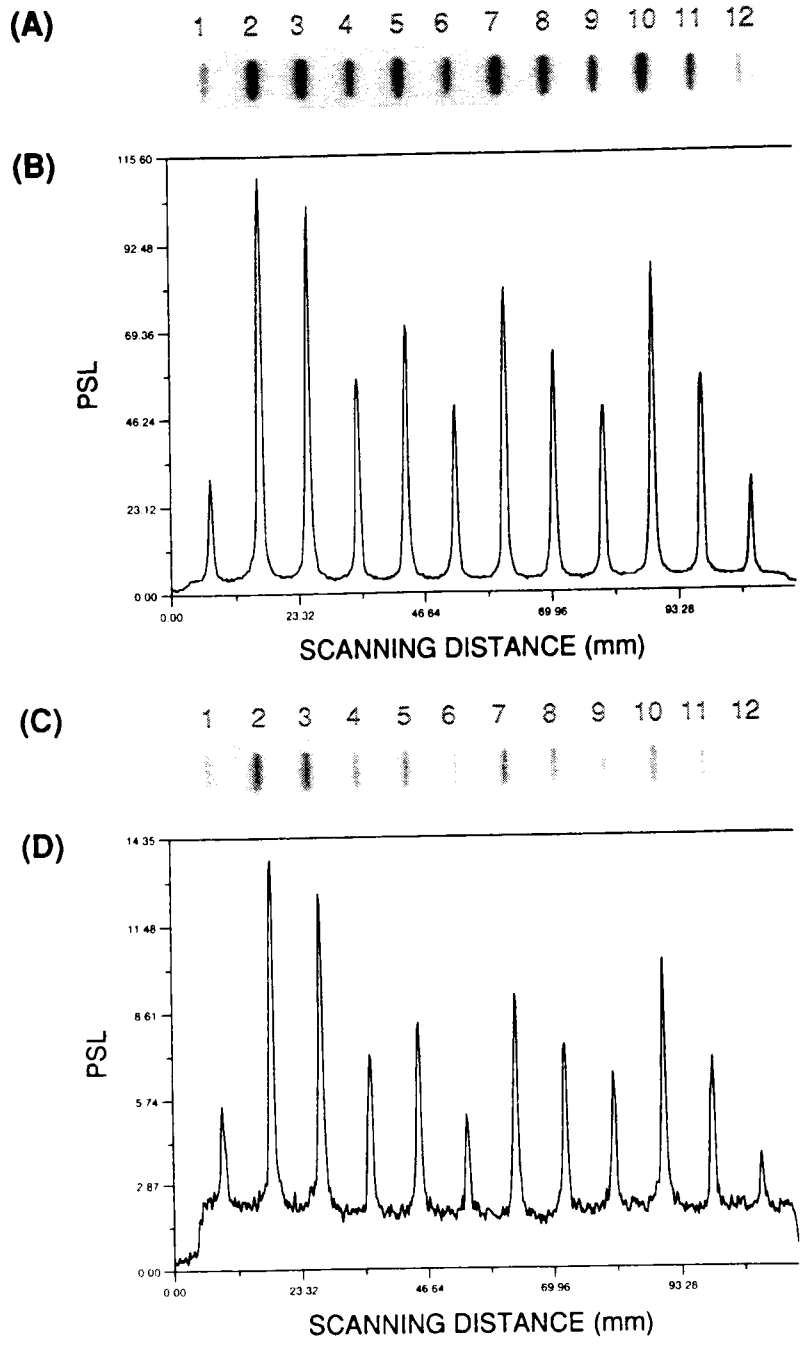
(B)

	Type A ES cell (Random Integration)	Type B ES cell (Homologous Recombination)
probe E	2	2
probe I	2	1

β-gal/pGK-Neo cassette (data not shown). One targeted ES cell clone (clone #142), previously identified and confirmed by Southern blot hybridization and PCR analysis, was used as a targeted ES cell control. In the analysis, equal amounts (1 μg) of genomic DNA from one wild-type and one targeted (clone#142) ES cell clone were used in lanes 1 and 12 respectively, while equal volumes of DNA from ten G418-resistant ES cell clones were loaded in lanes 2-11 as described in Materials and Methods. According to the hypothesis, ES cells of Type A and Type B both contain two copies of DNA fragments recognized by the external DNA probe. Therefore, the hybridization intensity detected by the external DNA probe should be proportional to the amount of DNA loaded in each lane. A similar level of hybridization was detected in lanes 1 and 12, as predicted in our hypothesis (Fig. 19A,B). Different hybridization intensities were detected in lanes 2-11, indicating that different amounts of DNA were present in each lane. One may estimate the amount of DNA in lane 2 to be approximate 4.5 μg because of the approximate 4.5 fold increase in the intensity in lane 2 compared to lane 1. The amount of DNA in lanes 3-11 can be estimated in the same manner using the intensity in lane 1 as a control. The use of the same volume rather than the same amount of DNA avoids the need for quantitation of DNA samples, which is an advantage when a large number of samples are screened. However, the amount of DNA that can be loaded in each well is limited due to the binding capacity of the nylon membrane and the size of the wells. A range of 0.5-5.0 μg of total genomic DNA was found to give the most reproducible results in our analysis (data not shown).

The same nylon membrane was then stripped and rehybridized with the internal DNA probe, which would only detect the wild-type alleles but not the targeted alleles (Fig. 19C,D). Here, the hybridization intensity in lane 12 was reduced by approximately 50% in comparison to lane 1. This is because the DNA used in lane 12

Fig. 19 Slot blot hybridization analysis of ES cell genomic DNA. A nylon membrane containing 12 DNA samples was hybridized to a specific external DNA probe (A,B), stripped, and rehybridized to a specific internal DNA probe (C,D). The results of hybridization are represented by the autoradiogrphy signal on X-ray films (A,C) as well as by the photostimulated luminescence (PSL) profiles detected by the phosphor imaging system (B,D).


(A)
1 2 3 4 5 6 7 8 9 10 11 12
(B)
115.60
92.48
69.36
46.24
23.12
0.00
PSL
0.00
23.32
46.64
69.96
93.28
SCANNING DISTANCE (mm)
(C)
1 2 3 4 5 6 7 8 9 10 11 12
(D)
14.35
11.48
8.61
5.74
2.87
0.00
PSL
0.00
23.32
46.64
69.96
93.28
SCANNING DISTANCE (mm)

is from a targeted ES cell (clone #142), which contains only one copy of the wild-type allele. This result demonstrates that comparison of the hybridization intensity of the external probe to the internal probe can indeed be used to identify correctly targeted ES cells. To facilitate the comparison, the hybridization intensities detected by the internal probe in lanes 2-11 were normalized to that in lane 1 (Fig. 20). After normalization, only the DNA sample in lane 6 has a 2:1 ratio of the external to the internal probe hybridization intensity, indicating that homologous recombination has occurred at the *Emr1/F4/80* locus in this ES cell clone. To confirm this result, Southern blot analysis was performed. The 12 DNA samples were digested with *Hind* III and hybridized with the external DNA probe (Fig. 21). This probe recognizes an endogenous (wild-type) band of 10 kb and a targeted allele of 14 kb. The targeted allele is detected only in lanes 6 and 12, confirming the effectiveness of slot blot hybridization analysis for identifying targeted ES cells.

This paper offers an alternative method for identifying targeted ES cells. The use of slot blot hybridization analysis has at least three advantages over the more conventional Southern blot or PCR-based method. First, the method is economical. Expensive reagents such as agarose, restriction endonuclease, and *Taq* DNA polymerase are not needed. Second, it saves time. There is no need for restriction enzyme digestion, gel electrophoresis, and gel blotting. Using various slot and/or dot blot apparatuses (48, 96 or 192 wells format), a large number of samples can be applied on one nylon membrane and multiple membranes can be processed at the same time. Third, it is efficient. Putative targeted ES cell clones can be easily identified by simply comparing the hybridization intensities detected by the specific internal and external DNA probes.

This new screening method also has other possible applications. It may be utilized in experiments where different types of targeting constructs are used. For

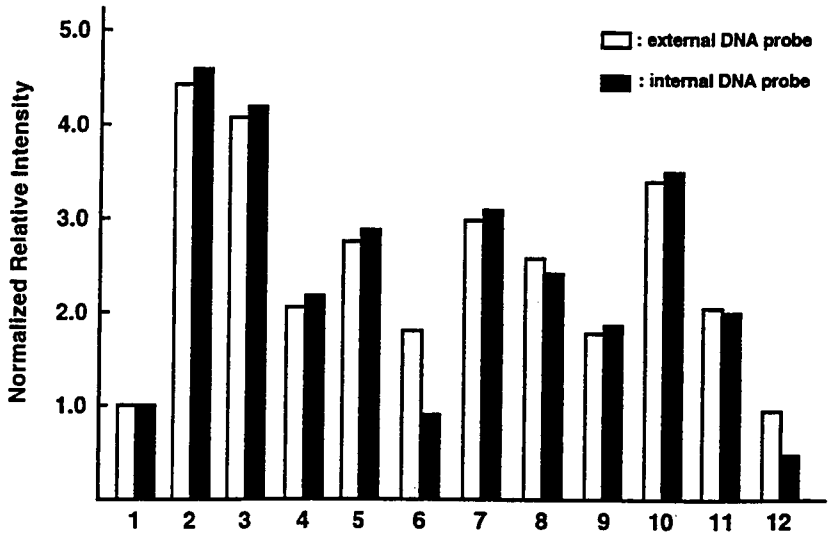


Fig. 20 Comparison of normalized hybridization intensities. The hybridization intensity was normalized using the intensity detected in wild-type DNA (lane 1) as a control. Numbers in the Y-axis represent arbitrary units. The white and black bars represent the normalized hybridization intensities detected by the external DNA probe and the internal DNA probe, respectively.

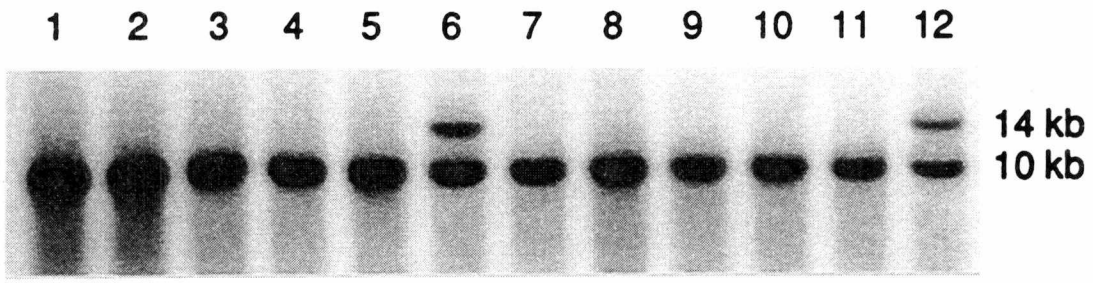


Fig. 21 Southern blot hybridization analysis of putatively targeted ES cell clones. Genomic DNA of the 12 ES cell clones was analyzed for the presence of the targeted allele. Bands of 10 kb and 14 kb represent the wild-type and the targeted allele detected by the external DNA probe, respectively.

example, another commonly used targeting strategy is to insert a selectable marker gene cassette into one of the exons of the gene of interest (143). In this case, no internal DNA probes are available because there is no deletion of genomic DNA sequences. However, one may be able to adapt this method using a specific oligonucleotide probe spanning the disrupted DNA region. This methodology would also be amenable to automation. Using robotic sample preparation workstations, multichannel pipets, and multiwell slot/dot blot apparatuses, a high through-put screening process is possible for hundreds or thousands of ES cell clones simultaneously. Overall, this new screening method will allow for rapid identification of targeted ES cell clones, thereby increasing the efficiency of gene targeting experiments.

Epilogue

The experimental results presented in the preceding chapters described : (i) the use of the *Cmyb* -/- mice as a model system for the identification of novel genes that may be involved in hematopoiesis, and the identification of a seven transmembrane hormone receptor using differential display (Chapters 1 and 2); (ii) the generation and characterization of *Emr1/F4/80* homozygous mutant mice (Chapter 3); and (iii) a new screening strategy for selecting targeted ES cell clones (Chapter 4). Several remarks should be considered in the final conclusion of this study.

The application of differential display technique in the *Cmyb* -/- mouse model system has proven to be a successful approach for the identification of novel genes that may be involved in hematopoiesis. The high percentage of differentially expressed genes (22/38) and putative novel genes (8/12) identified in this study strongly suggests that continued application of differential display in the *Cmyb* -/- mouse model system in the future should identify more novel genes. To more quickly identify putative novel genes in the future, new screening strategies may be conducted. For example, the utilization of reverse Northern blot analysis (144), which uses the original DD-PCR products as hybridization probes to simultaneously screen several individual clones that are derived from one differential cDNA band isolated from DD sequencing gels, may facilitate the confirmation of differential expression of cloned genes. Another method, which makes use of the cDNA probes synthesized from amplified RNA in a differential screening strategy (145), may also help to identify putative positive cDNA clones more efficiently.

More restricted sampling of experimental materials may offer other selective advantages. For instance, the use of cells of specific hematopoietic lineage, sorted by the expression of lineage-specific cell surface markers, instead of cells of the

whole fetal liver as the starting materials should eliminate many false positives caused by the hypocellularity of the *Cmyb* -/- fetal livers. In addition, the use of more stringent starting materials, such as cells of erythroid or myeloid lineage, could be coupled with more restricted screening methods, such as Northern blot analysis using RNA isolated from cells of specific hematopoietic lineage at different developmental stages. These strategies may allow one to identify genes that are involved in specific hematopoietic lineages more efficiently.

Since Myb is a transcription factor, the expression levels of genes that are modulated by the Myb protein are expected to be affected by the null mutation of the *Cmyb* gene generated by targeted mutagenesis. Therefore, one could also explore whether the differentially expressed genes identified in this study are the target genes of Myb protein. For example, one could analyze the promoter region of the gene of the interest to determine whether there is any putative Myb-binding site. If Myb-binding sites are recognized, one could determine whether the Myb protein can bind to and modulate the activity of the promoter of the gene of interest by using electrophoresis mobility shift assay (146) and reporter activity assay in cell culture system (147), respectively. Overall, the model system presented in this study represents a unique resource in the future for the identification of genes involved in hematopoiesis and genes that are modulated by Myb protein.

The cloning of *Emr1/F4/80* gene reveals that the gene encodes a Egf-like motif containing 7TM receptor with unknown biological function(s) (110, 112). The most striking feature of Emr1/F4/80 gene product is its unusual protein structure, which was found in only two other proteins, human EMR1 and human CD97 leukocyte antigen (95, 139). Due to its protein structure and the strong homology to the peptide hormone receptor family, Emr1/F4/80 is believed to be involved in protein-protein interaction and/or ligand binding (116). However, no known ligands or interacting

proteins of Emr1/F4/80 protein have been identified. It is therefore of interest in the future to characterize such putative ligand(s) and/or interacting protein(s). Several methods, including the direct-expression cloning strategy, which uses a fusion protein consisting of the extracellular domain of the receptor of interest linked to the constant domain of a immunoglobulin G molecule as a probe to detect and isolate the putative cellular ligand (148), or the generation and the identification of the cellular target of MAb that specifically blocks the cellular adhesion between cells expressing the ligand and cells expressing the receptor (118), have been used successfully to identify the cellular ligands of specific cell surface receptors. One may adapt these methods in future studies.

Another interesting feature of *Emr1/F4/80* gene is the finding of alternatively spliced transcripts. At least four alternatively spliced *Emr1/F4/80* RNA species were identified by RT-PCR (112). These RNA transcripts differ only in the number and/or combination of Egf-like motifs, resulting in four predicted protein isoforms containing the same 7TM region but different extracellular domains. Questions, such as why are there so many different Emr1/F4/80 protein isoforms, and do these protein isoforms bind to the same ligand and/or interacting protein need to be addressed. Additional questions concerning the temporal and spacial expression of each isoform await further analyses. Again, identification of ligand(s) and/or interacting protein(s) of Emr1/F4/80 would be the first step toward resolving some of the questions. Generation of MAbs specific for the individual Emr1/F4/80 protein isoforms should also help to determine the ligand-binding affinity and the tissue distribution pattern of different Emr1/F4/80 protein isoforms.

The biological function of *Emr1/F4/80* gene was assessed by the generation of *Emr1* -/- mice. No obvious phenotypes were identified in the *Emr1* -/- mice by the analyses performed in this study. The possibility of leaky expression of Emr1/F4/80

protein in the mutant mice was ruled out, since the *Emr1* -/- mice generated here are null mutation, which were demonstrated by analyses at DNA, RNA, and protein levels (Chapter 3). Therefore, as mentioned in Chapter 3, further analyses, including more *in vitro* functional assays for Mϕ, and *in vivo* immunoresponses against pathogenic bacteria and virus infections, are needed to determine whether the Mϕ-specific Emr1/F4/80 protein is involved in immunoreactions. However, it should be pointed out that it is also possible that the Emr1/F4/80 protein has no any real biological function, thus the disruption of *Emr1/F4/80* gene *in vivo* would result in no phenotypes.

On the other hand, the apparent normal phenotype of the *Emr1* -/- mice might result from the complement of the function of Emr1/F4/80 protein by other cell surface receptors. Therefore, it would be worthwhile to attempt to study whether there are other proteins with similar protein structure that may have overlapping function with the Emr1/F4/80 protein. The similar protein structure and the close proximity of both the human *EMR1* and *CD97* genes in chrmosome 19p13 suggests that there may be a EGF-TM7 gene family in this region (140). If this turns out to be true, it is likely that other EGF-TM7 proteins may have similar function as the Emr1/F4/80 protein, and thus complement the loss of Emr1/F4/80 protein in the Emr1 -/- mice. Since the *Emr1* -/- Mϕ do not express any *Emr1/F4/80* RNA and protein, the *Emr1* -/- mice would be an excellent model system for the cloning of the putative *Emr1/F4/80*-related genes. To do that, one may take advantage of the highly homologous amino acid sequence found in the 7TM regions of EGF-TM7 proteins, and design degenerate oligonucleotide primers for RT-PCR amplification of the putative *Emr1/F4/80* related cDNA using RNA isolated from *Emr1* -/- mice. The identification of such putative *Emr1/F4/80*-related genes would help to elucidate additional information on the evolution of EGF-TM7 gene family and the possible biological function(s) of the Emr1/F4/80 protein.

In this study, a new screening strategy for selecting targeted ES cell clones was also presented. This method takes advantage of the sensitivity of slot blot hybridization analysis and provides a rapid alternative for the identification of homologous recombination events in ES cell clones. Since targeted mutagenesis is currently the method of choice among many investigators for the study of the biological functions of specific genes *in vivo*, this new method shows promise in increasing the efficiency of the study of gene function. Nevertheless, improvements for the method may be made to further refine the experimental procedures in the future. For example, the ES cell genomic DNA used in this method is isolated from single ES cell suspension trypsinized from cell culture dishes which is time-consuming and rate-limiting. One may try to isolate the ES cell genomic DNA directly from ES cell clones cultured in 24- or 96- well dishes as described by Ramirez-Solis et al (149) to accelerate the entire analysis. All the points presented here still need to be addressed. Overall, the *Cmyb* -/- mouse model system will remain an excellent system for the study of novel genes involved in hematopoiesis for the years to come.

REFERENCES

References

1. Wood, W.G. 1982. Erythropoiesis and haemoglobin production during development. *In* Biochemical development of the fetus and neonate. Jones. C.T., editor. Elsevier Biomedical Press, New York. 127-162.

2. Dzierzak, E., and A. Medvinsky. 1995. Mouse embryonic hematopoiesis. *Trends Genet* 11, no. 9:359-366.

3. Morris, L., C.F. Graham, and S. Gordon. 1991. Macrophages in haemopoietic and other tissues of the developing mouse detected by the monoclonal antibody F4/80. *Development* 112, no. 2:517-526.

4. Fantoni, A., A. Bank, and P.A. Marks. 1967. Globin composition and synthesis of hemoglobins in developing fetal mice erythroid cells. *Science* 157, no. 794:1327-1329.

5. Hansen, J.N., D.A. Konkel, and P. Leder. 1982. The sequence of a mouse embryonic beta-globin gene. Evolution of the gene and its signal region. *J Biol Chem* 257, no. 2:1048-1052.

6. Brotherton, T.W., D.H. Chui, J. Gauldie, and M. Patterson. 1979. Hemoglobin ontogeny during normal mouse fetal development. *Proc Natl Acad Sci U S A* 76, no. 6:2853-2857.

7. Wong, P.M., S.W. Chung, J.S. White, S.M. Reicheld, M. Patterson, B.J. Clarke, and D.H. Chui. 1983. Adult hemoglobins are synthesized in murine fetal hepatic erythropoietic cells. *Blood* 62, no. 6:1280-1288.

8. Medvinsky, A.L., N.L. Samoylina, A.M. Muller, and E.A. Dzierzak. 1993. An early pre-liver intraembryonic source of CFU-S in the developing mouse. *Nature* 364, no. 6432:64-67.

9. Moore, M.A., and D. Metcalf. 1970. Ontogeny of the haemopoietic system: yolk sac origin of in vivo and in vitro colony forming cells in the developing mouse embryo. *Br J Haematol* 18, no. 3:279-296.

10. Orkin, S.H. 1992. GATA-binding transcription factors in hematopoietic cells. *Blood* 80, no. 3:575-581.

11. Copp, A.J. 1995. Death before birth: clues from gene knockouts and mutations. *Trends Genet* 11, no. 3:87-93.

12. Mucenski, M., K. McLain, A. Kier, S. Swerdlow, C. Schreiner, T. Miller, D. Pietryga, J. Scott WJ, and S. Potter. 1991. A functional c-myb gene is required for normal murine fetal hepatic hematopoiesis. *Cell* 65, no. 4:677-89.

13. Tsai, F.Y., G. Keller, F.C. Kuo, M. Weiss, J. Chen, M. Rosenblatt, F.W. Alt, and S.H. Orkin. 1994. An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature* 371, no. 6494:221-226.

14. Scott, E.W., M.C. Simon, J. Anastasi, and H. Singh. 1994. Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science* 265, no. 5178:1573-1577.

15. Shen-Ong, G.L. 1990. The myb oncogene. *Biochim Biophys Acta* 1032, no. 1:39-52.

16. Moscovici, C. 1975. Leukemic transformation with avian myeloblastosis virus: present status. *Curr Top Microbiol Immunol* 71:79-101.

17. Radke, K., H. Beug, S. Kornfeld, and T. Graf. 1982. Transformation of both erythroid and myeloid cells by E26, an avian leukemia virus that contains the myb gene. *Cell* 31, no. 3 Pt 2:643-653.

18. Leprince, D., A. Gegonne, J. Coll, C. de Taisne, A. Schneeberger, C. Lagrou, and D. Stehelin. 1983. A putative second cell-derived oncogene of the avian leukaemia retrovirus E26. *Nature* 306, no. 5941:395-397.

19. Nunn, M.F., P.H. Seeburg, C. Moscovici, and P.H. Duesberg. 1983. Tripartite structure of the avian erythroblastosis virus E26 transforming gene. *Nature* 306, no. 5941:391-395.

20. Rushlow, K.E., J.A. Lautenberger, T.S. Papas, M.A. Baluda, B. Perbal, J.G. Chirikjian, and E.P. Reddy. 1982. Nucleotide sequence of the transforming gene of avian myeloblastosis virus. *Science* 216, no. 4553:1421-1423.

21. Kanter, M.R., R.E. Smith, and W.S. Hayward. 1988. Rapid induction of B-cell lymphomas: insertional activation of c-myb by avian leukosis virus. *J Virol* 62, no. 4:1423-1432.

22. Luscher, B., and R.N. Eisenman. 1990. New light on Myc and Myb. Part II. Myb. *Genes Dev* 4, no. 12B:2235-2241.

23. Shen-Ong, G.L., M. Potter, J.F. Mushinski, S. Lavu, and E.P. Reddy. 1984. Activation of the c-myb locus by viral insertional mutagenesis in plasmacytoid lymphosarcomas. *Science* 226, no. 4678:1077-1080.

24. Biedenkapp, H., U. Borgmeyer, A.E. Sippel, and K.H. Klempnauer. 1988. Viral myb oncogene encodes a sequence-specific DNA-binding activity. *Nature* 335, no. 6193:835-837.

25. Ibanez, C.E., and J.S. Lipsick. 1990. trans activation of gene expression by v-myb. *Mol Cell Biol* 10, no. 5:2285-2293.

26. Sakura, H., C. Kanei-Ishii, T. Nagase, H. Nakagoshi, T.J. Gonda, and S. Ishii. 1989. Delineation of three functional domains of the transcriptional activator encoded by the c-myb protooncogene. *Proc Natl Acad Sci U S A* 86, no. 15:5758-5762.

27. Weston, K., and J.M. Bishop. 1989. Transcriptional activation by the v-myb oncogene and its cellular progenitor, c-myb. *Cell* 58, no. 1:85-93.

28. Gonda, T.J., and D. Metcalf. 1984. Expression of myb, myc and fos proto-oncogenes during the differentiation of a murine myeloid leukaemia. *Nature* 310, no. 5974:249-251.

29. Sheiness, D., and M. Gardinier. 1984. Expression of a proto-oncogene (proto-myb) in hemopoietic tissues of mice. *Mol Cell Biol* 4, no. 7:1206-1212.

30. Anfossi, G., A.M. Gewirtz, and B. Calabretta. 1989. An oligomer complementary to c-myb-encoded mRNA inhibits proliferation of human myeloid leukemia cell lines. *Proc Natl Acad Sci U S A* 86, no. 9:3379-3383.

31. Gewirtz, A.M., and B. Calabretta. 1988. A c-myb antisense oligodeoxynucleotide inhibits normal human hematopoiesis in vitro. *Science* 242, no. 4883:1303-1306.

32. Gewirtz, A.M., G. Anfossi, D. Venturelli, S. Valpreda, R. Sims, and B. Calabretta. 1989. G1/S transition in normal human T-lymphocytes requires the nuclear protein encoded by c-myb. *Science* 245, no. 4914:180-183.

33. Venturelli, D., M.T. Mariano, C. Szczylik, M. Valtieri, B. Lange, W. Crist, M. Link, and B. Calabretta. 1990. Down-regulated c-myb expression inhibits DNA synthesis of T-leukemia cells in most patients. *Cancer Res* 50, no. 22:7371-7375.

34. Clarke, M.F., J.F. Kukowska-Latallo, E. Westin, M. Smith, and E.V. Prochownik. 1988. Constitutive expression of a c-myb cDNA blocks Friend murine erythroleukemia cell differentiation. *Mol Cell Biol* 8, no. 2:884-892.

35. McMahon, J., K.M. Howe, and R.J. Watson. 1988. The induction of Friend erythroleukaemia differentiation is markedly affected by expression of a transfected c-myb cDNA. *Oncogene* 3, no. 6:717-720.

36. Todokoro, K., R.J. Watson, H. Higo, H. Amanuma, S. Kuramochi, H. Yanagisawa, and Y. Ikawa. 1988. Down-regulation of c-myb gene expression is a prerequisite for erythropoietin-induced erythroid differentiation. *Proc Natl Acad Sci U S A* 85, no. 23:8900-8904.

37. Weber, B.L., E.H. Westin, and M.F. Clarke. 1990. Differentiation of mouse erythroleukemia cells enhanced by alternatively spliced c-myb mRNA. *Science* 249, no. 4974:1291-1293.

38. Lin, H., D. Sternfeld, S. Shinpock, R. Popp, and M. Mucenski. 1996. Functional analysis of the c-myb proto-oncogene. *Curr Top Microbiol Immunol* 211:79-87.

39. Ness, S.A., A. Marknell, and T. Graf. 1989. The v-myb oncogene product binds to and activates the promyelocyte-specific mim-1 gene. *Cell* 59, no. 6:1115-1125.

40. Nakagoshi, H., T. Nagase, C. Kanei-Ishii, Y. Ueno, and S. Ishii. 1990. Binding of the c-myb proto-oncogene product to the simian virus 40 enhancer stimulates transcription. *J Biol Chem* 265, no. 6:3479-3483.

41. Aurigemma, R.E., D.G. Blair, and S.K. Ruscetti. 1992. Transactivation of erythroid transcription factor GATA-1 by a myb-ets-containing retrovirus. *J Virol* 66, no. 5:3056-3061.

42. Melotti, P., and B. Calabretta. 1996. Induction of hematopoietic commitment and erythromyeloid differentiation in embryonal stem cells constitutively expressing c-myb. *Blood* 87, no. 6:2221-2234.

43. Melotti, P., D.H. Ku, and B. Calabretta. 1994. Regulation of the expression of the hematopoietic stem cell antigen CD34: role of c-myb. *J Exp Med* 179, no. 3:1023-1028.

44. Siu, G., A.L. Wurster, J.S. Lipsick, and S.M. Hedrick. 1992. Expression of the CD4 gene requires a Myb transcription factor. *Mol Cell Biol* 12, no. 4:1592-1604.

45. Lee, S.W., C. Tomasetto, and R. Sager. 1991. Positive selection of candidate tumor-suppressor genes by subtractive hybridization. *Proc Natl Acad Sci U S A* 88, no. 7:2825-2829.

46. Scott, M.R., K.H. Westphal, and P.W. Rigby. 1983. Activation of mouse genes in transformed cells. *Cell* 34, no. 2:557-567.

47. Liang, P., and A.B. Pardee. 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction [see comments]. *Science* 257, no. 5072:967-971.

48. Sager, R., A. Anisowicz, M. Neveu, P. Liang, and G. Sotiropoulou. 1993. Identification by differential display of alpha 6 integrin as a candidate tumor suppressor gene. *FASEB J* 7, no. 10:964-970.

49. Okamoto, K., and D. Beach. 1994. Cyclin G is a transcriptional target of the p53 tumor suppressor protein. *EMBO J* 13, no. 20:4816-4822.

50. Fung-Leung, W.P., and T.W. Mak. 1992. Embryonic stem cells and homologous recombination. *Curr Opin Immunol* 4, no. 2:189-194.

51. Banks, T., and M. Mucenski. 1994. Gene targeting of the immune system: the surprises continue. *Curr Opin Biotechnol* 5, no. 6:604-10.

52. Evans, M.J., and M.H. Kaufman. 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature* 292, no. 5819:154-156.

53. Martin, G.R. 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A* 78, no. 12:7634-7638.

54. Chakrabarti, S., and M.M. Seidman. 1986. Intramolecular recombination between transfected repeated sequences in mammalian cells is nonconservative [published erratum appears in Mol Cell Biol 1986 Oct;6(10):3568]. *Mol Cell Biol* 6, no. 7:2520-2526.

55. Lin, F.L., K. Sperle, and N. Sternberg. 1984. Model for homologous recombination during transfer of DNA into mouse L cells: role for DNA ends in the recombination process. *Mol Cell Biol* 4, no. 6:1020-1034.

56. Wake, C.T., F. Vernaleone, and J.H. Wilson. 1985. Topological requirements for homologous recombination among DNA molecules transfected into mammalian cells. *Mol Cell Biol* 5, no. 8:2080-2089.

57. Capecchi, M.R. 1989. The new mouse genetics: altering the genome by gene targeting. *Trends Genet* 5, no. 3:70-76.

58. Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 162, no. 1:156-159.

59. Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. . Second edition *In* Molecular Cloning: A Laboratory Manual. C. Nolan, editor. Cold Spring Harbor Laboratory Press, New York.

60. Doyle, J., S. Hoffman, C. Ucla, W. Reith, B. Mach, and L. Stubbs. 1996. Locations of human and mouse genes encoding the RFX1 and RFX2 transcription factor proteins. *Genomics* 35, no. 1:227-230.

61. Manly, K.F. 1993. A Macintosh program for storage and analysis of experimental genetic mapping data. *Mamm Genome* 4, no. 6:303-313.

62. Silver, J. 1985. Confidence limits for estimates of gene linkage based on analysis of recombinant inbred strains. *J Hered* 76, no. 6:436-440.

63. Thomas, M.L. 1989. The leukocyte common antigen family. *Annu Rev Immunol* 7:339-369.

64. Begley, C.G., P.D. Aplan, S.M. Denning, B.F. Haynes, T.A. Waldmann, and I.R. Kirsch. 1989. The gene SCL is expressed during early hematopoiesis and encodes a differentiation-related DNA-binding motif. *Proc Natl Acad Sci U S A* 86, no. 24:10128-10132.

65. Pevny, L., M.C. Simon, E. Robertson, W.H. Klein, S.F. Tsai, V. DqAgati, S.H. Orkin, and F. Costantini. 1991. Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature* 349, no. 6306:257-260.

66. Andrews, N.C., H. Erdjument-Bromage, M.B. Davidson, P. Tempst, and S.H. Orkin. 1993. Erythroid transcription factor NF-E2 is a haematopoietic-specific basic-leucine zipper protein. *Nature* 362, no. 6422:722-728.

67. Evans, T., and G. Felsenfeld. 1989. The erythroid-specific transcription factor Eryf1: a new finger protein. *Cell* 58, no. 5:877-885.

68. Weiss, M.J., and S.H. Orkin. 1995. GATA transcription factors: key regulators of hematopoiesis. *Exp Hematol* 23, no. 2:99-107.

69. Tsai, S.F., D.I. Martin, L.I. Zon, A.D. DqAndrea, G.G. Wong, and S.H. Orkin. 1989. Cloning of cDNA for the major DNA-binding protein of the erythroid lineage through expression in mammalian cells. *Nature* 339, no. 6224:446-451.

70. Yamamoto, M., L.J. Ko, M.W. Leonard, H. Beug, S.H. Orkin, and J.D. Engel. 1990. Activity and tissue-specific expression of the transcription factor NF-E1 multigene family. *Genes Dev* 4, no. 10:1650-1662.

71. Martin, D.I., L.I. Zon, G. Mutter, and S.H. Orkin. 1990. Expression of an erythroid transcription factor in megakaryocytic and mast cell lineages. *Nature* 344, no. 6265:444-447.

72. Michaud, E.J., M.J. van Vugt, S.J. Bultman, H.O. Sweet, M.T. Davisson, and R.P. Woychik. 1994. Differential expression of a new dominant agouti allele (Aiapy) is correlated with methylation state and is influenced by parental lineage. *Genes Dev* 8, no. 12:1463-1472.

73. Lin, C.S., S.K. Lim, V. DqAgati, and F. Costantini. 1996. Differential effects of an erythropoietin receptor gene disruption on primitive and definitive erythropoiesis. *Genes Dev* 10, no. 2:154-164.

74. Begley, C.G., J. Visvader, A.R. Green, P.D. Aplan, D. Metcalf, I.R. Kirsch, and N.M. Gough. 1991. Molecular cloning and chromosomal localization of the murine homolog of the human helix-loop-helix gene SCL. *Proc Natl Acad Sci U S A* 88, no. 3:869-873.

75. Green, A.R., E. Salvaris, and C.G. Begley. 1991. Erythroid expression of the 'helix-loop-helix' gene, SCL. *Oncogene* 6, no. 3:475-479.

76. Mignotte, V., J.F. Eleouet, N. Raich, and P.H. Romeo. 1989. Cis- and trans-acting elements involved in the regulation of the erythroid promoter of the human porphobilinogen deaminase gene. *Proc Natl Acad Sci U S A* 86, no. 17:6548-6552.

77. Taketani, S., J. Inazawa, Y. Nakahashi, T. Abe, and R. Tokunaga. 1992. Structure of the human ferrochelatase gene. Exon/intron gene organization and location of the gene to chromosome 18. *Eur J Biochem* 205, no. 1:217-222.

78. Bauer, D., H. Muller, J. Reich, H. Riedel, V. Ahrenkiel, P. Warthoe, and M. Strauss. 1993. Identification of differentially expressed mRNA species by an improved display technique (DDRT-PCR). *Nucleic Acids Res* 21, no. 18:4272-4280.

79. Banfield, D.K., and R.T. MacGillivray. 1992. Partial characterization of vertebrate prothrombin cDNAs: amplification and sequence analysis of the B chain of thrombin from nine different species. *Proc Natl Acad Sci U S A* 89, no. 7:2779-2783.

80. Hempel, W.M., and A.L. DeFranco. 1991. Expression of phospholipase C isozymes by murine B lymphocytes. *J Immunol* 146, no. 11:3713-3720.

81. Tsui, F.W., H.W. Tsui, S. Mok, I. Mlinaric, N.G. Copeland, D.J. Gilbert, N.A. Jenkins, and K.A. Siminovitch. 1993. Molecular characterization and mapping of murine genes encoding three members of the stefin family of cysteine proteinase inhibitors. *Genomics* 15, no. 3:507-514.

82. Robert, B., P. Barton, A. Minty, P. Daubas, A. Weydert, F. Bonhomme, J. Catalan, D. Chazottes, J.L. Guenet, and M. Buckingham. 1985. Investigation of genetic linkage between myosin and actin genes using an interspecific mouse back-cross. *Nature* 314, no. 6007:181-183.

83. Porcher, C., W. Swat, K. Rockwell, Y. Fujiwara, F.W. Alt, and S.H. Orkin. 1996. The T cell leukemia oncoprotein SCL/tal-1 is essential for development of all hematopoietic lineages. *Cell* 86, no. 1:47-57.

84. Shivdasani, R.A., M.F. Rosenblatt, D. Zucker-Franklin, C.W. Jackson, P. Hunt, C.J. Saris, and S.H. Orkin. 1995. Transcription factor NF-E2 is required for platelet formation independent of the actions of thrombopoietin/MGDF in megakaryocyte development. *Cell* 81, no. 5:695-704.

85. Shivdasani, R.A., and S.H. Orkin. 1995. Erythropoiesis and globin gene expression in mice lacking the transcription factor NF-E2. *Proc Natl Acad Sci U S A* 92, no. 19:8690-8694.

86. Wu, H., X. Liu, R. Jaenisch, and H.F. Lodish. 1995. Generation of committed erythroid BFU-E and CFU-E progenitors does not require erythropoietin or the erythropoietin receptor. *Cell* 83, no. 1:59-67.

87. Pevny, L., C.S. Lin, V. DqAgati, M.C. Simon, S.H. Orkin, and F. Costantini. 1995. Development of hematopoietic cells lacking transcription factor GATA-1. *Development* 121, no. 1:163-172.

88. Chiba, T., Y. Ikawa, and K. Todokoro. 1991. GATA-1 transactivates erythropoietin receptor gene, and erythropoietin receptor-mediated signals enhance GATA-1 gene expression. *Nucleic Acids Res* 19, no. 14:3843-3848.

89. Liang, P., L. Averboukh, K. Keyomarsi, R. Sager, and A.B. Pardee. 1992. Differential display and cloning of messenger RNAs from human breast cancer versus mammary epithelial cells. *Cancer Res* 52, no. 24:6966-6968.

90. Zon, L.I. 1995. Developmental biology of hematopoiesis. *Blood* 86, no. 8:2876-2891.

91. Metcalf, D. 1989. The molecular control of cell division, differentiation commitment and maturation in haemopoietic cells. *Nature* 339:27-30.

92. Balkwill, F.R., and F. Burke. 1989. The cytokine network [see comments]. *Immunol Today* 10, no. 9:299-304.

93. Westin, E.H., R.C. Gallo, S.K. Arya, A. Eva, L.M. Souza, M.A. Baluda, S.A. Aaronson, and F. Wong-Staal. 1982. Differential expression of the amv gene in human hematopoietic cells. *Proc Natl Acad Sci U S A* 79, no. 7:2194-2198.

94. Liang, P., L. Averboukh, and A.B. Pardee. 1993. Distribution and cloning of eukaryotic mRNAs by means of differential display: refinements and optimization. *Nucleic Acids Res* 21, no. 14:3269-3275.

95. Baud, V., S.L. Chissoe, E. Viegas-Pequignot, S. Diriong, V.C. NqGuyen, B.A. Roe, and M. Lipinski. 1995. EMR1, an unusual member in the family of hormone receptors with seven transmembrane segments. *Genomics* 26, no. 2:334-344.

96. Mucenski, M.L., B.A. Taylor, J.N. Ihle, J.W. Hartley, d. Morse HC, N.A. Jenkins, and N.G. Copeland. 1988. Identification of a common ecotropic viral integration site, Evi-1, in the DNA of AKXD murine myeloid tumors. *Mol Cell Biol* 8, no. 1:301-308.

97. McClinton, D., J. Stafford, L. Brents, T.P. Bender, and W.M. Kuehl. 1990. Differentiation of mouse erythroleukemia cells is blocked by late up-regulation of a c-myb transgene. *Mol Cell Biol* 10, no. 2:705-710.

98. Evans, J.L., T.L. Moore, W.M. Kuehl, T. Bender, and J.P. Ting. 1990. Functional analysis of c-Myb protein in T-lymphocytic cell lines shows that it trans-activates the c-myc promoter. *Mol Cell Biol* 10, no. 11:5747-5752.

99. Tanaka, T., M. Tsudo, H. Karasuyama, F. Kitamura, T. Kono, M. Hatakeyama, T. Taniguchi, and M. Miyasaka. 1991. A novel monoclonal antibody against murine IL-2 receptor beta-chain. Characterization of receptor expression in normal lymphoid cells and EL-4 cells. *J Immunol* 147, no. 7:2222-2228.

100. Kozak, M. 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res* 15, no. 20:8125-8148.

101. Maslen, C.L., G.M. Corson, B.K. Maddox, R.W. Glanville, and L.Y. Sakai. 1991. Partial sequence of a candidate gene for the Marfan syndrome [see comments]. *Nature* 352, no. 6333:334-337.

102. Pan, T.C., T. Sasaki, R.Z. Zhang, R. Fassler, R. Timpl, and M.L. Chu. 1993. Structure and expression of fibulin-2, a novel extracellular matrix protein with multiple EGF-like repeats and consensus motifs for calcium binding. *J Cell Biol* 123, no. 5:1269-1277.

103. Moren, A., A. Olofsson, G. Stenman, P. Sahlin, T. Kanzaki, L. Claesson-Welsh, P. ten Dijke, K. Miyazono, and C.H. Heldin. 1994. Identification and characterization of LTBP-2, a novel latent transforming growth factor-beta-binding protein. *J Biol Chem* 269, no. 51:32469-32478.

104. Stenflo, J., A.K. Ohlin, W.G. Owen, and W.J. Schneider. 1988. beta-Hydroxyaspartic acid or beta-hydroxyasparagine in bovine low density lipoprotein receptor and in bovine thrombomodulin. *J Biol Chem* 263, no. 1:21-24.

105. Segre, G.V., and S.R. Goldring. 1993. Receptors for secretin, calcitonin, parathyroid hormone (PTH)/PTH-related peptide, vasoactive intestinal peptide,

glucagon-like peptide 1, growth hormone-releasing hormone, and glucagon belong to a newly discovered G-protein-linked receptor family. *Trends Endocrinol. Metab.* 4:309-314.

106. Kyte, J., and R.F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157, no. 1:105-132.

107. Leytus, S.P., D.C. Foster, K. Kurachi, and E.W. Davie. 1986. Gene for human factor X: a blood coagulation factor whose gene organization is essentially identical with that of factor IX and protein C. *Biochemistry* 25, no. 18:5098-5102.

108. Foster, D.C., S. Yoshitake, and E.W. Davie. 1985. The nucleotide sequence of the gene for human protein C. *Proc Natl Acad Sci U S A* 82, no. 14:4673-4677.

109. Rebay, I., R.J. Fleming, R.G. Fehon, L. Cherbas, P. Cherbas, and S. Artavanis-Tsakonas. 1991. Specific EGF repeats of Notch mediate interactions with Delta and Serrate: implications for Notch as a multifunctional receptor. *Cell* 67, no. 4:687-699.

110. McKnight, A.J., A.J. Macfarlane, P. Dri, L. Turley, A.C. Willis, and S. Gordon. 1996. Molecular cloning of F4/80, a murine macrophage-restricted cell surface glycoprotein with homology to the G-protein-linked transmembrane 7 hormone receptor family. *J Biol Chem* 271, no. 1:486-489.

111. Austyn, J.M., and S. Gordon. 1981. F4/80, a monoclonal antibody directed specifically against the mouse macrophage. *Eur J Immunol* 11, no. 10:805-815.

112. Lin, H., L. Stubbs, and M. Mucenski. 1997. Identification and characterization of a seven transmembrane hormone receptor using differential display. *Genomics* 41, no. 3:301-8.

113. Gordon, S. 1995. The macrophage. *Bioessays* 17, no. 11:977-986.

114. Ezekowitz, R.A., J. Austyn, P.D. Stahl, and S. Gordon. 1981. Surface properties of bacillus Calmette-Guérin-activated mouse macrophages. Reduced expression of mannose-specific endocytosis, Fc receptors, and antigen F4/80 accompanies induction of Ia. *J Exp Med* 154, no. 1:60-76.

115. Ezekowitz, R.A., and S. Gordon. 1982. Down-regulation of mannosyl receptor-mediated endocytosis and antigen F4/80 in bacillus Calmette-Guérin-activated mouse macrophages. Role of T lymphocytes and lymphokines. *J Exp Med* 155, no. 6:1623-1637.

116. McKnight, A.J., and S. Gordon. 1996. EGF-TM7: a novel subfamily of seven-transmembrane-region leukocyte cell-surface molecules. *Immunol Today* 17, no. 6:283-287.

117. Gordon, S., L. Lawson, S. Rabinowitz, P.R. Crocker, L. Morris, and V.H. Perry. 1992. Antigen markers of macrophage differentiation in murine tissues. *Curr Top Microbiol Immunol* 181:1-37.

118. Hamann, J., B. Vogel, G.M. van Schijndel, and R.A. van Lier. 1996. The seven-span transmembrane receptor CD97 has a cellular ligand (CD55, DAF). *J Exp Med* 184, no. 3:1185-1189.

119. Gray, A., T.J. Dull, and A. Ullrich. 1983. Nucleotide sequence of epidermal growth factor cDNA predicts a 128,000-molecular weight protein precursor. *Nature* 303, no. 5919:722-725.

120. Selander-Sunnerhagen, M., M. Ullner, E. Persson, O. Teleman, J. Stenflo, and T. Drakenberg. 1992. How an epidermal growth factor (EGF)-like domain binds calcium. High resolution NMR structure of the calcium form of the NH2-terminal EGF-like domain in coagulation factor X. *J Biol Chem* 267, no. 27:19642-19649.

121. Clapham, D.E. 1993. Mutations in G protein-linked receptors: novel insights on disease [comment]. *Cell* 75, no. 7:1237-1239.

122. Spiegel, A.M., L.S. Weinstein, and A. Shenker. 1993. Abnormalities in G protein-coupled signal transduction pathways in human disease. *J Clin Invest* 92, no. 3:1119-1125.

123. Dohlman, H.G., J. Thorner, M.G. Caron, and R.J. Lefkowitz. 1991. Model systems for the study of seven-transmembrane-segment receptors. *Annu Rev Biochem* 60:653-688.

124. Reid, K.B. 1986. Activation and control of the complement system. *Essays Biochem* 22:27-68.

125. Stahl, P.D. 1992. The mannose receptor and other macrophage lectins. *Curr Opin Immunol* 4, no. 1:49-52.

126. Friedrich, G., and P. Soriano. 1991. Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev* 5, no. 9:1513-1523.

127. Swiatek, P.J., and T. Gridley. 1993. Perinatal lethality and defects in hindbrain development in mice homozygous for a targeted mutation of the zinc finger gene Krox20. *Genes Dev* 7, no. 11:2071-2084.

128. Crocker, P.R., S. Kelm, C. Dubois, B. Martin, A.S. McWilliam, D.M. Shotton, J.C. Paulson, and S. Gordon. 1991. Purification and properties of sialoadhesin, a sialic acid-binding receptor of murine tissue macrophages. *EMBO J* 10, no. 7:1661-1669.

129. Fraser, I., D. Hughes, and S. Gordon. 1993. Divalent cation-independent macrophage adhesion inhibited by monoclonal antibody to murine scavenger receptor. *Nature* 364, no. 6435:343-346.

130. Rosen, H., and S. Gordon. 1987. Monoclonal antibody to the murine type 3 complement receptor inhibits adhesion of myelomonocytic cells in vitro and inflammatory cell recruitment in vivo. *J Exp Med* 166, no. 6:1685-1701.

131. Rabinowitz, S.S., and S. Gordon. 1991. Macrosialin, a macrophage-restricted membrane sialoprotein differentially glycosylated in response to inflammatory stimuli [published erratum appears in J Exp Med 1992 Jan 1;175(1):309]. *J Exp Med* 174, no. 4:827-836.

132. Lowell, C.A., P. Soriano, and H.E. Varmus. 1994. Functional overlap in the src gene family: inactivation of hck and fgr impairs natural immunity. *Genes Dev* 8, no. 4:387-398.

133. Martinez-Pomares, L., M. Kosco-Vilbois, E. Darley, P. Tree, S. Herren, J.Y. Bonnefoy, and S. Gordon. 1996. Fc chimeric protein containing the cysteine-rich

domain of the murine mannose receptor binds to macrophages from splenic marginal zone and lymph node subcapsular sinus and to germinal centers. *J Exp Med* 184, no. 5:1927-1937.

134. von Heijne, G. 1986. A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res* 14, no. 11:4683-4690.

135. Mendelsohn, C., E. Ruberte, M. LeMeur, G. Morriss-Kay, and P. Chambon. 1991. Developmental analysis of the retinoic acid-inducible RAR-beta 2 promoter in transgenic animals. *Development* 113, no. 3:723-734.

136. Rosenstreich, D.L., and S.B. Mizel. 1978. The participation of macrophages and macrophage cell lines in the activation of T lymphocytes by mitogens. *Immunol Rev* 40:102-135.

137. Crocker, P.R., and S. Gordon. 1989. Mouse macrophage hemagglutinin (sheep erythrocyte receptor) with specificity for sialylated glycoconjugates characterized by a monoclonal antibody. *J Exp Med* 169, no. 4:1333-1346.

138. Hughes, D.A., I.P. Fraser, and S. Gordon. 1995. Murine macrophage scavenger receptor: in vivo expression and function as receptor for macrophage adhesion in lymphoid and non-lymphoid organs. *Eur J Immunol* 25, no. 2:466-473.

139. Hamann, J., W. Eichler, D. Hamann, H.M. Kerstens, P.J. Poddighe, J.M. Hoovers, E. Hartmann, M. Strauss, and R.A. van Lier. 1995. Expression cloning and chromosomal mapping of the leukocyte activation antigen CD97, a new seven-span transmembrane molecule of the secretion receptor superfamily with an unusual extracellular domain. *J Immunol* 155, no. 4:1942-1950.

140. Hamann, J., E. Hartmann, and R.A. van Lier. 1996. Structure of the human CD97 gene: exon shuffling has generated a new type of seven-span transmembrane molecule related to the secretin receptor superfamily. *Genomics* 32, no. 1:144-147.

141. Yi, T., J.L. Cleveland, and J.N. Ihle. 1991. Identification of novel protein tyrosine phosphatases of hematopoietic cells by polymerase chain reaction amplification. *Blood* 78, no. 9:2222-2228.

142. Walter, U.M., L.M. Ayer, B.A. Wolitzky, D.D. Wagner, R.O. Hynes, A.M. Manning, and A.C. Issekutz. 1997. Characterization of a novel adhesion function blocking monoclonal antibody to rat/mouse P-selectin generated in the P-selectin-deficient mouse. *Hybridoma* 16, no. 3:249-257.

143. Hasty, P., and A. Bradley. 1993. Gene targeting vectors for mammalian cells. *In* Gene Targeting : A Practical Approach. A.L. Joyner, editor. Oxford University Press. 1-31.

144. Vogeli-Lange, R., N. Burckert, T. Boller, and A. Wiemken. 1996. Rapid selection and classification of positive clones generated by mRNA differential display. *Nucleic Acids Res* 24, no. 7:1385-1386.

145. Poirier, G.M., J. Pyati, J.S. Wan, and M.G. Erlander. 1997. Screening differentially expressed cDNA clones obtained by differential display using amplified RNA. *Nucleic Acids Res* 25, no. 4:913-914.

146. Wilson, D.B., D.M. Dorfman, and S.H. Orkin. 1990. A nonerythroid GATA-binding protein is required for function of the human preproendothelin-1 promoter in endothelial cells. *Mol Cell Biol* 10, no. 9:4854-4862.

147. Foos, G., S. Grimm, and K.H. Klempnauer. 1994. The chicken A-myb protein is a transcriptional activator. *Oncogene* 9, no. 9:2481-2488.

148. Goodwin, R.G., M.R. Alderson, C.A. Smith, R.J. Armitage, T. VandenBos, R. Jerzy, T.W. Tough, M.A. Schoenborn, T. Davis-Smith, K. Hennen, and a. et. 1993. Molecular and biological characterization of a ligand for CD27 defines a new family of cytokines with homology to tumor necrosis factor. *Cell* 73, no. 3:447-456.

149. Ramirez-Solis, R., J. Rivera-Perez, J.D. Wallace, M. Wims, H. Zheng, and A. Bradley. 1992. Genomic DNA microextraction: a method to screen numerous samples. *Anal Biochem* 201, no. 2:331-335.

VITA

Hsi-Hsien Lin was born in Nantou, Taiwan on August 25, 1965. He graduated from Taichung First High School, Taichung, Taiwan in 1983. He then attended National Cheng-Kung University where he received a Bachelor of Science degree in Biology in the summer of 1987. Then he enrolled into National Taiwan University to attend Graduate School. He graduated with a Master of Science degree in Biochemistry in June of 1989. He then worked as a research assistant at the Department of Medical Research in Veterans General Hospital - Taipei for three years. He came to United States in 1992 to enter the Ph.D. program in Biomedical Sciences located in the Biology Division of Oak Ridge National Laboratory, Oak Ridge, Tennessee. He received his Ph.D. degree in Biomedical Sciences in December of 1997 for research performed under the direction of Dr. Michael L. Mucenski.