1994

Identification and expression of a novel homeobox gene in a megakaryocytic leukemia cell line

Leonard John Landesberg
Yale University

IDENTIFICATION AND EXPRESSION OF A NOVEL HOMEOBOX GENE IN A MEGAKARYOCYTIC LEUKEMIA CELL LINE

LEONARD JOHN LANDESBERG

Yale University

1994

HARVEY CUSHING / JOHN HAY WHITNEY
MEDICAL LIBRARY

MANUSCRIPT THESES

Unpublished theses submitted for the Master's and Doctor's degrees and deposited in the Medical Library are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but passages must not be copied without permission of the authors, and without proper credit being given in subsequent written or published work.

This thesis by has been used by the following persons, whose signatures attest their acceptance of the above restrictions.

NAME AND ADDRESS DATE

IDENTIFICATION AND EXPRESSION OF A NOVEL HOMEOBOX GENE IN A MEGAKARYOCYTIC LEUKEMIA CELL LINE

A Thesis Submitted to the Yale University School of Medicine in Partial Fulfillment of the Requirements for the Degree of Doctor of Medicine

by

Leonard John Landesberg

1994

Permission for photocoping or microfilming of "Identification and Expression of a Novel Homeobox Gene in a Myeloproliferative Leukemia Cell Line"

(Title of thesis)

for the purpose of individual scholarly consultation or reference is hereby granted by the author. This permission is not to be interpreted as affecting publication of this work or otherwise placing it in the public domain, and the author reserves all rights of ownership guaranteed under common law protection of unpublished manuscripts.

Leonard Landolfi

Signature of Author

3/30/94

Date

General Acknowledgments

I would like to dedicate this thesis to my grandfather, Barnett J. Gold, MD, who passed away on March 8, 1994. He was one of my principal role models, both as a physician and as a person.

In February, 1990, I met Dr. Kenichi Takeshita, whose studies of homeobox genes in hematopoiesis sparked my interest. However, while Dr. Takeshita was on sabbatical in Japan, I spent the first summer of medical school under the able direction of Dr. Jacques Bollekens in the Ruddle lab at KBT. Jacques is an excellent teacher who has always been available for advice, whether it is on the latest PCR technique or on how to solve my computer graphics problems.

Once Dr. Takeshita returned from Japan, I was able to move back to the medical school and resume my work in the Hematology section. This location was more convenient for me as I began my clinical education. It is probably one of the elements which makes a Yale education unique, but I can recall many days when I would finish my work on the wards and then run up to the lab to run a gel, develop a film, or set up some other experiments. Luckily, I only ruined a few ties that way.

As I look back on some of the choices I made during medical school, my decision to work with Ken Takeshita was my best. Ken has been both my mentor and friend, and he exhibits all the qualities which I believe make a great academic physician. His patience, optimism, and soft-spoken nature will attract many future students as he advances in his career. I can only say "thank you" for all the time and effort he has generously spent in making me both a better scientist and future physician.

Helene, thank you for being there for me and for putting up with my various moods during our phone calls. You really did cheer me up. At this point, you probably know all there is to know about molecular biology.

Last, I would like to thank my family: Mom, Dad, Jeff, and Grandma. You are all very important to me. I could not have succeeded in medical school without your love and support.

IDENTIFICATION AND EXPRESSION OF A NOVEL HOMEOBOX GENE IN A MEGAKARYOCYTIC LEUKEMIA CELL LINE. Leonard J. Landesberg and Kenichi Takeshita. Section of Hematology, Department of Internal Medicine, Yale University School of Medicine, New Haven, CT.

The homeobox-containing genes encode transcription factors responsible for regulating cytodifferentiation and developmental events including those of hematopoiesis. Utilizing a set of degenerate primers designed to detect different types of divergent homeodomains by reverse transcriptase-PCR (RT-PCR), a novel homeobox gene (preliminarily named PAX-TS) has been identified in a megakaryocytic cell line (TS 9;22) derived from a patient in chronic myelogenous leukemia (CML) blast crisis. The RNase-protection assay against five leukemia cell lines confirmed that PAX-TS is only expressed TS 9;22. Sequence analysis demonstrates 100% amino acid homology to the homeobox of the recently described Cart-1 gene, which is expressed exclusively in rat chondrocytic cells, and 69% homology to PAX-3, a member of the paired-type homeobox-containing genes. Expression analysis by RT-PCR with Southern blot of other organs and tissues indicates expression in normal human bone marrow, in the day 11 mouse embryo, and in the adult rat kidney, suggesting that PAX-TS is either the same gene as Cart-1 or is a new gene which is expressed during normal embryogenesis, hematopoiesis, and in adult tissues. Similar studies, using chronic-phase and blast-crisis bone marrow total RNA samples from the CML patient whose bone marrow was used to establish the TS 9;22 cell line failed to demonstrate PAX-TS expression. In summary, we have isolated a new homeobox gene, PAX-TS, and conclude that 1) in TS 9;22, PAX-TS is expressed either ectopically or in the small subset of bone marrow cells which may have been preferentially selected in the megakaryocytic cell line and 2) that it is expressed in some non-hematopoietic, non-transformed cells.

INTRODUCTION

Cell differentiation is governed by strict molecular controls. During embryogenesis, a single celled zygote divides and differentiates into the cell groups destined to assume the characteristics of all tissue types. In many adult organs, such as the bone marrow and gastrointestinal tract, multipotent stem cell populations differentiate to provide a source of cells for mature, functioning organs. The mechanisms responsible for the precise organization and control of these events at the cellular and molecular levels remain undiscovered. When the events of differentiation are effectively carried out, the result is a phenotypically normal organism or a normal functioning adult organ. However, when these mechanisms go awry, the result may be a phenotypically abnormal or nonviable organism, an inappropriately functioning organ system, or the production of a cell population (i.e. a tumor) whose growth and differentiation characteristics are out of control.

Normal hematopoiesis

Hematopoiesis, the formation of white and red cells and platelets in the bone marrow, is an excellent model for the study of both normal and abnormal cell differentiation. It may be studied within a single compartment (the bone marrow) and the end products of differentiation are relatively few in number and need not conform to the defined spatial organization of other tissues. Many clinically important entities, such as the leukemias and aplastic anemias, represent abnormalities in hematopoiesis.

Hematopoiesis serves to continuously provide an organism with the full complement of functional cells which make up the peripheral blood. These cells include erythrocytes, neutrophils, eosinophils, basophils, lymphocytes, monocyte/macrophages, and the megakaryocyte-derived platelets. It is an extraordinarily flexible system with a capacity for the production of 2.4 million red blood cells per second in a normal adult, and with the ability to effect rapid and controlled fluctuations in specific blood populations in response to stresses such as infection, hemorrhage, and allergic conditions (1). With the

advent of newer biochemical and molecular techniques, many important hematopoietic humoral growth factors and cell receptors have been discovered. These factors have been found to directly influence the proliferation and differentiation of bone marrow elements. However, the control and coordination of hematopoiesis at the molecular level remains largely unknown.

The current view of hematopoiesis is reflected by the stem cell model. This model underlies our current understanding of both normal hematopoiesis and pathologic hematologic disease states. In this system, all hematopoietic cells derive from a relatively small pool (estimated at approximately 1 out of every 10^6 nucleated bone marrow cells) of undifferentiated, pluripotent hematopoietic stem cells. These cells would have the ability to differentiate along any of the lineages of mature peripheral blood elements and for self-renewal in order to replenish the source of future hematopoiesis (2). Under the rubric of the stem cell model, hematopoiesis is divided into three major steps: 1) commitment, in which the pluripotent stem cell is somehow instructed to become limited to a specific subset of hematopoietic lineage possibilities (e.g. erythroid), 2) proliferation, in which committed, progenitor cells divide and increase their numbers in response to both humoral hematopoietic growth factors and specific intracellular transcription factors, and 3) maturation, in which each committed progenitor cell assumes the phenotypic properties of the lineage to which it has committed (Figure 1).

Evidence of a pluripotent stem cell was demonstrated in experiments involving bone marrow transplantation into mice whose marrow had been ablated by irradiation. The transplanted bone marrow rescued these mice (3) and replaced the hematopoietic process. It was also noted that the spleens of these mice contained large colonies of maturing hematopoietic cells (CFU-S) reflecting all the different lineage possibilities. When these colonies were extracted and injected into another irradiated mouse, they, in turn, were able to reconstitute hematopoiesis (4). Subsequent chromosomal analysis proved that the CFU-S-derived colonies were clonal in origin, suggesting that the cells in these colonies derived

from a pluripotent stem cell (5, 6). When the mitotic activity of the CFU-S-derived colonies was analyzed, it was shown that only 5% of the cells in a single colony were actively dividing. This data supported the hypothesis that a small subset of quiescent stem cells divided infrequently in order replenish the supply of multilineage precursors (7).

As stated above, during the process of commitment, the pluripotent stem cell is somehow restricted into assuming the fate of one cell lineage. At this stage, the cells are no longer pluripotent. They are now only multipotent and have a limited capacity for proliferation; they probably divide only during the process of their own differentiation in response to growth factors (Figure 1). The first stage of commitment involves restriction to either a lymphoid or myeloid lineage (8). These multipotential progenitor cells can be detected experimentally as colony forming units (CFUs) (such as CFU-L and CFU-GEMM, which includes the granulocyte, erythrocyte, eosinophil, basophil, monocyte, and megakaryocyte lineages) in the colony assay. In this procedure, bone marrow progenitor cells of the desired lineage(s) are detected by colony formation in a semi-solid medium, such as methylcellulose in response to a specific growth factor(s). Colonies (50-100 cells) or bursts (>100 cells) are clonal in origin and derive from a single CFU (or BFU) progenitor cell. Those cells which are unresponsive to the given growth factor are nonviable (9).

The multipotential progenitor cells further differentiate into unipotential progenitor cells of the specific lineages. These unipotent progenitor cells which can also be detected in the colony assay (CFU- E, MK, G, M, Eo, Ba, B, T and BFU-E) (8) (see Figure 1). Interestingly, while the action of the humoral growth factors is required for differentiation, the lineage choices made by progenitor cells appear to be random. Ogawa et al. showed that when CFU-GEMM progenitor cells were exposed to GM-CSF (which permits terminal differentiation along granulocytic, monocytic, eosinophilic, erythroid, and megakaryocytic lineages) *in vitro*, the commitment to a specific lineage was made randomly. Even daughter cells isolated from a single progenitor cell division did not differentiate in any predictable

fashion (10).

Although hematopoietic cells look morphologically alike at the progenitor stage, it is possible to isolate lineage-committed populations based upon the presence or absence of certain surface antigens known as "cluster determinants" (CD). Over the course of differentiation, the pattern of CD expression on the cell surface changes. These proteins are defined for the most part by monoclonal antibodies raised against human hematopoietic cells in mice and have been used to classify and fractionate cells of various lineages. For example, CD33 is present on the granulocytic, erythroid, monocytic and megakaryocytic progenitor cells. CD13 is a marker present on granulocytic and monocytic progenitors and CD14 is specific for mature monocytes. CD34 is a marker of primitive hematopoietic cells including the pluripotent stem cells and most progenitors (11).

It is currently thought that pluripotent stem cells are $CD34^{+}HLA\text{-}DR^{-}CD33^{-}CD38^{-}$. $CD34^{+}HLA\text{-}DR^{-}$ bone marrow cells contain the pluripotent stem cells capable of reconstituting lethally irradiated bone marrow (12, 13). The $CD34^{+}$ cell fraction may also be further subdivided to isolate committed progenitors ($CD34^{+}/CD38^{+}$) from the more primitive hematopoietic cells ($CD34^{+}/CD38^{-}$) (Figure 2) (14).

Humoral factors

Hematopoietic cell proliferation and terminal differentiation into the mature phenotype are controlled by the actions of humoral growth factors and as yet undiscovered transcriptional regulators. The human hematopoietic growth factors include a group of at least 12 glycoprotein hormones. The major growth factors, or colony stimulating factors (CSF), are summarized in Table 1. These include: granulocyte-macrophage CSF (GM-CSF), granulocyte CSF (G-CSF), macrophage CSF (M-CSF), erythropoietin (Epo), and the interleukins (IL-1,3,4, 5, 6, 7, 8, 9 (IL-3 is also known as multi-CSF)). They all are relatively small, acidic glycoproteins (18,000-90,000kD) and are active at extremely low concentrations (picomolar range *in vitro*). In general, they are not lineage specific, and are produced by several different cell types. In addition to their proliferative effects, these

proteins also are important for the maintenance of hematopoietic progenitor cell viability (1).

The four "classical" colony stimulating factors, GM-CSF, G-CSF, M-CSF, and IL-3 also act on mature blood elements such as granulocytes and monocytes in order to promote the functional activities of these cells. They prolong cell survival, induce membrane receptor changes, prime for oxidative bursts, are chemotactic, and enhance the ability for phagocytosis, bacterial cell kill, and cytotoxicity. Among the CSFs, GM-CSF acts on eosinophils, neutrophils and monocyte/macrophages; IL-3 acts on eosinophils and monocytes; G-CSF acts on neutrophils; M-CSF acts on monocyte/macrophages. These CSFs are complemented by erythropoietin (Epo) which supports the proliferation and differentiation of late erythroid progenitors (CFU-E) *in vitro*. Growth of earlier progenitors (BFU-E) depends on the burst promoting activity of IL-3 and GM-CSF. Epo is required later for proliferation and hemoglobinization (Table 2) (15).

It is important to note that, although the CSFs bear little structural resemblance to each other, the growth factor receptors do share significant homology in the amino acid sequence. In the case of multi-subunit receptors, they share one of the receptor subunits . They are grouped together in the cytokine receptor family which includes the erythropoietin, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, G-CSF, and GM-CSF receptors (16). This similarity at the receptor level is important for understanding the mechanisms by which growth factor stimulation is transmitted to the cell's transcriptional regulatory apparatus which, in turn, controls cell proliferation and/or differentiation.

Several types of receptor-mediated systems are known to function in cell signaling. These include: 1) the neurotransmitter-gated ion channels which open or close briefly to alter the electrical excitability of the cell, 2) the G-protein linked receptors which directly activate or deactivate plasma-membrane-bound enzymes or ion channels by hydrolyzing their bound GTP, 3) the catalytic receptors which are mainly tyrosine kinases that directly phosphorylate specific target cell proteins on tyrosine residues, and 4) receptors which

activate intracellular tyrosine kinases (17).

The signal-transducing events that are activated by stimulation of hematopoietic growth factor receptors appear to involve tyrosine phosphorylation. It has been shown that some hematopoietic cell lines containing constitutively activated tyrosine kinases do not require CSFs for survival and growth (18, 19). Although the cytokine receptors do not contain a consensus tyrosine kinase catalytic domain, it has been noted that tyrosine phosphorylation is rapidly induced upon treatment with several growth factors. For example, there is data that IL-3-mediated cell activation occurs via tyrosine kinase activation (20). Also, studies of the erythropoietin/erythropoietin receptor interaction have demonstrated the induction of phosphorylation and kinase activity of the c-fps/fes proto-oncogene product in Epo-responsive cells (21). Mutagenesis experiments have demonstrated that disruption of specific regions of the erythropoietin receptor leads to impairment of the induction of tyrosine kinase activity, although the Epo-receptor itself has no kinase activity (22). Therefore, although the CSFs vary in structure, the homology of the cytokine receptors has a functional correlate. The binding of cytokines to their receptor results in intracellular biochemical changes, such as the activation of tyrosine kinase phosphorylation of proteins. Recent reports suggest that some of these tyrosine kinases may phosphorylate transcription factors (23, 24, 25). These studies provide a valuable link between the extracellular actions of the CSFs and the intracellular events which ultimately determine cell proliferation and differentiation.

It may be generalized that hematopoiesis involves the regulation of two primary processes: 1) the intracellular events which result in changes in gene expression and the extracellular signals which alter these events. The external signals are, in part, transduced via cytokine receptors, which activate biochemical changes such as phosphorylation. In this signal transduction pathway, transcription factors affecting gene expression could be modified by phosphorylation. Thus, transcription factors are an important participant in the signal transduction cascade.

Transcription factors in hematopoiesis

Gene transcription, the process of copying a DNA template into messenger RNA, is performed in eukaryotes such as humans by RNA polymerase II. In addition to the binding of RNA polymerase to DNA at specific initiation sites, each gene contains regulatory sequences which are able to modify expression in response to the needs of the cell. These sequences are called promoters, enhancers, and silencers and have the capacity to bind specific regulatory proteins which act as transcription factors. The protein structures and the mechanisms by which several of these factors bind DNA are now known. It is possible to group these proteins into families which share common amino acid motifs. These motifs may either directly affect DNA-binding or indirectly affect it by enabling monomeric proteins to dimerize into an active form. Examples include the helix-turn-helix, helix-loop-helix, leucine zipper, and Zn-finger families.

The helix-turn-helix group of proteins is characterized by three contiguous alpha helical regions separated by single short turns. They dimerize and bind to DNA with an alpha helix from each monomer occupying the major groove of the DNA double helix. The leucine-zipper and helix-loop-helix "amphipathic" transcription factors are believed to also function as dimers. Specific secondary structural features of the protein monomers (e.g. sequentially spaced leucine residues) enable their assembly and hence, the proper association of the DNA-binding domains with the DNA. The zinc-finger proteins require zinc (Zn) for DNA-binding activity. Sequence analysis demonstrates that these proteins contain cysteine and histidine residues at fixed intervals along the amino acid chain. Cysteine and histidine are known ligands for Zn in proteins. A putative structure involves Zn ions interacting with the cysteine and histidine residues in order to support the folded DNA-binding domain in a "finger-like" formation (26).

Several transcription factors have been identified whose patterns of expression suggest that they are involved in hematopoiesis. Some of these genes, such as Id and c-myb are believed to function as general transcription factors with important effects on

hematopoiesis. C-myb is a DNA-binding, cellular proto-oncogene whose expression has been reported in a number of human malignancies such as colon, lung, and breast carcinoma. In addition, it has been isolated from several non-human retroviruses involved in hematologic malignancies. While it appears that c-myb is a general transcription factor and is expressed in many tissues during embryogenesis as well as during cellular proliferation in general, Mucenski et al. reported that this gene is required for normal hematopoiesis. By disrupting c-myb in chimeric mice by homologous recombination in mouse embryonic stem cells and then interbreeding to select for homozygous c-myb- progeny, it was shown that the c-myb- mice became severely anemic during gestation. Embryonic erythropoiesis occurring in the yolk sac was normal, but the subsequent adult-type erythropoiesis occurring in the liver was severely impaired (27).

The Id gene codes for a helix-loop-helix protein which acts as a negative transcription factor. Id lacks a DNA-binding domain but is able to heterodimerize with other helix-loop-helix transcription factors and thus, inactivate them. Kreider et al. studied Id expression during myeloid differentiation. It is not expressed in lymphoid or erythroid cells. They noted a decrease in the expression of Id as a myeloid precursor cell line (32D c13(G)) was induced to differentiate with G-CSF. When Id was constitutively expressed it blocked the ability of 32D c13(G) to differentiate. However, Id expression is not limited to hematopoietic cells (28). Thus, Id may act as a negative regulator by heterodimerization and blockade of the action of other helix-loop-helix transcription factors during myelopoiesis and possibly other tissue types.

Among the transcription factors whose expression patterns are more restricted to hematopoietic cells, GATA-1 and tal-1 have been especially well-studied. GATA-1 is a Zn-finger transcription factor which binds to a consensus DNA element in the regulatory regions of the alpha- and beta-globin gene clusters and other erythroid cell-specific genes (29, 30, 31, 32, 33, 34). It is also expressed in megakaryocytes and mast cells but, importantly, not in other blood cell lineages or in non-hematopoietic cells (35, 36, 37).

Several studies have shown that GATA-1 is involved in erythroid cell maturation. Pevny et al. showed that when GATA-1 was disrupted by homologous recombination in a murine embryonic stem cell line and introduced into chimeric mice, the progeny mutant cells contributed to all cell tissue types except erythrocytes. This study proved that GATA-1 was required for normal erythropoiesis (38).

The tal-1 gene, also known as SCL, has also been implicated as a hematopoietic transcription factor. In normal hematopoiesis, tal-1 has been shown to be expressed in early myeloid, erythroid, megakaryocytic, and mast cell lines, but not in the lymphocytic lineages. Among the non-hematopoietic tissues, tal-1 is only expressed in the developing brain. tal-1 is a member of the family of amphipathic helix-loop-helix transcription factors, (39, 40), a group which includes several genes that are clearly implicated in lymphoid malignancies because of their involvement in chromosomal translocations (e.g. the t(8;14) translocation involving the MYC gene in Burkitt's lymphoma and the t(1;19) translocation resulting in the E2A-PBX1 hybrid in B-cell ALL). Tal-1 is deregulated via chromosomal derangements (either translocations or deletions) in 30% of acute T-cell leukemia (T-ALL). In these cases, tal-1 expression is coupled to the regulatory regions of genes which are expressed in normal T-cells. Somehow, tal-1 overexpression changes the course of hematopoiesis along an abnormal lymphoid lineage (39, 41, 42, 43).

A recent study has demonstrated that both tal-1 and GATA-1 are downregulated when the K562 human erythroleukemia cell line is induced to differentiate toward a macrophage-like morphology by treatment with hexamethylene bisacetamide (HMBA). Northern blot analysis also demonstrated a down-regulation of both the alpha- and ß-globin genes accompanying HMBA-induced differentiation away from an erythroid phenotype (44). A prior study using antisense disruption of tal-1 in the same cell line, K562, showed that this gene is also required for cell proliferation (45).

Mouthon et al. have performed a similar study of GATA-1 and tal-1 in which their expression was analyzed in normal hematopoietic cells rather than in leukemic cell lines.

They showed that tal-1 and GATA-1 were coexpressed in the committed progenitor (CD34+/CD38+) cell fraction whereas they were undetectable in the more primitive (CD34+/CD38-) cell fraction (see Figure 2). When GATA-1 and tal-1 expression were studied by *in situ* hybridization to CD34+, BFU-E-derived-colonies and cultured, immature megakaryocyte colonies, it was noted that both genes were coordinately downregulated during erythroid and megakaryocytic terminal maturation (46).

It is evident that the activity of certain DNA-binding proteins are important for the regulation of normal hematopoiesis even in the presence of the appropriate growth factors. Some have also been implicated in certain hematologic malignancies. The expression of certain transcription factors such as GATA-1 and tal-1 appear to be more restricted to hematopoietic and a few other cell types. These, and other as yet undiscovered genes are clearly essential for proliferation and maturation of mature blood elements. Further studies may show that they also play an vital role in the less understood process of lineage commitment from the pluripotent stem cell.

The homeobox genes

The homeobox-containing genes, a family of interrelated DNA-binding proteins, have also been implicated as transcription factors in the hematopoietic process (47). A homeotic transformation is either a duplication of an anterior/posterior body structure or the replacement of one body structure with another during embryogenesis. In the early 1980s it was found that a homeotic transformation leading to the formation of legs in place of antennae in the fruit fly, *Drosophila melanogaster*, could be explained by a mutation in a single gene. This gene was characterized and named Antennapedia (Antp.) (Figure 4). Other such mutations in *Drosophila* led to the discovery of a group of genes which are believed to control body pattern formation during embryonic development. Comparison of these gene sequences with Antp. demonstrated that they share a highly conserved 180-183 nucleotide sequence named the homeobox (48, 49, 50). It was postulated that these genes functioned at the level of transcription in order to influence the production of effector genes

which determine cell type fate and tissue identity (51, 52, 53). Now known as the homeobox family, these genes are characterized by an evolutionarily conserved 60-61 amino acid domain, known as the homeodomain, which was shown by X-ray crystallography to form a helix-turn-helix motif capable of binding to DNA (54, 55, 56) (Figure 3). A diagram of the Antp. homeobox and its three conserved alpha-helical regions is shown in Figure 4.

The homeobox has been identified in 25 *Drosophila* genes which are grouped together in a cluster called HOM-C. The HOM-C cluster is characterized by Antp.-like homeoboxes and these are known as the class I homeobox genes. Although these genes differ considerably outside the homeobox, the homeodomain itself is remarkably similar. It has also been shown that these genes lie in the same order in the homeobox cluster as does their expression pattern along the anterior/posterior axis of the the developing insect embryo, suggesting a formal relationship between chromosomal architecture and embryogenesis (57). There are also several genes which contain homeodomains which are related to but are divergent from Antp. A divergent homeobox is one in which the sequence of the homeodomain varies greatly from Antp., but does retain certain highly conserved amino acid residues at the 3' end of the homeodomain, namely the amino acid sequence, IWFQNRR, which is extremely well conserved in almost all homeobox genes. These divergent genes are not necessarily confined to the HOM-C cluster and are located throughout the genome in the *Drosophila*.

Mammalian homeobox genes

The homeodomain has been conserved among diverse species including the nematodes, annelids, arthropods, plants, and vertebrates including humans (58, 59, 60, 61). This conservation suggests an evolutionarily functional significance. As in the *Drosophila*, the mammalian body plan also appears to be governed at least in part by the actions of homeobox genes. Expression analysis has demonstrated activity in all embryonic tissues including the nervous system, the developing extremities, and the genital

region (62, 63, 64, 65). Recent studies with mice have revealed gross developmental anomalies after the homozygous deletion of individual homeobox genes. For example, mice with a homozygous deletion of Hox a3 died shortly after birth but presented with congenital absence of the thyroid, parathyroid, and thymus glands as well as with anomalies in the craniofacial, submaxillary and cardiac regions, suggesting a similarity to the human DiGeorge's syndrome (66). Similar "knock-out" experiments with other homeobox genes have demonstrated homeotic transformations resulting in: 1) delayed hindbrain neural tube closure, absence of portions cranial nerves VII, VIII, IX and X, and malformed inner ear and parieto-occipital bony structures with Hox a1 deletion (67, 68), and 2) transformations of thoracic skeletal structures from deletion of Hox c8 (69). These defects are related to the spatial patterning of the embryo rather than to single tissues, organs or cell types, reflecting the putative role of the homeobox genes as body plan organizers.

In humans, 38 class I homeobox genes have been identified. As in *Drosophila*, the class I homeobox genes are designated because of their similarity to Antp. and are clustered into distinct chromosomal regions. However, in humans they are spread among four separate chromosomes in four clusters called HOXA, HOXB, HOXC, and HOXD located on chromosomes 7, 17, 12, and 2, respectively (70). See Figure 5 for the chromosomal organization of the human and murine class I homeobox genes. Each cluster contains between 9 and 11 genes which are numbered to correspond to both their linear organization on the chromosome and to their pattern of expression . This new nomenclature replaces a former system in which the human HOX clusters were numbered 1-4 with each gene designated by a capital letter corresponding to its order of discovery . As in the human, the murine equivalents (Hox) of the class I genes are clustered on four chromosomes in loci named Hoxa, Hoxb, Hoxc, and Hoxd and are also designated with numerical appellations (71, 72). On the basis of comparative amino acid sequences within the homeodomain, the HOX genes can be organized into 13 "paralogous" groups between the chromosomes (for

example HOXA4, B4, C4, and D4) (Figure 5). It is significant that the paralogues occupy the same positions in their respective chromosomal HOX clusters so that the linear organization of each cluster is the same. This paralagous relationship also corresponds to the members of the *Drosophila* HOM-C cluster (Figure 5). Based on this organization of paralogous clusters, a consensus of opinion holds that, over the course of the evolution of higher animals, these four chromosomal homeobox complexes arose from successive duplications and mutations of a primordial homeobox cluster (73).

The empty gene spaces within each HOX cluster were originally left for the predicted paralogues which had not yet been discovered (70, 74) (Figure 5). However, these regions have been cloned and sequenced and do not contain any genes. The significance of this finding is unclear.

In both mice and humans, as in *Drosophila*, the expression of homeobox genes within a cluster parallels the spatial patterns of embryogenesis. In mice, the anterior-posterior pattern of homeobox expression corresponds to the 5' to 3' position in the Hox cluster (62). In other words, one can define class I homeobox genes as "anterior" or "posterior" based on their position within the cluster, with the expression domain of the 5' genes being represented more anteriorly than their 3' neighbors. This spatial hierarchical arrangement has been demonstrated in transgenic mice when the ectopic expression of Hox d4 outside of its native pattern resulted in phenotypic anomalies whereby the occipital bones were transformed into cervical vertebrae. In essence, a normally rostral area was replaced by an ectopically placed caudal structure (75). In humans, it has been shown that retinoic acid is able to induce HOX gene expression in teratocarcinoma cells in a sequential manner such that expression favors the more 3' ends of each cluster (76, 77).

The complexity of homeobox functioning is magnified by the discovery that the expression of single genes is modified by alternative splicing of the mRNA transcripts. In other words, one gene can encode several proteins. The alternative splicing products include transcripts with or without the homeodomain, and transcripts with varied regulatory

regions (78, 79, 80, 81, 82, 83) . This added dimension suggests another level of regulation which acts to "fine-tune" the effects of single or groups of homeobox genes.

Divergent homeobox genes

As in *Drosophila*, the human genome contains a subset of divergent homeodomain-containing genes which, as discussed above, do retain some highly conserved amino acid residues in the homeobox and are not confined within the HOX clusters. Some of these genes are direct homologues of the *Drosophila* divergent homeoboxes; others have been identified which have no known fruit fly equivalent. It is possible to group these divergent homeobox genes into families which share significant sequence homology in the first ten amino acids of the homeobox (84). Among these, are the evx family (related to the even-skipped *Drosophila* gene), the en family (related to the engrailed *Drosophila* gene), and the octamer family.

Figure 4 contains a schematic representation of several of the known types of homeodomain-containing genes. Some divergent homeobox genes contain additional, conserved DNA-binding peptide domains. An unexpected finding was the ZF class of divergent homeobox genes which contain some combination of several Zn-finger motifs and extremely divergent homeodomains (85, 86). In addition to a homeodomain, the Pax family contains a common protein domain homologous to the *Drosophila* paired box. The paired-box motif also has DNA-binding properties (87). This family's homology lies in the paired domain rather than in the homeobox which may be variably included or even non-existent (88) Only three of the eight known Pax genes contain a homeodomain (89). The LIM homeodomain-containing proteins bear a cysteine-histidine-rich domain which may enable it to bind DNA. The POU family of genes was originally discovered as a group of transcription factors responsible for tissue specific gene expression. Examples include the Oct-2, LFB1 and GHF-1 proteins which are expressed in B lymphocytes, the liver, and the pituitary, respectively (90, 91, 92, 93, 94). These expression patterns suggested that homeobox genes play a role in maintaining the differentiated state of adult

cells in addition those involved in embryogenesis.

Homeobox genes in hematopoiesis

In the last five years, several groups have identified homeobox gene expression in hematopoietic cells. Most of the initial studies concentrated on leukemia cell lines whose phenotypic pattern suggested specific lineage commitments. Newer data incorporates the use of normal hematopoietic cells selected by cell surface antigen characteristics.

HOX genes in hematopoiesis

The earliest studies reporting homeobox gene expression in hematopoiesis utilized Northern blot analysis against leukemia cell lines of varied phenotype (95, 96, 97). Although some of the preliminary conclusions have been modified, these experiments suggested that the homeobox gene transcription factors were active in hematopoietic cells and predicted that certain ones were lineage specific.

Further studies attempted a more targeted analysis of HOX expression. Shen et al. screened four members of the HOX B locus in 18 leukemia cell lines representing erythroid, granulocyte-macrophage, macrophage, and B- and T-lymphoid lineages. It was concluded that HOX B4, B5, B6, and B7 are expressed predominantly in the erythroid cell lines. This study also noted that as two erythroid cell lines were induced to differentiate, the levels of HOX B4 transcription increased while the levels of one species of the HOX B7 transcript decreased (96).

Since it appeared that the HOX genes were lineage specific and that groups of HOX genes within certain loci were restricted in expression pattern, larger studies examined the expression patterns of all class I homeobox genes in leukemia cell lines. Magli et al. looked at the expression of the 38 class I genes in four cell lines representing erythroid, myeloid and lymphoid lineages. They noted that: 1) the HOX D locus was not active in any of the cell lines, 2) the HOX B locus was restricted to the erythroid lines, 3) in the HOX A locus, four of the 5' genes (HOX A9, A10, A11, A12) were not expressed in the erythroid lines but were expressed in the myeloid and lymphoid lines, and 4) in the HOX C

locus, four of the 5' genes (HOX C10, C11, C12, C13) were expressed in the erythroid lines but not expressed in the myeloid and lymphoid lines. HOX C5, C6, C8, and C9 were expressed in all four lines. They concluded that not only is class I gene expression in hematopoietic cells lineage specific but that it may be coordinated by switching entire or parts of HOX loci on and off (98). In support of this study and utilizing 19 cell lines of varied lineage, Matthews et al. showed that with the exception of HOX B7, HOX B locus expression was restricted to erythroid cell lines (99).

As the methods of reverse-transcriptase polymerase chain reaction (RT-PCR) became more widely practiced, several studies reported HOX expression patterns which conflicted with some of the earlier data utilizing Northern blotting (100, 101). These results are likely to be due to the higher sensitivity of RT-PCR versus Northern blot for detecting low levels of expression. Also, Deguchi et al. showed that HOX B7 levels were rapidly increased during activation of mature T-cells (97). Taken together, these data probably reflect the varying stages of homeobox expression in the specific "windows" of hematopoiesis reflected in the available leukemia cell lines and not in normal cells. Future data will need to identify homeobox expression in normal bone marrow cells selected for specific lineage characteristics by antigen-typing.

Possible function of HOX genes in hematopoiesis

Recent experiments have utilized antisense oligonucleotides to block homeobox expression by RNase-H degradation of the mRNA/oligo duplex *in vitro*. Shen and Detmer used an antisense HOX B6 expression vector to downregulate endogenous expression HEL and found that the erythroid features of the cell increased. Overexpression of the same gene in the K562 myeloerythroblastic cell line resulted in decreased erythroid characteristics such as alpha globin synthesis (102). Takeshita et al. identified the expression of several HOX C cluster genes in a mouse erythroleukemia cell line by RT-PCR. It was shown that HOX C6 transcription increased as the cells were induced to differentiate. Antisense oligonucleotide down-regulation of HOX C6 expression led to

decrease in proliferation of CFU-E-derived colonies but left the proliferation of BFU-E-derived and myeloid colonies unaffected. Terminal maturation indices such as hemoglobinization were not affected (84) These studies support the earlier Northern blot expression data and demonstrate that certain homeobox genes have precise roles in specific aspects of hematopoiesis, namely that both induced expression and down-regulation are vital to hematopoietic differentiation.

Divergent homeobox genes in hematopoietic cells

As discussed above, the divergent homeobox genes retain the highly conserved region of the homeobox, but may be grouped into families which share particular homologous sequences within the homeodomain. Table 2 contains a summary of the divergent homeobox genes expressed in hematopoietic cells.

The first evidence that divergent homeobox genes were involved in hematopoiesis was the isolation of the S8 homeobox from a mouse spleen library by Kongsuwan et al. However, it was noted that S8 most likely originated from supporting stromal cells since it was expressed in a fibroblastic cell line and could not be detected in any hematopoietic cell lines by Northern analysis (95).

Currently, several divergent homeobox genes are known to be expressed in hematopoietic cells and cell lines. HB9 and HB24 were isolated from a B-lymphocyte cDNA library (103). The HB24 gene was found to be homologous to the murine Hlx gene and is now known as HLX (104). HLX and HB9 are located on chromosome 1 and are therefore separate from the HOX loci. Both HLX and HB9 were found to be strongly expressed in the CD34+ fraction of bone marrow cells and silent in the CD34- fraction. Inhibition of expression of HLX by antisense oligonucleotide-directed down-regulation interfered with the proliferative potential of CD34+ cells treated with IL-3 and GM-CSF. In addition, the levels of other early transcription factors such as c-fos, c-myc, and c-myb were decreased (103). Takeshita et al. reported expression in myeloid cells but found that within the CD34+ fraction of hematopoietic cells, HLX was expressed in the

CD34+HLA+DR+ but not in the CD34+HLA-DR- fractions. Since the CD34+HLA-DR- cells contain pluripotent stem cells and can replete ablated bone marrow, HLX seems to be activated at an early point during hematopoiesis. Additionally, antisense experiments showed that down-regulation of HLX inhibited myeloid but did not influence erythroid progenitor colony formation (105). These data on HLX expression suggest a role for HLX in early myeloid commitment and that its expression is linked to the expression of other known transcription factors.

Xu et al. isolated LH-2, a member of the LIM family of homeobox genes, while screening for markers of early B-lymphocyte differentiation. It was expressed in hematopoietic cell lines in a lineage-specific manner. Specifically, expression was identified in several pre-B-cell-stage cell lines and only at much lower levels in those cell lines reflecting more mature B-lymphocyte development. In T-lymphocytes, its expression did not appear to be restricted to any particular stage of differentiation (106).

Several groups have used RT-PCR with divergent oligonucleotide primers to probe for new homeodomain-containing genes in hematopoietic cells. Using this protocol, the human equivalents of the murine Msx-2 and Dlx-1 genes were isolated from two myeloid leukemia cell line cDNAs. These genes are now known as MSX-2 and DLX-1 and had not been previously described in human hematopoietic cell lines (105). The human equivalent of the murine Cdx gene was isolated from the Jurkat human T-cell leukemia line (107). Cdx has subsequently been shown to be expressed in the murine myeloid leukemia cell line, B6-SUT (unpublished data). PRH (also known as HEX) was isolated from chicken myeloid leukemia cells, the human promyelocytic HL60 cell line, normal human bone marrow, and peripheral blood. Northern blot expression analysis determined PRH to be expressed in multilineage progenitor cell lines and specifically in myeloid and some B-lymphocyte cell lines. Among non-hematopoietic tissues, it was found only in the liver (108, 109); (110). GBX1 has recently been cloned from a human genomic DNA library and has been mapped to chromosome 7 near the region occupied by the EN-2 Homeobox

gene. Northern blotting confirmed expression within the hematopoietic compartment and specifically in the myelo/erythroblastic K562 and Daudi cell lines (111). The Oct 11 gene, a member of the POU family, was found to be expressed in a mouse myeloma cell line (112).

The ultimate strategy for examining homeobox gene expression in hematopoiesis would involve lineage-specific cells from normal bone marrow during stem cell commitment . These experiments would validate the earlier studies and eliminate the chance that normal expression patterns are altered in transformed cells or during the establishment of the leukemia cell lines. Preliminary experiments have noted the expression of certain HOX B genes in normal bone marrow by Northern blot analysis (96, 99, 97).

Homeobox genes and hematologic malignancies

A logical extension of the discussion of transcription factors in general and of homeobox genes in particular would be their potential role in hematologic malignancies. Since the activity of these genes has clearly been implicated in the differentiation of hematopoietic tissues, three mechanisms for dysregulating expression could account for leukemogenesis. A normally expressed homeobox gene or group of genes could be over-expressed or under-expressed leading to a shift in the delicate balance of the factors controlling cell growth characteristics and subsequent cell transformation. Also, a homeobox gene could become expressed ectopically either as a direct result of a specific genetic insult or by induction along an entire cascade of dysregulated transcription factors.

Several examples of these phenomena have already been noted in the literature. The murine myelomonocytic leukemia cell line, WEHI-3B, was shown to have a retroviral insertion in the 5' non-coding region of Hox b8 leading to abnormal expression of the gene. This cell line also has a retroviral insertion upstream of the IL-3 gene. Overexpression of IL-3 alone did not induce leukemia, but co-infection of retroviruses carrying IL-3 or IL-3 plus Hox b8 into murine bone marrow resulted in myelomonocytic

leukemia only with the combination of IL-3 and Hox b8 (113, 114, 115). When Hox b8 was overexpressed in a myeloid cell line, terminal differentiation was inhibited (116). These results suggest that Hox b8 may function as an oncogene and prevents terminal differentiation of myeloid cells in response to growth factors.

Two specific chromosomal translocations in human leukemias led to the discovery of the divergent homeoboxes, HOX 11 and PBX1. 5-7% of T-cell ALL have a t(10;14)(q24,q11) translocation which fuses a portion of the T-cell receptor (TCR) delta with the coding region of HOX 11 leading to its overexpression in these cells. A rarer translocation, t(7;10)(q35,q24), places HOX11 immediately downstream to the TCR-ß gene. Normal T-cells express HOX 11 at much lower levels (117, 118, 119). Subsequent overexpression of HOX11 in transfected NIH3T3 cells resulted in cell transformation (120).

25% of B-cell ALL contain the translocation t(1;19)(q23,p13.3) which fuses the a portion of the E2A gene with a divergent homeobox gene now called PBX1. E2A protein normally binds to the enhancer of the immunoglobulin kappa light chain gene in B-cells. PBX1 is normally expressed in several adult tissues, but not in normal developing B-cells (121, 122, 123, 124). This fusion protein contains both a homeodomain and an activator for immunoglobulin transcription and may act to incorrectly target E2A action on regions which are normally responsive only to PBX1. As with other homeobox genes, the E2A-PBX1 fusion gene results in several alternatively spliced transcripts. The proteins encoded by two of these transcripts from the fusion gene have been shown to result in tumors in nude mice injected with transfected NIH3T3 cells (125).

In addition, HLX has been implicated in leukemogenesis (103, 126). Overexpression of this gene in a human T-cell line (Jurkat) resulted in increased cell proliferation and also led to the formation of tumors when injected into nude mice (127, 128).

<u>Approach to the study of divergent homeobox gene expression in hematopoiesis</u>

These results suggest that, as a group of transcription factors with DNA-binding ability, the homeobox genes may control hematopoiesis and serve as protooncogenes and, when deregulated, are able to control certain events crucial to leukemogenesis. With this in mind, I studied homeobox gene expression in a new cell line, TS9;22, which was established from a CML patient in megakaryocytic blast crisis. The aim of the study was to identify homeobox genes which may be involved in leukemogenesis and/or megakaryocytopoiesis.

As an established, stable cell line TS9;22 contains both myelo/monocytoid and megakaryocytic characteristics. The immunophenotypic composition (by CD and HLA antigens) of both the patient and TS9;22 are presented in Table 3. Although TS9;22 is a stable cell line, it contains a heterogeneous population of cells which show characteristics of several individual cell types. Also, it is important to note that the immunophenotype of the cells assayed changed over a period of two years and after establishment of the cell line (129).

Utilizing RT-PCR with divergent primers designed to identify all classes of known divergent homeoboxes, I isolated a new homeobox gene in the TS9;22 cell line and studied its expression in various leukemia cell lines and non-hematopoietic cells.

METHODS AND MATERIALS

Cell lines

The six cell lines used and their defining characteristics are summarized in Table 4.

Cell culture and RNA extraction

TS9;22, ES, HL60, MEL, KG1 and Meg01 cell lines were grown as described (130, 131); (132, 133); (134). The TS9;22 cell line is cultured in RPMI 1640 +15% FCS (K. Ohyashiki (unpublished data)). Large-scale preparations of total RNA were made using the acid guanidine procedure (135). Poly-A+ RNA was isolated by hybridization of 500µg total RNA (estimated by OD_{260}) with biotinylated oligo-dT followed by selection with MagneSphere streptavidin paramagnetic particles (PolyATtract mRNA Isolation Systems, Promega) according to the manufacturer's protocol . The poly-A+-mRNA was then extensively treated with RNase-free DNase (Gibco-BRL) prior to cDNA synthesis and subsequent RT-PCR screening for divergent homeobox expression. These DNased poly-A+-mRNA samples were then tested with the PCR primer pair, 2551-2552 (84), to certify that all contaminating genomic DNA was removed. The PCR conditions were 94˚C for 1 min., 47˚C for 1.5 min., and 72˚C for 2 min. for 35 cycles, using 1µg of each primer and 1µl of the DNased poly-A+-mRNA.

First-strand cDNA synthesis

First-strand cDNA was synthesized from the DNase-treated polyA+ mRNA with M-MLV reverse transcriptase according to the directions in the Superscript Preamplification System (Gibco-BRL). cDNAs were prepared with both random hexamer and oligo-dT primers. The random hexamer derived cDNAs were then tested with the PCR primer pair, 2551 and 2552 (84) (see Table 5), to determine if the cDNA was intact and if it contained HOX homeodomains. The PCR conditions were 94˚C for 1 min., 47˚C for 1.5 min., and 72˚C for 2 min. for 35 cycles, using 1µg of each primer and 1µl of the cDNA.

RT-PCR

The TS9;22- and ES-derived cDNAs were initially selected for study. RT-PCR

was performed exactly as previously described (84, 136) except that the 5' primers shown in Table 5 were used in order to detect the amino acid sequences characteristic of 19 types of divergent homeobox genes in addition to the class I HOX genes. The PCR conditions were 94°C for 1 min., 47°C for 1.5 min., and 72°C for 2 min. for 35 cycles, using 1µg of each primer and 1µl of the cDNA synthesized with the random hexamer primer. The PCR products were viewed on a 3% agarose gel .

Subcloning of PCR fragments

The TS9;22 PCR products were selected for study. The reactions which produced a visible ~150bp band were reamplified under the same conditions. These fragments were then ligated into the TA Cloning vector (Invitrogen Corp.) exactly as described by the manufacturer's protocol.

Competent cells were transformed with the ligated DNA, and positive clones were picked by ß-galactosidase blue-white selection. Nine homeobox clones were characterized by dideoxy sequencing using Sequenase (US Biochemicals). Of the three different homeoboxes identified, a single clone containing a homeobox derived from the TS9;22 poly A+ mRNA (PAX-TS) was selected for further analysis.

RNA expression analysis

The RNase protection assay for PAX-TS was performed as follows: 1µg of the plasmid containing the PAX-TS homeobox DNA was linearized with XhoI (see Figure 11 for orientation of insert in TA Cloning vector). Antisense mRNA was radiolabeled by runoff transcription with SP6 RNA polymerase (Promega) in the presence of 100µCi of alpha-^{32}P CTP at a specific activity of 600Ci/mmol. The reaction mixture was treated with RNase-free DNase (Gibco-BRL), extracted with phenol-chloroform, and gel purified from a denaturing 5% polyacrylamide gel to eliminate minor products of transcription.

For the RNase-protection assay, 500,000cpm of the radiolabeled antisense probe were used per reaction. 15µg of each of TS9;22, HL60, Meg01, KG1, and MEL total RNA suspension were co-precipitated along with the antisense probe. As a positive

control, a cold, sense mRNA probe from the same plasmid linearized with HindIII was prepared by runoff transcription with T7 RNA polymerase. 1µl of a 1:500 dilution of the cold probe was used per reaction. Sterile water served as a negative control assay. The pellet was washed with 70% ETOH and resuspended in 30µl 5X PIPES (200mM PIPES (pH6.7), 2M NaCl, 1mM EDTA) diluted to 1X with formamide. The RNA was denatured at 85°C for 5 min. and then allowed to hybridize at 45°C overnight. After hybridization, the mixture was cooled to room temperature and digested with 300µl RNase digestion mix (10mM Tris (pH7.4), 5mM EDTA, 300mM NaCl, 40µg/ml RNaseA, 1µg/ml RNaseT1) at 30°C for 60 min. to remove the non-hybridized, single-stranded RNA . Following removal of the RNases and protein with proteinase-K in 1% SDS, the resulting reaction mixture was extracted with phenol-chloroform, ETOH precipitated, and loaded onto a denaturing 5% polyacrylamide gel (1200V, 40mA, 2hrs). The gel was autoradiographed with an intensifying screen.

Oligonucleotide preparation

Oligonucleotides specific for the TS9;22 homeobox under consideration (PAX-TS) were ordered (Figure 13). These oligonucleotides correspond to the base sequence internal to the divergent PCR primers 3766 and 2552. The sequences of the oligonucleotides were: TGCAGCTAGAGGAGCTGGAG, TS-1; GCTAGAGGAGCTGGAG, TS-2; GATGTGTATGTCAGAGAACA, TS-3; TGAGGACAGAGCTCACTGAG, TS-4; corresponding to the sense strand and CTGGGACCGGAGTCACTCG, TS-RA; CAGTGAGCTCTGTCCTCAG, TS-RB; corresponding to the antisense strand.

Determination of the specificity of primers

RT-PCR with primers TS-1 and TS-RA was performed on total RNA, DNased total RNA, and first-strand cDNA derived from the TS9;22, Meg01, ES, HL60, KG1, and MEL cell lines described above, followed by Southern blotting. $ß_2$-microglobulin-specific primers and DLX-1-specific primers were used as positive cDNA controls (105).

The concentration of total RNA in each sample was estimated by OD_{260} and

adjusted so that the amounts were approximately equal prior to DNase treatment and cDNA synthesis. The PCR conditions were 94°C for 1 min., 55°C for 1.5 min., and 72°C for 2 min. for 30 cycles, using 1μg of each primer and 1μl of the cDNA synthesized with the oligo-dT primer. The PCR products were viewed on a 3% agarose gel. These samples were then Southern blotted and hybridized with a radiolabeled TS-3 probe. For Southern blotting hybridization, 10ng of the oligonucleotide were labeled with 50μCi gamma-^{32}P ATP in the presence of polynucleotide kinase (NE Biolabs). After prehybridization of the filter (50% formamide, 1M NaCl, 1% SDS, 10% dextran sulfate) in sealable plastic bags at 50°C for 3 hours, the radiolabeled oligonucleotide was suspended in salmon sperm DNA (150μg/ml) and added to the prehybridization mix. Hybridization was carried out at 43°C overnight. The filter was rinsed for 1 min. in 2X SSC, and washed in TMAC (3M tetramethylammonium chloride, 2mM EDTA, 50mM Tris-HCl (pH 8.0)) at room temperature for 30 min at room temperature. It was then washed in TMAC, 0.2% SDS at 53°C for 10 minutes, and finally re-rinsed in TMAC for 10 min. at room temperature (137).

The ~150bp band formed by PCR of TS9;22 cDNA with the TS-1 and TS-RA primers was reamplified and cloned into the TA Cloning vector as described above. DNA was isolated from small scale cultures and sequenced by the dideoxy method to verify the specificity of the TS-1 and TS-RA primers for PAX-TS.

PAX-TS expression in bone marrow and in CML

RT-PCR and Southern blotting was additionally performed on 200ng of cDNA from two adult human bone marrow samples to assess expression in hematopoietic cells. Total RNA extracts from the patient with CML whose bone marrow in blast crisis was used to establish the TS9;22 cell line described above were also obtained. Both chronic phase and subsequent blast crisis total RNA were available for analysis. In order to assess expression of PAX-TS in this patient, 3.2μg of total RNA from the blast crisis (BC) and chronic phase (CP) were treated with RNase-free DNase and then used as a template for

first-strand cDNA synthesis with oligo-dT as described above for RT-PCR. The PCR conditions were 94°C for 1 min., 55°C for 1 min., and 72°C for 2 min. for 30 cycles, using 1μg of each primer, 1μl of DNased total RNA, and 1μl of the cDNA . The PCR products of the DNased total RNA and cDNA samples were viewed on a 3% agarose gel, Southern blotted and probed with radiolabeled TS-3 oligonucleotide as described above.

RT-PCR expression analysis of other tissues and organs

RT-PCR analysis was performed on mouse genomic DNA to identify the murine equivalent of PAX-TS, and on first-strand cDNA prepared from DNased-total RNA from an eleven day mouse embryo to assess expression in developing tissues. A comparison of the TS-1, TS-2, TS-RA, and TS-RB primer pairs for PCR recognition of PAX-TS in the human and mouse was made. Each of the four combinations of primers was used: TS-1 and TS-RA, TS-1 and TS-RB, TS-2 and TS-RA, and TS-2 and TS-RB. The PCR conditions were 94°C for 1 min., 62°C for 1.5 min., and 72°C for 2 min. for 30 cycles, using 1μg of each primer and 1μg and 1.3μg of the mouse and human genomic DNAs, respectively. The PCR products were viewed on a 3% agarose gel, Southern blotted and hybridized with radiolabeled TS-3 oligonucleotide as described above.

Organ lysates (lung, kidney, liver, spleen, colon, stomach, small bowel, and skeletal muscle) from an adult rat were obtained. Total RNA was extracted from 1ml of each lysate by the acid guanidine procedure as described above (135). The concentration of total RNA in each sample was estimated by OD_{260} and adjusted so that the amounts were approximately equal. Total RNA, DNased total RNA, and first-strand cDNA synthesized with oligo-dT primer were prepared. The PCR conditions were 94°C for 1 min., 55°C for 1 min., and 72°C for 2 min. for 30 cycles, using 1μg of each primer (TS-1 and TS-RA) and 1μl of the cDNA . The PCR products from the DNased total RNA and cDNA samples were then run on a 3% agarose gel, Southern blotted, and hybridized with the radiolabeled TS-3 probe as described above.

The ~150bp fragments from the PCR experiments on the mouse genomic DNA and

on the rat kidney cDNA were cloned into the TA Cloning vector as described above. DNA was isolated from small scale cultures and then sequenced by the dideoxy method to compare the human, rat, and mouse homologues of the PAX-TS homeobox.

PAX-TS screening in genomic and cDNA libraries

Murine cDNA (Clonetech) and human genomic DNA libraries were screened unsuccessfully by the standard method using both the radiolabeled TS-3 oligonucleotide and the cloned PAX-TS cDNA PCR fragment as probes. Possible explanations for this difficulty will be addressed below.

Isolation of the PAX-TS genomic clone

The four primers TS-1, TS-2, TS-RA, and TS-RB were submitted to Genome Systems, Inc. for genomic library screening by PCR (138, 139, 140). A ~70kb genomic clone packaged in bacteriophage P1 was isolated. The control P1 phagemid and P1 phagemid containing the genomic insert were isolated from 250ml cultures in LB + kanamycin 25µg/ml induced with IPTG. They were digested with BamHI, EcoRI, and HindIII, run on a 0.8% gel, Southern blotted and hybridized with radiolabeled TS-3 oligonucleotide as described above.

Subcloning of DNA fragments from P1 genomic clone

In order to subclone the 10kb EcoRI fragment which hybridized to the TS-3 probe, 10µg of the P1 clone was digested with EcoRI overnight in the presence of excess RNase. The digested DNA was phenol-extracted, ETOH-precipitated and ligated overnight into predigested Lambda Zap II/EcoRI/CIAP phage arms according to the manufacturers protocol (Stratagene). In this manner, we created a phage library containing ~20 different EcoRI inserts from the 100kb genomic DNA. The ligation was packaged into phage using the Gigapack II packaging extract exactly as described by the manufacturer. The phage was stored in 500µl SM buffer in the presence of 20µl chloroform.

The phage library was then screened by PCR using the TS-1 and TS-RA primers for detection of positive clones. The ligation efficiency was determined by ß-galactosidase

blue-white selection (~70% clear plaques) and the phage titer was measured by serial dilutions. Approximately 500 phage particles were plated onto ten 100mm diam. LB+10mM $MgSO_4$ plates (for a total of 5000 phage particles) in the presence of 200µl XL1-Blue MRF' cells (OD_{600}=0.5) (Stratagene) and 3ml top agar. The phage were eluted by adding 3ml SM buffer directly onto the plates and incubating for 2 hours at room temperature on a shaker. The eluates were transferred to eppendorf tubes, spun in the microfuge for 2 minutes to pellet any bacteria and agar, and then transferred to fresh eppendorf tubes for PCR analysis. The PCR conditions were 94°C for 1 min., 60°C for 1.5 min., and 72°C for 2 min. for 30 cycles, using 0.5µg of each primer and 1µl of the phage eluate. Eluates which produced "positive" PCR bands were saved.

In the next screening step, approximately 30 phage particles from a "positive" eluate were plated onto twelve 20mm diam. LB+10mM $MgSO_4$ plates (for a total of 500 phage particles), eluted, and analyzed by PCR as described above. The eluate which produced a "positive" PCR band was retitered and plated onto a single 100mm diam. LB+10mM $MgSO_4$ plate for a total of ~50 phage plaques using the ß-galactosidase blue-white selection to screen for phage containing inserts. Individual plaques were selected, eluted in 200µl SM buffer as described above, and analyzed by PCR. The eluate which resulted a "positive" PCR band represents a single clone containing the 10kb PAX-TS genomic insert.

The 10kb PAX-TS genomic insert was rescued by *in vivo* excision using the ExAssist /SOLR system (Stratagene) which places the fragment into the EcoRI site of the Bluescript plasmid. 16 individual colonies were picked. The plasmid DNA from these colonies was then digested with EcoRI to isolate the PAX-TS fragment.

Isolation of the PAX-TS cDNA clone

The murine cDNA library (Clonetech, Inc.) was screened similarly to the genomic library. The cDNA library itself was screened using the PCR assay described above with

primers TS-1 and TS-RA . This initial screening confirmed that this cDNA library contained the cDNA clone of interest. Screening of this library by standard filter hybridization methods produced no clones. Thus, $1x10^6$ phage clones were plated onto 20 plates (150 mm diam., 50,000 plaques per plate), and the bacteriophage eluted from the agar plate into 7 ml of SM phage buffer. 1µl of the phage eluate was used in a PCR reaction using the primers TS-1 and TS-RA in order to localize the plate which contained the desired cDNA clone. 50,000 plaques from this fraction were then re-screened by the PCR assay. By successive screening of smaller and smaller fractions, a positive clone was isolated.

After isolation of a single clone by PCR, a high titer stock of this phage clone was prepared as a liquid lysate and passed over a Qiagen column as per manufacturer's directions to prepare phage DNA. The DNA was digested with EcoRI to release the PAX-TS cDNA fragment, isolated from a 1% agarose gel by electroelution and passed over a Elutip-D column (Schleicher and Schuell) exactly as described by the manufacturer. The purified EcoRI fragment was ligated into an EcoRI-linearized Bluescript KS^- vector in the presence of T4 DNA ligase (NEB) and then transformed into competent E. coli. cells.

RESULTS

Identification of homeobox gene expression in cell lines

Figure 6 shows the 3% agarose gel of the initial PCR experiment in which the cDNAs derived from the ES, TS9;22, KG1 and MEL cell lines and DNased poly-A+-mRNAs were tested with the 2551 5'-class I homeobox primer. Compared to the 1kb ladder DNA size marker, there is a ~180bp band in the ES, TS9;22, KG, and MEL lanes using the $ß_2$-microglobulin primers, suggesting that the cDNA preparations were intact. There are also ~150bp bands in the ES, TS9;22, and KG1 treated with the HOX primers indicating that class I homeobox genes are expressed in these cell lines. The DNased poly-A+-mRNAs treated with the same primers failed to produce similar PCR products. Therefore, the DNase treatment of the mRNA successfully eliminated genomic DNA contamination from the reverse-transcriptase reaction substrates.

Having shown that the first-strand cDNAs described above were intact, the TS9;22 cell line was selected for RT-PCR analysis with the 19 divergent homeobox primers (Table 5). The 3% agarose gels containing the PCR screening are shown in Figure 7. A summary this data is contained in Table 6. "+" results were recorded for a 150bp band with the class I (2551) primer and for a 160 bp band with the divergent primers. The divergent primers recognize a specific region of the homeobox which is upstream from the 2551 primer site and therefore, they produce a slightly larger PCR band on the 3% agarose gel. "-" results include those primers which failed to produce a 160bp band or which produced large or multiple bands.

Homeobox sequences

The primers which yielded "+" PCR products from the TS9;22 cDNA included classI-HOX, Hox7.1/msh, and the xhox3/eve. These PCR products were cloned and sequenced and are shown in Figures 8 and 9. In Figure 10, the amino acid sequences of the three cloned homeoboxes are compared to the Antp. homeobox. The primer regions of each clone are in lower case letters since the actual gene may be degenerate within the

region of the primers and would not be represented in this PCR amplified product. The classI-HOX (2551) and msh primers identified the previously known genes, HOX B9 and MSX-2, respectively. Note that the classI-HOX-derived homeobox fragment from HOX B9 is shorter than the others since primer 2551 is designed to recognize the evolutionarily conserved sequence 'ELEKEF' of the class I genes which is 3' to the regions recognized by the Hox7.1/msh (3066) and xhox3/eve (3766) primers (84, 141, 142, 143, 144). Primer 2551 would not have recognized the msh- or xhox3/eve-derived homeobox fragments shown.

Clones derived from the PCR reaction with the xhox3/eve primer appeared to be a previously undescribed gene. Using the sequence of these clones, a FASTA search of EMBL + Genbank databases showed that it represented a new gene in humans that is almost identical to a very recently described gene, Cart-1, which was isolated from a rat chondrosarcoma cell line (145). The two genes demonstrated 88% identity in a 138bp overlap and 100% identity in a 35 amino acid overlap (Figure 9). This gene also bears significant homology to the paired-type homeobox genes, such as Pax-3 which has 72% identity in a 113bp overlap and 67% identity in a 36 amino acid overlap (Figure 9). Based on this initial sequence information, and since the sequences 5' or 3' to the homeobox are currently under investigation, this new gene was preliminarily named "PAX-TS."

Expression analysis

In order to prove PAX-TS expression in the TS9;22 cell line, the RNase protection assay was performed (Figures 11, 12). In Lane 1, the gel-purified, 290bp, radiolabeled run-off transcript is shown. It comprises 116bp of the TA cloning vector plus the 174bp PCR insert which contains 66bp of the two primers and 111bp of actual coding sequence. The positive control, Lane 2, shows the RNase protection product of the radiolabeled probe and the cold, 296bp, opposite strand transcript. The overlap between the two transcripts is 260bp, but the band in Lane 2 is only 215bp. This apparent discrepancy is most likely due to the fact that single stranded RNA products move faster than the denatured double

stranded DNA fragments of the pBR322/MspI marker on the polyacrylamide gel. The negative control, Lane 3, has no RNase protection product since sterile water was used instead of mRNA template and the transcript was completely digested. Lanes 4, 5, 6 7, and 8 show the RNase protection products of the radiolabeled transcript against TS9;22, HL60, Meg01, KG1, and MEL total RNA. There are two bands, 123bp and 130bp in the TS9;22 lane. It is important to emphasize that the PAX-TS clone contains sequences which are not likely to be represented in the actual gene. These sequences correspond to the two degenerate primers at the ends of the clone (especially since one primer (3766) was designed to detect the Evx gene family). Thus, the actual bands are smaller than the 174bp insert. This fact may also explain why two degradation products appear on the gel, perhaps representing some secondary structure formation which interferes with digestion in a fraction of the PAX-TS-antisense probe population or a mismatch near the end of the insert. In any case, these results suggest that PAX-TS is expressed in the TS9;22 cell line at readily detectable levels, and not in any of the other cell lines studied.

Having proven PAX-TS expression by the RNase-protection assay, RT-PCR was used to facilitate further expression studies. Figure 14 shows the RT-PCR results of screening the TS9;22 and Meg01 total RNA, DNased total RNA, and first-strand cDNA with the PAX-TS-specific primers, TS-1 and TS-RA. Each sample is presented in a group of three after PCR with primers TS-1 and TS-RA, and with primers specific for the DLX and $ß_2$-microglobulin ($ß_2$) genes. A human genomic DNA template was also used for comparison. The larger, minor bands in the PAX, DLX and $ß_2$ lanes are probably artifacts due to the high concentration of DNA in the reaction mix which leads to some non-specific binding. In lanes 5-13, the total RNA, DNased total RNA, and cDNA from TS9;22 were used as templates. In lanes 5-7, the PCR products show the same banding pattern as with the human genomic DNA, reflecting the genomic DNA contamination of the total RNA. After DNase treatment in lanes 8-10, these bands do not reappear. After reverse transcription in lanes 11-13, the 104bp PAX-TS product reappears and a shorter, single

PCR band appears in the DLX and $ß_2$ lanes, proving that cDNA is intact. The results generated in Meg01 (Lanes 16-24), as well as in the HL60, KG1, MEL, and ES cell lines, were similar, except that the 104bp PAX-TS PCR product was not produced with these cDNA templates.

Southern blot data from the RT-PCR experiments are presented in Figures 15. Only the human-derived cell lines (TS, Meg01, HL60 and KG1) were tested. The $ß_2$ probe recognized its gene product in all four cell lines, but the radiolabeled TS-3 probe only hybridized to the PCR product in the TS9;22 and human genomic DNA lanes. The TS-3 probe did not produce any signal in the DNased total RNA PCR lanes (data not shown). Dideoxy sequencing of the TS-1/TS-RA PCR product ligated into the TA Cloning vector demonstrated that it contains the identical sequence as was cloned using the xhox3/eve degenerate primer (data not shown).

In summary, the results of the above RT-PCR, Southern blot and sequencing data suggest that TS-1 and TS-RA are specific for PAX-TS and may be used to screen for its expression by RT-PCR, and that PAX-TS is expressed preferentially in TS9;22 which is in concurrence with the RNase protection data. Since the Southern blot data agrees with the RNase-protection data, RT-PCR is a suitable instrument for evaluating PAX-TS expression.

PAX-TS expression in the normal bone marrow

The 3% agarose gel and Southern blot data from RT-PCR of two normal human bone marrow RNA samples is shown in Figure 16. On the gel, in lanes 2 and 3, the faint 104bp bands corresponding to PAX-TS can be seen in the cDNA from the two bone marrow samples. Directly below, the $ß_2$-microglobulin controls for cDNA are shown. The corresponding Southern blot shows two faint bands corresponding to PAX-TS in both bone marrow samples, suggesting a very low level of PAX-TS expression in normal bone marrow.

PAX-TS expression in a patient in CML blast crisis

Total RNA samples from the CML patient from which the TS9;22 cell line was derived were obtained. Both the chronic phase and blast crisis RNA samples were subjected to RT-PCR using the PAX-TS-specific primers. The 3% agarose gel and Southern blot data are presented in Figure 17. PAX-TS was only found to be expressed in the TS9;22 cDNA.

PAX-TS in embryogenesis

Using the TS-1 and TS-RA oligonucleotide primers and RT-PCR, we are able to demonstrate PAX-TS expression in eleven-day mouse embryo RNA (Figure 18). This result suggests that PAX-TS does exist in the mouse and that it is expressed during embryogenesis.

Murine homologue of PAX-TS

We then attempted to identify the murine equivalent of PAX-TS by PCR. The agarose gel and subsequent Southern blot data for the PCR on human and mouse genomic DNA with the four combinations of primers (TS-1 and TS-RA, TS-1 and TS-RB, TS-2 and TS-RA, TS-2 and TS-RB) are shown in Figure 19. As would be expected, the size of the PCR products decreases as the primer pairs move closer together. Note that in the mouse lanes (6-9), the PCR product vanishes as the TS-2 and TS-RB primers are used. On the Southern blot, only the human PAX-TS PCR products hybridize to TS-3 under the wash conditions used. These data suggest that while TS-1, TS-2, TS-RA, and TS-RB recognize and are specific for the human PAX-TS, the murine homologue has some degree of degeneracy in that it does not hybridize as well to TS-2, TS-RB, and TS-3 under the PCR and Southern blot conditions described.

A comparison of the human and mouse PAX-TS homeobox sequence data is presented in Figure 20. In the region of the homeobox between the TS-1 and the TS-RA primers, there is 85.9% DNA homology but only a single amino acid difference.

PAX-TS expression in rat organs

The RT-PCR data on the rat organ lysates is shown in Figure 21. Of the organ systems

tested, PAX-TS expression was localized to the kidney (lane 2). The Southern blot made from a similar gel, with the DNased total RNA samples as controls to rule out genomic DNA contamination and with TS9;22 and Meg01 cDNA as positive and negative expression controls, demonstrates a positive signal only in the positive control lane, suggesting that the oligonucleotide probe does not recognize the rat homologue in the region of the TS-3 primer under the conditions used. However, direct cloning and sequencing of this PCR product confirmed its identity as PAX-TS.

A comparison of the human and rat PAX-TS homeodomain sequence data is presented in Figure 20. In the region of the homeobox between the TS-1 and TS-RA primers, there is 85.9% DNA homology but 100% amino acid conservation.

Stringency of the TMAC wash conditions

It is important to note that, in the region of the radiolabeled primer, GATGTGTATGTCAGAGAACA, there is only a single base change in both the rat and mouse which interferes with its ability to hybridize to TS-3 under the conditions we used for the Southern blot wash with TMAC. The fact that TMAC is highly sequence specific and can distinguish single base differences explains why the rat and mouse homologues of PAX-TS were not identified in the Southern blot data (137).

PAX-TS genomic clone

We were unsuccessful in obtaining a PAX-TS genomic clone by the standard method of filter hybridization. Thus, a P1 phagemid clone carrying ~70kb of genomic DNA around PAX-TS was obtained from Genome Systems, Inc., who used the PAX-TS-specific primers to isolate this clone. The HindIII, BamHI, and EcoRI digests of the 70kb genomic DNA insert in the P1 plasmid are shown in Figure 22. The uncut, supercoiled plasmid is in lane 6. Southern blot hybridization with the PAX-TS-specific TS-3 probe demonstrates ~8kb HindIII, ~5kb BamHI, and ~9kb EcoRI fragments in lanes 7-9 (Figure 23).

Several strategies were employed to subclone the P1 clone. Random cloning with

colony hybridization and direct cloning using gel-purified DNA fragments did not yield any clones containing PAX-TS. Random subcloning into a phage vector created a "mini-library" which was successfully screened by PCR, using the PAX-TS-specific primers.

Figure 24 shows the secondary PCR screening step in which the number of phage particles containing the PAX-TS genomic insert was reduced from 1 in 500 to 1 in 50. A similar screening step isolated a single phage particle containing PAX-TS from a plate of 50 plaques. However the rescued Bluescript plasmid did not contain the correct insert. The insert in every clone was only 3kb and it was negative for PAX-TS by PCR analysis. Sequencing the ends of this insert from the T3 and T7 promoter binding sites in Bluescript did not yield any recognizable genes after a EMBL + Genbank databank search (data not shown).

PAX-TS cDNA clone

A phage clone containing PAX-TS sequences was isolated from a mouse day 11 embryo library using a PCR screening technique as described in Methods. This clone contained a 1.2kb EcoRI fragment, as confirmed by the Southern blot (Figure ; this Figure is currently in preparation), showing that a band of this size hybridized to the radiolabeled PAX-TS probe.

DISCUSSION

As a family of transcription factors, the homeobox genes have been implicated in the regulation of cell differentiation in both embryogenesis and adult tissues. In addition, it has been shown that dysregulated homeobox gene expression may function in the development of both solid tumors and hematologic malignancies. Studies on the homeobox genes and other transcription factors indicate that the molecular events occurring during embryogenesis and carcinogenesis may share many common features. Hematopoiesis provides an excellent model for studying normal and abnormal cell differentiation.

Since very little is known about hematopoiesis at the level of transcription, and since both class I and divergent homeobox genes are believed to contribute to this process, my aim was to identify new divergent homeobox genes in hematopoietic cells. My long-term goal was to show that these genes were essential to the normal development of hematopoietic cells or were important in leukemogenesis.

The use of degenerate primers to identify novel divergent homeobox genes

An RT-PCR strategy using degenerate oligonucleotide primers designed to detect new members of the known divergent homeobox gene families has been successful in identifying the expression of several genes in hematopoietic cells including MSX-2, DLX-1, CDX, and GBX (105, 107, 111). This strategy has proven to be simpler than former methods of cDNA library screening which required the use of radiolabeled probes to conserved sequences under low stringency wash conditions.

I used this RT-PCR strategy to identify a new homeobox gene in a megakaryocytic leukemia cell line, TS9;22. In addition, the MSX-2 and HOX B9 homeoboxes were also cloned from this cell line. The new gene, named PAX-TS was cloned using a PCR primer designed to detect the evx family of divergent genes which contains the amino acid sequence "RRYRTAFT" at the 5' end of the homeobox. However, for the PCR screening of TS9;22, a low annealing temperature of 48°C was used in order to decrease the

specificity of the primer's binding. This strategy allowed for the detection of related, but slightly different sequences. The sequence of the PAX-TS homeobox is only 30% homologous to the mouse Evx-1 gene so it will be of interest to analyze the sequence of the entire PAX-TS gene from either a genomic or cDNA library to determine if the actual gene contains the evx-like 5' region, although this is unlikely.

A. ANALYSIS OF THE PAX-TS GENE

PAX-TS has a "paired-like" homeodomain

DNA and amino acid sequence homology analysis suggest that PAX-TS is a member of the "paired-like" homeobox genes (Figure 9). The "paired-like" grouping includes the six of the eight members of the Pax family which, as discussed earlier in the Introduction (page 15), contains the highly conserved "paired" amino acid DNA-binding motif upstream of a complete or partial homeodomain. Pax-1 and Pax-5 do not contain a homeodomain (88). Additional members of the "paired-like" group include S8, ceh-10, otd, gsbBSH, and smox-3 (isolated from Schistosoma mansoni) which retain between 50 to 75% homology to the cloned PAX-TS homeobox (146, 147, 148, 149, 150). These genes do not contain paired domains and are considered distinct from the Pax family.

Based on the existence of these three types of genes: paired domain without a homeodomain, paired domain with a homeodomain, and homeodomain without a paired domain, it has been suggested that the Pax family arose via chromosomal crossovers between ancestral paired and homeobox genes with subsequent divergence and duplication (88). Therefore, the Pax genes without homeodomains (Pax-1 and Pax-5) and the "paired-like" genes may represent more ancient evolutionary relatives of the current Pax genes. PAX-TS may fit into this classification.

Very recently, during the course of my studies, Zhao et al. published the sequence of a homeobox gene, Cart-1, from a rat chondrosarcoma cDNA library. At the nucleotide level, this gene is identical to the rat homologue of the PAX-TS gene over a stretch of about 100 base pairs (145) (Figure 25). I have isolated human genomic and cDNA PAX-TS

clones and restriction mapping and sequence data is pending at this time. It is currently not known whether Cart-1 and PAX-TS are the same gene, although the identity between the rat PAX-TS and Cart-1 homeoboxes suggests that they are. However, even within a single species, many homeobox genes are virtually identical at the amino acid or nucleotide level within the homeodomain but differ in the remainder of the gene. The peptide sequences of the paralogous HOX A5 and HOXB5 genes and the divergent *Xenopus* msx-1 and msx-2 genes are identical within the homeobox (151). Until the sequence upstream of the PAX-TS homeobox is obtained, I cannot exclude the possibility that the gene varies from Cart-1, and defines a new gene which may contain a paired domain.

The Cart-1 cDNA clone includes a putative 326 amino acid coding region with a 460bp 5' untranslated region. Only a single (1.3kb) mRNA transcript for Cart-1 was detected, and in analyzing overlapping cDNAs, no evidence for alternatively-spliced transcripts was demonstrated. However, this does not exclude the possibility that an alternative transcript which includes a paired domain does exist in some cell types or in certain diseases.

Interestingly, the sequence of Cart-1 in the region of the 5' PCR evx primer is "RRHRTTFT." A comparison of the sequences at the nucleotide level demonstrates only two mismatches out of twenty. This motif is similar enough to the evx primer so that it would be detected under the "low-stringency" PCR conditions used to screen TS9;22. It is also quite similar to the 5' amino acid region shared by several genes including Pax-3 and Pax-7 which contain "RRSRTTFT" and smox3 which contains "RRIRTTFT" (Figure 9). If PAX-TS does, in fact, contain the RRHRTTFT 5' motif, it would have to be classified as a closer relative to the "paired-like" smox3 than to the Pax genes.

Human-rat-mouse comparisons

Figure 20 shows the DNA and amino acid sequence comparisons of the human, mouse, and rat homologues of PAX-TS. Except for a single amino acid change in the mouse, the peptide sequences of PAX-TS in the three species are identical. The DNA

variability is located in the "wobble" positions of each codon such that the amino acid specified is the same. This data supports the notion that homeobox gene sequences are highly conserved between distinct evolutionary species. This identity will also be helpful for future studies since it will be easier to draw conclusions about the behavior of PAX-TS in humans from experiments involving expression in the mouse.

Cloning of the PAX-TS human genomic clone

It was not possible to isolate a genomic PAX-TS clone using the standard method of screening a genomic phage library with the PAX-TS cDNA insert (174bp). Two explanations for the failure of this method are a potentially low specific-activity of the labeled, linearized plasmid probe and the small length of the probe itself. Under the wash conditions used, this probe may not have been sufficient to provide a strong enough signal. Thus, I obtained a P1 phage clone carrying a 70kb insert which contained the PAX-TS gene. Direct subcloning into the Bluescript plasmid was attempted, using Geneclean II to gel purify the 10kb EcoRI P1-fragment in Figure 23. Unfortunately, owing to the large size of the P1 phagemid, it was difficult to isolate enough insert from the EcoRI digest to facilitate a ligation. In addition, the large size of the insert (between 5 and 10kb) may have reduced the efficiency of the ligation reaction into the Bluescript vector.

As a result, I resorted to constructing a phage "mini-library" from the EcoRI digest of the P1 clone, and utilized a technique of screening successive phage titers by PCR with the PAX-TS-specific primers, TS-1 and TS-RA, as described in Methods. The screening method was essentially similar to the method recently described in (152). Unfortunately, even though this method appeared to successfully isolate a single phage containing the PAX-TS 10kb insert, none of the rescued Bluescript plasmids contained the proper insert. It appears that the PAX-TS insert is stable in phage but is rearranged during transfer to Bluescript during the ExAssist/SOLR rescue resulting in loss of the PAX-TS gene. It is possible that some repetitive sequence in the chromosomal DNA flanks the PAX-TS homeobox. Given the large size of the PAX-TS genomic insert, the likelihood of a

crossover event would be increased. Alternatively, given the low copy number of P1 plasmids compared to the high copy number of Bluescript, the rescued Bluescript clones may be unstable (138, 139) .

Since the PAX-TS clone appears to be stable in the lambda phage, I now plan to amplify the single PAX-TS-containing phage in culture, isolate the phage DNA, and digest the 10kb insert into smaller fragments which will then be subcloned into plasmids and characterized. This strategy may eliminate the possibility of DNA rearrangement.

Cloning of the PAX-TS cDNA clone

A 1.2kb EcoRI fragment was isolated after screening the mouse cDNA library by PCR. Southern blot data suggests that this fragment contains the PAX-TS sequence. Restriction mapping and sequence data are pending at this time. This data will confirm whether PAX-TS is the same gene as Cart-1 or represents a new gene whose homeobox is homologous to that of Cart-1.

B. EXPRESSION STUDIES

Expression pattern in leukemia cells and in normal bone marrow cells

The RNase-protection data demonstrates that PAX-TS is expressed at measurable levels in TS9;22 but not in any of the other cell lines assayed, including those with predominantly erythroid (MEL), myeloid (HL60, KG-1), and megakaryoblastic (Meg01) characteristics. From this limited selection of cell lines it is not possible to hypothesize as to whether PAX-TS displays lineage-restricted expression.

If PAX-TS is identical to Cart-1, it is possible to explain the sizes of the RNase-protection products. It was postulated that, since the PAX-TS homeobox clone was generated with PCR using degenerate primers, the actual gene may differ in the region of the two primers. Comparison of the PAX-TS and Cart-1 sequences demonstrates that they do differ at both primer sites. The evx primer sequence is "RRYRTAFT" whereas Cart-1 contains "RRHRTTFT." The 3' primer sequence is "IWFQNRR" whereas Cart-1 contains "VWFQNRR." Therefore, the RNase-protection assay would have eliminated up

to an additional 18 nucleotides (6 amino acids) from the 5' end and 21 nucleotides (7 amino acids) from the 3' end. Since the PAX-TS homeobox measures 174bp, subtracting an additional 39 nucleotides during the RNase digestion phase would yield a 135 nucleotide product which is in general agreement with the autoradiograph data.

It was also shown that PAX-TS is expressed at very low levels in two normal bone marrow samples by RT-PCR (Figure 16). Zhao et al. did not test for Cart-1 expression in normal bone marrow. Therefore, unless PAX-TS is expressed ectopically in TS9;22, one can presume that it is expressed in some fraction of the hematopoietic cells of the bone marrow. As a caveat, it must be noted that Northern blot data is being compared to RT-PCR data which have both different sensitivities and false-positive rates. It is also possible that there is a small degree of "chondrocytic contamination" in the two normal bone marrow samples (considering that the biopsy needle must pass through bone) which would be detectable by PCR and not by Northern blotting. However, it is highly unlikely that there is any similar contamination in the TS9;22 cell line.

Expression pattern in the CML patient whose bone marrow was the source TS9;22

Total RNA samples from the patient whose bone marrow was the source of the TS9;22 cell line during the chronic phase and megakaryocytic blast crisis phase of the disease were obtained for RT-PCR analysis. PAX-TS expression was not detected in either sample even though the gene is expressed in the cell line derived from the blast crisis bone marrow (Figure 17). Two explanations may account for this finding. First, since we have already demonstrated the low level of expression in normal bone marrow, it is probable that PAX-TS is only expressed in a small subset of the hematopoietic cells. This subset may be reduced even further in the leukemic bone marrow. A comparison of the immunophenotype characteristics of the CML patient and the TS9;22 cell line shows that the relative numbers of each cell type are different (Table 3). The population of cells in TS9;22 are distinctly less lymphoid and more myeloid, monocytoid, and megakaryocytic than the patient's leukemia cells two years earlier. The fact that pelleted TS9;22 cells

appear pink-brown suggests hemoglobinization and some erythroid features as well (personal observation). The cell type which normally expresses PAX-TS may have been selected during establishment of the TS9;22 cell line. Second, PAX-TS may have been ectopically activated over baseline bone marrow levels during establishment of the cell line. It has been noted that mutations in the tumor suppressor gene, p53, are more frequent in myeloid leukemia cell lines than they are in patients with AML. This finding suggests that, since cells with abnormal growth characteristics are selected, alterations in oncogene expression may be a common feature of cell line establishment (153). Both of these explanations could account for the observed level of PAX-TS expression in TS9;22 relative to the low expression levels measured in the two bone marrow samples.

Expression patterns of PAX-TS in rat organs

RT-PCR expression data indicates that PAX-TS is expressed in the rat kidney (Figure 21). The gene was not expressed in any other organs although we did not test the heart, bone or central nervous system. This data is in disagreement with that of Zhao et al. who only found PAX-TS expression in chondrocytic cells and not in the kidney. It was mentioned above that "chondrocytic contamination" of the bone marrow aspirates may have occurred. However, for the kidney, this is highly unlikely. The appropriate controls were used to eliminate the possibility of genomic DNA contamination of the cDNA samples. Since these studies utilized RT-PCR while Zhao et al. used Northern blotting, it is possible that baseline PAX-TS expression in the kidney is too low for detection by Northern blotting, given the different sensitivities and specificities of the two methods. Therefore, I have concluded that PAX-TS does display some degree of renal expression.

Expression of PAX-TS in the day 11 mouse embryo

PAX-TS is expressed during embryogenesis. RT-PCR of total RNA from a day 11 mouse embryo demonstrates PAX-TS expression (Figure 18). If PAX-TS is identical to Cart-1, this data is in agreement with Zhao et al. who demonstrate Cart-1 expression in the developing rat embryo by Northern blotting.

C. IMPLICATIONS OF PAX-TS EXPRESSION DATA

The current data indicates that the homeobox gene PAX-TS has a pattern of expression consistent with embryogenesis, hematopoiesis, and functional adult cells. Although PAX-TS may be identical to the chondrocytic Cart-1 gene, we have localized its expression additionally to include the kidney, bone marrow and a megakaryocytic leukemia cell line. If it does contain a paired domain, and becomes the ninth member of the Pax family, it will be of interest to explain its expression relative to the rest of the Pax genes.

As mentioned above, the Pax family currently contains eight members (Pax 1-8). Human counterparts have only been named for Pax-1 (HuP48), Pax-3 (HuP2), Pax-6 (AN), and Pax-7 (HuP1). Of the eight genes, natural mutations in Pax-1, Pax-3, Pax-6, and Pax-8 have been found to result in gross phenotypic anomalies in mice. Of these, the Pax-3 (splotch) and the Pax-6 (small eye) have human correlates (88).

Waardenburg syndrome is inherited in an autosomal dominant fashion and is characterized by pigmentary disturbances, lateral displacement of the inner canthus of the eyes, occasional deafness, and mental retardation. The disorder has been mapped to chromosome 2 in the region of HuP2 and several affected families have been shown to carry mutations in this gene, which is homologous to Pax-3 (154, 155, 156). Aniridia is a congenital, bilateral panocular disorder characterized by complete or partial absence of the iris resulting in impaired vision. It may be associated with anomalies such as foveal and optic nerve hypoplasia and may deteriorate over time and result in cataracts, glaucoma, and corneal opacification. The incidence is rare (between 1/64000 and 1/96000) and it is believed to be inherited as an autosomal dominant trait (157, 158). The cDNA for Aniridia (AN) has been cloned and is homologous to Pax-6 which results in a similar phenotype in mice (159).

It is surprising that four known mouse disorders (and two human syndromes) result from mutations in a family of genes which only comprises eight members. In fact there are no pre-existing mutants for the entire Hox gene family. Two hypotheses have

been offered to explain this phenomenon. First, since the Pax genes are quite large, the chance of mutation is increased relative to the Hox genes. Second, since the known mouse mutations have been shown to exhibit a dominant effect, it has been suggested that Pax genes are required in minimum concentrations for phenotypic expression. The loss of one allele could bring the concentration of the protein in the cell to below the critical threshold, making the mutants easily recognizable (88). This critical concentration level does not appear to be required for the Hox genes and, therefore, the loss of a single allele would not result in a phenotypic change (66, 68). It will be of interest to determine if the PAX-TS gene is involved in human disease.

The oncogenic potential of the Pax genes has also been examined. Maulbecker and Gruss were able to induce tumor formation in mice by overexpression of the Pax-1,2, 3, and 8 genes. Since Pax-2 and Pax-8 contain incomplete homeodomains, they localized the transforming ability to the paired domain. In support of this hypothesis, they showed that a mutated paired domain in Pax-1 interferes with DNA binding and results in defective tumor formation (160).

In support of these empiric studies, it has also been shown that Pax-3 is consistently involved in the t(2;13)(q35;q14) translocation found in the pediatric solid tumor, alveolar rhabdomyosarcoma (161). In this case, the chromosomal rearrangement fuses the 5' portion of Pax-3 with the 3' portion of the ALV gene, a member of the "forkhead" family of transcription factors. While the fusion protein contains an intact paired domain and homeodomain, the ALV DNA-binding domain is disrupted. Similar to the mechanism proposed for the E2A-PBX1 fusion gene, it is hypothesized that ALV donates an ectopic transcriptional activator to Pax-3 which incorrectly targets its action and enhances its ability to promote cell transformation (162). The expression of PAX-TS in a leukemia cell line raises the possibility that it may be involved in leukemogenesis even though it was not detected in the RNA samples obtained from the TS9;22 patient.

Given its apparent expression in the rat kidney, it is important to note that PAX-TS

may, in fact, belong to the Pax family. Both Pax-2 and Pax-8 have been found to be expressed in the embryonic kidney. Pax-2 is expressed in the ureter epithelium, the condensing mesenchyme cells and their early epithelial derivatives. As differentiation occurs, the levels of Pax-2 are down-regulated. It was also shown that Pax-2 was expressed in elevated levels in the undifferentiated renal epithelial cells of human Wilms' tumor, suggesting a correlation between kidney development and down-regulation of Pax-2 (163). Overexpression of Pax-2 in transgenic mice resulted in a histologically abnormal and dysfunctional renal epithelium which behaved like the congenital nephrotic syndrome. This data supported earlier studies in that while early development of the kidney proceeded normally, the epithelial maturation phase which requires down-regulation of Pax-2 was abnormal (164). Pax-8 has a similar expression pattern in the developing mouse kidney, but has also been identified in the adult kidney by Northern blotting (165). Pax-2 and Pax-8 are related in that they are the two members of the Pax family that contain truncated homeodomains which suggests a common evolutionary origin. Since PAX-TS contains a complete homeobox, it is not as closely related to Pax-2 and Pax-8; yet it appears to be expressed in the adult kidney.

Speculations on the potential role of the PAX-TS gene in hematopoiesis

From the immunophenotype profile, it is clear that the TS9;22 cell line contains a heterogeneous population of hematopoietic cells (Table 3). The large percentage (43.7%) of cells which express the platelet/megakaryocyte-specific antigen, CD41, classify TS9;22 as a megakaryocytic cell line. Currently, we are unable to conclude definitively that the expression of PAX-TS bears any relationship to the process of megakaryocytopoiesis. In fact, very little is known about the unique process of megakaryocyte differentiation, from progenitor cell to polyploidization, at the molecular level. We were not able to demonstrate PAX-TS expression in another megakaryocytic cell line, Meg01. While both cell lines have megakaryocytic characteristics, it is impossible to compare them as representatives of hematopoiesis in general or megakaryocytopoiesis in particular. Rather, they serve as a

guide to future study.

D. FUTURE DIRECTIONS OF STUDY

The isolation and characterization of the cDNA and human genomic clones of PAX-TS will be important for several reasons. The genomic and cDNA sequences will resolve the identity of the human PAX-TS gene, and will determine whether it is a Pax gene or whether it is identical to the recently described Cart-1 homeobox gene. Then, the cDNA or genomic clone will be used to map the human chromosomal location of the gene. As with the translocation involving Pax-3 in rhabdomyosarcoma, it would be very interesting if PAX-TS was associated with the Philadelphia chromosome in TS9;22.

Up to now, the lack of a cDNA or genomic clone has made PAX-TS expression analysis very difficult. My immediate goal is to obtain one of these clones which would facilitate its use as a probe for Northern blot analysis. In this way, the PAX-TS expression analysis in TS9;22, the normal bone marrow, and the kidney would be confirmed. Consideration could then be extended to include many different hematopoietic cell lines and RNA samples from other leukemic patients. In this way, the expression pattern of PAX-TS could be more accurately defined in the hematopoietic compartment.

In order to precisely determine the hematopoietic source of PAX-TS expression it will be necessary to use purified normal bone marrow cells of restricted lineages. In this way, one could use the colony assay to screen for PAX-TS expression in single-lineage hematopoietic precursors. Once the "window" of expression has been narrowed, future studies should include antisense "knock out" assays to assess the effects of downregulation on hematopoiesis. Conversely, overexpression studies will provide additional data on its hematopoietic effects and may even suggest a role for PAX-TS as a proto-oncogene. Ultimately, this data will be used in order to perform both *in situ* hybridization studies on developing embryos and experiments in transgenic mice.

E. CONCLUSION

I have succeeded in cloning a novel homeodomain-containing gene named PAX-TS

from the human megakaryocytic TS9;22 cell line. The rat and mouse homologues have also been cloned and sequenced. Based on RNase-protection and RT-PCR data, I conclude that PAX-TS is expressed in TS9;22 but not in any of the other cell lines assayed, including those with predominantly erythroid (MEL), myeloid (HL60, KG-1), and megakaryoblastic (Meg01) characteristics. RT-PCR data also demonstrates expression in normal bone marrow, an eleven day mouse embryo, and in the rat kidney, suggesting that, in addition to functioning in some hematopoietic cells, PAX-TS is expressed both during embryogenesis and in non-hematopoietic, adult tissues. These and future studies will help to elucidate how homeobox genes participate in the control of cell differentiation.

TABLES AND FIGURES

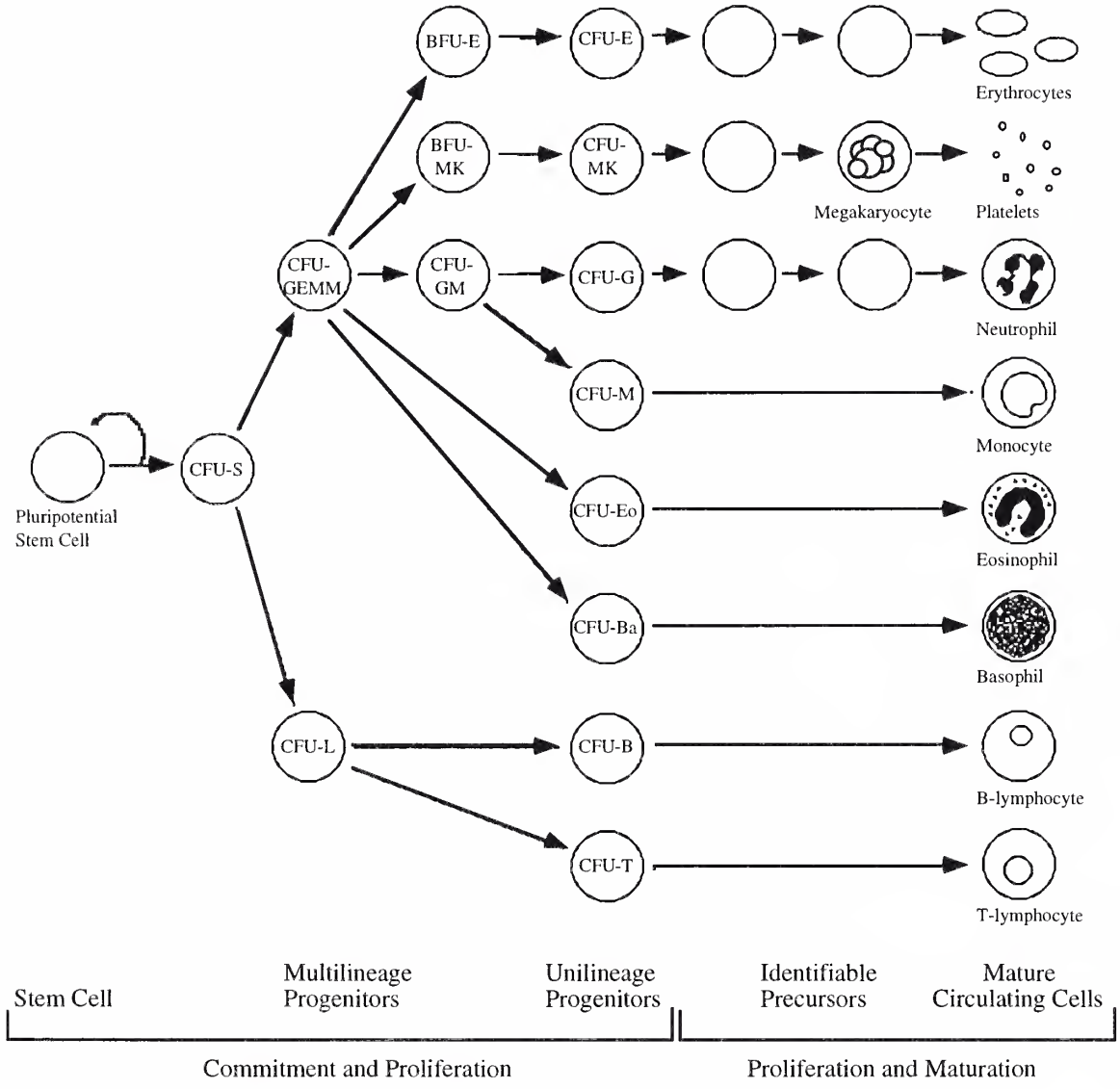


Figure 1: Stem Cell Model of Hematopoiesis

Quiescent pluripotential stem cells periodically undergo a balanced, asymmetric mitosis. In this process, one daughter cell returns to the resting state, while the other becomes committed to further division and differentiation. With each subsequent division, the resulting progenitor cells become progessively restricted to fewer and fewer hematopoietic lineages. Finally, restricted progenitor cells divide to produce mature blood cells of each of the lineages (2).

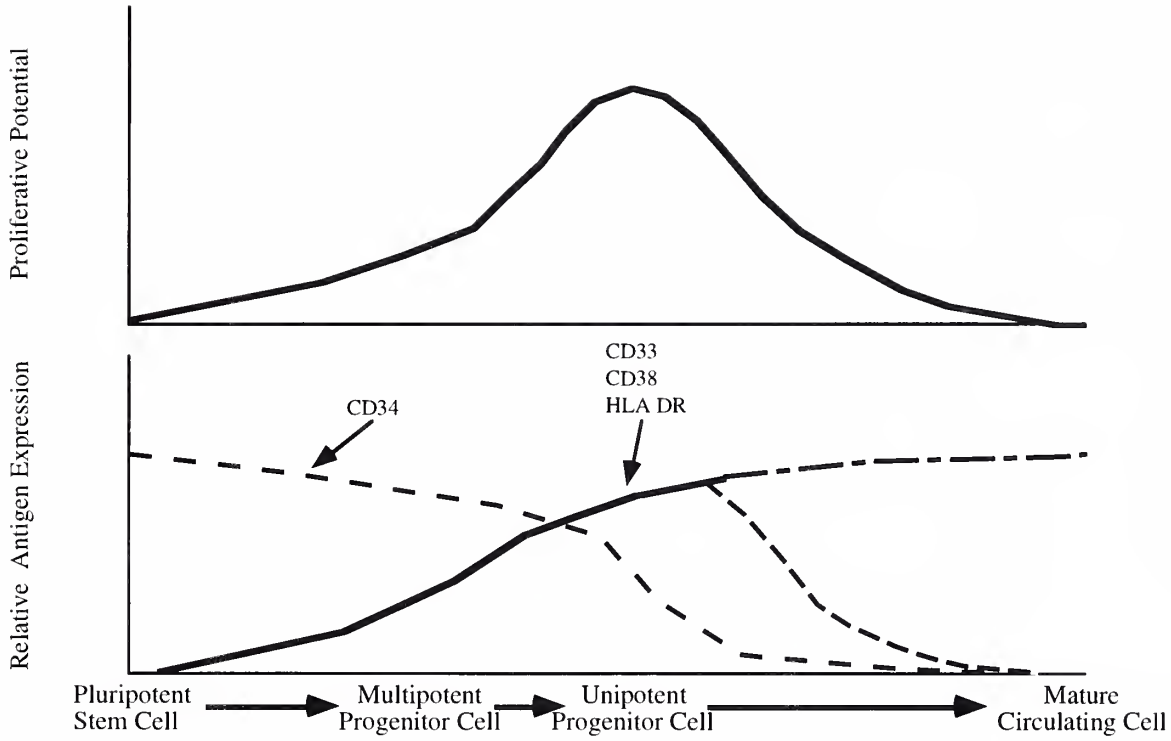


Figure 2: Dynamics of hematopoiesis

Throughout hematopoiesis, the characteristics of the differentiating cell population changes with regard to proliferative ability and surface antigen expression. The most primitive pluripotent cells are CD34+/CD33-/CD38-/HLA-DR- and have the capacity for self-renewal but proliferate at a slow rate. With the onset of hematopoiesis, these stem cells acquire the CD33, CD38, and HLA-DR antigens and proliferate in response to cytokines. This simplified view of hematopoiesis does not include the many other CD antigens whose expression are restricted to specific lineages. Over the course of terminal maturation, the capacity for proliferation declines.

Hematopoietic Growth Factors

Growth Factor	Molecular Weight (Kd)	Chromosome Location	Source	Major Target Cells	
				Progenitor	Mature
Erythropoietin	39	7	Kidney	CFU-E	Nucleated Erythroblasts
GM-CSF	18-30	5	T-cells, monocytes, Endothelial cells, Fibroblasts	CFU-GM, -Eo CFU-GEMM BFU-E CFU-Meg	Monocytes Neutrophils Eosinophils
IL-3	15-25	5	T-cells	CFU-GEMM BFU-E CFU-Eo, -GM	Monocytes Eosinophils
G-CSF	20	17	Monocytes, Fibroblasts Endothelial cells	CFU-G, -GM (CFU-GEMM, -GM -M, at high conc.)	Neutrophils
M-CSF	70-90 (dimer)	5	Monocytes, Placenta, Endothelial cells, (human urine)	CFU-M	Monocytes
IL-1	22	2	Monocytes, Leukocytes		Monocytes Neutrophils Endothelial cells Fibroblasts
IL-4	16-20	5	T-cells		
IL-5	46 (dimer)	5	T-cells	CFU-Eo	Eosinophils
IL-6	19-21	7	Fibroblasts, Leukocytes, Epithelial cells	cofactor for blast colonies	
IL-7			Leukocytes		B-cells, Megakaryocytes
IL-8			Leukocytes		Granulocytes
IL-9			Lymphocytes		Megakaryocytes

Table 1: Hematopoietic growth factors

Note that, with the exception of G-CSF, the CSFs are not located on any of the chromosomes associated with the HOX loci (adapted from (1) and (15)).

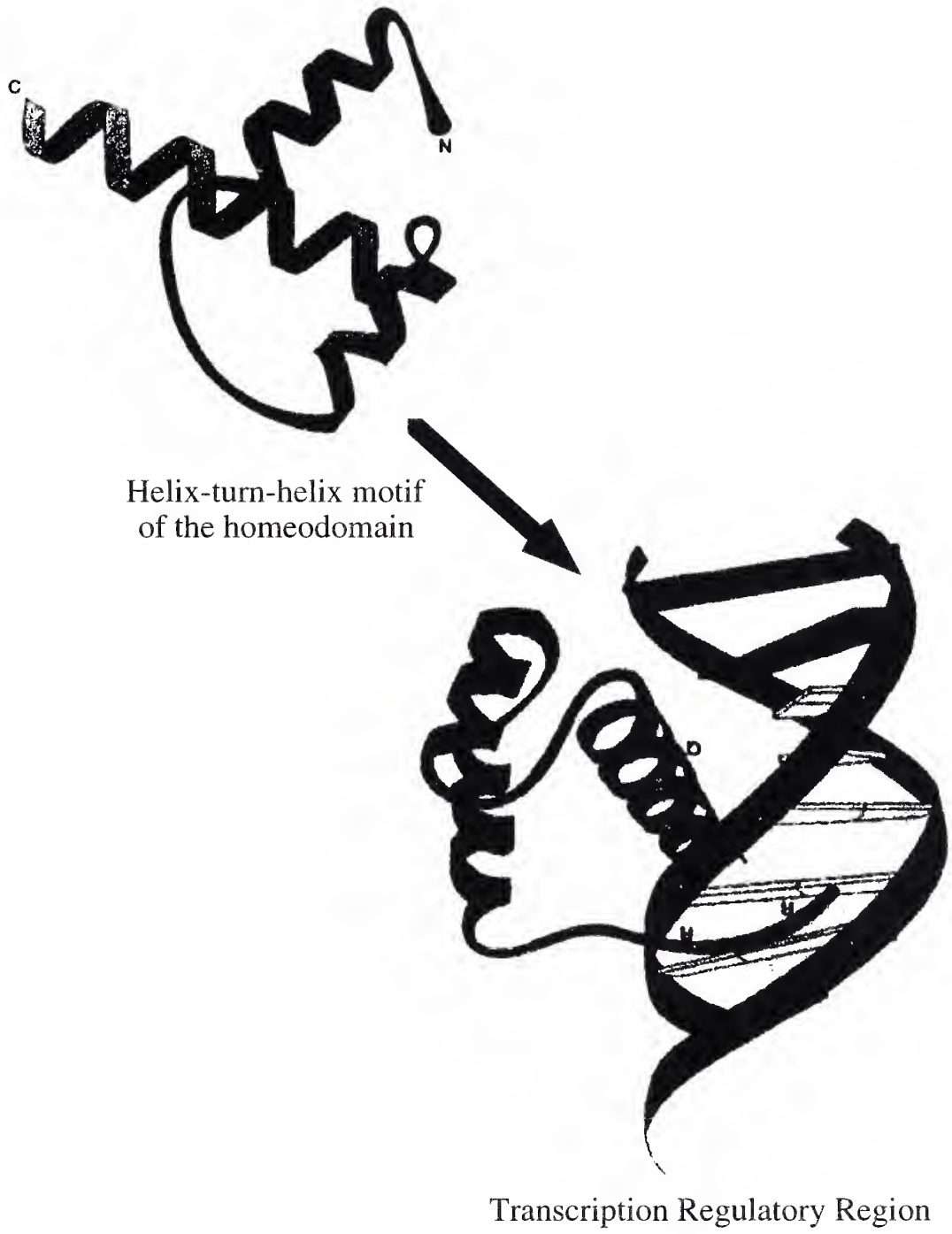


Figure 3: The helix-turn-helix motif of the homeodomain binds to transcription regulatory regions of DNA

The helix-turn-helix motif binds the major groove of the DNA double helix in transcription regulatory regions. These structures have been confirmed by X-ray crystallography (adapted from (54)).

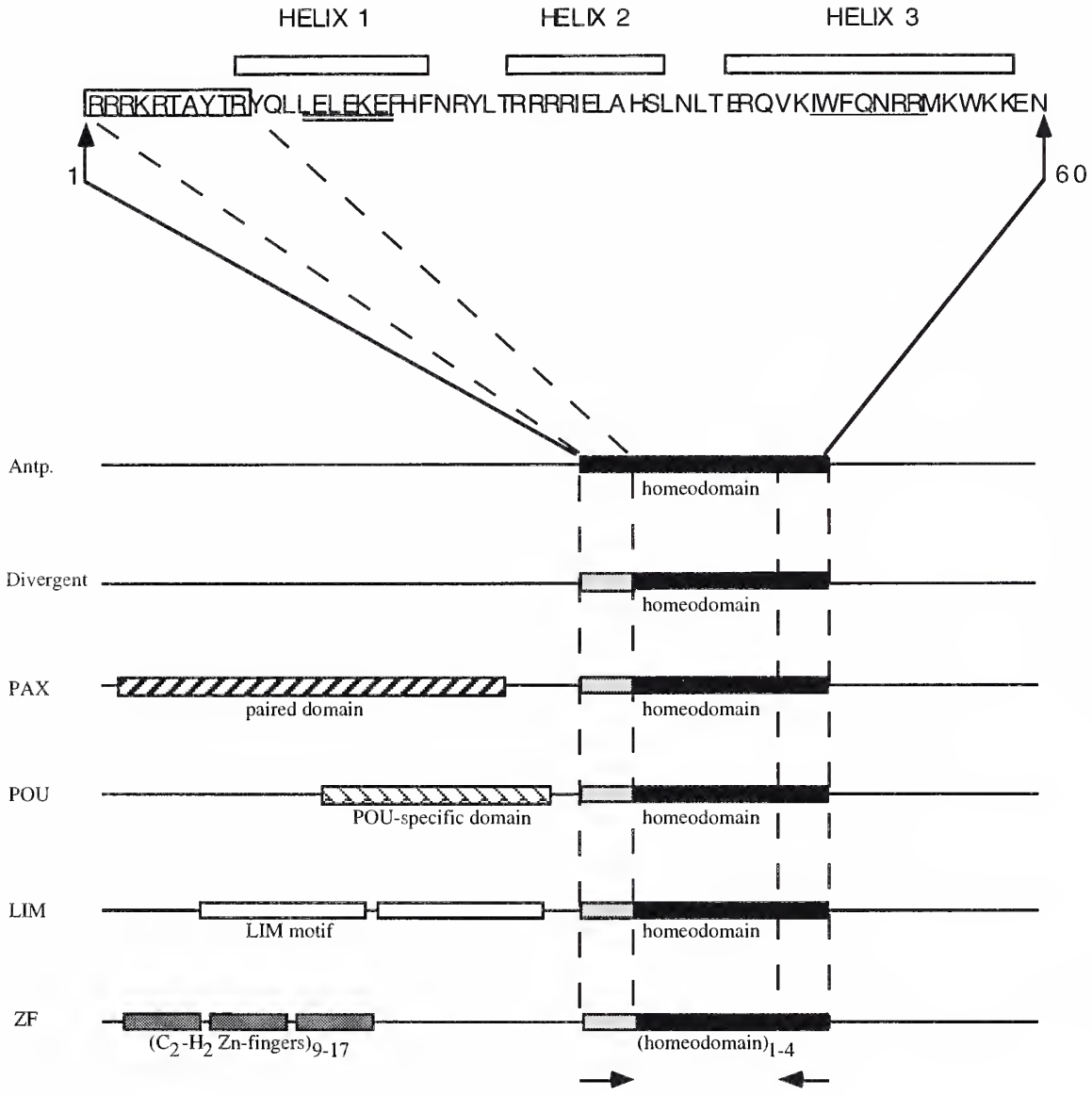


Figure 4: A Comparison Between Antp. and Several Families of Divergent Homeobox Genes

The amino acid sequence (in single letter code) of the Antp. homeodomain is listed at the top of the figure. This consensus sequence contains three alpha helical regions separated by short turns. The most highly conserved amino acid sequence is underlined; the second most is double-underlined. Note that these sequences are in the alpha-helical regions which are vital for DNA-binding (see Figure 3).

Schematic representations of the Antp. gene compared with some of the divergent homeobox genes including those containing other conserved DNA-binding motifs (Pax, POU, LIM, ZF families) are diagrammed in the lower half of the figure. The first ten amino acids in the homeobox are relatively conserved within a given divergent homeobox gene family. This homology was used to create the divergent homeobox primers for the RT-PCR assay (see Table 5). The regions corresponding to the primers are denoted with arrows at the bottom of the figure.

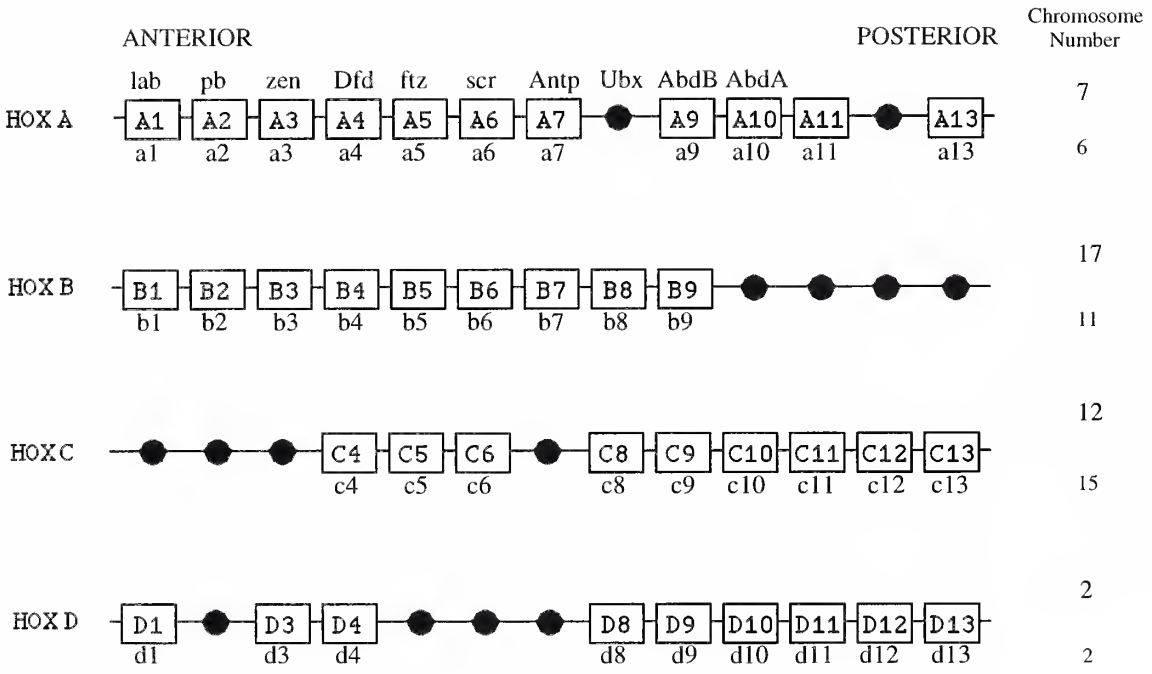


Figure 5: Organization of the HOX Clusters in Humans

The four HOX loci A, B, C, and D, are located on chromosomes 7, 17, 12, and 2, respectively. The current nomenclature designates the class I homeoboxes with numerals corresponding to their anterior-posterior position within the HOX cluster. The anterior-posterior designation is derived from the locations of Hox gene expression during *Drosophila* and murine embryogeneis.

The murine equivalents are located on chromosomes 6, 11, 15, and 2. These are listed beneath their human counterparts. The related *Drosophila* genes are listed at the top of the figure.

Divergent Homeobox Genes in Hematopoietic Cells

Gene	Family	Human Chromosome	Known Expression in Hematopoietic Cells	Reference
HLX	H2.0	1	B-lymphocytes Myeloid cells ? Overexpression in some AML patients	(104)
HB9	HB9	1	B-lymphocytes and myeloid cells	(103)
LH-2	LIM		Early B-lymphocytes T-lymphocytes	(106)
MSX-2	msh	5	Many myeloid leukemia cell lines	(105)
DLX-1	dll	2	Myeloid leukemia cell lines (KG1, F36E)	(105)
CDX/JRX	cad		Human T-cell leukemia cell line (Jurkat) Murine myeloid leukemia cell line (B6-SUT)	(107)
PRH (HEX)	PRH (HEX)	10	Myelocytic leukemia cell line (HL60) Normal bone marrow Peripheral blood	(108, 109, 110)
GBX-1	GBX-1	7	K562 Myeloerythroblastic cell line Daudi B-cell line	(111)
Oct 11	POU		Murine myeloma cell line	(112)
HOX 11	H2.0	10	Ectopically in t(10;14)(q24,q11) T-cell ALL low levels in normal T-cells	(117, 118 119)
PBX-1	mat-alpha2	1	Ectopically in t(1;19)(q23,p13.3) B-cell ALL	(121, 122 123, 124)

Table 2: Known Expression of Divergent Homeobox Genes in Hematopoietic Cells

This table lists all the divergent homeobox genes known to be expressed in hematopoietic cells as of February, 1994. Each gene is presented with its family name, chromosomal location and expression patterns (if known).

	Patient 2/3/91	TS9;22 2/12/93
Peroxidase	NEG	NEG
T-lymphoid-associated antigens		
CD1a	0.0%	0.0%
CD2	13.5	1.7
CD3	3.5	1.6
CD4	85.9	8.2
CD5	7.9	3.4
CD7	13.7	4.3
CD8	5.2	0.9
CD25	51.3	25.7
B-lymphoid-associated antigens		
CD10	2.9	5.0
CD19	3.3	1.9
CD20	1.6	0.0
CD22	1.3	0.0
Myeloid/Monocytoid-associated antigens		
CD11b	8.1	1.5
CD13	22.7	90.8
CD14	2.2	0.0
CD33	88.1	98.4
CD36	50.3	100.0
Platelet/Megakaryocyte-associated antigens		
CD41	13.4	43.7
CD42b	2.3	0.0
Others		
CD56	0.9	0.0
CD57	0.3	
HLA-DR	8.7	92.7
CD34		29.1

Table 3: Immunophenotype of the CML patient (during accelerated phase of CML in 1991) and of the TS9;22 cell line derived from this patient's bone marrow cells during megakaryocytic blast crisis.

The numbers for each cell surface characteristic represent the percentage of cells assayed which expressed the antigen. The presence or absence of peroxidase activity was also assayed. These data provided by Dr. K. Ohyashiki, Tokyo Medical College, Tokyo, Japan (unpublished cell line).

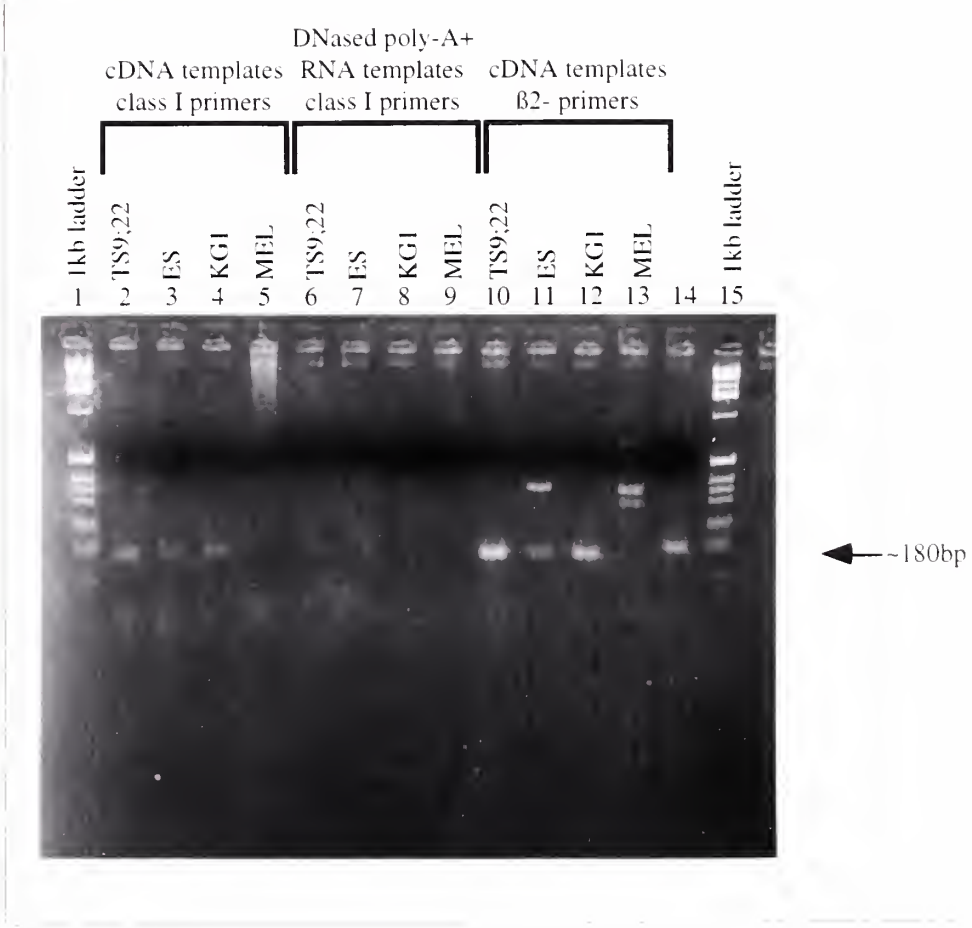


Figure 6: Identification of Homeobox Expression in Four Leukemia Cell Lines.

3% agarose gel of the initial PCR experiment in which TS9;22, ES, KG1, and MEL cDNAs (Lanes 2-5) and DNased poly-A+-mRNAs (Lanes 6-9) were tested with the 5'-class I homeobox primer (2551). ß2-microglobulin was used as a cDNA control (Lanes 10-13). The murine-derived ES and MEL cDNAs do not yield the appropriate PCR signal with the ß2-microglobulin primers because of mismatches in the DNA sequence of the primer which is based on the human sequence (Lanes 11, 13). The template for lane 14 is a plasmid with a known class I homeobox sequence.

Leukemia Cell Lines

Cell Line	Name	Source	Reference
ES	Embryonic Stem Cell	Pluripotent cell line derived from murine blatocyst cells	(133)
KG-1		Myeloblastic cell line derived from a patient with AML	(130)
TS9;22		Japanese patient with CML in blast crisis of megakaryocytic lineage	(K. Ohyashiki, (unpublished cell line))
MEL	Mouse erythroleukemia	Murine erythroid cell line	(134)
HL60		Myeloid cell line derived from a patient with AML type M2	(131)
Meg01		Megakaryoblastic cell line derived from a patient with CML in blast crisis	(132)

Table 4: Cell Lines Used for RT-PCR Assays

These cell lines contain transformed hematopoietic cells representing the myeloid, erythroid, and megakaryocytic lineages. The ES cells are not hematopoietic cells but are pluripotent and may be induced to differentiate into many, if not all, tissue-types.

5' Oligonucleotide Primers Designed to Detect the Class I and 19 Types of Divergent Homeoboxes

5' Primer Name	Divergent Gene Family[1]	Amino Acid Sequence[2]	Primer Sequence[3]	Reference
3067	chox3	.....TAFTYEQL	ACIGCITT(CT)ACITA(CT)GA(AG)CA (AG)CT	
3066	Hox7.1/msh	.RKPRTPFT....	(CA)GIAA(GA)GIACICCITT(TC)AC	(141,142)
3582	LIM	.RGPRTTI.....	(CA)GIGGICCI(CA)GIACIACIAT	(166)
3583	H2.0	.SWSRAVF.....	(AT)(GC)ITGG(AT)(GC)I(CA)GIGCI TITT	(167)
3689	dll	.RKPRTIY.....	(CA)GIAA(GA)CCI(CA)GIACIAT(ACT) TA	(168)
3690	ceh 5	.KRPRTDN.....	AA(GA)(CA)GICCI(CA)GIACIGA(CT)AA	(169)
3691	TTF/NK 2	RRKRRVLF..... K	(CA)(GA)I(CA)GIAA(GA)(CA)GI(CA)G IGTICTITT	(170)
3692	ceh 7	.PRRRTTF.....	CI(CA)GI(CA)GI(CA)GIACIACITT	(59)
3693	chox 7	...RRTAFTS... K	(CA)(GA)I(CA)GIACIGCITT(CT)ACITC	(171)
3694	cut	.KKQRVLF.....	AA(GA)AA(GA)CA(GA)(CA)GIGTICTITT	(172)
3695	htr	.KRTRTAY.....	AA(GA)(CA)GIACI(CA)GIACIGCITA	
3696	NK 4	KRKPRVLF.....	AA(GA)(CA)GIAA(GA)CCI(CA)GIGTICT ITT	(170)
3697	isl	.TRVRTVL.....	ACI(CA)GIGT(CA)GIACIGTICT	(166)
3763	NK 3	.KRSRAVF.....	AA(GA)(CA)GI(AT)(GC)I(CA)GIGCIGC ITT	(170)
3765	chox 1	.GSLRTSF.....	GGI(TA)(CG)I(CT)TI(CA)GIACI(TA) (CG)ITT	(173)
3766	xhox 3/eve	.RRYRTAFT....	GI(CA)GITA(TC)(CA)GIACIGCITT(CT) AC	(143,144)
3767	ceh 3	.DKYRMVY.....	GA(CT)AA(GA)TA(TC)(CA)GIATGGTITA	(169)
3768	zen 2	.KRSRTAF.....	AA(GA)(CA)GI(TA)(GC)I(CA)GIACIGC ITT	
3769	e4e5/ems	.KRVRTAF..... I	AA(GA)(CA)GI(GA)TI(CA)GIACIGCITT	(174)
2551	Class I	ELEKFLF[4]	GA(AG)(CT)TIGA(AG)AA(AG)GA(AG)TT (CT)C(AT)IT(AT)(CT)	(84)

Table 5
(adapted from (84))
1 "Gene Family" indicates a representative member of the divergent gene family
2 "Amino Acid Sequence" is the first ten amino acids of the homeobox sequence characteristic of the divergent homeobox gene family. Amino acids not represented in the primer are depicted with a period (.). In some cases, there is degeneracy in the amino acid sequence of a divergent gene family.
3 All the primers used here are preceded by the sequence, GCGAATTC, which contains the restriction site for EcoRI at the 5' end in order to facilitate subcloning of the PCR fragment.
4 This seven amino acid sequence is conserved in most of the class I homeobox genes.

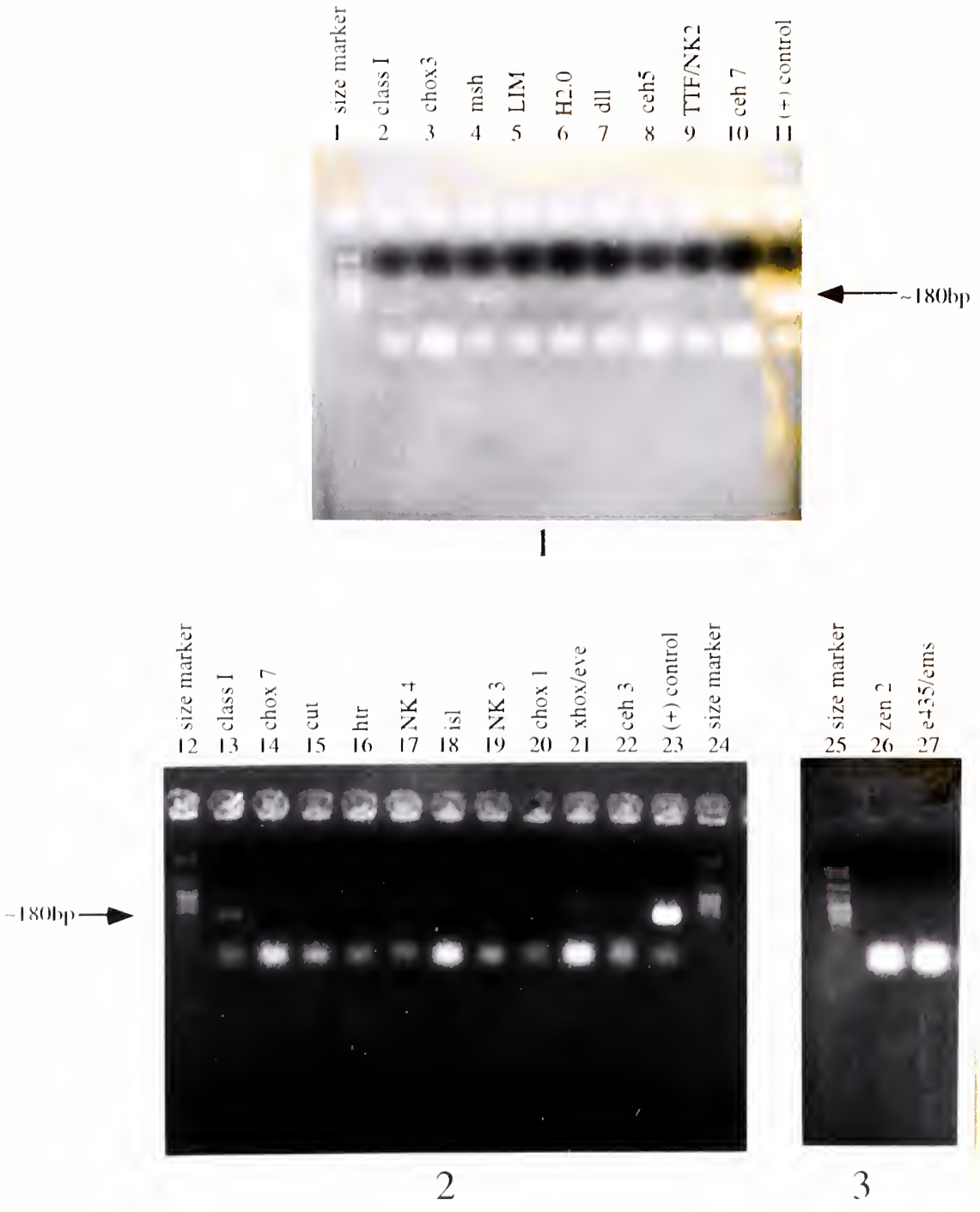


Figure 7: Screening of TS9;22 cDNA With Divergent Homeobox Primers

Ethidium stain of 3% agarose gels containing the PCR products from TS9;22 cDNA using degenerate oligonucleotide primers shown in Table 5. A positive signal was generated with the class I, msh, and xhox/eve primers (Lanes 2 and 13, 4, 21, respectively). Lanes 11 and 24 contain a class I homeobox "positive control" plasmid. Note that the msh and xhox/eve PCR fragments are larger than the class I-derived fragments, reflecting the more 5' position of the divergent primers within the homeobox for this set. "Size marker" indicates the MspI digest of plasmid pBR322.

Primer	Result (+/-)
HOX class I	+
3067	-
msh	+
LIM	-
H2.0	-
dll	-
ceh5	-
TTF	-
ceh7	-
chox7	-
cut	-
htr	-
NK4	-
isl	-
NK3	-
chox1	-
xhox3/eve	+
ceh3	-
zen2	-
e4/e5	-

Table 6: Summary of RT-PCR Screening of TS9;22.

Summary of RT-PCR screening of the TS9;22 cell line with 19 divergent homeobox primers shown in Table 5. A "+" indicates that an appropriate sized PCR product was visible on the ethidium stain of the 3% agarose gel.

PAX-TS

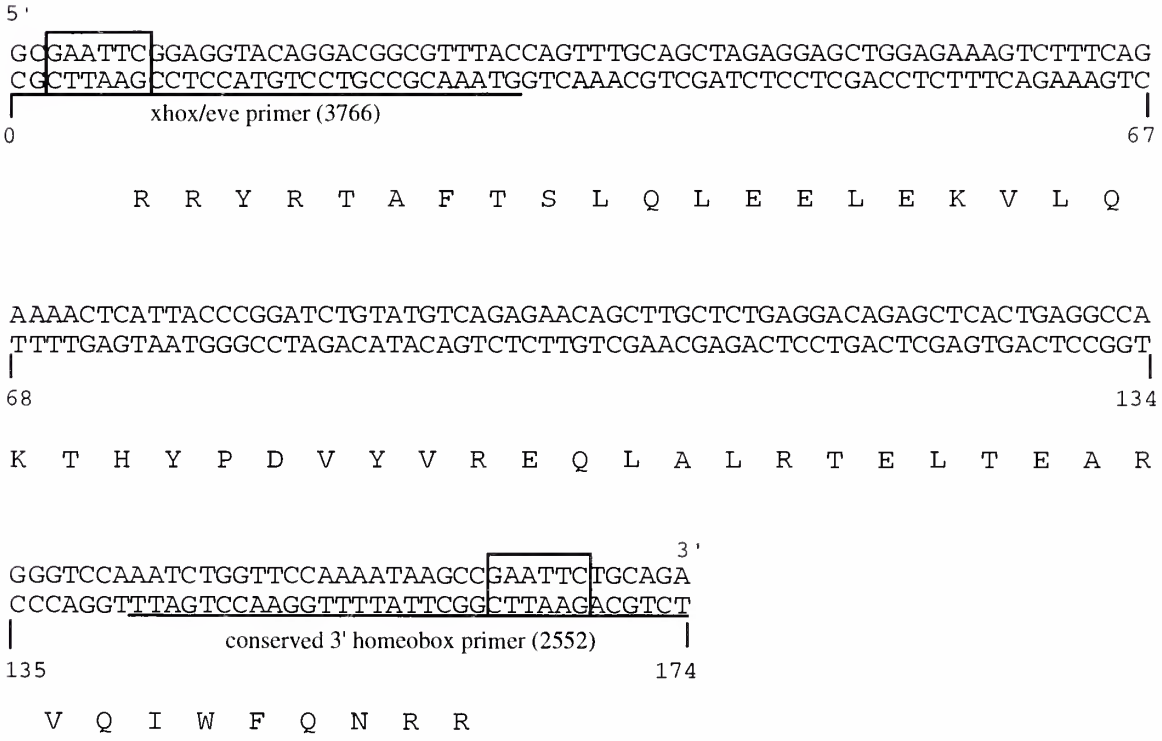


Figure 8: Sequence of PAX-TS RT-PCR Clone

DNA nucleotide sequence of the PAX-TS homeobox clone. The degenerate primer regions are underlined with the EcoRI cloning sites boxed. The deduced amino acid sequence is listed below in single letter code.

Comparison of Cloned TS9;22 Homeoboxes with Published Amino Acid Sequences

Primers Used	Sequence	Homology (%)
1) Class I (2551)	ELEKEFLFNMYLTRDRRHEVARLLNLSERQVKIWFQNRR	
HOX B9	----------	100
2) Hox 7.1/msh (3066)	RKPRTPFTTSQLLALERKFRQKQYLSIAERAEFSSSLNL	
MSX-2	----------	100
	TETQVKIWFQNRR	
	----------	
3) xhox 3/eve (3766) human	SLQLEELEKVFQKTHYPDVYVREQLALRTELTEARV	
Cart-1 (rat)	----------	100
smox-3 (S. mansoni)	----K---RA--E-----I-T--D----ID------	75
Pax-3 (mouse)	AE------RA-ER-----I-T--E--Q-AK------	67
paired (fly)	AS--D---RA-ER-Q---I-T--E--Q--N-----I	61
Pax-7 (mouse)	AE-------A-ER-----I-T--E--Q--K----F-	69
S8 (mouse)	-S--QA--R--ER-----AF---E--R-VN-S----	64
ceh-10 (mouse)	QY-ID----A--DS----I-A--V--GK---Q-D-I	58
gsbBSH (Drosophila)	AE---A--RA-SR-Q-----T--E--QT-A-----I	61
otd (Drosophila)	RA--DV--AL-G--R---IFM--EV--KIN-P-S--	50
Evx-1 (mouse)	RE-IAR---E-YREN-VSRPR-CE--AALN-P-TTI	30

Figure 9: Amino Acid Sequences of Cloned TS9;22 Homeobox Genes and Comparison to Other Known Homeodomains

The three homeoboxes cloned from TS9;22 derive from HOX B9, MSX-2, and a previously unidentified gene. A Genbank search revealed that the amino acid sequence of this homeobox is identical to the rat Cart-1 gene. It also bears close resemblance to the "paired-like" genes including smox-3 and Pax-7 and shares significant homology with several other divergent homeobox genes (151).

Comparison of Homeoboxes Cloned from TS9;22 to Antennapedia Consensus Sequence

```
Antennapedia  RKRGRQTYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKENK

Hox B9                          elekeflfNMYLTRDRRHEVARLLNLSERQVKiwfqnrr

MSX-2                rkprtpfttSQLLALERKFRQKQYLSIAERAEFSSSLNLTETQVKiwfqnrr

PAX-TS               rryrtaftSLQLEELEKVFQKTHYPDVYVREQLALRTELTEARVQiwfqnrr
```

Figure 10: Amino Acid Sequence Deduced from the Nucleotide Sequence of the Cloned PCR Products from Figure 9

Amino acids are in single letter code. Lower case letters represent the PCR primer sequences. Capital letters represent the cloned peptide sequence within PCR primers. Antp. is used as a reference sequence.

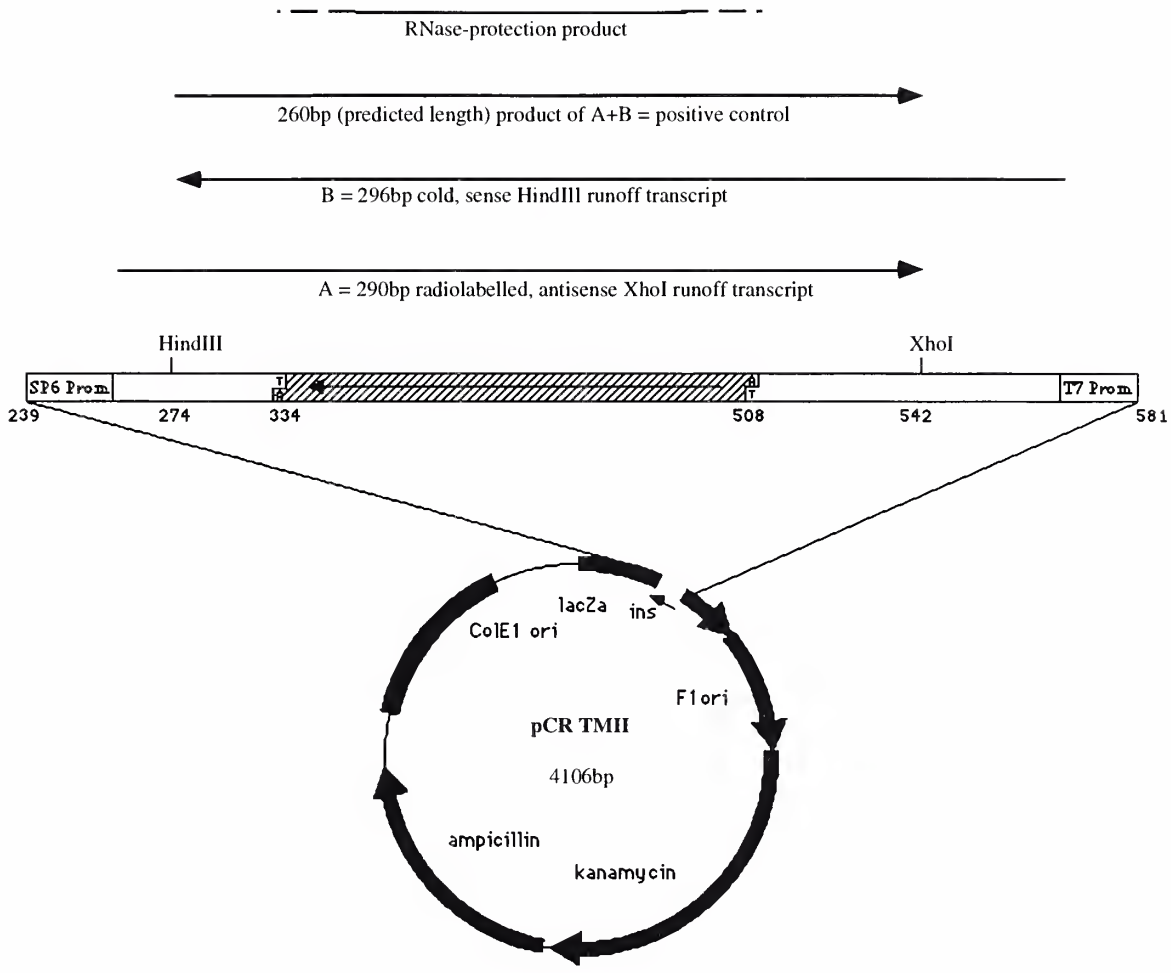


Figure 11: TA Cloning Vector Containing the 174bp TS9;22 PCR Product

TA Cloning vector containing the 174bp RT-PCR product using TS9;22 first-strand cDNA as template and primers 3766 and 2552 (xhox3/eve family). The orientation of the coding sequence of the insert within the plasmid is shown with an arrow. The enlarged view of the multiple cloning site and the PCR insert shows the two promoters and restriction sites used to generate the radiolabeled antisense (A) and cold sense (B) RNA transcripts used in the RNase-protection assay. The orientation of both transcripts and the two RNase-protection products are depicted above. The dotted lines at the ends of the RNase-protection product indicates the potential for mismatches between the degenerate primer and the true gene sequence.

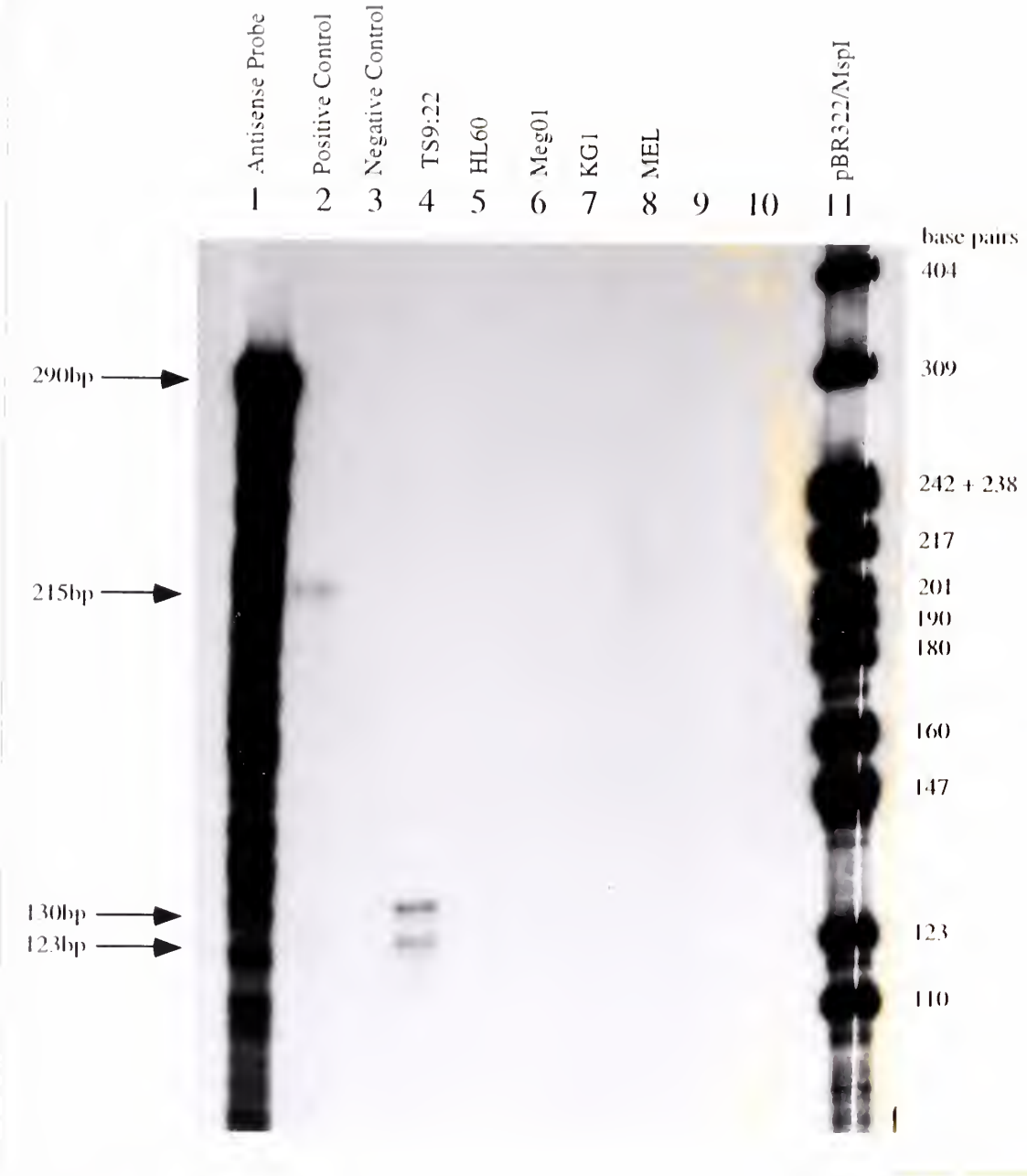
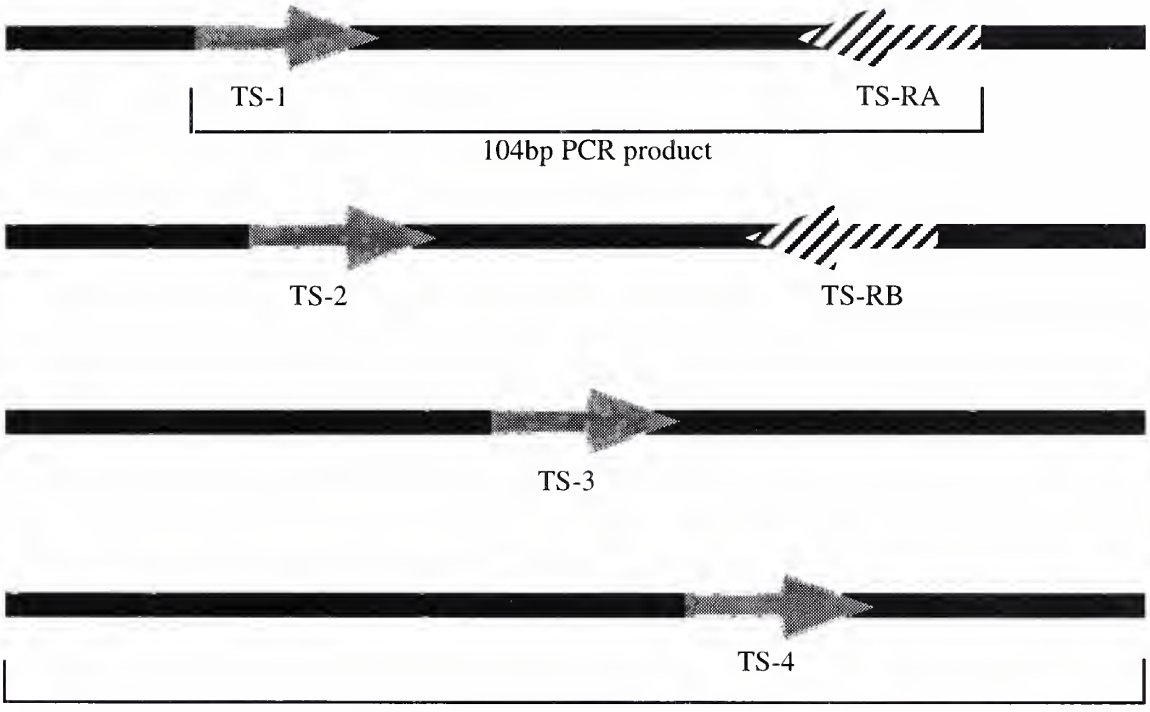


Figure 12: RNase-protection Assay to Determine PAX-TS Expression in the TS9;22, HL60, Meg01, KG1, and MEL Cell Lines

PAX-TS-specific, radiolabelled antisense RNA was hybridized to the total RNA samples indicated, digested with ribonuclease and electrophoresed. Lane 1 contains the 290bp radiolabelled "antisense" transcript. Lane 2 contains the positive control RNase-protection product of the radiolabelled "antisense" transcript and a cold sense transcript. Lane 3 contains the negative control product of the radiolabeled RNA probe and sterile water as a template. Lanes 4-8 contain the RNase-protection products of the radiolabeled "antisense" probe and total RNA from the five cell lines. Only TS9;22 demonstrates PAX-TS expression. See Figure 11 for the predicted fragment sizes. Note that the RNA products run faster on the gel than DNA fragments of similar length. Thus, the apparent base-pair sizes of the RNase-protection products are smaller than predicted.

PAX-TS Homeobox



TS-1 - TGCAGCTAGAGGAGCTGGAG

TS-2 - GCTAGAGGAGCTGGAGAAAG

TS-3 - GATGTGTATGTCAGAGAACA

TS-4 - TGAGGACAGAGCTCACTGAG

TS-RA - CTGGGACCGGAGTCACTCG

TS-RB - CAGTGAGCTCTGTCCTCAG

Figure 13: PAX-TS Primers Listed in This Thesis

The six primers shown above are specific for the human PAX-TS homeobox and are based on the sequence data of Figure 8. The rightward arrows indicate primers corresponding to the top strand (as in Figure 8) while the leftward arrows indicate the lower strand. TS-1 and TS-RA were used in the majority of the RT-PCR assays. TS-3 was labeled for use as a probe in the Southern blot experiments.

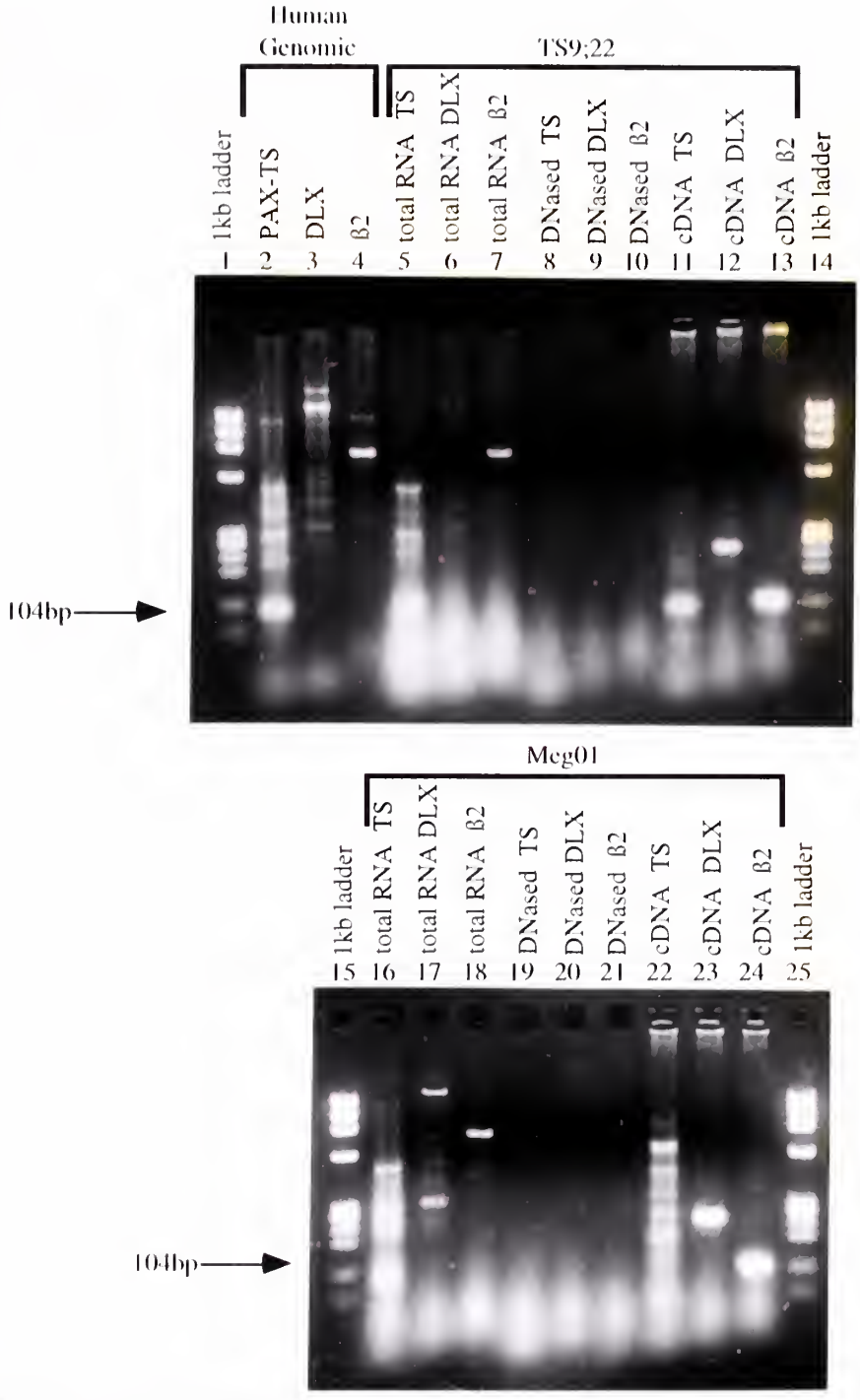


Figure 14: RT-PCR Assay for PAX-TS Expression in the TS9;22 and Meg01 Cell Lines

The TS9;22 and Meg01 cell lines were screened for PAX-TS expression by the RT-PCR assay. The screening for each of the two cell lines includes total RNA, DNased total RNA, and first-strand cDNA. ß2-microglobulin- (ß2) and DLX-specific primers were used as positive cDNA controls. A human genomic DNA template was used for comparison (Lanes 2-4). The 104bp PAX-TS band was only present in the human genomic DNA and the TS9;22 cDNA.

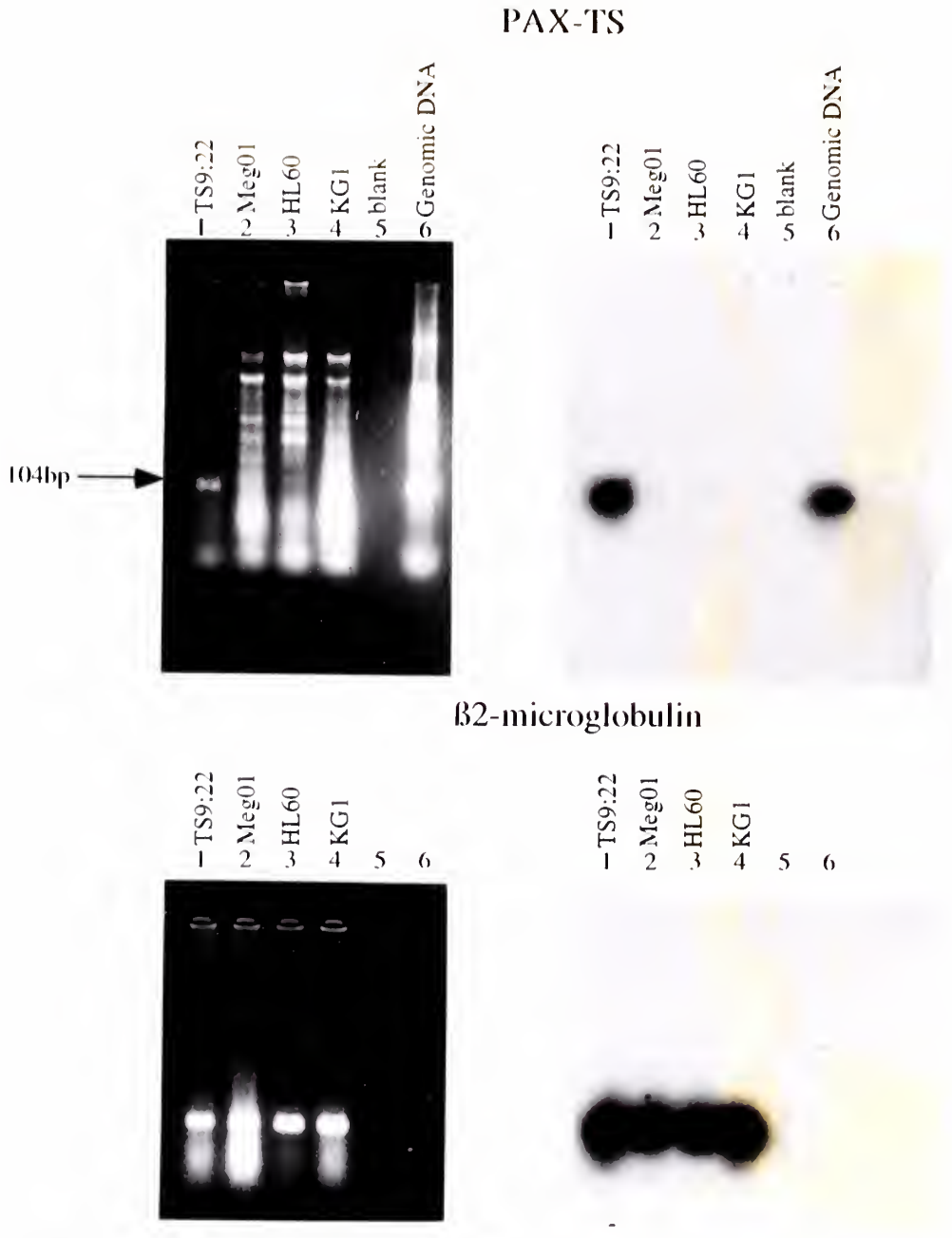


Figure 15: PAX-TS Expression Determined by RT-PCR Analysis Followed by Southern blot

The left panel contains the ethidium-stained gels of the PCR products generated using cDNA templates synthesized from TS9;22, Meg01, HL60, and KG1 and assayed for PAX-TS and ß2-microglobulin (ß2) expression. Human genomic DNA was used as a positive control (Lane 6, above). The right panel shows a Southern blot of the gels on the left using radiolabeled TS-3 or ß2 primers as probes.

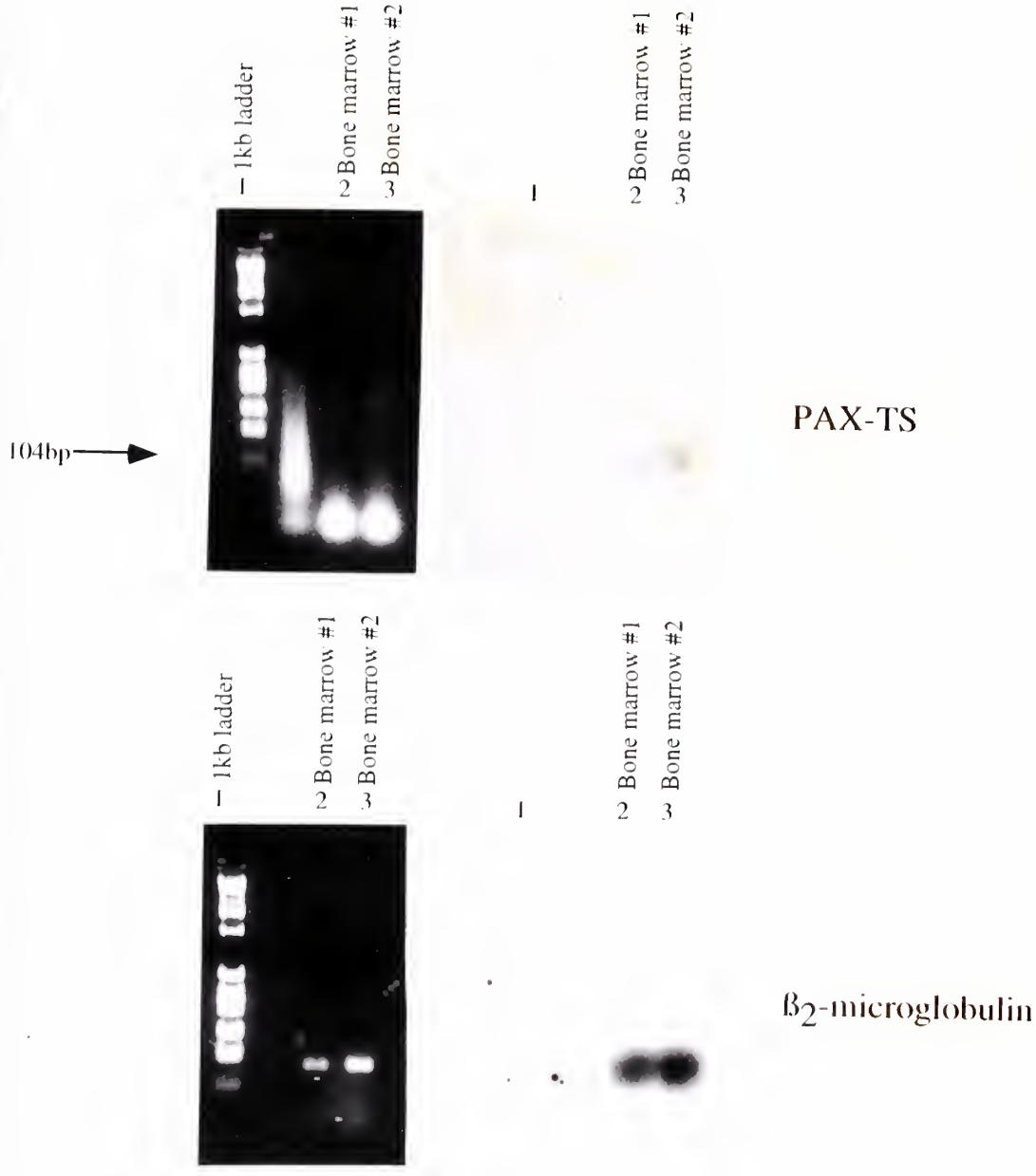


Figure 16: PAX-TS Expression in Normal Bone Marrow

RT-PCR analysis was performed on RNA samples obtained from normal bone marrow cells. The upper left panel contains the ethidium-stained gels of the RT-PCR assay for PAX-TS expression in two healthy human bone marrow samples. ß2-microglobulin (ß2) serves as a comparison of the cDNA templates (lower left). The right panel contains the Southern blot using the radiolabeled TS-3 or ß2 oligonucleotides as probes.

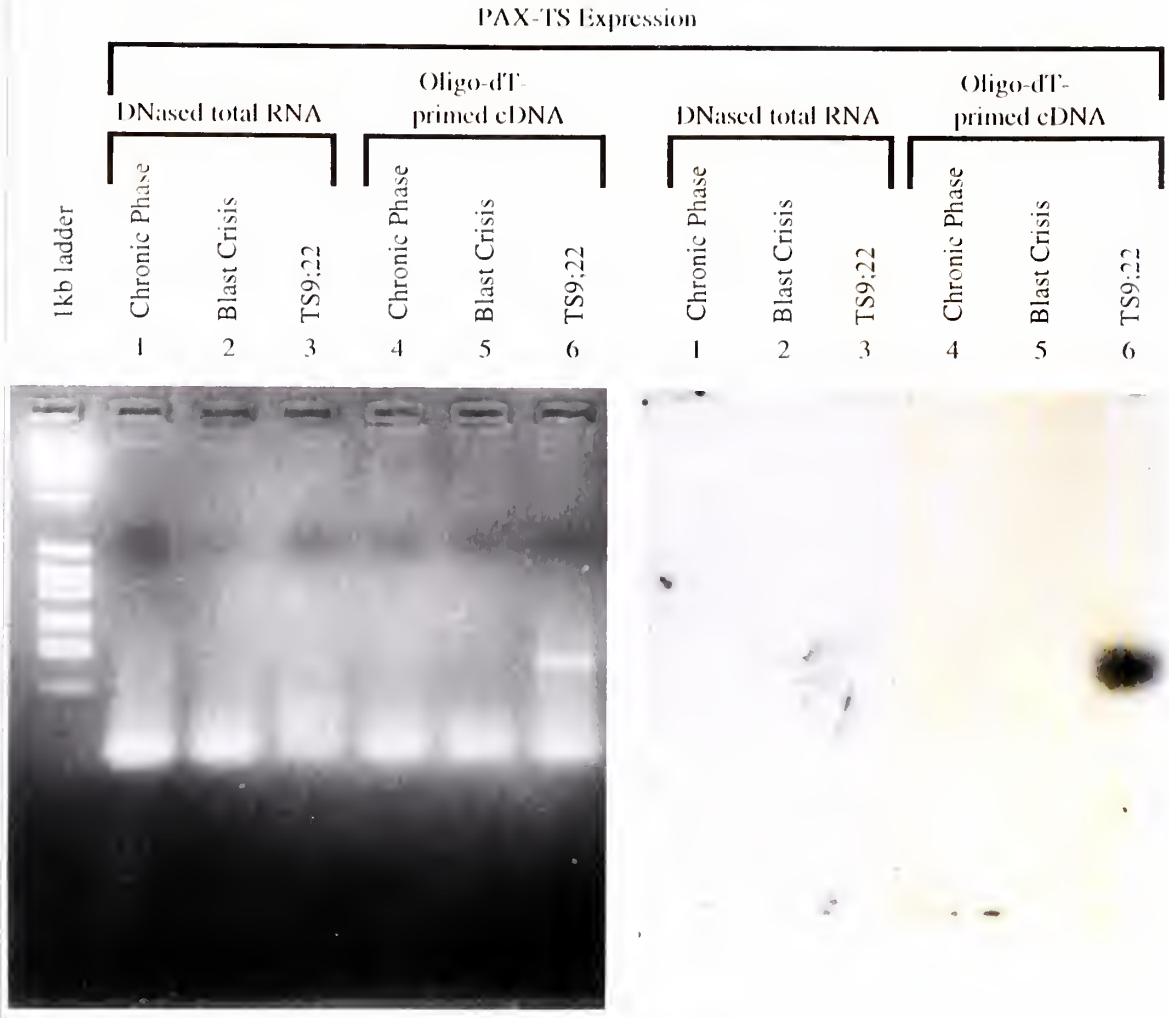


Figure 17: RT-PCR Analysis of TS9;22, Chronic Phase CML, and Megakaryocytic Blast Crisis RNA

The left panel contains the ethidium-stained gel of RT-PCR analysis of TS9;22, chronic phase CML, and megakaryocytic blast crisis total RNA. The right panel contains the Southern blot of this gel. Lanes 1-3 are the DNased total RNA samples; lanes 4-6 are the first strand oligo-dT cDNA samples.

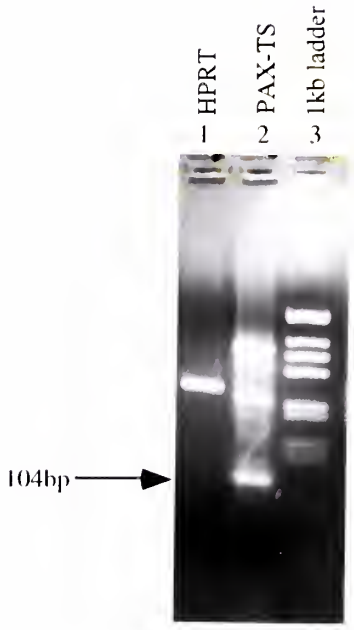


Figure 18: PAX-TS Expression in the 11 Day Mouse Embryo

Ethidium-gel of RT-PCR of 11 day mouse embryo total RNA. cDNA was synthesized from day 11 mouse embryo total RNA and used for PCR analysis. HPRTase expression was used as a cDNA control (Lane 1). PAX-TS expression is shown in lane 2.

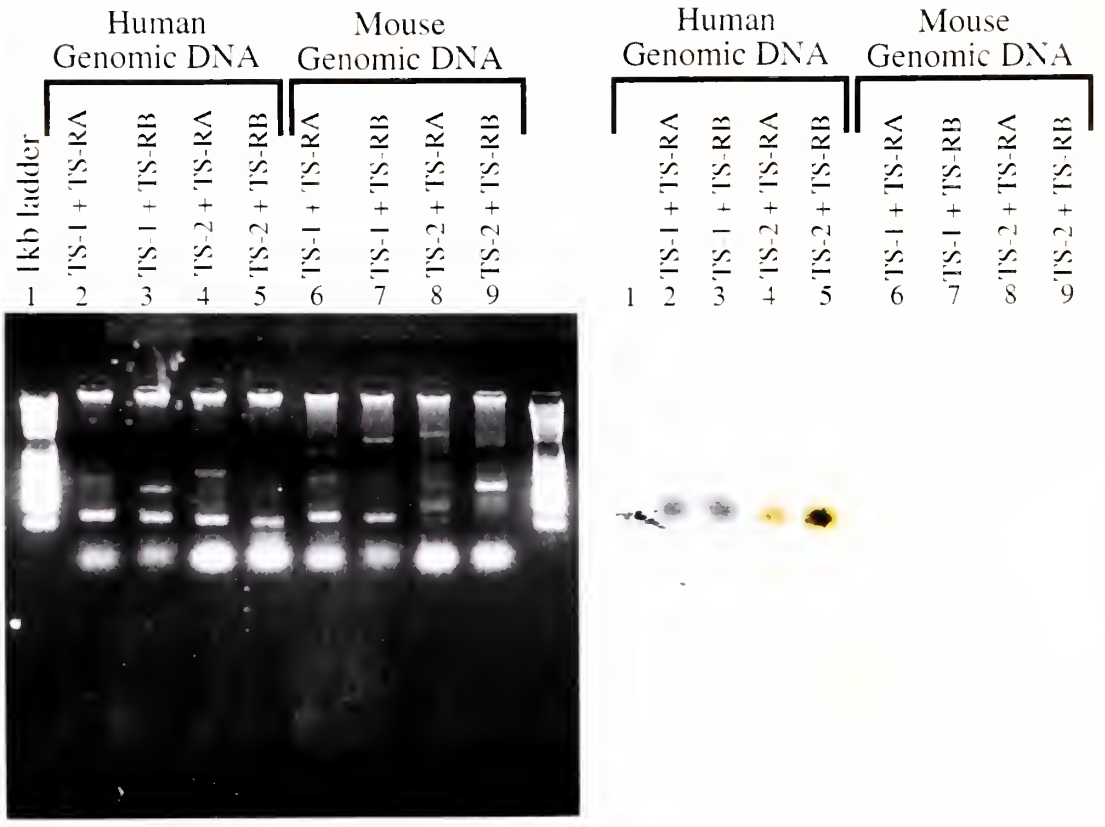


Figure 19: A Comparison of Mouse and Human Genomic DNA Templates for PCR Using Human PAX-TS-Specific primers

The left panel shows the ethidium-stained gel of the PCR analysis of human (Lanes 2-5) and mouse (Lanes 6-9) genomic DNA templates using the four primer pairs: TS-1 and TS-RA, TS-1 and TS-RB, TS-2 and TS-RA, and TS-2 and TS-RB. Refer to Figure # for information on the PAX-TS-specific primers. The right panel contains the Southern blot using the radiolabeled TS-3 oligonucleotide as a probe. Although the TS-3 probe did not hybridize to the murine PCR products, these bands were confirmed as PAX-TS on the basis of direct cloning and sequencing (see Figure 20).

Comparison of DNA and Amino Acid Sequences Between Human, Mouse, and Rat Homologues of PAX-TS

DNA AND AMINO ACID SEQUENCES

```
        [    TS-1 Primer   ]                                                                  [  TS-RA Primer  ]

HUMAN   TTTGCAGCTAGAGGAGCTGGAGAAAGTCTTTCAGAAAACTCATTACCCGGATGTGTATGTCAGAGAACAGCTTGCTCTGAGGACAGAGCTCACTGAGGCCAGGGTC
          L  Q  L  E  E  L  E  K  V  L  Q  K  T  H  Y  P  D  V  Y  V  R  E  Q  L  A  L  R  T  E  L  T  E  A  R  V

MOUSE   TTTGCAGCTAGAGGAGCTGGAGAAAGTTTTGCAGAAAACCCATTACCCAGATGTATATGTCAGAGAACAACTTGCCCTGAGAACGGAGCTCACTGAGGCCAGGGTC
          L  Q  L  E  E  L  E  K  V  L  Q  K  T  H  Y  P  D  V  Y  V  R  E  Q  L  A  L  R  T  E  L  T  E  A  R  V

RAT     TTTGCAGCTAGAGGAGCTGGAGAAGGTCTTCCAAAAAACGCATTACCCGGATGTATATGTCAGAGAACAGCTTGCACTGAGAACGGAGCTCACTGAGGCCAGGGTC
          L  Q  L  E  E  L  E  K  V  L  Q  K  T  H  Y  P  D  V  Y  V  R  E  Q  L  A  L  R  T  E  L  T  E  A  R  V
```

AMINO ACID SUMMARY

```
HUMAN LQLEELEKVFQKTHYPDVYVREQLALRTELTEARV

MOUSE LQLEELEKVLQKTHYPDVYVREQLALRTELTEARV

RAT   LQLEELEKVFQKTHYPDVYVREQLALRTELTEARV
```

Figure 20: Comparison of the Rat, Mouse, and Human Homologues of PAX-TS.
A comparison of the DNA and amino acid sequences of the human, rat, and mouse PAX-TS homeoboxes demonstrates 85.9% DNA homology and a single amino acid difference between the human and mouse (indicated in bold face) and 85.9% DNA homology but 100% amino acid conservation between the human and rat. The rat clone came from the RT-PCR analysis of the rat kidney total RNA (see Figure 21) and the mouse clone came from PCR of mouse genomic DNA (see Figure 19).

DNased Total RNA
1 1kb ladder
2 TS9;22
3 Meg01
4 Lung
5 Kidney
6 Liver
7 Colon
8 Spleen
9 Skeletal Muscle
10 Stomach
11 Small Bowel
cDNA
104bp
12 1kb ladder
13 TS9;22
14 Meg01
15 Lung
16 Kidney
17 Liver
18 Colon
19 Spleen
20 Skeletal Muscle
21 Stomach
22 Small Bowel

Figure 21

Figure 21: PAX-TS Expression in Rat Organs by RT-PCR Analysis

In the left panel, the ethidium-stained gel contains the RT-PCR products from RNA extracted from the organs of an adult rat. Total RNA was extracted from lung, stomach, small bowel, large bowel, spleen, liver, kidney and skeletal muscle and PCR analysis was performed with the PAX-TS specific primers, TS-1 and TS-RA. The PCR products from the DNased total RNA are in the upper row (Lanes 2-11) and the products from the first-strand cDNA samples are in the lower row (Lanes 13-22). TS9;22 and Meg01 served as positive and negative controls, respectively (Lanes 2 + 3, and 13 + 14). In the rat organ systems tested, only the kidney demonstrates PAX-TS expression. On the accompanying Southern blot, only the TS9;22 PCR product hybridized to the probe. The rat kidney fragment does not hybridize under the wash conditions of this Southern blot, suggesting that there is some mismatch in the rat compared to the human sequence in the region of the radiolabelled oligonucleotide. This fragment was confirmed as PAX-TS by direct cloning and sequencing (see Figure 20).

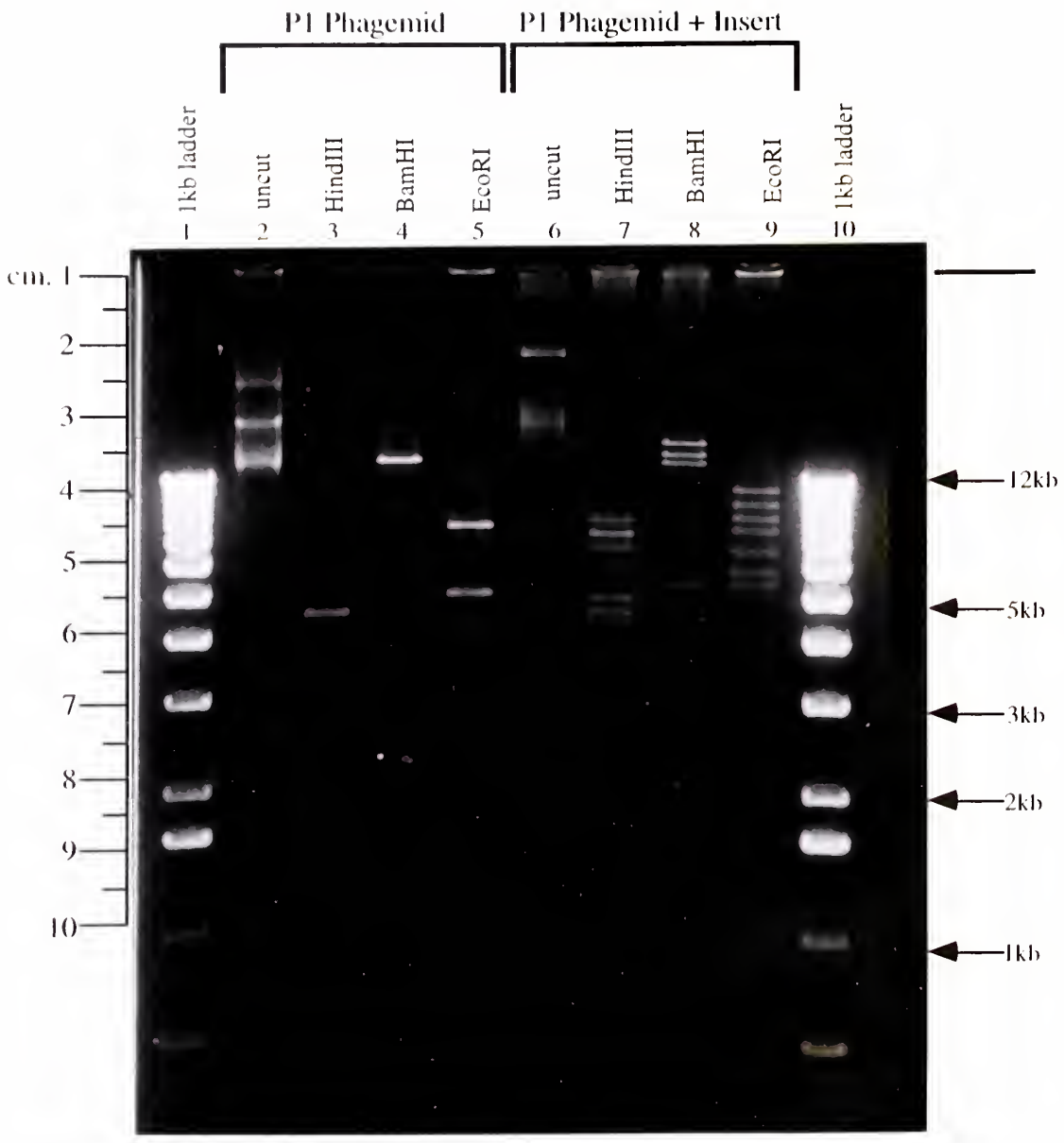


Figure 22: Restriction Digest of the P1 Phagemid Containing a ~70kb PAX-TS Insert

Ethidium-stained gel of restriction digests of the P1 phagemid with and without a genomic insert. Lane 1 contains the uncut P1 phage vector. Lanes 2-4 contain the P1 phage vector digested with HindIII, BamHI, and EcoRI. Lane 5 contains the uncut P1 phagemid clone carrying the ~70kb PAX-TS genomic insert. Lanes 6-8 contain the P1 phagemid PAX-TS clone digested with HindIII, BamHI, and EcoRI.

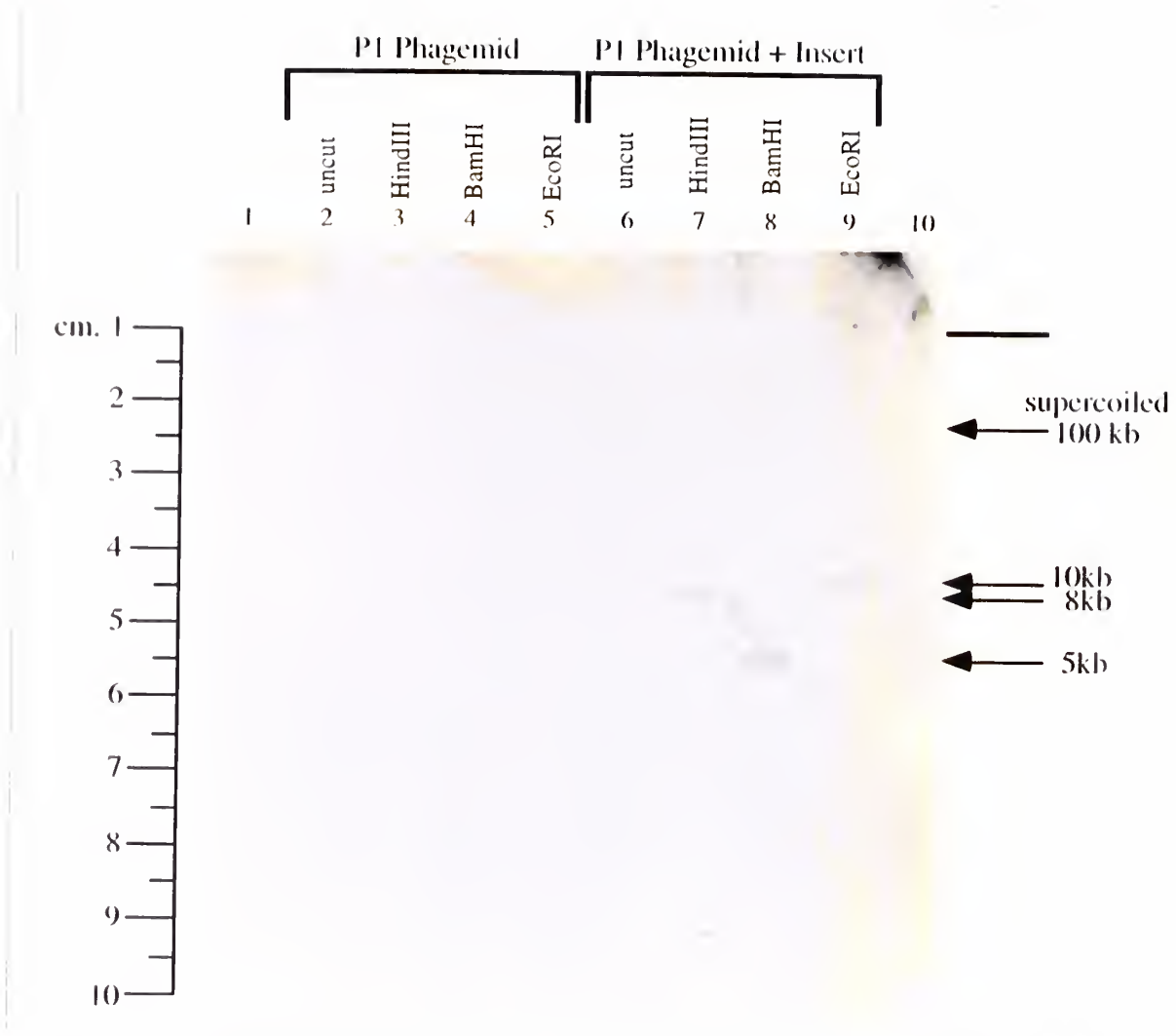


Figure 23: P1 Southern Blot of P1 Phagemid Clone Carrying the PAX-TS Genomic Sequence

The Southern blot of the P1 phagemid restriction digest using the radiolabelled TS-3 oligonucleotide as a probe. Lanes 2-5 contain the uncut and digested control P1 phagemid without insert. Lane 6 contains the uncut P1 phagemid carrying the genomic insert. Lanes 7, 8, and 9 depict the ~8kb, ~5kb, and ~10kb fragments containing the PAX-TS sequence after digestion with HindIII, BamHI, and EcoRI, respectively.

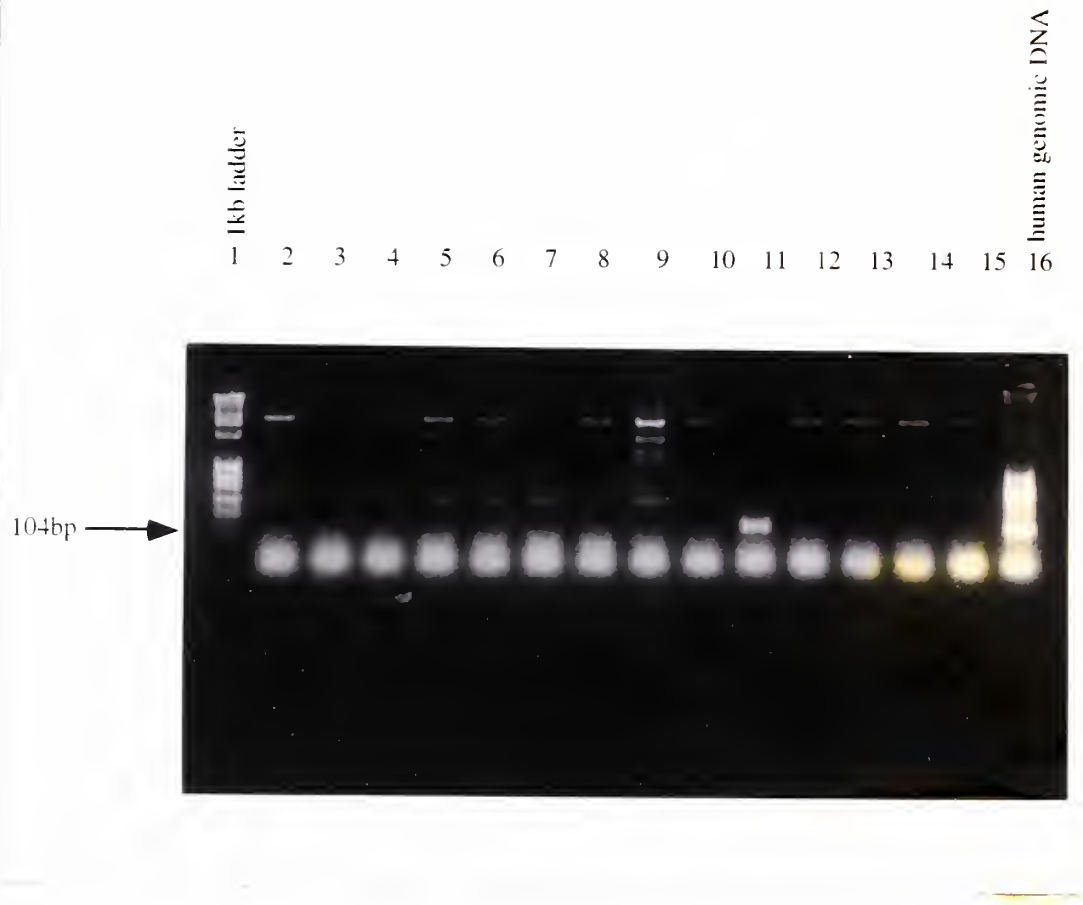


Figure 24: Secondary PCR screening of P1 genomic phage mini-library
Fourteen plates containing 50 phage each were screened by the PCR assay using the PAX-TS-specific primers, TS-1 and TS-RA, after elution into SM buffer. Lane 11 contains the single positive plate. Lane 16 contains a human genomic DNA template as a positive control.

Figure 25: A Comparison of the Nucleotide Sequences of the PAX-TS and Cart-1 Homeoboxes

There is 84% homology between Cart-1 and PAX-TS. However, most of the mismatches occur in the regions of the two degenerate primers (underlined) which may not reflect the actual gene sequence. All of the other mismatches are conservative such that the amino acid sequences are identical.

BIBLIOGRAPHY

1. Robinson, B. E., & Quesenberry, P. J. (1990). Review: Hematopoietic growth factors: Overview and Clinical Applications, Part I. The American Journal of the Medical Sciences, 300, 163-170.
2. Emerson, S. G., (1991). "The Stem Cell Model of Hematopoiesis", (pp. pp. 72-81). In Hoffman, R. (ed): Hematology: Basic Principles and Practice: Churchill Livingstone, New York.
3. Ford, L. C., Hamerton, J. L., Barnes, D. W. H., & Loutit, J. F. (1956). Cytological identification of radiation chimeras. Nature, 177, 452.
4. Till, J. E., & McCulloch, E. A. (1961). Direct measurement of the radiation sensitivity of normal mouse bone marrow cells. Radiation Research, 14, 213.
5. Becker, A. J., McCulloch, E. A., & Till, J. E. (1963). Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells. Nature, 197, 452.
6. Wu, A. M., Till, J. E., Siminovitch, L., & McCulloch, E. A. (1967). A cytological study of the capacity for differentiation of normal hemopoietic colony-forming cells . J. Cell Physiol., 69, 177.
7. Lajtha, L. G., Possi, L. V., Schofield, R., & Fox, M. (1969). Kinetic properties of hematopoietic stem cells. Cell Tissue Kinetics, 2, 39.
8. Calabretta, B., & Baserga, R., (1991). "Control of Cell Growth and Differentiation", (pp. pp. 65-72). In Hoffman, R. (ed): Hematology: Basic principles and practice: Churchill Livingstone: New York.
9. Heyworth, C. M., & Spooncer, E. (1993). "In vitro clonal assays for murine multipotential and lineage restricted myeloid progenitor cells". In Testa, N.G. (ed): Haemopoiesis: A practical approach: Oxford University Press: New York.
10. Nakahata, T., & Gross, A. J. (1982). Stochastic model of self-renewal and commitment to differentiation of the primitive hemopoietic stem cells in culture. J. Cell Physiol, 113, 455.
11. Lübbert, M., Herrmann, F., & Koeffler, H. P. (1991). Expression and Regulation of myeloid-specific genes in normal and leukemic myeloid cells. Blood, 77, 909.
12. Berenson, R. J., Andrews, R. G., Bensinger, W. I., Kalamasz, D., Knitter, G., Buckner, C. D., & Bernstein, I. D. (1988). Antigen CD34 positive marrow cells engraft lethally irradiated baboons. J. Clin. Invest., 81, 951.
13. Bell, A. J., Hamblin, T. J., & Oscier, D. G. (1987). Peripheral blood stem cell autografting. Hematol. Oncol., 5, 45.
14. Terstappen, L. W. M., Huang, S., Safford, M., Landsdorp, P. M., & Loken, M. R. (1991). Sequential generations of hematopoietic colonies derived from single nonlineage-committed CD34+CD38- progenitor cells. Blood, 77, 1218-1227.
15. Platzer, E. (1989). Human hemopoietic growth factors. Eur J Haematol, 42, 1.
16. Barber, D. L., & D'Andrea, A. D. (1992). The erythropoietin receptor and the molecular basis of signal transduction. Seminars in Hematology, 29, 293.
17. Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., & Watson, J. (1989). "Cell Signaling". Molecular Biology of the Cell: 2nd ed.: Garland Publishing, Inc.: New York.
18. Mathey-Prevot, B., Nabel, G., Palacios, R., & Baltimore, D. (1986). Abelson virus abrogation of interleukin-3 dependence in a lymphoid cell line. Mol. Cell Biol., 6, 4133.
19. Metcalf, D. (1989). The molecular control of cell division, differentiation commitment and maturation in hemopoietic cells. Nature, 339, 27.
20. Morla, A. O., & Schreurs, J. (1988). Hematopoietic growth factors activate the tyrosine phosphorylation of distinct sets of proteins in interleukin-3-dependent murine cell lines. Mol. Cell Biol., 8, 2214.

21. Hanazono, Y., Chiba, S., Sasaki, K., Mano, H., Yazaki, Y., & Hirai, H. (1993). Erythropoietin induces tyrosine phosphorylation and kinase activity of the c-fps/fes proto-oncogene product in human erythropoietin-responsive cells. Blood, 81, 3193.
22. Miura, O., Cleveland, J. L., & Ihle, J. N. (1993). Inactivation of erythropoietin receptor function by point mutations in a region having homology with other cytokine receptors. Mol. Cell Biol., 13, 1788.
23. Aziz, N., Wu, J., Dubendorff, J. W., Lipsick, J. S., Sturgill, T. W., & Bender, T. P. (1993). c-myb and v-myb are differentially phosphorylated by p42mapk in vitro. Oncogene, 8, 2259.
24. Igarashi, K., David, M., Larner, A. C., & Finbloom, D. S. (1993). In vitro activation of a transcription factor by gamma interferon requires a membrane-associated tyrosine kinase and is mimicked by vanadate. Mol. Cell Biol., 13, 3984.
25. Li, S., & Sedivy, J. M. (1993). Raf-1 protein kinase activates the NF-kappa B transcription factor by dissociating the cytoplasmic NF-kappa B-I kappa B complex. Proc. Natl. Acad. Sci. U.S.A., 90, 9247.
26. Darnell, J., Lodish, H., & Baltimore, D. (1990). In: Molecular Cell Biology, 2nd ed.: Scientific American Books, W.H. Freeman and Co., New York.
27. Mucenski, M. L., McLain, K., Kier, A. B., Swerdlow, S. H., Schreiner, C. M., Miller, T. A., Pietryga, D. W., Scott, W. J. J., & Potter, S. S. (1991). A functional c-myb gene is required for normal murine fetal hepatic hematopoiesis. Cell, 65, 677-89.
28. Kreider, B., Benezra, R., Rovera, G., & Kadesch, T. (1992). Inhibition of myeloid differentiation by the helix-loop-helix protein Id. Science, 255, 1700.
29. Martin, D. I. K., Tsai, S. -., & Orkin, S. H. (1989). Increased gamma globin expression in a nondeletion HPFH mediated by an erythroid-specific DNA binding factor. Nature, 338, 435.
30. Tsai, S., Martin, D. K., Zon, L. I., D'andrea, A. D., Wong, G. G., & Orkin, S. H. (1989). Cloning of cDNA for the major DNA-binding protein of the erythroid lineage through expression in mammalian cells. Nature, 339, 446.
31. Wall, L., deBoer, E., & Grosveld, F. (1989). The human ß-globin gene 3' enhancer contains multiple binding sites for an erythroid-specific protein. Genes Dev., 2, 1089.
32. Evans, T., Reitman, M., & Felsenfeld, G. (1988). An erythrocyte-specific DNA-binding factor recognizes a regulatory sequence common to all chicken globin genes. Proc. Natl. Acad. Sci. U.S.A., 85, 5976.
33. Mignotte, V., Wall, L., deBoer, E., Grosveld, F., & Romeo, P. (1989). Two tissue-specific factors bind the erythroid promoter of the human porphobilinogen deaminase gene. Nucleic Acids Res., 17, 37.
34. Youssoufian, H., Zon, L., Orkin, S. H., D'Andrea, A. D., & Lodish, H. F. (1990). Structure and transcription of the mouse erythropoietin receptor gene. Molec. Cell Biol., 10, 3675.
35. Martin, D. I. K., & Orkin, S. H. (1990). Transcriptional activation and DNA binding by the erythroid factor GF1/NF-E1/Eryf1. Genes Dev., 4, 1886.
36. Martin, D. I. K., Zon, L. I., Mutter, G., & Orkin, S. (1990). Expression of an erythroid transcription factor in megakaryocytic and mast cell lineages. Nature, 344, 444.
37. Romeo, P. H., Prandini, M. H., Joulin, V., Mignotte, V., Prenant, M., Vainchecker, W., Marguerie, G., & Uzan, G. (1990). Megakaryocytic and erythrocytic lineages share specific transcription factors. Nature, 344, 447.
38. Pevny, L., Simon, M. C., Robertson, E., Klein, W. H., Tsai, S. F., D'Agati, V., Orkin, S. H., & Constantini, F. (1991). Erythroid differentiation in chimeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1.

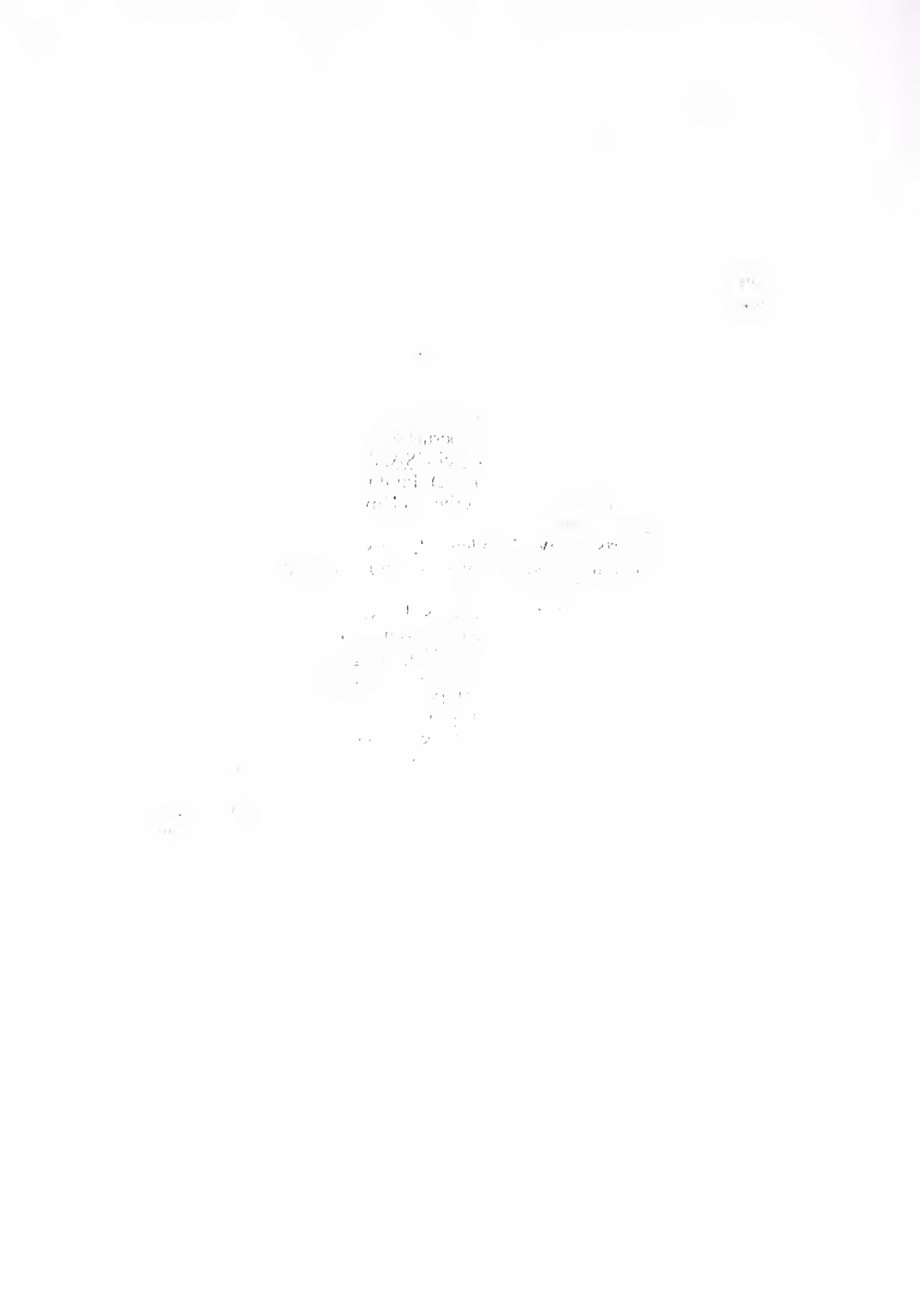
Nature, 349, 257-260.
39. Bernard, O., Guglielmi, P., Jonveaux, P., Cherif, D., Gisselbrecht, S., Mauchauffé, M., Larsen, C. J., Berger, R., & Mathieu-Mahul, D. (1990). Two distinct mechanisms for SCL gene activation in the t(1;14) translocation of T-cell leukemia. Genes Chromosomes Cancer, 1, 194.
40. Chen, Q., Cheng, Q. T., Tsai, L. H., Schneider, N., Buchanan, G., Carroll, A., Crist, W., Ozanne, B., Siciliano, M. J., & Baer, R. (1990). The tal gene undergoes chromosome translocation in T cell leukemia and potentially encodes a helix-loop-helix protein. EMBO Journal, 9, 145.
41. Begley, C. G., Aplan, P. D., Denning, S. M., Haynes, B. F., Waldmann, T. A., & Kirsch, I. R. (1989). The gene SCL is expressed during early hematopoiesis and encodes a differentiation-related DNA-binding motif. Proc Natl Acad Sci USA, 86, 10128-32.
42. Green, A. R., Salvaris, E., & Begley, G. C. (1991). Erythroid expression of the helix-loop-helix SCL gene. Oncogene, 6, 475.
43. Green, A. R., LInts, J. T., Visvader, J., Harvey, R., & Begley, C. G. (1992). SCL is co-expressed with GATA-1 in hematopoietic cells but is also expressed in developing brain. Oncogene, 7, 653.
44. Green, A. R., Rockman, S., DeLuca, E., & Begley, C. G. (1993). Induced myeloid differentiation of K562 cells with downregulation of erythroid and megakaryocytic transcription factors: a novel experimental model for hemopoietic lineage restriction. Experimental Hematology, 21, 525.
45. Green, A. R., & Begley, C. G. (1992). SCL and related hemopoietic helix-loop-helix transcription factors. Int. J. Cell Cloning, 10, 269.
46. Mouthon, M. A., Bernard, O., Mitlavila, M. T., Romeo, P. H., Vainchenker, W., & Mathieu-Mahul, D. (1993). Expression of tal-1 and GATA-binding proteins during human hematopoiesis. Blood, 81, 647-655.
47. Lawrence, H. J., & Largman, C. (1992). Homeobox genes in normal hematopoiesis and leukemia. . Blood, 80, 2445-53.
48. Gehring, W. J. (1985). The homeobox: a key to the understanding of development. Cell, 40, 3.
49. McGinnis, W., Levine, M. S., Hafen, E., Kuroiwa, A., & Gehring, W. J. (1984). A conserved DNA sequence in homeotic genes of the Drosophila Antennapedia and bithorax complexes. Nature, 308, 428.
50. Scott, M. P., & Weiner, A. J. (1984). Structural relationships among genes that control development: sequence homology between the Antennapedia, Ultrabithorax and fushi tarazu loci of Drosophila. Proc. Natl. Acad. Sci. U.S.A., 81, 4115.
51. Desplan, C., Theis, J., & O'Farrell, P. H. (1988). The sequence specificity of homeodomain-DNA interaction. Cell, 54, 1081-90.
52. Gehring, W. J. (1987). Homeoboxes in the study of development. Science, 236, 1245.
53. Hoey, T., & Levine, M. (1988). Homeobox proteins as sequence-specific transcription factors. Cell, 55, 537.
54. Harrison, S. C. (1991). A structural taxonomy of DNA-binding domains. Nature, 353, 715.
55. Billeter, M., Qian, Y., Otting, G., Muller, M., Gehring, W. J., & Wuthrich, K. (1990). Determination of the three-dimensional structure of the Antennapedia homeodomain from Drosophila in solution by 1H nuclear magnetic resonance spectroscopy. J Molecular Biology, 214, 183-97.
56. Kissinger, C. R., Liu, B. S., Martin-Blanco, E., Kornberg, T. B., & Pabo, C. O. (1990). Crystal structure of an engrailed homeodomain-DNA complex at 2.8 A resolution: a framework for understanding homeodomain-DNA interactions. Cell, 63, 579-90.
57. Kaufman, T. C., Seeger, M. A., & Olsen, G. (1990). Molecular and genetic

organization of the Antennapedia gene complex of Drosophila melanogaster. Adv. Genet., 27, 309.
58. Scott, M. P., Tamkun, J. W., & Hartzell, G. W. (1989). The structure and function of the homeodomain. Biochim Biophys Acta, 989, 25-48.
59. Bürglin, T. R., Finney, M., Coulson, A., & Ruvkun, G. (1989). Caenorhabditis elegans has scores of homeobox-containing genes. Nature, 341, 239.
60. Weeden, C. J., Kostriken, R. G., Matsumura, I., & Weisblat, D. A. (1990). Evidence for a new family of evolutionarily conserved homeobox genes. Nucleic Acids Res., 18, 1908.
61. Volbrecht, E., Veit, B., Sinha, N., & Hake, S. (1991). The developmental gene Knotted-1 is a member of a maize homeobox gene family. Nature, 350, 241.
62. Krumlauf, R., Holland, P. W., McVey, J. H., & Hogan, B. L. (1987). Developmental and spatial patterns of expression of the mouse homeobox gene, Hox 2.1. Development, 99, 603-17.
63. Dollé, P., Izpisua-Belmonte, J. C., Falkenstein, H., Renucci, A., & Duboule, D. (1989). Coordinate expression of the murine Hox-5 complex homeobox-containing genes during limb pattern formation. Nature, 324, 767.
64. Yokouchi, Y., Sasaki, H., & Kuroiwa, A. (1991). Homeobox gene expression correlated with the bifurcation process of limb cartilage development. Nature, 353, 443-5.
65. Dollé, P., Izpisua-Belmonte, J. C., Brown, J. M., Tickle, C., & Duboule, D. (1991). HOX-4 genes and the morphogenesis of mammalian genitalia. Genes Dev., 5, 1767.
66. Chisaka, O., & Capecchi, M. R. (1991). Regionally restricted developmental defects resulting from targeted disruption of the mouse homeobox gene hox-1.5. Nature, 350, 473-9.
67. Chisaka, O., Musci, T. S., & Capecchi, M. R. (1992). Developmental defects of the ear, cranial nerves and hindbrain resulting from targeted disruption of the mouse homeobox gene Hox-1.6 [see comments. Nature, 355, 516-20.
68. Lufkin, T., Dierich, A., LeMeur, M., Mark, M., & Chambon, P. (1991). Disruption of the Hox-1.6 homeobox gene results in defects in a region corresponding to its rostral domain of expression. Cell, 66, 1105-19.
69. Le, Mouellic. H., Lallemand, Y., & Brulet, P. (1992). Homeosis in the mouse induced by a null mutation in the Hox-3.1 gene. Cell, 69, 251-64.
70. Boncinelli, E., Acampora, D., Pannese, M., D'Esposito, M., Somma, R., Gaudino, G., Stornaiuolo, A., Cafiero, M., Faiella, A., & Simeone, A. (1989). Organization of human class I homeobox genes. Genome, 31, 745-56.
71. McAlpine, P. J., & Shows, T. B. (1990). Nomenclature for human homeobox genes [letter. Genomics, 7, 460.
72. Scott, M. P. (1992). Vertebrate homeobox gene nomenclature [letter. Cell, 71, 551-3.
73. Schughart, K., Kappen, C., & Ruddle, F. H. (1989). Duplication of large genomic regions during the evolution of vertebrate homeobox genes. Proc Natl Acad Sci USA, 86, 7067-71.
74. Acampora, D., Simeone, A., & Boncinelli, E. (1991). The human HOX homeobox genes. Oxford Surv. Euk. Genes, 7, 1.
75. Lufkin, T., Mark, M., Hart, C. P., Dolle, P., LeMeur, M., & Chambon, P. (1992). Homeotic transformation of the occipital bones of the skull by ectopic expression of a homeobox gene. Nature, 359, 835-41.
76. Stornaiuolo, A., Acampora, D., Pannese, M., D'Esposito, M., Morelli, F., Migliaccio, E., Rambaldi, M., Faiella, A., Nigro, V., Simeone, A., & et, a. (1990). Human HOX genes are differentially activated by retinoic acid in embryonal carcinoma cells according to their position within the four loci. Cell Differentiation Development, 31, 119-27.

77. Simeone, A., Acampora, D., Nigro, V., Faiella, A., D'Esposito, M., Stornaiuolo, A., Mavilio, F., & Boncinelli, E. (1991). Differential regulation by retinoic acid of the homeobox genes of the four HOX loci in human embryonal carcinoma cells. Mechanisms of Development, 33, 215-27.
78. Shen, W. F., Detmer, K., Simonitch-Eason, T. A., Lawrence, H. J., & Largman, C. (1991, February 11). Alternative splicing of the HOX 2.2 homeobox gene in human hematopoietic cells and murine embryonic and adult tissues. Nucleic Acids Research, 19(3), 539-45.
79. Lowney, P., Corral, J., LeBeau, M. M., Deaven, L., Lawrence, H. J., & Largman, C. (1991). A human Hox 1 homeobox gene exhibits myeloid-specific expression of alternative transcripts in human hematopoietic cells. Nucleic Acids Res., 19, 3443.
80. O'Connor, M. B., Binari, R., Perkins, L. A., & Bender, W. (1988). Alternative RNA products from the ultrabithorax domain of the bithorax complex. EMBO J., 7, 435.
81. Cho, K. W., Goetz, J., Wright, C. V., Fritz, A., Hardwicke, J., & DeRobertis EM. (1988). Differential utilization of the same reading frame in a Xenopus homeobox gene encodes two related proteins sharing the same DNA-binding specificity. EMBO J, 7, 2139-49.
82. LaRosa, G. J., & Gudas, L. J. (1988). Early retinoic acid-induced F9 teratocarcinoma stem cell gene ERA-1: alternate splicing creates transcripts for a homeobox-containing protein and one lacking the homeobox. Mol Cell Biol, 8, 3906-17.
83. Condie, B. G., Brivanlou, A. H., & Harland, R. M. (1990). Most of the homeobox-containing Xhox 36 transcripts in early Xenopus embryos cannot encode a homeodomain protein. Mol Cell Biol, 10, 3376-85.
84. Takeshita, K., Bollekens, J. A., Hijiya, N., Ratajczak, M., Ruddle, F. H., & Gewirtz, A. M. (1993). A homeobox gene of the Antennapedia class is required for human adult erythropoiesis. Proc Natl Acad Sci USA, 90, 3535-8.
85. Fortini, M. E., & Rubin, G. M. (1991). The Drosophila zfh-1 and zfh-2 genes encode novel proteins containing both zinc-finger and homeodomain motifs. Mech. Dev., 34, 113.
86. Morinaga, T., Yasuda, H., Hashimoto, T., Higashio, K., & Tamaoki, T. (1991). A human alpha-fetoprotein enhancer-binding protein, ATBF1, contains four homeodomains and seventeen zinc fingers. Mol. Cell. Biol., 11, 6041.
87. Bopp, D., Burri, M., Baumgartner, S., Frigerio, G., & Noll, M. (1986). Conservation of a large protein domain in the segmentation gene paired and in functionally related genes of Drosophila. Cell, 47, 1033.
88. Gruss, P., & Walther, C. (1992). Pax in development. Cell, 69, 719.
89. Deschamps, J., & Meijlink, F. (1992). Mammalian homeobox genes in normal development and neoplasia. . Critical Reviews in Oncogenesis, 3, 117-73.
90. Parslow, T. G., Blair, D. L., Murphy, W. J., & Granner, D. K. (1984). Structure of the 5' ends of immunoglobulin genes: a nove conserved sequence. Proc. Natl. Acad. Sci. U.S.A., 81, 2650.
91. Ko, H. S., Fast, P., McBride, W., & Staudt, L. M. (1988). A human protein specific for the immunoglobulin octamer DNA motif contains a functional homeobox domain. Cell, 55, 135-44.
92. Frain, M., Swart, G., Monaci, P., Nicosia, A., Stampfli, S., Frank, R., & Cortese, R. (1989). The liver-specific transcription factor LF-B1 contains a highly diverged homeobox DNA binding domain. Cell, 59, 145-57.
93. Bodner, M., Castrillo, J. L., Theill, L. E., Deerinck, T., Ellisman, M., & Karin, M. (1988). The pituitary-specific transcription factor GHF-1 is a homeobox-containing protein. Cell, 55, 505-18.
94. Ingraham, H. A., Chen, R. P., Mangalam, H. J., Elsholtz, H. P., Flynn, S. E.,

Lin, C. R., Simmons, D. M., Swanson, L., & Rosenfeld, M. G. (1988). A tissue-specific transcription factor containing a homeodomain specifies a pituitary phenotype. Cell, 55, 519-29.

95. Kongsuwan, K., Webb, E., Housiaux, P., & Adams, J. M. (1988). Expression of multiple homeobox genes within diverse mammalian haemopoietic lineages. EMBO J, 7, 2131-8.
96. Shen, W. F., Largman, C., Lowney, P., Corral, J. C., Detmer, K., Hauser, C. A., Simonitch, T. A., Hack, F. M., & Lawrence, H. J. (1989). Lineage-restricted expression of homeobox-containing genes in human hematopoietic cell lines. Proc Natl Acad Sci USA, 86, 8536-40.
97. Deguchi, Y., Moroney, J. F., Wilson, G. L., Fox, C. H., Winter, H. S., & Kehrl, J. H. (1991). Cloning of a human homeobox that resembles a diverged Drosophila homeobox gene and is expressed in activated lymphocytes. New Biol, 3, 353-363.
98. Magli, M. C., Barba, P., Celetti, A., Devita, G., Cillo, C., & Boncinelli, E. (1991). Coordinate regulation of HOX genes in human hematopoietic cells. Proc. Natl. Acad. Sci. U.S.A., 88, 6348-6352.
99. Matthews, C. H. E., Detmer, K., Boncinelli, E., Lawrence, H. J., & Largman, C. (1991). Erythroid-restricted expression of homeobox genes of the human HOX 2 locus. Blood, 78, 2248.
100. Vielle-Grosjean, I., Roullot, V., & Courtois, G. (1992). Lineage and stage specific expression of HOX 1 genes in the human hematopoietic system. Biochem. Biophys. Res. Comm., 183, 1124.
101. Petrini, M., Quaranta, M. T., Testa, U., Samoggia, P., Tritarelli, E., Care, A., Cianetti, L., Valtieri, M., Barletta, C., & Peschle, C. (1992). Expression of selected human HOX-2 genes in B/T acute lymphoid leukemia and interleukin-2/interleukin-1 beta-stimulated natural killer lymphocytes. Blood, 80, 185-93.
102. Shen, W. F., Detmer, K., Mathews, C. H., Hack, F. M., Morgan, D. A., Largman, C., & Lawrence, H. J. (1992). Modulation of homeobox gene expression alters the phenotype of human hematopoietic cell lines. EMBO J, 11, 983-9.
103. Deguchi, Y., & Kehrl, J. H. (1991). Selective expression of two homeobox genes in CD34-positive cells from human bone marrow. Blood, 78, 323-8.
104. Allen, J. D., Lints, T., Jenkins, N. A., Copeland, N. G., Strasser, A., Harvey, R. P., & Adams, J. M. (1991). Novel murine homeobox gene on chromosome 1 expressed in specific hematopoietic lineages and during embryogenesis. Genes Dev, 5, 509-520.
105. Takeshita, K., Bollekens, J. A., Landesberg, L., Graubert, T., Berliner, N., Hoffman, R., Ruddle, F. H., Yazaki, Y., & H. Hirai, H. (1994). Function of divergent homeobox genes during myelopoieis. submitted,
106. Xu, Y., Baldassare, M., Fisher, P., Rathbun, G., Oltz, E. M., Yancopoulos, G. D., Jessell, T. M., & Alt, F. W. (1993, January 1). LH-2: a LIM/homeodomain gene expressed in developing lymphocytes and neural cells. Proc Natl Acad Sci USA, 90(1), 227-31.
107. Inamori, K., Takeshita, K., Chiba, S., Yazaki, Y., & Hirai, H. (1993). Identification of homeobox genes expressed in human T-lymphocytes. Biochem Biophys Res Commun, 196, 203-208.
108. Hromas R., Radich, J., & Collins, S. (1993). PCR cloning of an orphan homeobox gene (PRH) preferentially expressed in myeloid and liver cells. Biochem. and Biophys. Res. Comm., 195, 976.
109. Bedford, F. K., Ashworth, A., Enver, T., & Wiedemann, L. M. (1993). HEX: a novel homeobox gene expressed during haematopoiesis and conserved between mouse and human. Nucleic Acids Research, 21, 1245-9.
110. Crompton, M. R., Bartlett, T. J., MacGregor, A. D., Manfioletti, G., Buratti, E.,



Giancotti, V., & Goodwin, G. H. (1992). Identification of a novel vertebrate homeobox gene expressed in hematopoietic cells. Nucl Acids Res, 20, 5661-5667.
111. Matsui, T., Hirai, M., Wakita, M., Hirano, M., & Kurosawa, Y. (1993, May 10). Expression of a novel human homeobox-containing gene that maps to chromosome 7q36.1 in hematopoietic cells. FEBS Letters, 322(2), 181-5.
112. Goldsborough, A. S., Healy, L. E., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., Willison, K. R., & Ashworth, A. (1993). Cloning, chromosomal localization and expression pattern of the POU domain gene Oct-11. Nucleic Acid Res., 21, 127.
113. Kongsuwan, K., Allen, J., & Adams, J. M. (1989). Expression of Hox 2.4 gene directed by proviral insertion in myeloid leukemia. Nucl Acids Res, 17, 1881-1892.
114. Ben-David, L., Aberdam, D., Sachs, L., & Blatt, C. (1991). A deletion and a rearrangement distinguish between the intracisternal A-particle of Hox-2.4 and that of interleukin-3 in the same leukemic cells. Virology, 182, 382-7.
115. Perkins, A., Kongsuwan, K., Visvader, J., Adams, J. M., & Cory, S. (1990). Homeobox gene expression plus autocrine growth factor production elicits myeloid leukemia. Proc Natl Acad Sci USA, 87, 8398-402.
116. Blatt, C., Lotem, J., & Sachs, L. (1992). Inhibition of specific pathways of myeloid cell differentiation by an activated Hox-2.4 homeobox gene. Cell Growth Differentiation, 3, 671-6.
117. Hatano, M., Roberts, C. W., Minden, M., Crist, W. M., & Korsmeyer, S. J. (1991). Deregulation of a homeobox gene, HOX11, by the t(10;14) in T cell leukemia. Science, 253, 79-82.
118. Kennedy, M. A., Gonzalez, S. R., Kees, U. R., Lampert, F., Dear, N., Boehm, T., & Rabbits, T. H. (1991). HOX 11, a homeobox-containing T-cell oncogene on human chromosome 10q24. Proc. Natl. Acad. Sci. U.S.A., 88, 8900.
119. Lu, M., Zhang, N., & Ho, A. D. (1992). Genomic organization of the putative human homeobox proto-oncogene HOX-11 (TCL-3) and its endogenous expression in T cells. Oncogene, 7, 1325-.
120. Dube, I. D., Noble, L., & Kamel-Reid, S. (1992). Functional studies of HOX 11, the homeobox gene involved in the (10;14) translocation of T cell neoplasias. J. Cell Biochem., 16C, 73.
121. Carroll, A. J., Crist, W. M., Parmley, R. T., Roper, M., Cooper, M. D., & Finley, W. M. (1984). Pre-B cell leukemia associated with chromosome translocation. Blood, 63, 721.
122. Williams, D. L., Look, A. T., Melvin, S. L., Roberson, P. K., Dahl, G., Flake, T., & Stass, S. (1984). New chromosomal translocations correlate with specific immunophenotypes of childhood acute lymphoblastic leukemia. Cell, 36, 101.
123. Nourse, J., Mellentin, J. D., Galili, N., Wilkinson, J., Stanbridge, E., Smith, S. D., & Cleary, M. L. (1990). Chromosomal translocation t(1;19) results in synthesis of a homeobox fusion mRNA that codes for a potential chimeric transcription factor. Cell, 60, 535-45.
124. Kamps, M. P., Murre, C., Sun, X. H., & Baltimore, D. (1990). A new homeobox gene contributes the DNA binding domain of the t(1;19) translocation protein in pre-B ALL. Cell, 60, 547-55.
125. Kamps, M. P., Look, A. T., & Baltimore, D. (1991). The human t(1;19) translocation in pre-B ALL produces multiple nuclear E2A-Pbx1 fusion proteins with differing transforming potentials. Genes Dev., 5, 358.
126. Deguchi, Y., Kirschenbaum, A., & Kehrl, J. H. (1992). A diverged homeobox gene is involved in the proliferation and lineage commitment of human hematopoietic progenitors and highly expressed in acute myelogenous leukemia. Blood, 79, 2841-8.
127. Deguchi, Y., Thevenin, C., & Kehrl, J. H. (1992). Stable expression of HB24, a

diverged human homeobox gene, in T lymphocytes induces genes involved in T cell activation and growth. J Biol Chem, 267, 8222-8229.
128. Deguchi, Y., & Kehrl, J. H. (1993). High level expression of the homeobox gene HB24 in a human T-cell line confers the ability to form tumors in nude mice. Cancer Research, 53, 373-7.
129. Ohyashiki, K. (1993). unpublished.
130. Koeffler, H. P., & Golde, D. W. (1978). Acute myelogenous leukemia: A human cell line responsive to colony stimulating activity. Science, 200, 1153-1154.
131. Collins, S. J., Gallo, R. C., & Gallagher, R. E. (1977). Continuous growth and differentiation of human myeloid cells in suspension. Nature, 270, 347-349.
132. Ogura, M., Morishima, Y., Ohno, R., Hirabayashi, N., Nagura, H., & Saito, H. (1985). Establishment of a novel human megakaryoblastic leukemia cell line, Meg-01, with positive Philadelphia chromosome. Blood, 66, 1384.
133. Matsui, Y., Zsebo, K., & Hogan, B. L. (1992). Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. Cell, 70, 841-7.
134. Friend, C., Scher, W., Holland, J. G., & Sato, T. (1971). Hemoglobin synthesis in murine virus-induced leukemic cells in vitro: stimulation of erythroid differentiation by dimethylsulfoxide. Proc. Natl. Acad. Sci. U.S.A., 68, 378.
135. Chomczynski, P., & Sacchi, N. (1987). Single step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. Anal Biochem, 162, 156-159.
136. Murtha, M. T., Leckman, J. F., & Ruddle, F. H. (1991). Detection of homeobox genes in development and evolution. Proc Natl Acad Sci USA, 88, 10711-10715.
137. DiLella, A. G., & Woo, S. L. C. (1987). Hybridization of genomic DNA to oligonucleotide probes in the presence of tetramethylammonium chloride. Methods in Enzymology, 152, 447-451.
138. Smoller, D. A., Petrov, D., & Hartl, D. L. (1991). Charaterization of bacteriophage P1 library containing inserts of Drosophila DNA of 75-100 kilobase pairs. Chromosoma, 100, 487-494.
139. Pierce, J. C., Sternberg, N., & Sauer, B. (1992). A mouse genomic library in the bacteriophage P! cloning system: Organization and characterization. Mammalian Genome, 3, 550-558.
140. Sternberg, N. L. (1992). Cloning high molecular weight DNA fragments by the bacteriophage P1 system. Trends Genetics, 8, 10-16.
141. Mackenzie, A., Leeming, G. L., Jowett, A. K., Ferguson, M. W., & Sharpe, P. T. (1991). The homeobox gene Hox 7.1 has specific regional and temporal expression patterns during early murine craniofacial embryogenesis, especially tooth development in vivo and in vitro. Development, 111, 269-85.
142. Hill, R. E., Jones, P. F., Rees, A. R., Sime, C. M., Justice, M. J., Copeland, N. G., Jenkins, N. A., Graham, E., & Davidson, D. R. (1989). A new family of mouse homeo box-containing genes: Molecular structure, chromosomal location and developmental expression of HOX 7.1. Genes Dev, 3, 26-37.
143. Dush, M. K., & Martin, G. R. (1992). Analysis of mouse Evx genes: Evx-1 displays graded expression in the primitive streak. Developmental Biology, 151, 273-87.
144. Bastian, H., & Gruss, P. (1990). A murine even-skipped homologue, Evx 1, is expressed during early embryogenesis and neurogenesis in a biphasic manner. EMBO J, 9, 1839-52.
145. Zhao, G., Zhou, X., Eberspaecher, H., Solursh, M., & DeCrombrugghe, B. (1993). Cartilage homeoprotein 1, a homeoprotein selectively expressed in chondrocytes. Proc. Natl. Acad. Sci. U.S.A., 90, 8633.
146. Opstelten, D. J., Vogels, R., Robert, B., Kalkhoven, E., Zwartkruis, F., de, L. a. L., Destree, O. H., Deschamps, J., Lawson, K. A., & Meijlink, F. (1991). The

mouse homeobox gene, S8, is expressed during embryogenesis predominantly in mesenchyme. Mechanisms of Development, 34, 29-41.
147. Hawkins, N. C., & McGhee, J. D. (1990). Homeobox-containing genes in the nematode Caenorhabditis. Nucl Acids Res, 18, 6101-6106.
148. Finkelstein, R., & Perrimon, N. (1990). The orthodenticle gene is regulated by bicoid and torso and specifies Drosophila head development [see comments. Nature, 346, 485-8.
149. Baumgartner, S. D., Bopp, D., Burri, M., & Noll, M. (1987). Structure of two genes at the gooseberry locus related to the paired gene and their spatial expression during Drosophila embryogenesis. Genes Dev., 1, 1247.
150. Webster, P. J., & Mansour, T. E. (1992). Conserved classes of homeodomains in Schistosoma mansoni, an early bilateral metazoan. Mechanisms of Development, 38, 25-32.
151. Bürglin, T. R. (1994). The homeodomain phylum: A comprehensive survey of homeodomain sequences and associated motifs. (Unpublished).
152. Amaravadi, L., & King, M. W. (1994). A rapid and efficient nonradioactive method for screening recombinant DNA libraries. BioTechniques, 16, 98-103.
153. Sugimoto, K., Toyoshima, H., Sakai, R., & Hirai, H. (1992). Frequent mutations in the p53 gene in human myeloid leukemia cell lines. Blood, 79, 2378-2383.
154. Tassabehji, M., Read, A. P., Newton, V. E., Harris, R., Balling, R., Gruss, P., & Strachan, T. (1992). Waardenburg's syndrome patients have mutations in the human homologue of the Pax-3 paired box gene [see comments. Nature, 355, 635-6.
155. Baldwin, C. T., Hoth, C. F., Amos, J. A., da-Silva, E. O., & Milunsky, A. (1992). An exonic mutation in the HuP2 paired domain gene causes Waardenburg's syndrome. Nature, 355, 637.
156. Burri, M., Tromvoukis, Y., Bopp, D., Frigerio, G., & Noll, M. (1989). Conservation of the paired domain in metazoans and its structure in three isolated human genes. EMBO J., 8, 1183.
157. Nelson, L. B., Spaeth, G. L., Nowinski, T. S., Margo, C. E., & Jackson, L. (1984). Aniridia. A review. Surv. Opthalmol., 28, 621.
158. Shaw, M. W., Falls, H. F., & Neel, J. V. (1960). Congenital aniridia. Am. J. Hum. Genet., 12, 389.
159. Ton, C. C., Hirvonen, H., Miwa, H., Weil, M. M., Monaghan, P., Jordan, T., van, H. e. V., Hastie, N. D., Meijers-Heijboer, H., Drechsler, M., & et, a. (1991). Positional cloning and characterization of a paired box- and homeobox-containing gene from the aniridia region. Cell, 67, 1059-74.
160. Maulbecker, C. C., & Gruss, P. (1993). The oncogenic potential of Pax genes. EMBO J, 12, 2361-2367.
161. Barr, F. G., Galili, N., Holick, J., Biegel, J. A., Rovera, G., & Emanuel, B. S. (1993). Rearrangement of the PAX3 paired box gene in the paediatric solid tumour alveolar rhabdomyosarcoma. Nature Genetics, 3, 113-7.
162. Shapiro, D. N., Sublett, J. E., Baitao, L., Downing, J. R., & Naeve, C. W. (1993). Fusion of PAX 3 to a member of the Forkhead family of transcription factors in human alveolar rhabdomyosarcoma. Cancer Research, 53, 5108.
163. Dressler, G. R., & Douglass, E. C. (1992). Pax-2 is a DNA-binding protein expressed in embryonic kidney and Wilms tumor. Proc. Natl. Acad. Sci. U.S.A., 89, 1179.
164. Dressler, G. R., Wilkinson, J. E., Rothenpieler U.W., Patterson, L. T., Williams-Simons, L., & Westphal, H. (1993). Deregulation of Pax-2 expression in transgenic mice generates severe kidney abnormalities. Nature, 362, 65-7.
165. Plachov, D., Chowdhury, K., Walther, C., Simon, D., Guenet, J. -., & Gruss, P. (1990). Pax 8, a murine paired box gene expressed in the developing excretory system and thyroid gland. Development, 110, 643.

166. Karlsson, O., Thor, S., Norberg, T., Ohlsson H., & Edlund, T. (1990). Insulin gene enhancer binding protein Isl-1 is a member of a novel class of proteins containing both a homeo- and a Cys-His domain. Nature, 344, 879-882.
167. Barad, M., Jack, T., Chadwick, R., & McGinnis, W. (1988). A novel, tissue-specific, Drosophila homeobox gene. EMBO J, 7, 2151-61.
168. Cohen, S. M., Bronner, G., Kuttner, F., Jurgens, G., & Jackle, H. (1989). Distal-less encodes a homeodomain protein required for limb development in Drosophila. Nature, 338, 432-434.
169. Bürglin, T. R., Finney, M., Coulson, A., & Ruvkun, G. (1989). Caenorhabditis elegans has scores of homeobox-containing genes. Nature, 341, 239-243.
170. Kim, Y., & Nirenberg, M. (1989). Drosophila NK-homeobox genes. Proc Natl Acad Sci USA, 86, 7716-20.
171. Nohno, T., Noji, S., Koyama, E., Nishikawa, K., Myokai, F., Saito, T., & Taniguchi, S. (1992). Differential expression of two msh-related homeobox genes Chox-7 and Chox-8 during chick limb development. Biochem Biophys Res Commun, 182, 121-8.
172. Andres, V., Nadal-Ginard, B., & Mahdavi, V. (1992). Clox, a mammalian homeobox gene related to Drosophila cut, encodes DNA-binding regulatory proteins differentially expressed during development. Development, 116, 321-34.
173. Rangini, Z., Frumkin, A., Shani, G., Guttmann, M., Eyal-Giladi, H., Gruenbaum, Y., Fainsod, & A. (1989). The chicken homeo box genes CHox1 and CHox3: cloning, sequencing and expression during embryogenesis. Gene, 76, 61-74.
174. Dalton, D., Chadwick, R., & McGinnis, W. (1989). Expression and embryonic function of empty spiracles: A Drosophila homeobox gene with two patterning functions on the anterior-posterior axis of the embryo. Genes Dev, 3, 1940-1956.