(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
22 November 2007 (22.11.2007)

PCT

(10) International Publication Number
WO 2007/131280 A1

(51) International Patent Classification:
C12N 15/11 (2006.01) *A61K 38/17* (2006.01)
A61P 7/00 (2006.01) *C12N 5/10* (2006.01)

(21) International Application Number:
PCT/AU2007/000645

(22) International Filing Date: 14 May 2007 (14.05.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2006902556 12 May 2006 (12.05.2006) AU

(71) Applicant *(for all designated States except US)*: **THE WALTER AND ELIZA HALL INSTITUTE OF MEDICAL RESEARCH** [AU/AU]; 1G Royal Parade, Parkville, Victoria 3052 (AU).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **HILTON, Douglas, James** [AU/AU]; 244 Research Road, Warrandyte, Victoria 3113 (AU). **ALEXANDER, Warren, Scott** [AU/AU]; 13 Park Street, Moonee Ponds, Victoria 3039 (AU). **KILE, Benjamin, Thomas** [AU/AU]; 618 Rathdowne Street, Carlton, Victoria 3054 (AU). **LOUGHRAN, Stephen, John** [AU/AU]; 27 Blair Street, Brunswick, Victoria 3056 (AU).

(74) Agents: **HUTCHISON, Jane, E.** et al.; Davies Collison Cave, 1 Nicholson Street, Melbourne, Victoria 3000 (AU).

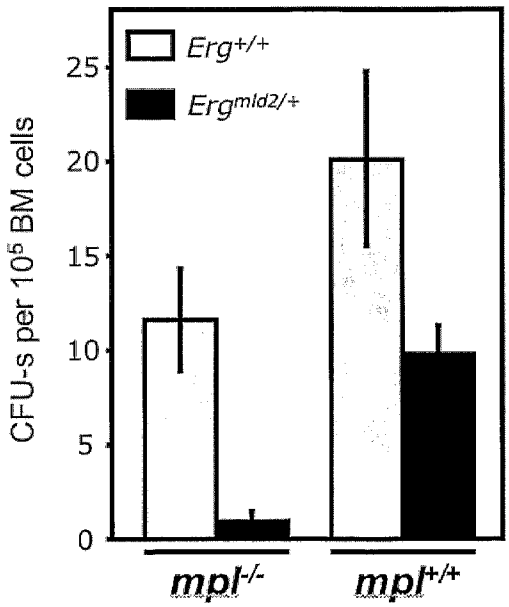
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

[Continued on next page]

(54) **Title:** HEMATOPOIESIS AND REGULATION THEREOF BY ETS RELATED GENE (Erg)



(57) **Abstract:** Methods are provided for modulating hematopoietic cell activity comprising modulating the level or activity of ets-related gene (Erg) polypeptide or a transcriptional target of Erg polypeptide or a downstream effector of Erg polypeptide activity. Hematopoietic cells also include erthrocytes and leukocytes. Hematopoietic cells include hematopoietic stem cells (HSCs) and hematopoietic progenitor cells. Hematopoietic cell activity includes one or more functions including homing, engrafting, survival, self-renewal and differentiation. Methods are also provided for screening for modulators of hematopoietic cell activity. Modulators based *inter alia* on *Erg* gene-silencing or the administration of Erg polypeptide-encoding sequences are also described.

WO 2007/131280 A1

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

HEMATOPOIESIS AND REGULATION THEREOF BY ETS RELATED GENE (Erg)

BACKGROUND OF THE INVENTION

FIELD OF THE INVENTION

The present invention relates generally to a model system providing methods and modulators suitable for use in modulating hematopoietic cell activity and for use in the treatment, prophylaxis and diagnosis of diseases or conditions associated with aberrant hematopoiesis and/or hematopoietic cell activity. Diseases or conditions including hematological disorders or genetic conditions and their sequelae including cancer and autoimmune conditions are particularly contemplated. The present invention provides modified cells or a non-human animal comprising modified cells as models of aberrant hematopoiesis and/or early blood cell function wherein the activity of an ETS-related transcription factor is modified in the cell or animal relative to the activity in an unmodified cell or animal. The present invention further provides agents and assays for the identification and development of agents that modulate the ets related gene (Erg) signalling pathway and Erg associated transcription suitable for use in a range of therapeutic, prophylactic and medical or veterinary applications.

DESCRIPTION OF THE PRIOR ART

Bibliographic details of references in the subject specification are also listed at the end of the specification.

The reference in this specification to any prior publication (or information derived from it), or to any matter which is known, is not, and should not be taken as an acknowledgment or admission or any form of suggestion that that prior publication (or information derived from it) or known matter forms part of the common general knowledge in the field of endeavour to which this specification relates.

Stem cells are undifferentiated cells that give rise to a succession of mature functional cells. Regulation of these events is critical to the body's ability to respond to its environment, to heal and thrive. Dysregulation of stem cell homeostatic mechanisms can lead *inter alia* to cancer.

Hematopoietic stem cells (HSC) and early multipotent hematopoietic progenitors are of great biomedical interest. Much of this is directly due to their role in the generation of all mature blood cells, which implicates them in hematological diseases including blood cell deficiencies and leukemias, and makes them of significant therapeutic value for transplantation. In addition, HSCs are the best characterised adult stem cell population, and therefore serve as a model system in which the regulation of stem cell self renewal and differentiation may be investigated. These functions are critical to medically significant processes such as cancer and tissue regeneration. However, the mechanisms which regulate the self-renewal of HSCs and their differentiation into multipotent progenitors are only vaguely characteristics.

Erg (ets related gene) encodes a member of the Erythroblast Transformation Specific (ETS) family of transcription factors, which bind to 5'-GGA(A/T)-3' core recognition motifs in enhancers and promoters to activate transcription (reviewed in Sharrocks, *Nat. Rev. Mol. Cell Biol., 2*:827-837, 2001). It contains two distinct domains. Towards the amino-terminus is a pointed domain, the specific function of which has not been determined in Erg, but which has been associated with protein-protein interaction in other ETS family members (Kim *et al.*, *EMBO J., 20*:4173-4182, 2001; Carrere *et al.*, *Oncogene, 16*:3261-3268, 1998). Near the carboxyl-terminus is an ETS domain, which is definitive for all ETS family members, and has a winged helix-turn-helix topology. The ETS domain binds DNA *via* its third helix, and is critical to the function of Erg (Carrere *et al.*, 1998 (*supra*)); Kodandapani *et al.*, *Nature. 380*:456-460, 1996). Erg may be phosphorylated at serine residues, but the mechanism and function of such phosphorylation is not yet known (Murakami *et al.*, *Oncogene, 8*:1559-1566, 1993).

Erg is expressed in various mesodermal cells and neural crest cells during embryogenesis, T cell precursors, endothelial cells, platelets and HSCs (Vlaeminck-Guillem *et al.*, *Mech. Dev., 91*:331-335, 2000; Anderson *et al.*, *Development, 126*:3131-3148, 1998; McLaughlin *et al.*, *J. Cell Sci., 112(Pt 24)*:4695-4703, 1999; Rainis *et al.*, *Cancer Res., 65*:7596-7602, 2005). To date, the biological function of Erg has not been characterised. One study showed that insertion of an *Erg* transgene into the developing limb buds of the chick was sufficient to maintain chondrocytes (cartilage-forming cells) in an immature state and prevent the replacement of cartilage with bone (Iwamoto *et al.*, *J. Cell Biol., 150*:27-40, 2000). This implicates Erg in the regulation of cartilage formation, consistent with its expression pattern in the embryo (Vlaeminck-Guillem *et al.*, 2000 (*supra*); Dhordain *et al.*, *Mech. Dev., 50*:17-28, 1995). Another study demonstrated that knocking down *Erg* expression in cultured endothelial cells altered the expression of key angiogenic genes and prevented the formation of tubular vessel structures, implicating Erg in angiogenesis (McLaughlin *et al.*, *Blood, 98*:3332-3339, 2001).

Erg has also been implicated in cancer. Overexpression of Erg in NIH3T3 cells allows growth in low serum and independent of adhesion to a surface, as well as giving them the capacity to form tumours in nude mice, demonstrating that Erg is a proto-oncogene with transforming capabilities (Hart *et al.*, *Oncogene, 10*:1423-1430, 1995). In 5-10% of patients with the paediatric bone cancer Ewing's sarcoma, a t(21;22) chromosomal translocation which fuses the *EWS* and *Erg* genes can be detected (Sorensen *et al.*, *Nat. Genet., 6*:146-151, 1994). In 80% of remaining Ewing's sarcomas, EWS is fused with FLI-1, the most closely related paralog of Erg. EWS is an RNA binding protein that can associate with the RNA polymerase II complex responsible for transcription, but its cellular function is unclear. The resulting EWS-Erg fusion protein consists of the N-terminal region of EWS and the C-terminal region of Erg, including the ETS domain. Carcinogenesis is thought to occur because the EWS region confers aberrantly strong transactivation activity while the ETS domain of Erg continues to bind DNA (reviewed in Arvand *et al.*, *Oncogene, 20*:5747-5754, 2001).

A similar chromosomal translocation found in several types of myeloid leukemia – t(16,21) – fuses TLS/FUS and Erg (Shimizu *et al., Proc. Natl. Acad. Sci. U. S. A., 90*:10280-10284, 1993; Ichikawa *et al., Cancer Res., 54*:2865-2868, 1994). TLS/FUS is closely related to EWS and also binds to RNA and RNA polymerase II (Yang *et al., Mol. Cell. Biol., 20*:3345-3354, 2000). The TLS-Erg fusion protein is sufficient to cause a leukemic phenotype when introduced to HSCs (Pereira *et al. Proc. Natl. Acad. Sci. U. S. A., 95*:8239-8244, 1998). There is evidence that this may occur due to the fusion protein causing inappropriate transactivation or by inhibiting RNA splicing (Yang *et al.*, 2000 (*supra*); Perrotti *et al., EMBO J., 17*:4442-4455, 1998). Further, in acute myeloid leukemias with complex karyotypes, hybridisation of genomic DNA to a bacterial artificial chromosome array demonstrated that the region of the genome encoding Erg was present in high copy number, and that *Erg* was overexpressed (Baldus *et al., Proc. Natl. Acad. Sci. U.S.A., 101*:3915-3920, 2004). *Erg* expression has also been observed in megakaryoblastic leukemia cell lines (Rainis *et al.*, 2005 (*supra*)).

Erg has been implicated generally in physiological hematopoiesis. One study examined the expression of Erg in cell populations purified from mice and found that *Erg* expression was induced during T lineage specification, then silenced following commitment, suggesting that Erg may play a role in T cell differentiation (Anderson *et al.*, 1998 (*supra*)). Another study used the K562 erythroleukemic cell line to demonstrate that Erg may influence lineage commitment in multipotential hematopoietic cells (Rainis *et al.*, 2005 (*supra*)). K562 cells usually differentiate down the erythroid lineage, but can be induced to undergo megakaryocytic differentiation with phorbol esters. *Erg* expression was induced during such megakaryocytic differentiation. Conversely, the overexpression of Erg in K562 cells was sufficient to activate the expression of the megakaryocytic genes gpIb, gpIIb and gpIIIa, and the megakaryocytic antigens CD41 and CD61, suggesting that Erg expression may influence differentiation from an erythroid to a megakaryocytic lineage. None of these studies indicate a critical role for Erg in regulation of early events in hematopoiesis or suggest an application for Erg in the treatment or prophylaxis of conditions characterised by multi-lineage defects. Indeed, loss of function studies have failed to demonstrate any affect of signalling pathways on early HSC activity.

There is a need in the art to identify the molecules that have a key regulatory role in hematopoiesis and cellular function and development in order to devise therapeutic and/or prophylactic strategies for use when these functions are impaired. By defining a single gene which, when modified in an animal model, causes disruption of very early events in blood cell development and particularly hematopoietic stem cell activity, the present invention provides screening methods and a class of molecules that are useful for modulating HSC activities, such as, self-renewal, homing and/or engrafting and/or modulating hematopoietic progenitor cell activity, or for the treatment and/or prophylaxis of conditions characterised by defective blood cell activity or associated with aberrant hematopoiesis.

SUMMARY OF THE INVENTION

Throughout this specification, unless the context requires otherwise, the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element or integer or group of elements or integers but not the exclusion of any other element or integer or group of elements or integers.

As used herein the singular forms "a", "an" and "the" include plural aspects unless the context clearly dictates otherwise. Thus, for example, reference to "a cell" includes a single cell, as well as two or more cells; reference to "an agent" includes one agent, as well as two or more agents; and so forth.

Nucleotide and amino acid sequences are referred to by a sequence identifier number (SEQ ID NO:). The SEQ ID NOs: correspond numerically to the sequence identifiers <400>1 (SEQ ID NO:1), <400>2 (SEQ ID NO:2), etc. A summary of sequence identifiers is provided in Table 1. A sequence listing is provided after the claims.

Genes and other genetic material (eg mRNA, nucleic acid constructs etc) are represented herein in italics while their proteinaceous expression products are represented in non-italicised form. Thus, Erg polypeptide is the expression product of *Erg* nucleic acid sequences. The terms Erg or *Erg* encompass all homologs in any animal species including human homologs. Representative examples of the nucleic acid and amino acid sequences of Erg molecules are provided in the sequence listing further described in Table 1. As the skilled addressee will appreciate Erg protein in mouse is highly homologous to Erg in human (96% identity). Due to this close similarity and the close evolutionary relationship between Erg homologs, the present invention encompasses human and other mammalian species.

The present invention is predicated, in part, on the identification of an E26 (ETS) family transcription factor, Erg (ets related gene) as a key molecule required for early hematopoietic cell activity in a mammalian animal model. Specifically, reduction of Erg

activity in a mammalian animal resulted in animals exhibiting *inter alia* thrombocytopenia (low platelet (thrombocyte) counts). Low platelet counts are associated with an increased risk of haemorrhage and as shown herein mammals with one mutant *Erg* allele showed reduced survival, internal haemorrhaging under the skin, around the joints, and skill (see Figure 14). In a sensitised animal model system, disruption of *Erg* resulted in animals exhibiting *inter alia* pancytopenia and bone marrow failure through lack of early hematopoietic progenitor cells. In the absence of a functional Mpl polypeptide, mice with one mutant *Erg* allele become anaemic, leukopenic and thrombocytopenic (see Figure 10). These mice are also profoundly deficient in the multipotent early hematopoietic progenitors that form spleen colonies indicating that without an adequate presence of Erg, hematopoiesis is disrupted through a profound deficiency in HSCs and/or multipotent progenitors (see Figure 16). That this deficiency is manifest at the earliest stages of hematopoiesis is shown by study of the yolk sac and fetal livers in mammals exhibiting reduced Erg functional activity. As shown in Tables 8 and 9, these mammals showed reduced hematopoietic progenitors consistent with a hematopoietic stem cell defect. In the presence of Mpl, the platelet deficiency is obscured but mice remain substantially thrombocytopenic and leukopenic.

Bone marrow (BM) transplant of Erg-defective BM cells shows that the HSCs and hematopoietic progenitors are substantially defective in one or more of their functional activities such as, but not limited to, homing, engrafting, survival (i.e. not undergoing apoptosis) survival, self-renewal and differentiation (see Figure 17 and Brief Description of the Figures). In contrast, Erg-normal BM cells, even on a background of no thrombopoietin facilitated normal hematopoiesis in myeloablated subjects (see Figure 18 and Brief Description of the Figures).

In one aspect, the present invention provides agents (modulators) that modulate the Erg signalling pathway. In a particular embodiment, the subject agents are useful for modulating the development and activity of multipotent progenitor cells including hematopoietic stem cells (HSC). Hematopoietic stem cells are undifferentiated cells that give rise to a succession of mature functional blood cells as found in normal adults. This

process is described as definitive hematopoiesis. In adults, HSC reside in the bone marrow, peripheral blood, liver, spleen and other organs. HSC are the first in a hierarchy of progenitor cells. They are capable of long-term self renewal (long term (LT)-HSCs). LT-HSCs differentiate into short-term multipotent HSCs, (ST-HSCs) that retain the ability to produce all blood types but only proliferate for a relatively short time. Next, lymphoid progenitors arise that ultimately produce immune cells, and myeloid progenitors arise that ultimately produce mainly red blood cells and platelets and some innate immune cells. These progenitor cells have various abilities to proliferate and differentiate and from these cells ultimately arise terminally differentiated cells. Accordingly, reference to HSC and hematopoietic progenitors include all the above mentioned progenitor cells and reference to hematopoietic or blood cells include any of their terminally differentiated descendants. In other embodiments, the agents are useful to treat or prevent aberrant hematopoiesis including aberrant haematopoietic cell activity.

In some embodiments, the modulator enhances or down regulates the activity of Erg polypeptide, or a transcriptional target of Erg polypeptide or a down system effector of Erg polypeptide activity in a cell. In some embodiments, a transcriptional target is a gene to which Erg polypeptide binds or its expression product. In other embodiments, the transcriptional target is the regulatory region of a gene to which Erg polypeptide binds to modulate transcription.

In some embodiments, upregulation of the activity of Erg in a subject using Erg polypeptide or a variant thereof or an agent from which an Erg polypeptide or variant thereof is producible or an agent that effectively enhances Erg activity are proposed to be useful for enhancing hematopoietic cell activity or for treating or preventing conditions associated with defective hematopoietic cell activity. In some embodiments, the hematopoietic cell is an erythrocyte, in other embodiments, a leukocyte. In a preferred embodiment, the cell is a HSC or a hematopoietic progenitor cell.

In other embodiments, agents that down regulate the functional activity of Erg in a cell are proposed for use in lowering platelet levels in a subject. Elevated platelet levels

(thrombocytosis) may increase the risk of blot clots and is observed after surgery, in iron deficient anaemia and myeloproliferative diseases.

In some embodiments, agents that down regulate the functional activity of Erg are proposed for use in treating or preventing clonal hemopathies including those that result from overactivity or dysregulated replication of stem cells such as found in cancer.

In some embodiments, the subject agents bind to *Erg* i.e. all or part of coding, non-coding or regulatory regions in *Erg* DNA or RNA and modulate gene expression at transcriptional or post-transcriptional, including translational stages. In other embodiments, the agent binds to a transcriptional target of Erg and modulates its activity. Genetic agents which reduce the activity of Erg polypeptides or a transcription target of Erg in a cell include genetic agents (i.e. comprising a nucleic acid molecule) which inhibit production of Erg in a cell at any stage including, for example, post-transcriptional silencing mediated by RNAi (see for example, United States Publication No. 20070042983, International Publication No. WO 01/68836 and International Publication No. WO 03/064626). Such nucleic acids can be chemically synthesised, expressed from a vector or enzymatically synthesised as known in the art.

The terms "agents", "agents", "modulator" or "modulators" are used interchangeably and refer to compositions which effectively modulate, directly or indirectly, the activity of Erg polypeptide or one or more of its transcriptional targets in a cell. As the skilled artisan will appreciate, a broad range of modulators will be effective in modulating Erg activity or the activity of its transcriptional targets.

In another aspect, the present invention provides methods for screening and testing agents for their ability to modulate Erg activity including Erg signalling pathways and/or Erg-associated transcription. In some embodiments, Erg regulates the expression of genes encoding hematopoietic signalling molecules such as cytokine, chemokine, hormones or their cellular or nuclear receptors or transcription factors. In some embodiments, Erg regulates the expression of genes encoding signalling molecules or their receptors or

transcription factors that are expressed in the early stages of hematopoietic stem cell development and differentiation. Exemplary transcriptional targets include without limitation, Tpo, flt3L, SCF, Mpl, scl, IL-6, IL-11, TGF-β and the genetic region of the genes encoding those cytokines to which Erg polypeptide binds.

Modulation of HSC and/or hematopoietic progenitor cell activity may be effected *in vitro* or *in vivo*.

In some embodiments, the present invention provides agents capable of modulating the activity of Erg polypeptide for use in the treatment of conditions associated with an over supply or an under supply of hematopoietic stem cells and/or progenitor cells.

In one embodiment, the present invention provides a composition comprising Erg or an agent from which Erg is producible or a variant of either of these which enhances the activity of Erg polypeptide or *Erg* genetic sequences or Erg transcriptional targets. In some embodiments, such compositions or agents are for use in modulating hematopoiesis. In some embodiments, modulation is potentiation or upregulation.

The present invention also contemplates a composition comprising an agent which down regulates the activity of *Erg* in a cell or subject. In some embodiments, such compositions or agents are for use in modulating hematopoietic cell activity or for treating or preventing clonal hemopathies. Efforts to reduce platelet numbers, for example, may be indicated to reduce the risk of a blood clot say after surgery, or in a subject with a myeloproliferative disorder, anaemia or cancer. Reduction in stem cell activity may be used to treat such conditions.

In some embodiments of present invention, enhancement of the activity of Erg in a cell or subject permits the normal or enhanced production i.e., up-regulates the production of platelets from megakaryocytes. In another embodiment, enhancement of activity of Erg potentiates the development, survival, or proliferation (collectively referred to as "activity") of early hematopoietic cells such as HSC and their progeny. Elevated platelet

numbers may be required in the treatment of autoimmune conditions (such as systemic lupus or autoimmune hemolytic anaemia), cancer and blood clotting disorders. The ability to modulate stem cell levels in a subject or *in vitro* has a wide range of applications in conditions associated with an under supply or activity of stem cells or an over supply or activity of stem cells.

Accordingly, in some embodiments, the present invention provides a method of modulating early hematopoietic cell activity in a mammalian subject comprising administering to the subject an effective amount of an agent that modulates the activity of Erg and/or its transcriptional targets in the subject. In particular embodiments, the early hematopoietic cell is a HSC and/or a hematopoietic progenitor cell. In other embodiments the cell is a mature blood cell including B and T lymphocytes, natural killer (NK) cells, granulocytes, monocytes, macrophages, erythrocytes and/or platelets.

In other embodiments, the present invention provides a method of increasing the activity of hematopoietic cells, HSC or progenitor cells *in vitro* comprising contacting a cell with an effective amount of an agent that enhances the activity of Erg in the cell. In a particular embodiment, the hematopoietic cell is a HSC cell and/or a hematopoietic progenitor cell.

Accordingly, in some embodiments the present invention provides a preparation of hematopoietic cells in a medium comprising an agent that modulates Erg activity. In some embodiments, the hematopoietic cell is a HSC or a multipotent or committed hematopoietic progenitor cell.

The agents and compositions of the present invention include small or large chemical molecules, peptides, polypeptides, proteins or nucleic acid molecules including antisense or other gene silencing molecules, and their precursors or derivatives. Furthermore, the terms agent or modulating include isolated cellular agents, cells (including genetically modified HSC or hematopoietic progenitor cells), plasmids or vectors comprising these agents.

The present invention further provides methods of screening or diagnosis to determine whether or not a subject has a hematological disorder such as a thrombocytopenia or a stem cell defect associated with a loss of function of Erg or is susceptible to developing same comprising screening a sample from the subject for a loss of function mutation in *Erg*.

The above summary is not and should not be seen in any way as an exhaustive recitation of all embodiments of the present invention.

BRIEF DESCRIPTION OF THE FIGURES

Some figures contain colour representations or entities. Coloured versions of the figures are available from the Patentee upon request or from an appropriate Patent Office. A fee may be imposed if obtained from a Patent Office.

Figure 1 is a photographic representation depicting the results of automated hematological analysis to determine platelet, white blood cell and red blood cell counts in ENU treated mice. This led to the identification of a mouse with a dominant multilineage deficiency phenotype, carrying a mutation designated as *mld2*. Each spot represents the peripheral blood cell counts of a single mouse. The results from the same mice are shown in each graph. Each mouse contains a unique set of mutations. Counts from the mouse carrying *mld2* are shown in filled in grey circle.

Figure 2 is a photographic representation depicting the results of automated hematological analysis to determine cell counts in mice. The *mld2* mutation causes deficiencies in all lineages in $Mpl^{-/-}$ mice. *mld2*/+ $Mpl^{-/-}$ mice were mated to +/+ $Mpl^{-/-}$ mice. The offspring were expected to include both *mld2*/+ $Mpl^{-/-}$ and +/+ $Mpl^{-/-}$. The RBC counts of the progeny divided into two clearly separated populations. Mice with low RBC counts were assumed to carry *mld2* (*mld2*/+ $Mpl^{-/-}$, blue), whereas those with high RBC counts were assumed not to (+/+ $Mpl^{-/-}$, green). Based on these genotypes, *mld2*/+ $Mpl^{-/-}$ mice were found to have significantly fewer mature blood cells for all the lineages examined. Each spot indicates the peripheral blood count of an individual mouse, the black bar indicates the mean for each genotype and the error bars indicate the standard deviation.

Figure 3 is a graphical representation depicting survivorship of *mld2*/+ $Mpl^{-/-}$ mice. *mld2*/+ $Mpl^{-/-}$ mice show decreased survival. All of the mice with RBC counts less than 7 per pL, assumed to be *mld2*/+ $Mpl^{-/-}$ (n =13), were found dead or were sacrificed due to illness. Conversely, none of those mice with RBC counts over 9 per pL, assumed to be +/+ $Mpl^{-/-}$ mice (n = 8) died due to illness over the same time period, although several were sacrificed because they were in excess of our requirements (indicated by triangles).

Figure 4 is a photographic representation depicting the results of automated hematological analysis to determine platelet, white blood cell and red blood cell counts in *mld2/+ Mpl*$^{-/-}$ mice. Mpl$^{+/-}$ mice which carry *mld2* have moderate deficiencies in WBCs and platelets. The *mld2* genotype of Mpl$^{+/-}$ mice was determined by SSLP genotyping in the *mld2* candidate interval. Mice were bled at seven weeks. Each spot represents the peripheral blood count from a single mouse. The black bar shows the mean for each genotype, and the error bars show the standard error of the mean.

Figure 5 is a graphical representation of sequence data showing that the mld2 mutation is a T to C base substitution in Erg. These DNA sequence trace files show both T and C peaks in the *mld2/+* heterozygote, but only a T peak at the same location in the wild type (+/+). The amino acid sequence of Erg encoded by this length of DNA is shown under the trace in single letter code. The S/P refers to the serine and proline residues encoded by the wild type and *mld2 Erg* genes respectively in the genome of mld2 heterozygotes.

Figure 6 is a representation of information regarding the source of Erg sequence aligned in Figures 7 and 8.

Figure 7 is a representation of an alignment of amino acid sequences of Erg homologs taken from sources set out in Figure 6.

Figure 8 is a representation of an alignment of nucleic acid sequences of Erg homologs taken from sources set out in Figure 6.

Figure 9 is a graphical representation of data showing the identification of the *mld2* mutation in a forward genetic screen on a sensitised *Mpl*$^{-/-}$ background. In order to screen for dominant hematopoietic phenotypes that may involve HSCs or multipotent progenitors, *Mpl*$^{-/-}$ C57BL/6 mice were injected with ENU as described previously and bred with isogenic females (Carpinelli *et al.*, *Proc. Natl. Acad. Sci. U. S. A.*, *101*:6553-6558, 2004). Blood was taken from their progeny at 7 weeks and analysed on an automated

hematological analyser. One mouse had low numbers of erythrocytes, leukocytes and platelets relative to *Mpl*$^{-/-}$ mice. The mutation responsible for this multilineage defect was designated *mld2*. Each circle represents the peripheral blood cell count of a single mouse, the results from the same mice are shown in each graph. The mouse carrying the *mld2* mutation is shown in red.

Figure 10 is a graphical representation of data showing that *mld2*/+ mice are anaemic, leukopenic and thrombocytopenic. The *mld2*/+ *Mpl*$^{-/-}$ founder mouse was mated to +/+ *Mpl*$^{-/-}$ mice to further investigate the *mld2* phenotype and its inheritance. The progeny were bled at seven weeks. The RBC counts of these progeny split into two clearly demarcated populations and genotypes were inferred on this basis (*mld2*/+ *Mpl*$^{-/-}$ for RBC < 8.0 per pL (blue), +/+ *Mpl*$^{-/-}$ for RBC > 8.0 per pL (green)). The *mld2*/+ *Mpl*$^{-/-}$ mice have an approximately 47% deficiency in RBCs. Using such inferred genotypes, *mld2*/+ *Mpl*$^{-/-}$ also appear to have deficiencies in all lineages of leukocytes (54% deficient overall) as well as platelets (80% deficient, Figure 2). Upon identification of the mutated gene (see below), data from accurately genotyped mice confirmed this finding. Each circle represents the peripheral blood cell count of a single mouse, the results from the same mice are shown in each graph. The means and standard deviations mice of each inferred genotype are also shown.

Figure 11 is a diagrammatical representation of data showing that the *mld2* mutation maps to a 1.6 megabase interval on chromosome 16. A positional cloning strategy was used to identify the region of the genome carrying the *mld2* mutation. Briefly, *mld2*/+ *Mpl*$^{-/-}$ C57BL/6 mice were crossed to wild type Balb/c mice. The G_1 progeny were then crossed to +/+ *Mpl*$^{-/-}$ C57BL/6 mice. Therefore, in the resultant G_2 mice, the genotypes of short sequence length polymorphic (SSLP) genetic markers nearest to the *mld2* mutation will be homozygous C57BL/6 in all mice carrying *mld2*, and conversely they will be heterozygous in those not carrying it. The *mld2* mutation can be localised by finding the smallest region of markers for which this remains true (indicated by a red box). G_2 *Mpl*$^{-/-}$ mice could easily be shown to be affected by the *mld2* phenotype by the presence of the characteristic multilineage defect, and the capacity to pass this defect to their progeny. Using this

strategy, the *mld2* mutation was localised to a 1.6 megabase (Mb) interval on the distal end of chromosome 16. This interval contains 8 genes.

The data represents the haplotypes of mice used to define the *mld2* candidate interval. The number of mice affected and unaffected by the *mld2* mutation with each haplotype is shown. Markers used and their positions on the April 2006 University of California, Santa Cruz (UCSC) mouse genome are indicated. C, homozygous C57BL/6 genotype; H, heterozygous C57BL/6 and Balb/c genotype; Mb, megabases.

Figure 12 is a graphical representation of data showing that the *mld2* mutation is predicted to cause a serine to proline substitution in the DNA binding domain of the ETS transcription factor Erg. The 1.6 megabase *mld2* candidate interval contained 8 genes. One of these genes encodes the ETS transcription factor Erg. Several ETS transcription factors have been implicated in hematopoiesis, so the coding exons and intron-exon boundaries of *Erg* were sequenced. A T to C nucleotide substitution was identified in exon 12 of *Erg* in all *mld2*/+ mice, but in no +/+ mice. This is predicted to cause a serine to proline substitution in the ETS domain of *Erg*. This mutation lies immediately distal to an alpha helix highly conserved amongst the 26 ETS family genes in the mouse and so may be predicted to disrupt one or more functions of this DNA binding domain. The data shown represents: **(A)** Electropherograms showing the nucleotide sequence of exon 12 of *Erg* from a +/+ and *mld2*/+ mouse, along with the predicted amino acid sequence of Erg. **(B)** The predicted amino acid sequences of the ETS domain of Erg in different species. The residue affected by the *mld2* mutation is highlighted in yellow.

Figure 13 is a diagrammatical representation of data showing that $Erg^{mld2/+}$ mice are leukopenic and thrombocytopenic on a $Mpl^{+/+}$ genetic background. $Erg^{mld2/+}$ $Mpl^{-/-}$ C57BL/6 mice were successively mated with $Erg^{+/+}$ $Mpl^{+/+}$ C57BL/6 in order to generate $Erg^{mld2/+}$ $Mpl^{+/+}$ mice. At 7 weeks of age, blood from these mice and $Erg^{+/+}$ $Mpl^{+/+}$ littermates was analysed. $Erg^{mld2/+}$ $Mpl^{+/+}$ mice showed no deficiency in erythrocytes, exhibited an approximately 40% reduction in leukocytes and platelets. The data shown

represents the means ± standard deviations of the blood cell counts from five mice of each genotype. The data in each of the three graphs is from the same five mice.

Figure 14 is a graphical representation of data showing $Erg^{mld2/+}$ $Mpl^{-/-}$ mice exhibit decreased survival. Within the colony of G_2 mapping mice (which have a mixed genetic background), $Erg^{mld2/+}$ $Mpl^{-/-}$ mice demonstrated reduced survival, while $Erg^{+/+}$ $Mpl^{-/-}$, $Erg^{mld2/+}$ $Mpl^{+/-}$ and $Erg^{+/+}$ $Mpl^{+/-}$ mice exhibited no decrease in survival. Therefore, mice with one mutant *Erg* allele only show decreased survival in the absence of functional Mpl. Approximately 200μL of blood was taken from all these mice at three weeks. Preliminary data from post mortem analyses suggests that this illness is associated with internal haemorrhaging under the skin, around the joints and in the skull. The data shown is a Kaplan-Meyer plot where a mouse being found dead, or discarded due to illness is represented by a drop in the curve, whereas the discarding of healthy mice from the colony due their being excess to our requirements is represented by vertical tick marks on the curve.

Figure 15 is a diagrammatical representation of data showing that $Erg^{mld2/+}$ $Mpl^{+/+}$ mice exhibit deficiencies in a range of lineage committed hemopoietic progenitors. The frequency of blast, neutrophil (G), neutrophil-macrophage (GM), macrophage (M), eosinophil (Eo) and megakaryocyte (Meg) colony forming cells (CFC) in bone marrow (BM) was determined by the clonal culture of hematopoietic precursors with the stimulation of stem cell factor, interleukin-3 and erythropoietin, as described previously (Alexander *et al., Blood, 87*:2162-2170, 1996). This demonstrated that $Erg^{mld2/+}$ $Mpl^{+/+}$ mice had an approximately 50% deficiency in lineage-committed progenitor cells from all of these hematopoietic lineages. The data shown represents the means and standard deviations of the frequency of CFCs in five mice of each genotype.

Figure 16 is a diagrammatical representation of data showing that $Erg^{mld2/+}$ mice exhibit a severe deficiency in multipotent progenitor activity. In order to investigate HSC and multipotent progenitor activity in $Erg^{mld2/+}$ mice, the frequency of spleen colony forming units (CFU-s) in the bone marrow (BM) was determined. 3.0 x 10^6 BM cells from

$Erg^{mld2/+}$ $Mpl^{-/-}$ or 1.5 x 10^6 from $Erg^{+/+}$ $Mpl^{-/-}$ mice were intravenously transplanted into four C57BL/6 recipient mice. Eight days later, the spleens were dissected from the recipients and macroscopic spleen colonies counted after fixation in Carnoy's fixative. $Erg^{mld2/+}$ $Mpl^{-/-}$ mice exhibited a profound reduction in the multipotent, early hematopoietic progenitors that form spleen colonies in comparison to $Erg^{+/+}$ $Mpl^{-/-}$ mice. 7.5 x 10^4 or 3 x 10^5 BM cells from $Erg^{mld2/+}$ $Mpl^{+/+}$ or 7.5 x 10^4 or 1.5 x 10^5 cells from $Erg^{+/+}$ $Mpl^{+/+}$ mice were transplanted into four C57BL/6 recipient mice (n = 5). CFU-s were counted as described above. $Erg^{mld2/+}$ $Mpl^{+/+}$ mice were approximately 50% deficient in CFU-s in comparison to $Erg^{+/+}$ $Mpl^{+/+}$ mice. This provides compelling evidence of a profound, bone marrow intrinsic deficiency in HSCs and/or multipotent progenitors in mice with one mutant *Erg* allele. The data shown represents the means and standard deviations of the frequency of CFU-s in bone marrow from three $Erg^{mld2/+}$ $Mpl^{-/-}$ mice, three $Erg^{+/+}$ $Mpl^{-/-}$ mice, five $Erg^{mld2/+}$ $Mpl^{+/+}$ mice and five $Erg^{mld2/+}$ $Mpl^{+/+}$ mice.

Figure 17 is a diagrammatical representation of data showing that $Erg^{mld2/+}$ bone marrow shows defects in the reconstitution of hematopoiesis after transplantation into myeloablated recipient mice. In order to investigate the function of $Erg^{mld2/+}$ HSCs, bone marrow (BM) transplants were performed. Following intravenous injection of 1 x 10^6 $Erg^{mld2/+}$ $Mpl^{-/-}$ $Ly5.1^-Ly5.2^+$ donor BM cells into five irradiated congenic C57BL/6 $Ly5.1^+Ly5.2^-$ recipients, all recipients died within three weeks (n =2), whereas all recipients receiving 1 x 10^6 $Erg^{+/+}$ $Mpl^{-/-}$ BM cells survived over the same period (n =1). This demonstrates that 1 x 10^6 $Erg^{+/+}$ $Mpl^{-/-}$ BM does not contain sufficient HSC and/or hematopoietic progenitor activity to rescue lethal myeloablation and reconstitute hematopoiesis.

Four lethally irradiated congenic C57BL/6 $Ly5.1^+Ly5.2^-$ recipients were reconstituted with 1 x 10^6 $Erg^{mld2/+}$ $Mpl^{+/+}$ or $Erg^{+/+}$ $Mpl^{+/+}$ $Ly5.1^-Ly5.2^+$ donor BM cells. Sixteen weeks post transplantation, the contribution of donor derived cells to the blood was determined by using flow cytometry. The reconstitution of hematopoiesis by donor cells, as measured by the contribution of donor derived cells to all measured hematopoietic lineages, was significantly reduced in recipients that received $Erg^{mld2/+}$ $Mpl^{+/+}$ BM. A similar deficiency

is observed across all cell lineages, including those bearing cell surface antigens specific to the B lymphocyte (B220+), T lymphocyte (CD4 & CD8) and myeloid (Gr-1 & Mac-1) lineages. In addition, the peripheral blood of mice reconstituted with $Erg^{mld2/+}$ $Mpl^{+/+}$ BM show leukopenia and thrombocytopenia indistinguishable to that observed in the peripheral blood of $Erg^{mld2/+}$ $Mpl^{+/+}$ mice. These data demonstrate that mutations in *Erg* cause bone marrow intrinsic deficiencies in HSC activity. Importantly, although $Erg^{mld2/+}$ $Mpl^{+/+}$ BM is defective in its ability to reconstitute all blood cell lineages relative to wild type cells, it *is* able to contribute to all these cell lineages. This shows that $Erg^{mld2/+}$ $Mpl^{+/+}$ BM can home, engraft, survive, self-renew and differentiate but that, relative to wild type cells, one or more of these functions is compromised by the mutation of *Erg*. This explains why $Erg^{mld2/+}$ $Mpl^{+/+}$ mice can survive normally with blood cells of all lineages, while remaining an excellent model in which HSC deficiencies and regulation may be investigated.

BM was transplanted from two $Erg^{mld2/+}$ $Mpl^{+/+}$ mice and two $Erg^{+/+}$ $Mpl^{+/+}$ mice. The data shown represent: **(A)** The means and standard deviations of the average proportion of blood cells derived from donor BM after reconstitution. The contributions to multiple blood cell lineages are shown. **(B)** The means and standard deviations of the average peripheral blood cell counts observed in the recipitents of BM. This data is from the same mice as shown in (A).

Figure 18 is a diagrammatical representation of data showing that $Erg^{mld2/+}$ bone marrow cannot compete with wild type bone marrow during the reconstitution of hematopoiesis. In order to further compare the function of $Erg^{mld2/+}$ and wild type HSCs, competitive transplants were performed. 1 x 10^6 $Erg^{mld2/+}$ $Mpl^{-/-}$ or $Erg^{+/+}$ $Mpl^{-/-}$ $Ly5.1^-Ly5.2^+$ test bone marrow (BM) cells were injected along with 1 x 10^6 $Erg^{+/+}$ $Mpl^{-/-}$ $Ly5.1^+Ly5.2^-$ competitor BM cells into C57BL/6 $Ly5.1^+Ly5.2^-$ recipients which had been myeloablated by irradiation. Similarly, 1 x 10^6 $Erg^{mld2/+}$ $Mpl^{+/+}$ or $Erg^{+/+}$ $Mpl^{+/+}$ $Ly5.1^-Ly5.2^+$ BM cells were injected with 1 x 10^6 $Erg^{+/+}$ $Mpl^{+/+}$ $Ly5.1^+Ly5.2^-$ BM cells into irradiated C57BL/6 $Ly5.1^+Ly5.2^-$ recipients. The contribution of test and competitor bone marrow to hematopoiesis was determined 16 weeks post transplant by comparing the relative

proportion of test donor derived $Ly5.1^{-}Ly5.2^{+}$ cells and competitor BM derived cells $Ly5.1^{+}Ly5.2^{-}$ in the blood by flow cytometry. On both the $Mpl^{-/-}$ and $Mpl^{+/+}$ genetic backgrounds, BM cells carrying a mutant *Erg* allele were severely deficient in their ability to contribute to hematopoiesis relative to $Erg^{+/+}$ BM. This indicates that $Erg^{mld2/+}$ HSCs have a severe functional deficiency.

Competitive transplants were performed using three $Erg^{mld2/+}$ $Mpl^{-/-}$, $Erg^{+/+}$ $Mpl^{-/-}$, $Erg^{mld2/+}$ $Mpl^{+/+}$ and $Erg^{+/+}$ $Mpl^{+/+}$ mice. mice and two $Erg^{+/+}$ $Mpl^{+/+}$ The data shown represents the means and standard deviations of the average proportion of test donor derived cells after reconstitution.

Figure 19 is a photographical representation of data showing that the *mld2* missense mutation does not effect the DNA binding of Erg. To investigate whether the *mld2* missense mutation affects the DNA binding function of Erg, electrophoretic mobility shift assays were performed (EMSA). FLAG epitope-tagged Erg and Erg^{mld2} were affinity purified from the lysates of 293T human embryonic kidney cells transfected with a mammalian expression vector to overexpress these proteins. Amounts of the resulting enriched Erg and Erg^{mld2} solutions containing equivalent amounts of each protein were mixed with 10ng of radiolabelled E74 probe, a double stranded oligonucleotide which contains a sequence that is specifically bound by ETS transcription factors, in DNA binding buffer (13mM Tris pH 7.5, 2.6mM $MgCl_2$, 1.3mM dithiothreitol, 260μM spermadine, 260μM EDTA) with water, 200ng unlabelled E74, 200ng unlabelled $E74^{mut}$ (in which has two nucleotides have been altered to ablate the ETS binding site), 2mg biotinylated M2 anti-FLAG antibody, or 2mg biotinylated mouse IgG_1. After 15 minutes incubation, probe/protein complexes were resolved from unbound probe by electrophorhesis through a 10% acrylamide-TBE gel.

Probe/protein complexes were detectable when either Erg or Erg^{mld2} solutions were mixed with labelled E74 indicating that protein in both the solutions could bind this oligonucleotide. The addition of a 20 times excess of unlabelled E74 disrupted probe/protein complexes, whereas the addition of the same mass of E74, which lacks a

canonical ETS binding site did not disrupt probe/protein complexes at all. These data indicate that protein in both the Erg and Erg^{mld2} solutions binds specifically to the ETS binding site of E74. Anti-FLAG antibody bound to the probe/protein complexes – making the complex migrate through the gel more slowly and 'supershifting' the resulting band, whereas a non-specific antibody of the same isotype had no effect. These data demonstrate that the protein that is binding to E74 is FLAG epitope-tagged Erg and Erg^{mld2}. In all comparable lanes, the bands representing the Erg/probe and Erg^{mld2}/probe complex are of similar intensity, indicating that the *mld2* mutation does not dramatically affect the DNA binding function of Erg.

The data shown represents: **(A)** The sequence of the E74 and $E74^{mut}$ oligonucleotides. The ETS binding site is shown in bold, the two nucleotides altered to remove the ETS binding site are highlighted yellow. **(B)** An autoradiograph of an EMSA gel. Each lane was loaded with E74 radiolabelled probe. Lane 2 was also loaded with affinity purified cell lysate from cells transfected with an empty mammalian expression vector (mock). Lanes 3, 5, 7, 9 and 11 were loaded with affinity purified cell lysate from cells transfected with a mammalian expression vector transfected to overexpress wild type (WT) Erg. Lanes 4, 6, 8, 10 and 12 were loaded with affinity purified cell lysate from cells transfected with a mammalian expression vector transfected to overexpress Erg^{mld2}. Lanes 5 and 6 were loaded with unlabelled E74. Lanes 7 and 8 were loaded with unlabelled $E74^{mut}$. Lanes 9 and 10 were also loaded with M2 anti-FLAG antibody. Lanes 11 and 12 were also loaded with mouse IgG_1. Underneath the autoradiograph, a silver stained SDS-PAGE gel shows the protein content of the enriched cell lysates loaded into each lane.

Figure 20 is a photographical representation of data showing that the *mld2* mutation disrupts the ability of Erg to transactivate transcription. In order to investigate the effect of the *mld2* mutation on the ability of Erg to drive transcription luciferase reporter assays were performed. COS monkey kidney cells were transfected with a firefly luciferase gene under the transcriptional control of a herpes simplex virus thymidine kinase promoter, which contains several canonical ETS transcription factor binding sites. Cells were cotransfected with one or more of the following – pEF-BOS, a mammalian expression

vector; pEF-FLAG-Erg, the same vector, but with a cDNA insert encoding FLAG epitope-tagged Erg, so that Erg is overexpressed; or pEF-FLAG-Erg^{mld2}, which results in the overexpression of FLAG epitope-tagged Erg^{mld2}. Luciferase activity was measured using a commercial luciferase assay kit (Promega). Cells expressing Erg showed a high level of luciferase activity, whereas cells expressing Erg^{mld2} did not show luciferase activity significantly different from cells transfected with the empty pEF-BOS plasmid. This demonstrates that the *mld2* missense mutation disrupts the function of Erg, preventing it from transactivating transcription. Cells cotransfected with Erg and Erg^{mld2} showed high luciferase activity, but this activity was significantly lower than that observed in cells with Erg alone ($p < 0.05$, Student's T-test). This shows that Erg^{mld2} partially interferes with the activity of Erg in the same cell. The data shown represent the means ± standard deviations of the luciferase activity of five wells of cells transfected with the plasmids shown. A Western blot probed with antibodies against the FLAG-epitope and Hsp70 (a housekeeping gene) is also shown. The bands representing Erg and Hsp70 on the blot are shown, as is the size of protein standards (in kDa).

Figure 21 is a graphical representation of data showing that the *mld2* mutation disrupts the biological activity of Erg. The effect of the *mld2* mutation on the biological activity of Erg was investigated by studying its expression in K562 eythroleukemia cells. The overexpression of Erg in K562 erythroleukemia cells has been shown to induce the megakaryocytic antigens CD41 and CD61 (Rainis *et al.*, 2005 (*supra*)). We transfected K562 cells with one of three vectors: pEF-Erg-IRES-GFP, a bicistronic plasmid that encodes Erg and green fluorescent protein (GFP); pEF-Erg^{mld2}-IRES-GFP, which encode Erg^{mld2} and GFP; or pEF-IRES-GFP, which encodes GFP alone. Three days post transfection, cells were analysed by flow cytometry. A large proportion of K562 cells overexpressing Erg were also expressing CD41 and CD61. Conversely, the proportion of Erg^{mld2} overexpressing cells that were positive for CD41 or CD61 was not significantly different to that of cells overexpressing GFP alone. These data demonstrate that the *mld2* mutation disrupts the biological activity of Erg. The data shown here are representative flow cytometry plots, along with the means ± standard deviations of the proportion of CD41$^+$ or CD61$^+$ cells from three experiments.

BRIEF DESCRIPTION OF THE TABLES

Table 1 provides a description of the SEQ ID NOs provided herein.

Table 2 provides an amino acid sub-classification.

Table 3 provides exemplary amino acid substitutions.

Table 4 provides a list of non-natural amino acids contemplated in the present invention.

Table 5 presents data demonstrating that the *mld2/+ Mpl*$^{-/-}$ genotype causes severe deficiency in HSC and/or multipotential progenitor activity. Cells derived from *mld2/+ Mpl*$^{-/-}$ marrow make a negligible contribution to hematopoiesis when in competition with +/+ *Mpl*$^{-/-}$ marrow. The mean percentage contribution ± standard deviation is shown. 1 x 10^6 Ly5.1$^-$Ly5.2$^+$ test cells were transplanted with 1 x 10^6 Ly5.1$^+$Ly5.2$^-$ +/+ *Mpl*$^{-/-}$ competitor cells into five Ly5.1$^+$Ly5.2$^+$ irradiated recipients. Data shown are percentages of total circulating nucleated cells.

Table 6 presents data indicating survivorship in progeny of heterozygous *mld2/+ Mpl*$^{-/-}$ mice. Mice with the mld2/mld2 homozygous genotype do not survive until birth. The table shows a number of progeny of a *mld2/+ Mpl*$^{+/-}$ intercross, genotyped at weaning by SSLP markers surrounding the known location of the *mld2* mutation.

Table 7 provides survivorship data for progeny of mice generated carrying the Mld2 mutation showing that the Mld2 mutation is lethal in homozygous form. While generating mice carrying the *mld2* mutation, we observed that no *Erg*$^{mld2/mld2}$ pups were present at weaning. To investigate this apparent lethality, we examined the embryonic progeny of *Erg*$^{mld2/+}$ *Mpl*$^{+/-}$ x *Erg*$^{mld2/+}$ *Mpl*$^{+/-}$ matings at a number of timepoints post conception, as this was expected to produce embryos of all possible *Erg* and *Mpl* genotypes. At embryonic day 9.5 (E9.5), all embryos were alive, and showed normal morphology. However, by E12.5, although embryos of all genotypes were observed at expected ratios,

almost all $Erg^{mld2/mld2}$ embryos were dead. Those few that survived showed severe developmental delays, suggesting they would die soon after. This indicates that the *mld2* mutation in both alleles of *Erg* is sufficient to cause death early in gestation.

Tables 8 and 9 provides data showing that mutations in *Erg* reduce the cellularity and number of colony forming cells in yolk sacs and fetal livers dissected from embryos at 10.5 days post conception. In order to further investigate the phenotype of embryos carrying mutations in Erg, embryonic hematopoiesis was analysed. One of the earliest sites of blood cell production in the embryo is the yolk sac, the membranous tissue surrounding the developing fetus. Later in embryogenesis hematopoiesis shifts to the fetal liver and then ultimately to the bone marrow and spleen.

Counts of embryonic hematopoietic progenitor cells were determined by an *in vitro* clonal assay of yolk sac cell suspensions and fetal liver cells from embryonic day 10.5 and 12.5 fetuses respectively, using cytokines to stimulate colony formation using previously described methods (Alexander *et al.*, 1996 (*supra*)). These analyses revealed that the overall cellularity of the fetal liver and yolk sac was considerably reduced and numbers of progenitor cells of all measured hematopoietic lineages were deficient in homozygous *mld2/mld2* mice. when compared with *mld2/+* or wild type mice. A trend towards reduced cellularity and deficient hematopoietic progenitor cell numbers was also observed in *mld2/+* tissues compared with wild type.

These data indicate that the hematopoietic deficiencies characteristic of the *mld2* mutation in Erg arise in the earliest stages of hematopoiesis. The multi-lineage deficiencies evident in mutant yolk sac and fetal liver are consistent with a hematopoietic stem cell defect. Correction of such defects by enhancing Erg activity or signaling or the activities of its transcriptional targets will therefore be useful in enhancing hematopoietic stem cell functional activity.

The data shown is the means ± standard deviations of the numbers of hematopoietic colony-forming cells in yolk sac and fetal liver suspensions cultured in stem cell factor +

interleukin-3 + erythropoietin for 7 days at 37°C in 5% CO_2 in air. G, granulocyte; GM, granulocyte-macrophage; M, macrophage; Eo, eosinophil; Meg, megakaryocyte; Ery, erythroid; Myel, myeloid (granulocyte, macrophage).

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The invention is predicated, in part, upon the determination that Erg polypeptide function is important for megakaryocytopoiesis (platelet production) and leukocytopoiesis and also for the proper or uncompromised activity of early hematopoietic cells such as hematopoietic stem cells and/or progenitor cells.

An Erg polypeptide as defined herein may be used in the expansion of hematopoietic stem cells and/or multipotent or committed hematopoietic progenitor cells of various blood lineages, and in the differentiation and/or proliferation of various hematopoietic cell types.

Given its role in early hematopoiesis, *Erg*, Erg and variants of these or agents that modulate Erg activity in a subject may also be used to treat or prevent various disorders or conditions associated with stem cell, progenitor cell or various different blood cell defects.

The ability of stem cells to self-renew is shared by neoplastic cells which can aberrantly replicate until they overwhelm the body. The regulatory mechanisms of cancer cell and stem cell proliferation may therefore have overlapping characteristics. In particular, a fully differentiated cell may take on the self-renewing phenotype of stem cells. Alternatively, or in addition, stem cells may assume a dysregulated self-renewal phenotype to form cancerous cells.

Reference herein to a hematological disorder or condition or similar grammatical expressions include, without limitation, thrombocytopenias, thrombocytosis, anaemias including drug induced anaemia, leukopenia, immunological disorders, autoimmune disorders, myeloproliferative disorders, cancer, conditions associated with stem cell defects or dysregulated proliferation or differentiation and bone marrow defects. Examples of particular conditions include without limitation: leukemia, Hodgkin's disease, non-Hogkin's lymphoma, acute lymphocytic anaemia, plasmacytomas, multiple mycloma Burkett's lymphoma, arthritis, asthma, AIDS, rheumatoid arthritis, granulomatous disease,

immune deficiency, inflammatory bowel disease, sepsis, neutropenia, neutrophilia, psoriasis, immune reaction to transplanted organ, SLE, hemophilia, hypercoagulation, diabetes, meningitis, lyme disease and allergies.

Autosomal recessive traits include: Fanconi Syndrome, Thrombocytopenia with absent radius (TAR) syndrome, Bernard-Soulier syndrome, Gray platelet syndrome, and Isolated thrombocytopenia. Autosomal dominant traits include: May-Hegglin anomaly, Alport syndrome variants, and Isolated thrombocytopenia. X-linked traits include: Wiskott-Adrich syndrome, and Isolated thrombocytopenia.

Other hematopoietic disorders are selected from the group consisting of: aplastic pancytopenias (traditionally known as aplastic anaemia), which result from aplasia or suppression of hematopoietic stem cells, including: Fanconi's anaemia, Shwachman-Diamond syndrome, Dyskeratosis congenita, Amegakaryocytic thrombocytopenia. Other genetic syndromes include Down's syndrome, Dubowitz's syndrome, Seckel's syndrome, Reticular dysgenesis, and Familial aplastic anaemia (non-Fanconi's).

The inhibition of Erg or its transcriptional targets or downstream effectors may be therapeutically useful in responding to the clonal hemopathies that result from the overactivity or replication of stem cells. These include: Preleukaemias (myelodysplasias) such as Idiopathic refractory sideroblastic anaemia, Idiopathic refractory nonosideroblastic anaemia, Pancytopenia with hyperplastic marrow, Monoclonal aplastic anaemia, and Paroxysmal nocturnal hemoglobinuria. Also, myeloproliferative diseases such as Chronic myeloproliferative diseases including Polycythemia vera, Chronic myelogenous leukemia (CML), Primary thrombocythemia, Idiopathic myelofibrosis, Subacute myeloproliferative disorders, Oligoblastic (smoldering) leukemias, Refractory leukemias with excess myeloblasts, Myelomnocytic leukemia, and Atypical myeloproliferative syndromes. Also, acute myeloproliferative disorders such as Acute myelogenous leukemia, Myeloblastic (granuloblastic), Promyelocytic: associated with intravascular coagulation, Myelomonocytic(granlomonoblastic), Monocytic: associated with tissue infiltration, Erythroid, Megakaryocytic: associated with marrow fibrosis, Eosinophilic: associated with

heart and lung fibrosis and Basophilic and Mast Cell. Also, acute biphenotypic (myeloid and lymphoid markers) leukemia and acute leukemia with lymphoid markers evolving from a prior clonal hemopathy.

In other embodiments, the present invention provides a method of modulating hematopoietic cell activity wherein the hematopoietic cell is selected from the group consisting B or T lymphocytes, natural killer (NK) cells, granulocytes, monocytes, macrophages, erythrocytes and platelets. The method is conducted *in vitro* or *in vivo*. In some embodiments, an agent that modulates the activity of Erg in a cell is administered to a subject in need thereof. In other embodiments, the cell is contacted *in vitro* with an effective amount of the agent sufficient to enhance the activity of Erg therein.

The present invention extends to animal and cellular models comprising Erg having a modified (including no) Erg activity such that the effects of modified Erg activity on blood cell development and function can be determined. Such models are also useful to screen for and test agents that have potential as therapeutic agents when these functions are impaired.

Accordingly, in one aspect the present invention provides an isolated cell, or a non-human animal comprising such cells, wherein Erg or *Erg* is modified to effectively modulate the functional effect of Erg in the cell or animal compared to a non-modified animal of the same species.

Cells may be derived from human or non-human animal sources. The term "derived" does not necessarily mean that the cells are directly obtained from a particular source.

Reference to a "cell" includes a system of cells such as a particular tissue or organ.

The term "modified" includes genetically modified but encompasses non-genetic or epigenetic modifications to affect Erg activity by, for example, the administration of an agent such as, without limitation, an organic or inorganic chemical agents, antibody,

enzyme, peptide, genetic or proteinaceous molecule to effectively modulate the functional activity of *Erg*.

Reference herein to "modulate" and "modulation" includes completely or partially inhibiting or reducing or down regulating all or part of Erg functional activity and enhancing or up regulating all or part its functional activity. Functional activity may be modulated by, for example, modulating *Erg* nucleic acid binding capabilities or transcriptional or translational activity, or its half-life. With regard to Erg polypeptide, its functional activity may be modulated by, for example, modulating its binding capabilities, its half-life, location in a cell. Thus, Erg level or activity may be modulated by modulating *Erg* expression, transcript stability, and the activity of its regulatory molecules.

Reference to the "activity" or "functional activity" of Erg encompasses any relevant, measurable activity or characteristic of the molecule in proteinaceous or genetic form and includes Erg polypeptide's specific DNA and protein binding abilities. Binding or transcriptional, translational or transactivational activity are preferred activities which are conveniently assessed using standard protocols known in the art as described in Sambrook, *Molecular Cloning: A Laboratory Manual*, 3rd Edition, CSHLP, CSH, NY, 2001; Ausubel (Ed) *Current Protocols in Molecular Biology*, 5th Edition, John Wiley & Sons, Inc, NY, 2002. For example, the ability of Erg to drive transcription is tested using luminescence reporter assays as shown in Figure 20. The ability of Erg to modulate cellular activities such as proliferation, development or survival can be assessed visually, spectroscopically, or using instrumentation to evaluate the activity or a molecular marker or reporter of the activity. Other activities assayed include HSC homing, engrafting, apoptosis and self-renewal. Functional activity, in some embodiments is assessed by analysing cells comprising Erg (such as K562, see Rainis *et al.*, 2005 (*supra*)) for expression of megakaryocytic or leukocytic markers antigens for example, by flow cytometry (see Figure 21). The ability of Erg to bind DNA *via* its C-terminal ETS DNA binding domain is tested as described in the Brief Description of the Figures for Figure 19. The ability of Erg polypeptide to bind to proteins *via* the N-terminal pointer domain is also tested. The Mld2 mutation (A, T and C nucleotide substitutions shown in Figure 12) does not affect

the direct DNA binding function of Erg although its functional activity is disrupted (see Figure 21 and Brief Description of the Figures). Erg polypeptide stimulates a characteristic gene expression profile which serves as a useful marker of Erg activity. Similarly, the gene expression profile of a cell when Erg polypeptide activity is down-regulated or inhibited serves as a useful marker of lack of Erg activity. Such assays may be conveniently adapted for high throughput evaluation, for example, cytometrically such as by flow cytometry, array technology such as microarray technology, antibody technology, chromatographic methods such as HPLC or thin layer chromatography or combinations of these. Binding is conveniently detected using antibodies. *In vitro* or *in vivo* assays can employ a wide range of markers or indicators of Erg activity using, for example, the methods exemplified herein. For example, platelet levels or turnover may be measured using automated hematological analysis; transcriptional activity may, for example, be assessed by measuring the level of specific RNA produced or may be assessed *via* measuring the activity of a reporter molecule.

Reference herein to the "activity" or "overactivity" and the like, in relation to cells include without limitation a reference to any one or more of the following: cellular development, proliferation, cellular differentiation, cell function such as homing, engraftment, self-renewal, survival differentiation, cell number and cell survival. As the skilled artisan will appreciate, the self renewal phenotype of stem cells is tightly regulated. As stem cells divide their daughter cells maintain a critical balance between two fates: either retaining stem cell function, or alternatively differentiating into mature effector cells. There are three potential outcomes or activities of stem cell division (i) extension symmetric division where both daughter cells retain stem cell function and which results into expansion of stem cell population (ii) maintenance asymmetric division which maintains the stem cell population by producing one daughter stem cell and one daughter cell committed to differentiation and (iii) committed symmetric division where both daughter cells are committed to differentiate. Accordingly, in relation to HSC reference to "activity" includes reference to extension symmetric division, maintenance asymmetric divisional and/or committed symmetric division.

The activity of Erg may be monitored using DNA or protein binding assays, reporter assays or direct or indirect assays of Erg activity including the use of antibodies or other proteinaceous or genetic agents in a number of assays which are well known to those of skill in the art. Antibodies, for example, may be used to detect Erg by Western Blotting, cytometric histochemical or ELISA procedures. As discussed herein below, such agents may also distinguish between active and inactive forms of the Erg or between mutant and normal forms of *Erg*. In accordance with the present invention, mutant forms of Erg are forms of *Erg* (found in a population of subjects) which are associated with aberrant hematopoiesis, such as thrombocytopenia, clonal hemopathies, anaemia, leukemia etc. Normal forms are forms of Erg which are not associated with these conditions. Mutant forms of *Erg* may also be conveniently be detected using nucleic acid based assays well know in the art and as described herein. Low levels of active polypeptide may be produced as a result of mutations in *Erg* leading to altered expression levels, altered transcript stability or altered functional activity. Thus, Erg activity may be monitored indirectly by monitoring RNA production and/or stability or the levels of regulatory molecules such as enhancers and repressors.

"Regulatory regions" include promoters, polyadenylation signals, transcriptional enhancers, translational enhancers, leader or trailing sequences that modulate mRNA stability and targeting sequences that direct a product encoded by a transcribed nucleic acid molecule to a particular location such as an intracellular compartment or to the extracellular environment.

The term "genetically modified" refers to changes at the genome level and refers herein to a cell or animal that contains within its genome a specific gene which has been altered. Alternations may be single base changes such as a point mutation or may comprise deletion of the entire gene such as by homologous recombination. Genetic modifications include alterations to regulatory regions, insertions of further copies of endogenous or heterologous genes, insertions or substitutions with heterologous genes or genetic regions etc. Alterations include, therefore, single of multiple nucleic acid insertions, deletions, substitutions or combinations thereof.

Cells and animals which carry a mutant *Erg* allele or where one or both alleles are deleted can be used as model systems to study the effects of Erg in hematopoietic cell development and function and/or to test for substances which have potential as therapeutic or teratogenic agents when these function are impaired. Animals for testing therapeutic agents can be selected after mutagenesis of whole animals or after treatment of germline cells or zygotes. Such treatments include insertion of mutant *Erg* alleles (including those carrying loxP flanking sequences), usually from a second animal of the same species, as well as insertion of disrupted homologous genes. Alternatively, the endogenous *Erg* gene of the animals may be disrupted by insertion or deletion mutation or other genetic alterations using conventional techniques. These animal models provide an extremely important testing vehicle for potential therapeutic products. The cells may be isolated from individuals with *Erg* mutations, either somatic or germline. Alternatively, the cell line can be engineered to carry the mutation in the *Erg* allele, as described above. After a test substance is applied to the cells, the phenotype of the cell is determined. Any trait of the cells can be assessed.

Thus a genetically modified animal or cell includes animals or cells from a transgenic animal, a "knock in" or knock out" animal, conditional variants or other mutants or cells or animals susceptible to co-suppression, gene silencing or induction of RNAi.

Conveniently, targeting constructs are initially used to generate the modified genetic sequences in the cell or organism. Targeting constructs generally but not exclusively modify a target sequence by homologous recombination. Alternatively, a modified genetic sequence may be introduced using artificial chromosomes. Targeting or other constructs including reporter constructs for screening potential Erg modulators are produced and introduced into target cells using methods well known in the art which are described in molecular biology laboratory manuals such as, for example, in Sambrook, 2001 (*supra*); Ausubel, 2002 (*supra*). Targeting constructs may be introduced into cells by any method such as electroporation, viral mediated transfer or microinjection. Selection markers are generally employed to initially identify cells which have successfully incorporated the targeting construct.

Genetically modified organisms are generated using techniques well known in the art such as described in Hogan *et al., Manipulating the Mouse Embryo: A Laboratory Manual*, Cold Spring Harbour Laboratory Press, CSH NY, 1986; Mansour *et al., Nature 336*:348-352, 1988; Pickert, *Transgenic Animal Technology: A Laboratory Handbook*, Academic Press, San Dingo, CA, 1994. Stem cells including embryonic stem cells (ES cells) are introduced into the embryo of a recipient organism at the blastocyst stage of development. There they are capable of integration into the inner cell mass where they develop and contribute to the germ line of the recipient organism. ES cells are conveniently obtained from pre-implantation embryos maintained *in vitro* (Robertson *et al., Nature 322*:445-448, 1986). Once correct targeting has been verified, modified cells are injected into the blastocyst or morula or other suitable developmental stage, to generate a chimeric organism. Alternatively, modified cells are allowed to aggregate with dissociated embryonic cells to form aggregation chimera. The chimeric organism is then implanted into a suitable female foster organism and the embryo allowed to develop to term. Chimeric progeny are bred to obtain offspring in which the genome of each cell contains the nucleotide sequences conferred by the targeting construct. Genetically modified organism may comprise a heterozygous modification or alternatively both alleles may be affected.

Another aspect of the present invention provides cells or animal comprising one, two or more genes or regions which are modified. For example, the genetically modified cells or animals may comprise a gene capable of functioning as a marker for detection of modified cells. Alternatively, the instant animals may be bred with other transgenic or mutant non-human animals to provide progeny some of which exhibit one or both traits or a modified trait/s. Chimeric animals are also contemplated.

The terms "genetic material", "genetic forms", "nucleic acids", "nucleotide" and "polynucleotide" include RNA, cDNA, genomic DNA, synthetic forms and mixed polymers, both sense and antisense strands, and may be chemically or biochemically modified or may contain non-natural or derivatized nucleotide bases, as will be readily

appreciated by those skilled in the art. Such modifications include, for example, labels, methylation, substitution of one or more of the naturally occurring nucleotides with an analog (such as the morpholine ring), internucleotide modifications such as uncharged linkages (e.g. methyl phosphonates, phosphotriesters, phosphoamidates, carbamates, etc.), charged linkages (e.g. phosphorothioates, phosphorodithioates, etc.), pendent moieties (e.g. polypeptides), intercalators (e.g. acridine, psoralen, etc.), chelators, alkylators and modified linkages (e.g. α-anomeric nucleic acids, etc.). Also included are synthetic molecules that mimic polynucleotides in their ability to bind to a designated sequence *via* hydrogen binding and other chemical interactions. Such molecules are known in the art and include, for example, those in which peptide linkages substitute for phosphate linkages in the backbone of the molecule.

The present invention further contemplates recombinant nucleic acids including a recombinant construct comprising all or part of *Erg*. The recombinant construct may be capable of replicating autonomously in a host cell. Alternatively, the recombinant construct may become integrated into the chromosonal DNA of the host cell. Such a recombinant polynucleotide comprises a polynucleotide of genomic, cDNA, semi-synthetic or synthetic origin which, by virtue of its origin or manipulation: (i) is not associated with all or a portion of a polynucleotide with which it is associated in nature; (ii) is linked to a polynucleotide other than that to which it is linked in nature; or (iii) does not occur in nature. Where nucleic acids according to the invention include RNA, reference to the sequence shown should be construed as reference to the RNA equivalent with U substituted for T. Such constructs are useful to elevate Erg levels or to down-regulate *Erg* levels such as *via* antisense means or RNAi-mediated gene silencing. As will be well known to those of skill in the art, such constructs are also useful in generating animal models carrying modified alleles of *Erg* and, as pharmaceutical compositions for modulating the activity of Erg in a subject *in vivo*.

Genetically modified cells or non-human organisms may be provided in the form of cells or embryos for transplantation. Cells and embryos are preferably maintained in a frozen state and may optionally be distributed or sold with instructions for use.

In a further aspect, the present invention provides a genetically modified cell, or non-human animal comprising such cells, wherein a *Erg* gene is modified and the cell or animal produces a substantially enhanced level or activity of Erg polypeptide, or substantially reduced level or activity of Erg polypeptide compared to a non-modified animal of the same species, or is substantially incapable of producing Erg polypeptides.

The genetically modified cells and non-human animals may be a non-human primate, livestock animal, companion animal, laboratory test animal, captive wild animal, reptile, amphibian, fish, bird or other organism. Preferably the genetically modified non-human animal is a murine animal.

In one aspect, the modified cell or non-human animal is genetically modified and produces a substantially reduced level of Erg or is substantially incapable of producing Erg or produces Erg having substantially reduced or no activity.

Preferably an *Erg* gene is modified. Modification may be in one or both alleles and may optionally be within a regulatory region of the gene.

In another embodiment, the genetic modification resulting in a cell or animal capable of exhibiting a modified level or activity of Erg comprises genetic modification outside the Erg gene to cause expression of genetic or proteinaceous molecules which effectively modulate the activity of Erg or *Erg*.

In another aspect, the modified cell or non-human animal is genetically modified and substantially overproduces Erg having normal or altered activity relative to an unmodified cell or animal of the same species.

In yet another aspect, the invention provides a method of screening for or testing an agent capable of complementing a phenotype shown by a cell or non-human animal comprising a modified *Erg* nucleic acid or Erg polypeptide and exhibiting a substantially modified level

or activity of Erg polypeptide. Preferably, the cell or animal is contacted with the agent and its effect on the activity of Erg or its transcriptional targets determined. In one aspect the method comprises screening for mutants which exhibit a complementing phenotype and then mapping and identifying the modifying gene. In another aspect the method comprises screening for agents which enhance the level or activity of Erg in a normal or modified cell. In some embodiments, small-molecule libraries are screened for agents which directly or indirectly modulate Erg polypeptide activity. One method is described by Stegmaier *et al.*, *PLOS Medicine, 4*(4):702-714, 2007. Here, expression profiles diagnostic of an Erg-on activity and an Erg-off activity are chosen, and the ability of small-molecules to produce either the Erg-on or the Erg-off profile is determined. Antisense knockdown strategies for selecting the Erg-off activity are routine in the art and include ShRNAs directed against the Erg transcript.

In further embodiment, the subject invention provides a use of a cell or non-human animal comprising a modified *Erg* or Erg and exhibiting a substantially reduced level or activity of Erg in screening for or testing agents for use in the treatment or prophylaxis of a hematological disorders as described further herein.

A substantially reduced level or activity of Erg is conveniently assessed in terms of a percent reduction relative to normal cells or animals or pre-treatment/pre-administration. A substantial reduction includes one which results in detectable thrombocytopenia in a subject or aberrant haematopoietic cell activity. Alternatively, a reduced level of gene expression of transcription targets or a reporter thereof is detected. Preferably, the reduction is at least 20% compared to normal cells, more preferably about 25%, still more preferably at least about 30% reduction, more preferably at least about 40% reduction in Erg level or activity. The reduction may of course be complete loss of Erg activity in a cell or animal. A "modified" level or activity includes enhanced levels of Erg activity relative to pre-treatment levels and may equate to or exceed the level or activity of Erg detectable in healthy subjects, control cells or cell-free systems or in subjects unlikely to develop thrombocytopenia, stem cell or bone marrow defects.

The present invention further provides a method for identifying agents useful in the treatment or prophylaxis of hematological disorders such as described herein comprising screening compounds for their ability to modulate the functional activity of Erg polypeptides.

In a further aspect, the present invention provides a composition comprising an agent which up regulates the level or activity of Erg in a cell for use in modulating hematopoietic cell levels including stem cells, early progenitor cells and platelet cells *in vivo* or *in vitro*. In another embodiment down regulation or Erg activity will be useful in the treatment or prevention of clonal hemopathies including those hereinbefore described.

The modulatory agents of the present invention may be chemical agents such small or large organic or inorganic chemical molecules, peptides, polypeptides including dominant negative forms, modified peptides such as constrained peptides, foldamers, peptidomimetics, cyclic peptidomimetics, proteins, lipids, carbohydrates or nucleic acid molecules including antisense or other gene silencing molecules. Small molecules generally have a molecular mass of less than 500 Daltons. Large molecules generally include whole polypeptides or other compounds having a molecular mass greater than 500 Daltons. Agents may comprise naturally occurring molecules, variants (including analogs) thereof as defined herein or non-naturally occurring molecules. Gene silencing agents (genetic agents) such as DNA (gDNA, cDNA), RNA (sense RNAs, antisense RNAs, mRNAs, tRNAs, rRNAs, small interfering RNAs (SiRNAs), short hairpin RNAs (ShRNAs), micro RNAs (miRNAs), small nucleolar RNAs (SnoRNAs, small nuclear (SnRNAs)) ribozymes, aptamers, DNAzymes or other ribonuclease-type complexes may be employed.

Agents in accordance with this aspect of the invention may directly interact with Erg. Here, for example, small molecule antibodies or peptides, peptidomimetics or analogs and other such molecules may be conveniently employed. Alternatively, genetic mechanisms are used to indirectly modulate the activity of Erg. Again, various strategies are well documented and include mechanisms for pre or post-transcriptional silencing. The

expression of antisense molecules or co-suppression or RNAi or siRNA strategies are particularly contemplated.

Agents which modulate the level or activity of Erg or *Erg* may be derived from Erg or *Erg*.

Alternatively, they may be identified in *in vitro* or *in vivo* screens. Natural products, combinatorial, synthetic/peptide/polypeptide or protein libraries or phage display technology are all available to screening for such agents. Natural products include those from coral, soil, plant, or the ocean or antarctic environments. Small molecule libraries are particularly convenient.

In each case the agent to be tested is contacted with a system comprising Erg or *Erg*. Then, the following may be assayed for: the presence of a complex between the agent and Erg or *Erg*; a change in the interaction between Erg and a target; a change in the activity of the target, or a change in the level or activity of an indicator of the activity of the target. Competitive binding assays and other high throughput screening methods are well known in the art and are described for example in International Publication Nos. WO 84/03564 and WO 97/02048).

In one embodiment, all or part of the *Erg* gene promoter is operatively linked to a reporter construct and engineered into an expression construct as known to those of skill in the art. For example, a pGL3-series reporter plasmid may be conveniently employed. Stable or transient transfection of cells may be used to generate cell lines capable of being tested with potential agents.

In a cell based approach an Erg responsive cell line is generated comprising an inducible *Erg* gene. Potential agents are tested for their ability to up-regulate or down-regulate expression differentiation markers when Erg is activated. For example, an Erg responsive cell line is the human cancer line K562 referred in Ceballos *et al., (Oncogene, 19*:2194-2204, 2000). Such techniques are well known in the art and are described, for example, in Sambrook, 2001 (*supra*) and Ausubel, 2002 (*supra*).

Antisense polynucleotide sequences are useful agents in preventing or reducing the expression of *Erg*. Alternatively, morpholines may be used as described by Summerton *et al.*, (*Antisense and Nucleic Acid Drug Development* 7:187-195, 1997). Antisense molecules may interfere with any function of a nucleic acid molecule. The functions of DNA to be interfered with can include replication and transcription. Replication and transcription, for example, can be from an endogenous cellular template, a vector, a plasmid construct or otherwise. The functions of RNA to be interfered with can include functions such as translocation of the RNA to a site of protein translation, translocation of the RNA to sites within the cell which are distant from the site of RNA synthesis, translation of protein from the RNA, splicing of the RNA to yield one or more RNA species, and catalytic activity or complex formation involving the RNA which may be engaged in or facilitated by the RNA. One preferred result of such interference with target nucleic acid function is modulation of the expression of the *Erg* gene.

While the preferred form of antisense compound is a single-stranded antisense oligonucleotide, in many species the introduction of double-stranded structures, such as double-stranded RNA (dsRNA) molecules, has been shown to induce potent and specific antisense-mediated reduction of the function of a gene or its associated gene products. This phenomenon occurs in both plants and animals.

In the context of the subject invention, the term "oligomeric compound" refers to a polymer or oligomer comprising a plurality of monomeric units. In the context of this invention, the term "oligonucleotide" refers to an oligomer or polymer of ribonucleic acid (RNA) or deoxyribonucleic acid (DNA) or mimetics, chimeras, analogs and homologs thereof. This term includes oligonucleotides composed of naturally occurring nucleobases, sugars and covalent internucleoside (backbone) linkages as well as oligonucleotides having non-naturally occurring portions which function similarly. Such modified or substituted oligonucleotides are often preferred over native forms because of desirable properties such as, for example, enhanced cellular uptake, enhanced affinity for a target nucleic acid and increased stability in the presence of nucleases.

While oligonucleotides are a preferred form of the compounds of this invention, the present invention comprehends other families of compounds as well, including but not limited to oligonucleotide analogs and mimetics such as those described herein.

The genetic agents or compositions in accordance with this invention preferably comprise from about 8 to about 80 nucleobases (i.e. from about 8 to about 80 linked nucleosides). One of ordinary skill in the art will appreciate that the invention embodies compounds of 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, or 80 nucleobases in length.

As mentioned previously herein, the agents or compositions of the present invention may be Erg or parts thereof, or *Erg* or parts thereof or complementary forms or molecules derived or designed from Erg or *Erg*.

Thus, the present invention provides a composition comprising Erg or a functional variant or Erg or an agent from which either or these is producible which substantially enhances the activity of Erg. In some embodiments the said composition effectively modulates hematopoietic cell level or activity such as enhancing stem cell or progenitor cell level or activity, platelet levels, red blood cell levels, white blood cell levels, immune cell levels such as macrophages and monocyte level or activity.

Any subject who could benefit from the present methods or compositions is encompassed. The term "subject" includes, without limitation, humans and non-human primates, livestock animals, companion animals, laboratory test animals, captive wild animals, reptiles and amphibians, fish, birds and any other organism. The most preferred subject of the present invention is a human subject. A subject, regardless of whether it is a human or non-human organism may be referred to as a patient, individual, subject, animal, host or recipient.

The term "composition" and terms such as "agent", "medicament", "active" and "drug" are used interchangeably herein to refer to a chemical compound or cellular composition which induces a desired pharmacological and/or physiological effect. The terms encompass pharmaceutically acceptable and pharmacologically active ingredients including but not limited to salts, esters, amides, pro-drugs, active metabolites, analogs and the like. The term includes genetic and proteinaceous or lipid molecules or analogs thereof as well as cellular compositions as previously mentioned. The instant compounds and compositions are suitable for the manufacture of a medicament for the treatment and/or prevention of conditions associates with early defects in blood cell development i.e. at the level of HSC and/or progenitor cell activity.

In relation to cellular compositions, the present invention extends to cellular compositions including genetically modified stem cells which are capable of regenerating tissues and/or organs of an animal subject *in situ* or *in vivo*. Stem cells or stem cell-like cells are preferably multipotent or pluripotent. Other cellular compositions comprise vectors such as viral vectors for delivery of nucleic acid constructs as described later herein.

In relation to *Erg*, the terms "functional form" or "variant", "functionally equivalent derivative" or "homolog" include molecules which selectively hybridize to *Erg* or a complementary form thereof over all or part of the genetic molecule under conditions of low or medium stringency at a defined temperature or range of conditions, or which have about 60% sequence identity to a nucleotide sequence encoding Erg polypeptides.

Exemplary *Erg* nucleotide sequences include those comprising nucleotide sequences set forth in SEQ ID NO: 1 (mouse *Erg* mRNA), SEQ ID NO: 3 (human *Erg-1* mRNA) or SEQ ID NO: 5 (human *Erg-2* mRNA) or their complements. For the avoidance of doubt however, it should be noted that the term "*Erg*" expressly encompasses all forms of the gene including regulatory regions and genomic forms or specific fragments and constructs comprising same, or parts thereof.

In relation to Erg, Erg polypeptides include all biologically active naturally occurring forms of Erg as well as biologically active portions thereof and variants and derivatives of these. Biological activity as determined herein includes enhancing hematopoietic stem cell and/or hematopoietic progenitor activity and potentiating transcription of transcriptional targets. The terms functional form or variant, functionally equivalent derivatives or homologs include polypeptides comprising a sequence of amino acids having about 60% sequence identity to Erg.

Derivatives and variants are molecules which exhibit at least one biologically relevant function of the naturally occurring polypeptide such as DNA binding (such as *via* the ETS domain) or protein binding (such as *via* the pointer domain).

Exemplary Erg amino acid sequences include those comprising sequences set forth in SEQ ID NO: 2 (mouse Erg), SEQ ID NO: 4 (human Erg-1) and SEQ ID NO: 6 (human Erg-2).

Reference herein to a "low stringency" includes and encompasses from at least about 0 to at least about 15% v/v formamide and from at least about 1 M to at least about 2 M salt for hybridization, and at least about 1 M to at least about 2 M salt for washing conditions. Generally, low stringency is at from about 25-30°C to about 42°C. The temperature may be altered and higher temperatures used to replace formamide and/or to give alternative stringency conditions. Alternative stringency conditions may be applied where necessary, such as "medium stringency", which includes and encompasses from at least about 16% v/v to at least about 30% v/v formamide and from at least about 0.5 M to at least about 0.9 M salt for hybridization, and at least about 0.5 M to at least about 0.9 M salt for washing conditions, or high stringency, which includes and encompasses from at least about 31% v/v to at least about 50% v/v formamide and from at least about 0.01 M to at least about 0.15 M salt for hybridization, and at least about 0.01 M to at least about 0.15 M salt for washing conditions. In general, washing is carried out $T_m = 69.3 + 0.41$ (G+C)% (Marmur *et al., J. Mol. Biol. 5:* 109, 1962). However, the T_m of a duplex DNA decreases by 1°C with every increase of 1% in the number of mismatch base pairs (Bonner *et al., Eur. J. Biochem. 46:* 83, 1974). Formamide is optional in these hybridization conditions.

Accordingly, particularly preferred levels of stringency are defined as follows: low stringency is 6 x SSC buffer, 0.1% w/v SDS at 25-42°C; a moderate stringency is 2 x SSC buffer, 0.1% w/v SDS at a temperature in the range 20°C to 65°C; high stringency is 0.1 x SSC buffer, 0.1% w/v SDS at a temperature of at least 65°C.

The terms "similarity" or "identity" as used herein includes exact identity between compared sequences at the nucleotide or amino acid level. Where there is non-identity at the nucleotide level, "similarity" includes differences between sequences which result in different amino acids that are nevertheless related to each other at the structural, functional, biochemical and/or conformational levels. Where there is non-identity at the amino acid level, "similarity" includes amino acids that are nevertheless related to each other at the structural, functional, biochemical and/or conformational levels. In a particularly preferred embodiment, nucleotide sequence comparisons are made at the level of identity and amino acid sequence comparisons are made at the level of similarity.

Preferably, the precent similarity between a particular amino sequence and a reference sequence is at least about 65% or at least about 70% or at least about 80% or at least about 85% or more preferably at least about 90% similarity or above such as at least about 95%, 96%, 97%, 98%, 99% or greater. Percent similarities between 60 and 100% are encompassed.

Preferably, the precent identity between a particular nucleotide sequence and a reference sequence is at least about 65% or at least about 70% or at least about 80% or at least about 85% or more preferably at least about 90% similarity or above such as at least about 95%, 96%, 97%, 98%, 99% or greater. Percent identities between 60 and 100% are encompassed.

A "reference sequence" is at least 12 but frequently 15 to 18 and often at least 25 or above, such as 30 monomer units, inclusive of nucleotides and amino acid residues, in length. Because two polynucleotides may each comprise (1) a sequence (i.e. only a portion of the complete polynucleotide sequence) that is similar between the two polynucleotides, and (2)

a sequence that is divergent between the two polynucleotides, sequence comparisons between two (or more) polynucleotides are typically performed by comparing sequences of the two polynucleotides over a "comparison window" to identify and compare local regions of sequence similarity. A "comparison window" refers to a conceptual segment of typically 12 contiguous residues that is compared to a reference sequence. The comparison window may comprise additions or deletions (i.e. gaps) of about 20% or less as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Optimal alignment of sequences for aligning a comparison window may be conducted by computerised implementations of algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package Release 7.0, Genetics Computer Group, 575 Science Drive Madison, WI, USA) or by inspection and the best alignment (i.e. resulting in the highest percentage homology over the comparison window) generated by any of the various methods selected. Reference also may be made to the BLAST family of programs as, for example, disclosed by Altschul *et al., Nucl. Acids Res. 25:* 3389, 1997. A detailed discussion of sequence analysis can be found in Unit 19.3 of Ausubel *et al., Current Protocols in Molecular Biology* John Wiley & Sons Inc, 1994-1998, Chapter 15).

A percentage of sequence identity between nucleotide sequences, for example, is calculated by comparing two optimally aligned sequences over the window of comparison, determining the number of positions at which the identical nucleic acid base (e.g. A, T, C, G, I) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (i.e., the window size), and multiplying the result by 100 to yield the percentage of sequence identity. For the purposes of the present invention, "sequence identity" will be understood to mean the "match percentage" calculated by the DNASIS computer program (Version 2.5 for windows; available from Hitachi Software engineering Co., Ltd., South San Francisco, California, USA) using standard defaults as used in the reference manual accompanying the software. Similar comments apply in relation to sequence similarity for amino acid sequences.

The term "derivative" or the plural "derivatives" whether in relation to genetic or proteinaceous molecules includes, as appropriate, parts, mutants, fragments, and analogues as well as hybrid, chimeric or fusion molecules and glycosylation variants. Particularly useful derivatives retain at least one functional activity of the parent molecule and comprise single or multiple amino acid substitutions, deletions and/or additions to the Erg amino acid sequence. Preferably, the derivatives have functional activity or alternatively, modulate Erg functional activity. The term "modulate" includes up modulate or up-regulate and down-modulate or down-regulate.

As used herein reference to a part, portion or fragment of *Erg* is defined as having a minimal size of at least about 10 nucleotides or preferably about 13 nucleotides or more preferably at least about 20 nucleotides and may have a minimal size of at least about 35 nucleotides. This definition includes all sizes in the range of 10 to 35 as well as greater than 35 nucleotides. Thus, this definition includes nucleic acids of 12, 15, 20, 25, 40, 60, 100, 200, 500 nucleotides of nucleic acid molecules having any number of nucleotides between 500 and the number shown in SEQ ID NO: 1, SEQ ID NO:3 or SEQ ID NO: 5 or a complementary form thereof. The same considerations apply *mutatis mutandis* to any reference herein to a part, portion or fragment of Erg.

Substitutional variants typically contain the exchange of one amino acid for another at one or more sites within the protein and may be designed to modulate one or more properties of the polypeptide such as stability against proteolytic cleavage without the loss of other functions or properties. Amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphipathic nature of the residues involved. Preferred substitutions are ones which are conservative, that is, one amino acid is replaced with one of similar shape and charge. Conservative substitutions are well known in the art and typically include substitutions within the following groups: glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid; asparagine, glutamine; serine, threonine; lysine, arginine; and tyrosine, phenylalanine (see Table 3).

Certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies or binding sites on substrate molecules or binding sites on proteins interacting with the Erg polypeptide. Since it is the interactive capacity and nature of a protein which defines that protein's biological functional activity, certain amino acid substitutions can be made in a protein sequence and its underlying DNA coding sequence and nevertheless obtain a protein with like properties. In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydrophobic amino acid index in conferring interactive biological function on a protein is generally understood in the art (Kyte *et al., J. Mol. Biol., 157:*105-132, 1982). Alternatively, the substitution of like amino acids can be made effectively on the basis of hydrophilicity. The importance of hydrophilicity in conferring interactive biological function of a protein is generally understood in the art (U.S. Patent No. 4,554,101). The use of the hydrophobic index or hydrophilicity in designing polypeptides is further discussed in U.S. Patent No. 5,691,198.

The term "homolog" or "homologs" refers herein broadly to functionally and/or structurally related molecules including those from other species.

Reference herein to "mimetics" includes nucleic acid or peptide mimetics and it intended to refer to a substance which has conformational features allowing the substance to perform as a functional analog. A peptide mimetic may be peptide containing molecules that mimic elements of protein secondary structure (Johnson *et al "Peptide Turn Mimetics" in Biotechnology and Pharmacy, Pezzuto et al* eds Chapman and Hall, New York, 1993). Peptide mimetics may be identified by screening random peptide libraries such as phage display libraries for peptide molecules which mimic the functional activity of Erg. Alternatively, mimetic design, synthesis and testing are employed.

Nucleic acid mimetics include, for example, RNA analogs containing N3'--P5' phosphoramidate internucleotide linkages which replace the naturally occurring RNA O3'--

P5' phosphodiester groups. Enzyme mimetics include catalytic antibodies or their encoding sequences, which may also be humanised.

Peptide or non-peptide mimetics can be developed as functional analogues of Erg by identifying those residues of the target molecule which are important for function. Modelling can be used to design molecules which interact with the target molecule and which have improved pharmacological properties. Rational drug design permits the production of structural analogs of biologically active polypeptides of interest or of small molecules with which they interact (e.g. agonists, antagonists, inhibitors or enhancers) in order to fashion drugs which are, for example, more active or stable forms of the polypeptide, or which, e.g. enhance or interfere with the function of a polypeptide *in vivo*. See, e.g. Hodgson (*Bio/Technology, 9:*19-21, 1991). In one approach, one first determines the three-dimensional structure of a protein of interest by x-ray crystallography, by computer modelling or most typically, by a combination of approaches. Useful information regarding the structure of a polypeptide may also be gained by modelling based on the structure of homologous proteins. An example of rational drug design is the development of HIV protease inhibitors (Erickson *et al.*, *Science 249:*527-533, 1990). In addition, target molecules may be analyzed by an alanine scan (Wells, *Methods Enzymol., 202:*2699-2705, 1991). In this technique, an amino acid residue is replaced by Ala and its effect on the peptide's activity is determined. Each of the amino acid residues of the peptide is analyzed in this manner to determine the important regions of the peptide.

It is also possible to isolate a target-specific antibody, selected by a functional assay and then to solve its crystal structure. In principle, this approach yields a pharmacore upon which subsequent drug design can be based. It is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original receptor. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced banks of peptides. Selected peptides would then act as the pharmacore.

As briefly described, it is possible to design or screen for mimetics which have enhanced activity or stability or are more readily and/or more economically obtained.

Analogs of Erg or other agents described herein preferably have enhanced stability and activity. They may also be designed in order to have an enhanced ability to cross biological membranes or to interact with only specific substrates. Thus, analogs may retain some functional attributes of the parent molecule but may posses a modified specificity or be able to perform new functions useful in the present context i.e., for administration to the nucleus, bone marrow, etc.

Analogs contemplated herein include but are not limited to modification to side chains, incorporating of unnatural amino acids and/or their derivatives during peptide, polypeptide or protein synthesis and the use of crosslinkers and other methods which impose conformational constraints on the proteinaceous molecule or their analogs.

Examples of side chain modifications contemplated by the present invention include modifications of amino groups such as by reductive alkylation by reaction with an aldehyde followed by reduction with $NaBH_4$; amidination with methylacetimidate; acylation with acetic anhydride; carbamoylation of amino groups with cyanate; trinitrobenzylation of amino groups with 2, 4, 6-trinitrobenzene sulphonic acid (TNBS); acylation of amino groups with succinic anhydride and tetrahydrophthalic anhydride; and pyridoxylation of lysine with pyridoxal-5-phosphate followed by reduction with $NaBH_4$.

The guanidine group of arginine residues may be modified by the formation of heterocyclic condensation products with reagents such as 2,3-butanedione, phenylglyoxal and glyoxal.

The carboxyl group may be modified by carbodiimide activation *via* O-acylisourea formation followed by subsequent derivitization, for example, to a corresponding amide.

Sulphydryl groups may be modified by methods such as carboxymethylation with iodoacetic acid or iodoacetamide; performic acid oxidation to cysteic acid; formation of a mixed disulphides with other thiol compounds; reaction with maleimide, maleic anhydride or other substituted maleimide; formation of mercurial derivatives using 4-chloromercuribenzoate, 4-chloromercuriphenylsulphonic acid, phenylmercury chloride, 2-chloromercuri-4-nitrophenol and other mercurials; carbamoylation with cyanate at alkaline pH.

Tryptophan residues may be modified by, for example, oxidation with N-bromosuccinimide or alkylation of the indole ring with 2-hydroxy-5-nitrobenzyl bromide or sulphenyl halides. Tyrosine residues on the other hand, may be altered by nitration with tetranitromethane to form a 3-nitrotyrosine derivative.

Modification of the imidazole ring of a histidine residue may be accomplished by alkylation with iodoacetic acid derivatives or N-carbethoxylation with diethylpyrocarbonate.

Examples of incorporating unnatural amino acids and derivatives during peptide synthesis include, but are not limited to, use of norleucine, 4-amino butyric acid, 4-amino-3-hydroxy-5-phenylpentanoic acid, 6-aminohexanoic acid, t-butylglycine, norvaline, phenylglycine, ornithine, sarcosine, 4-amino-3-hydroxy-6-methylheptanoic acid, 2-thienyl alanine and/or D-isomers of amino acids. A list of unnatural amino acid, contemplated herein is shown in Table 4.

Crosslinkers can be used, for example, to stabilize 3D conformations, using homo-bifunctional crosslinkers such as the bifunctional imido esters having $(CH_2)_n$ spacer groups with n=1 to n=6, glutaraldehyde, N-hydroxysuccinimide esters and hetero-bifunctional reagents which usually contain an amino-reactive moiety such as N-hydroxysuccinimide and another group specific-reactive moiety such as maleimido or dithio moiety (SH) or carbodiimide (COOH). In addition, peptides can be conformationally constrained by, for

example, incorporation of C_α and N $_\alpha$-methylamino acids and the introduction of double bonds between C_α and C_β atoms of amino acids.

The small or large chemicals, polypeptides, nucleic acids, antibodies, peptides, chemical analogs, or mimetics of the present invention can be formulated in pharmaceutic compositions which are prepared according to conventional pharmaceutical compounding techniques. See, for example, Remington's Pharmaceutical Sciences, (20th ed. Williams and Wilkins (2000)). The composition may contain the active agent or pharmaceutically acceptable salts of the active agent. These compositions may comprise, in addition to one of the active substances, a pharmaceutically acceptable excipient, carrier, buffer, stabilizer or other materials well known in the art. Such materials should be non-toxic and should not interfere with the efficacy of the active ingredient. The carrier may take a wide variety of forms depending on the form of preparation desired for administration, e.g. intravenous, oral, intrathecal, epineural or parenteral. Accordingly, pharmaceutical compositions are provided comprising an active agent which modulates the activity of Erg or its transcriptional targets for use or when used in modulating hematopoietic cell activity as defined herein. In another embodiment, the use of the herein described agent is expressly contemplated in the manufacture of a medicament for the treatment of conditions associated with HSC defects associated with Erg variants. In some embodiments, the subject is tested for Erg variants prior to administration.

For oral administration, the compounds can be formulated into solid or liquid preparations such as capsules, pills, tablets, lozenges, powders, suspensions or emulsions. In preparing the compositions in oral dosage form, any of the usual pharmaceutical media may be employed, such as, for example, water, glycols, oils, alcohols, flavoring agents, preservatives, coloring agents, suspending agents, and the like in the case of oral liquid preparations (such as, for example, suspensions, elixirs and solutions); or carriers such as starches, sugars, diluents, granulating agents, lubricants, binders, disintegrating agents and the like in the case of oral solid preparations (such as, for example, powders, capsules and tablets). Because of their ease in administration, tablets and capsules represent the most advantageous oral dosage unit form, in which case solid pharmaceutical carriers are

obviously employed. If desired, tablets may be sugar-coated or enteric-coated by standard techniques. The active agent can be encapsulated to make it stable to passage through the gastrointestinal tract while at the same time allowing for passage across the blood brain barrier. See for example, International Patent Publication No. WO 96/11698.

For parenteral administration, the compound may dissolved in a pharmaceutical carrier and administered as either a solution or a suspension. Illustrative of suitable carriers are water, saline, dextrose solutions, fructose solutions, ethanol, or oils of animal, vegetative or synthetic origin. The carrier may also contain other ingredients, for example, preservatives, suspending agents, solubilizing agents, buffers and the like. When the compounds are being administered intrathecally, they may also be dissolved in cerebrospinal fluid.

The active agent is preferably administered in a therapeutically effective amount. The actual amount administered and the rate and time-course of administration will depend on the nature and severity of the condition being treated. Prescription of treatment, e.g. decisions on dosage, timing, etc. is within the responsibility of general practitioners or specialists and typically takes account of the disorder to be treated, the condition of the individual patient, the site of delivery, the method of administration and other factors known to practitioners. Examples of techniques and protocols can be found in Remington's Pharmaceutical Sciences, 2000 (*supra*).

Alternatively, targeting therapies may be used to deliver the active agent more specifically to certain types of cell, by the use of targeting systems such as antibodies or cell specific ligands. Targeting may be desirable for a variety of reasons, e.g. if the agent is unacceptably toxic or if it would otherwise require too high a dosage or if it would not otherwise be able to enter the target cells.

Instead of administering these agents directly, they could be produced in the target cell, e.g. in a viral vector such as those described above or in a cell based delivery system such as described in U.S. Patent No. 5,550,050 and International Patent Publication Nos. WO 92/19195, WO 94/25503, WO 95/01203, WO 95/05452, WO 96/02286, WO 96/02646,

WO 96/40871, WO 96/40959 and WO 97/12635. The vector could be targeted to the target cells or expression of expression products could be limited to specific cells, stages of decelopment or cell cycle stages. The cell based delivery system is designed to be implanted in a patient's body at the desired target site and contains a coding sequence for the target agent. Alternatively, the agent could be administered in a precursor form for conversion to the active form by an activating agent produced in, or targeted to, the cells to be treated. See, for example, European Patent Application No. 0 425 731A and International Patent Publication No. WO 90/07936.

Treatment or prophylaxis of blood disorders or genetic conditions involving them as referred to herein by gene or cell therapy is also contemplated. Specifically, expression constructs are produced comprising all or part of *Erg* nucleic acid sequences as described herein or variants thereof as described herein.

In accordance with this aspect of the present invention, the cells of a subject may be tested to determine whether gene or cell therapy with an agent comprising *Erg* is indicated. The provision of wild type or enhanced Erg function to a cell which carries a mutant or altered form of *Erg* should in this situation complement the deficiency and result in an improvement in the subject. Alternatively, cells capable of providing normal or enhanced Erg function may be provided. The *Erg* allele may be introduced into a cell in a vector such that the gene remains extrachromosomally. Alternatively, artificial chromosomes may be used. Typically, the vector may combine with the host genome and be expressed therefrom.

Gene therapy would be carried out according to generally accepted methods, for example, as described by Friedman (In: *Therapy for Genetic Disease*, T. Friedman, Ed., Oxford University Press, pp. 105-121, 1991) or Culver (*Gene Therapy: A Primer for Physicians*, 2nd Ed., Mary Ann Liebert, 1996). Suitable vectors are known, such as disclosed in U.S. Patent No. 5,252,479, International Patent Publication No. WO 93/07282 and U.S. Patent No. 5,691,198. Gene transfer systems known in the art may be useful in the practice of the

gene therapy methods of the present invention. These include viral and non-viral transfer methods as well known in the art.

Non-viral gene transfer methods are also known in the art such as chemical techniques including calcium phosphate co-precipitation, mechanical techniques, for example, microinjection, membrane fusion-mediated transfer *via* liposomes and direct DNA uptake and receptor-mediated DNA transfer. Viral-mediated gene transfer can be combined with direct *in vivo* gene transfer using liposome delivery.

In an approach which combines biological and physical gene transfer methods, plasmid DNA of any size is combined with a polylysine-conjugated antibody specific to the adenovirus hexon protein and the resulting complex is bound to an adenovirus vector. The trimolecular complex is then used to infect cells. The adenovirus vector permits efficient binding, internalization and degradation of the endosome before the coupled DNA is damaged. For other techniques for the delivery of adenovirus based vectors, see U.S. Patent No. 5,691,198. Liposome/DNA complexes are also capable of mediating direct *in vivo* gene transfer.

Expression vectors in the context of gene therapy are meant to include those constructs containing sequences sufficient to express a polynucleotide that has been cloned therein. In viral expression vectors, the construct contains viral sequences sufficient to support packaging of the construct. If the polynucleotide encodes Erg, expression will produce Erg. If the polynucleotide encodes a sense or antisense polynucleotide or a ribozyme or DNAzyme, expression will produce the sense or antisense polynucleotide or ribozyme or DNAzyme. Thus, in this context, expression does not require that a protein product be synthesized. In addition to the polynucleotide cloned into the expression vector, the vector also contains a promoter functional in eukaryotic cells. The cloned polynucleotide sequence is under control of this promoter. Suitable eukaryotic promoters are routinely determined.

Receptor-mediated gene transfer may be achieved by conjugation of DNA to a protein ligand *via* polylysine. Ligands are chosen on the basis of the presence of the corresponding ligand receptors on the cell surface of the target cell/tissue type. These ligand-DNA conjugates can be injected directly into the blood if desired and are directed to the target tissue where receptor binding and internalization of the DNA-protein complex occurs. To overcome the problem of intracellular destruction of DNA, co-infection with adenovirus can be included to disrupt endosome function.

Accordingly, patients who carry an aberrant *Erg* allele are treated with a gene delivery vehicle such that some or all of their cells receive at least one additional copy of a functional normal Erg allele. Preferably only specific cells are targeted.

Alternatively, peptides or mimetics or other functional analogues which have Erg activity can be supplied to cells which carry aberrant *Erg* alleles. Protein can be produced by expression of the cDNA sequence in bacteria, for example, using known expression vectors. In addition, synthetic chemistry techniques can be employed to synthesize the instant active molecules. Active molecules can be introduced into cells by microinjection or by use of liposomes, for example. Alternatively, some active molecules may be taken up by cells, actively or by diffusion. In some embodiments, supply of molecules with Erg activity should lead to enhanced blood cell function, enhanced HSC and/or progenitor cell activity.

Diseases or a susceptibility to these diseases can now be diagnosed by monitoring subjects for modification in the level or activity of Erg or specific mutations or aberrations (such a methylation events) in *Erg*.

One particular mutation results in a substitution of proline for serine in the DNA binding domain of Erg.

A wide range of mutation detection screening methods are available as would be known to those skilled in the art. Any method which allows an accurate comparison between a test

and control nucleic acid sequence may be employed. Scanning methods include sequencing, denaturing gradient gel electrophoresis (DGGE), single-stranded conformational polymorphism (SSCP and rSSCP, REF-SSCP), chemical cleavage methods such as CCM, ECM, DHPLC and MALDI-TOF MS and DNA chip technology. Specific methods to screen for pre-determined mutations include allele specific oligonucleotides (ASO), allele specific amplification, competitive oligonucleotide priming, oligonucleotide ligation assay, base-specific primer extension, dot blot assays and RFLP-PCR. The strengths and weaknesses of these and further approaches are reviewed in Sambrook, *Chapter 13*, *Molecular Cloning*, 2001.

By identifying *Erg* as subject to mutations which affect the level or activity of Erg, the present invention provides methods of diagnosis of conditions associated with modified *Erg* and further provides genetic or protein based methods of determining the susceptibility of a subject to develop these conditions.

The diagnostic and prognostic methods of the present invention detect or assess an aberration in the wild type *Erg* gene or locus to determine if *Erg* will be produced or if it will be over-produced or under-produced. The term "aberration" in the *Erg* gene or locus encompasses all forms of mutations including deletions, insertions, point mutations and substitutions in the coding and non-coding regions of *Erg*. It also includes changes in methylation patterns of *Erg* or of an allele of *Erg*. Deletions may be of the entire gene or only a portion of the gene. Point mutations may result in stop codons, frameshift mutations or amino acid substitutions. Somatic mutations are those which occur only in certain tissues, e.g. in the tumor tissue and are not inherited in the germline. Germline mutations can be found in any of a body's tissues and are inherited. An *Erg* allele which is not deleted (e.g. that found on the sister chromosome to a chromosome carrying a *Erg* deletion) can be screened for other mutations such as insertions, small deletions, point mutations and changes in methylation pattern. It is considered in accordance with the present invention that many mutations found in cells such as hepatic cells are those leading to decreased or increased expression of the *Erg* gene.

Useful diagnostic techniques to detect aberrations in the *Erg* gene include but are not limited to fluorescent *in situ* hybridization (FISH), direct DNA sequencing, PFGE analysis, Southern blot analysis, single-stranded coformational analysis (SSCA), Rnase protection assay, allele-specific oligonucleotide (ASO hybridization), dot blot analysis and PCR-SSCP (see below). Also useful is DNA microchip technology.

Predisposition to conditions associated with stem cell defects can be ascertained by testing any tissue of a human or other mammal for mutations in a *Erg* gene. This can be determined by testing DNA from any tissue of a subject's body. In addition, pre-natal diagnosis can be accomplished by testing fetal cells, placental cells or amniotic fluid for mutations of the *Erg* gene. Alteration of a wild type allele whether, for example, by point mutation or by deletion or by methylation, can be detected by any number of means.

There are several methods that can be used to detect DNA sequence variation. Direct DNA sequencing, either manual sequencing or automated fluorescent sequencing, can detect sequence variation. Another approach is the single-stranded conformation polymorphism assay (SSCP) (Orita *et al., Proc. Nat. Acad. Sci. USA, 86:*2776-2770, 1989). This method can be optimized to detect most DNA sequence variation. The increased throughput possible with SSCP makes it an attractive, viable alternative to direct sequencing for mutation detection on a research basis. The fragments which have shifted mobility on SSCP gels are then sequenced to determine the exact nature of the DNA sequence variation. Other approaches based on the detection of mismatches between the two complementary DNA strands include clamped denaturing gel electrophoresis (CDGE) (Sheffield *et al., Am. J. Hum. Genet., 49:*699-706, 1991), heteroduplex analysis (HA) (White *et al., Genomics, 12:*301-306, 1992) and chemical mismatch cleavage (CMC) (Grompe *et al., Proc. Natl. Acad. Sci. USA, 86:*5855-5892, 1989). Other methods which might detect mutations in regulatory regions or which might comprise large deletions, duplications or insertions include the protein truncation assay or the asymmetric assay. A review of methods of detecting DNA sequence variation can be found in Grompe (*Proc. Natl. Acad. Sci. USA, 86:*5855-5892, 1993). Once a mutation is known, an allele specific detection approach such as allele specific oligonucleotide (ASO) hybridization can be

utilized to rapidly screen large numbers of other samples for that same mutation. Such a technique can utilize probes which are labeled with gold nanoparticles to yield a visual color result (Elghanian *et al.*, *Science 277:* 1078-1081, 1997).

Other tests for confirming the presence or absence of a wild type or mutant *Erg* allele include single-stranded conformation analysis (SSCA) (Orita *et al.*, 1989 (*supra*)); denaturing gradient gel electrophoresis (DGGE) (Wartell *et al.*, *Nucl. Acids Res., 18:*2699-2705, 1990; Sheffield *et al.*, *Proc. Natl. Acad. Sci. USA, 86:*232-236, 1989); RNase protection assays (Finkelstein *et al.*, *Genomics, 7:*167-172, 1990; Kinszler *et al.*, *Science, 251:*1366-1370, 1991); denaturing HPLC; allele-specific oligonucleotide (ASO hybridization) (Conner *et al.*, *Proc. Natl. Acad. Sci. USA, 80:*278-282, 1983); the use of proteins which recognize nucleotide mismatches such as the *E. coli* mutS protein (Modrich, *Ann. Rev. Genet., 25:*229-253, 1991) and allele-specific PCR (Ruano *et al.*, *Nucl. Acids. Res., 17:*8392, 1989). For allele-specific PCR, primers are used which hybridize at their 3' ends to a particular *Erg* mutation or to junctions of DNA caused by a deletion of *Erg*. If the particular *Erg* mutation is not present, an amplification product is not observed. Amplification Refractory Mutation System (ARMS) can also be used, as disclosed in European Patent Publication No. 0 332 435 and in Newtown *et al.* (*Nucl. Acids. Res., 17:*2503-2516, 1989). Insertions and deletions of genes can also be detected by cloning, sequencing and amplification.

Microchip technology is also applicable to the present invention. In this technique, thousands of distinct oligonucleotide or cDNA probes are built up in an array on a silicon chip or other solid support such as polymer films and glass slides. Nucleic acid to be analyzed is labelled with a reporter molecule (e.g. fluorescent label) and hybridized to the probes on the chip. It is also possible to study nucleic acid-protein interactions using these nucleic acid microchips. Using this technique, one can determine the presence of mutations or sequence the nucleic acid being analyzed or one can measure expression levels of a gene of interest or multiple genes of interest such as genes encoding products in a biochemical pathway. The technique is described in a range of publications including Hacia *et al.* (*Nature Genetics, 14:*441-447, 1996), Shoemaker *et al.* (*Nature Genetics, 14:*450-456,

1996), Chee *et al.* (*Science, 274:*610-614, 1996), Lockhart *et al.* (*Nature Biotechnology, 14:*1675-1680, 1996), DiRisi *et al.* (*Nature Genetics, 14:*457-460, 1996) and Lipshutz *et al.* (*Biotechniques, 19:*442-447, 1995).

Alteration of wild type *Erg* genes can also be detected by screening for alteration of wild type Erg proteins. For example, monoclonal antibodies immunoreactive with Erg can be used to screen a tissue. Lack of cognate antigen would indicate an Erg mutation. Antibodies specific for products of mutant alleles could also be used to detect mutant Erg gene product. Such immunological assays can be done in any convenient format known in the art. These include Western blots, immunohistochemical assays and ELISA assays.

The use of monoclonal antibodies in an immunoassay is particularly preferred because of the ability to produce them in large quantities and the homogeneity of the product. The preparation of hybridoma cell lines for monoclonal antibody production is derived by fusing an immortal cell line and lymphocytes sensitized against the immunogenic preparation (i.e. comprising Erg) or can be done by techniques which are well known to those who are skilled in the art. (See, for example, Douillard and Hoffman, Basic Facts about Hybridomas, in *Compendium of Immunology* Vol. II, ed. by Schwartz, 1981; Kohler *et al.*, *Nature, 256:*495-499, 1975; Kohler *et al., European Journal of Immunology. 6:*511-519, 1976).

Examples of primers used to amplify regions of *Erg* genetic sequence are routinely derived by the skilled addressee based on known sequences for *Erg*.

The present invention is further described by the following non-limiting Examples.

EXAMPLE 1

Identification of the mld2 mutation in a forward genetic screen on a sensitised $Mpl^{-/-}$ background

In order to screen for dominant hematopoietic phenotypes that may involve HSCs or multipotent progenitors, $Mpl^{-/-}$ C57BL/6 mice were injected with ENU as described previously and bred with isogenic females (Carpinelli *et al.*, *Proc. Natl. Acad. Sci. U. S. A.*, *101*:6553-6558, 2004). Blood was taken from their progeny at 7 weeks and analysed on an automated hematological analyser. One mouse had low numbers of RBCs, WBCs and platelets relative to $Mpl^{-/-}$ mice. The mutation was responsible for this multilineage defect was designated *mld2* (Figure 1).

EXAMPLE 2

mld2/+ mice are anaemic, leukopenic, thrombocytopenic and have reduced survival on a $Mpl^{-/-}$ genetic background

The *mld2*/+ $Mpl^{-/-}$ founder mouse was mated to +/+ $Mpl^{-/-}$ mice to further investigate the *mld2* phenotype and its inheritance. The progeny were bled at seven weeks. The RBC counts of these progeny split into two clearly demarcated populations and genotypes were inferred on this basis (*mld2*/+ $Mpl^{-/-}$ for RBC < 8.0 per pL, +/+ $Mpl^{-/-}$ for RBC > 8.0 per pL, Figure 2). The *mld2*/+ $Mpl^{-/-}$ mice have an approximately 47% deficiency in RBCs. Using such inferred genotypes, *mld2*/+ $Mpl^{-/-}$ also appear to have deficiencies in all lineages of leukocytes (54% deficient overall) as well as platelets (80% deficient, Figure 2). Upon identification of the mutated gene (see below), data from accurately genotyped mice confirmed this finding.

In addition to this multilineage deficiency, *mld2*/+ $Mpl^{-/-}$ mice also showed decreased survival (Figure 3). As above, this data is based on genotypes inferred from RBC counts. None of the 13 *mld2*/+ $Mpl^{-/-}$ progeny of the original mouse identified in the screen survived longer than 205 days, and only one survived longer than 110 days, before succumbing to illness. Preliminary data from post mortem analyses suggests that this

illness is associated with internal haemorrhaging under the skin, around the joints and in the skull. Conversely, none of the +/+ *Mpl*$^{-/-}$ littermates became ill. Preliminary data from accurately genotyped mice supports this data. A more comprehensive survival and pathology experiment is performed with a dedicated cohort of un-manipulated mice. A dominant lethal phenotype makes maintaining a colony difficult, however the original mouse identified in the screen bred, and sperm was frozen from an additional affected male. Moreover, the *mld2* lethality is dependent on a Mpl$^{-/-}$ phenotype (see below), and thus a colony of mice carrying the *mld2* mutation in the absence of lethal disease is maintained.

EXAMPLE 3

mpl$^{-/-}$ *mice heterozygous for the mld2 mutation have a severe deficiency in multipotent progenitor activity*

In order to investigate HSC and multipotent progenitor cell function, a CFU-S assay was performed. Bone marrow cells 1.5 x 10^6 from two *mld2*/+ *Mpl*$^{-/-}$ mice and one +/+ *Mpl*$^{-/-}$ mouse were injected into five irradiated recipients each. The spleens were taken after twelve days and colonies counted. The *mld2*/+ *Mpl*$^{-/-}$ bone marrow did not yield any detectable spleen colonies, whereas the *mld2*/+ *Mpl*$^{-/-}$ mice yielded 3.3 ± 2.1 colonies per 1.5 x 10^5 bone marrow cells. This provides compelling evidence of a profound, bone marrow intrinsic deficiency in HSCs and/or multipotent progenitors in *mld2*/+ *Mpl*$^{-/-}$ mice.

A deficiency in the HSC compartment of *mld2*/+ *Mpl*$^{-/-}$ mice is further supported by data from competitive reconstitution assays. Here, 1 x 10^6 Ly5.1$^-$Ly5.2$^+$ test cells (the same as those used for CFU-S assay) were transplanted with 1 x 10^6 Ly5.1$^+$Ly5.2$^-$ +/+ *Mpl*$^{-/-}$ competitor cells into five Ly5.1$^+$Ly5.2$^+$ irradiated recipients. In this system, the contribution of test, donor and endogenous cells to hematopoiesis may be easily quantitated by fluorescence activated cell sorting. The contribution of each cell population to hematopoiesis four weeks after transplantation was determined based on the number of peripheral blood cells with each Ly5 phenotype (Table 5). This demonstrated that *mld2*/+ *Mpl*$^{-/-}$ marrow was able to make only negligible contributions to hematopoiesis (2.1% and

1.2% of mature blood cells were derived from the two *mld2/+ Mpl*$^{-/-}$ donors) compared to a 22.4% contribution by +/+ *Mpl*$^{-/-}$ test marrow.

Together, these data suggest that *mld2/+ Mpl*$^{-/-}$ mice have a severe deficiency of HSCs and/or multipotential progenitor activity. More experiments will be performed to determine whether this is due to deficient numbers, functional activity or homing. Transplantation studies will be undertaken incorporating analysis of purified HSCs, as well as a full analysis of hematopoietic precursor cells, and semi-solid culture assays. These will confirm the effect of *mld2* on the activity of HSCs, multipotent progenitors and lineage committed progenitors.

EXAMPLE 4

The mld2 mutation exists in a 1.6 megabase interval on chromosome 16

The *mld2* mutation was localized using the same principles as for *mld3* (see above). Briefly, *mld2/+ Mpl*$^{-/-}$ C57BL/6 mice were crossed to wild type Balb/c mice. The G_1 progeny were then crossed to +/+ *Mpl*$^{-/-}$ C57BL/6 mice. Therefore, in the resultant G_2 mice, the genotypes of the genetic markers nearest to the *mld2* mutation will be homozygous C57BL/6 in all mice carrying *mld2*, and conversely they will be heterozygous in those not carrying it. The *mld2* mutation can be localised by finding the smallest region of markers for which this remains true. Which Mpl$^{-/-}$ G_2 mice carried *mld2* was easily determined the presence of the characteristic the multilineage defect, and the capacity to pass this defect to their progeny. Using this strategy, the *mld2* mutation was localised to a 1.6 megabase interval on the distal end of chromosome 16. This interval contains 8 genes which are all investigated.

The phenotype of mice heterozygous for the **mld2** *mutation varies with the presence of functional c-Mpl*

When *mld2/+ Mpl*$^{-/-}$ mice with multilineage deficiencies are crossed with wild type mice, none of the G_1 progeny have an obvious multilineage defect. However if these G_1 mice are

crossed with $Mpl^{-/-}$ mice, then approximately one quarter of the G_2 progeny show the multilineage defect. This suggests that the severe defect in *mld2/+* mice is dependent on a $Mpl^{-/-}$ mice genetic background.

Localization of *mld2* to a small candidate interval allowed the mixed background G_2 mice used for mapping to be genotyped for the *mld2* mutation. This allowed the comparison of *mld2/+* $Mpl^{+/-}$ and +/+ $Mpl^{+/-}$ phenotypes for a significant number of mice (Figure 4). *mld2/+* $Mpl^{-/-}$ mice demonstrated moderate but statistically significant deficiencies in WBCs (~16% deficiency) and platelets (~23% deficiency). Preliminary data suggests that there is no apparent difference between the *mld2/+* $Mpl^{+/-}$ and *mld2/+* $Mpl^{+/+}$ phenotypes.

There is a significant overlap between the counts of the two populations due to variation in WBC and platelet counts, making it impossible to predict the *mld2* genotype of an individual $Mpl^{+/-}$ mouse, even though there are clear deficiencies at the population level. This indicates the value of sensitised screens, where the increased severity of otherwise moderate phenotypes on a compromised background, makes the identification of mice with relevant mutations during screens more likely.

EXAMPLE 5

No mice homozygous for mld2 survive to birth

In order to determine the phenotype of mice homozygous for the mld2 mutation, G_1 mice from the mapping cross, which were *mld2/+* $Mpl^{+/-}$, were mated to one another. Because of the mixed background of these mice, their *mld2* genotype could be determined by SSLP genotyping of genetic markers surrounding the vicinity of the *mld2* mutation. Of 51 such progeny, genotyped at weaning, none carried two copies of the *mld2* mutation, although ~13 were expected to (Table 6). Instead, the ratio of genotypes is similar to that predicted if the *mld2/mld2* genotype is assumed to be lethal. The *mld2/mld2* genotype appears to be lethal regardless of the Mpl genotype (data not shown). Presumably the *mld2/mld2* mice are dying as embryos. The stage at which the embryos are dying is determined, and attempt to determine the cause of death, with a particular focus on hematopoietic defects.

EXAMPLE 6

The mld2 mutation causes a serine to proline substitution in the DNA binding domain of the ETS transcription factor Erg.

The 1.6 megabase interval known to contain the *mld2* mutation contained 8 genes. One of these genes is the ETS transcription factor *Erg*. Several ETS transcription factors have been implicated in hematopoiesis, so *Erg* was identified as a strong candidate gene to carry the *mld2* mutation. The coding exons and intron-exon boundaries of Erg were sequenced. A T to C nucleotide substitution was identified in exon 12 of *Erg* in all *mld2/+* mice, but in no +/+ mice (Figure 5). This causes a serine to proline substitution in the ETS domain of *Erg*. This mutation lies immediately distal to an alpha helix highly conserved amongst the 26 ETS family genes in the mouse, and so may be predicted to disrupt the function of this DNA binding domain.

The *mld2* gene and wild type *Erg* cDNA are cloned using standard procedure (Ausubel). Whether the *mld2* mutation prevents DNA binding will be determined using electrophoretic mobility shift assays. The effect of the mutation on function will be determined using reporter gene assays and biological assays such as the ability to induce megakaryocytic differentiation of K562 cells. The *mld2* multilineage phenotype will be correcting through the infection of wild type or mutant bone marrow with wild type or mutant *Erg*, or through the expression of wild type or mutant *Erg* in transgenic mice.

Because Erg is a transcription factor, it functions by modulating the expression of other genes. The mechanism by which the disruption of *Erg* by *mld2* causes the *mld2* phenotype, will be investigated by defining the target genes of *Erg* that are critical to the *mld2* phenotype. This is achieved using techniques such as chromatin immunoprecipitation and microarrays. The targets of Erg are themselves regulators of hematopoiesis, including HSCs and multipotent progenitor function. The *mld2/+ Mpl*$^{-/-}$ mice model system is used to determine the function of candidate hematopoietic regulators and their variants.

The identification of nucleic acid sequences bound by Erg is conveniently assessed using genome-wide location analyses such as ChIP-on Gene analysis such as described by Horak *et al.*, (*Methods Enzymol., 350*:469-483, 2002). In some embodiments Erg regulates the expression of hematopoietic signally molecules such as cytokine, hormone and chemokine. In a particular embodiment, the signalling molecules or their cellular or nuclear receptors are expressed early in hematopoiesis. Exemplary molecules include thrombopoietin and/or Mpl and/or Scl.

To identify transcriptional targets of Erg, the expression profile of normal hematopoietic cells, and cells that carry the *mld2* mutation in Erg are compared, for example, by microarray analysis. Interesting genes implicated in the Erg pathway are confirmed by real-time RT-PCR and followed up *in vivo* by transgenic and knockout mouse approaches. Another method uses genome-wide mapping of protein-DNA interactions by chromatin immunoprecipitation and DNA microarray hybridisation (ChIP-on-Chip) sequence location analyses. Chromatin immunoprecipitation (ChIP) using an antibody to Erg separates Erg from a cell lysate together with nucleic acid sequences to which Erg is bound. Then, by purifying the DNA, and running it across a microarray (Chip), the sequence of the bound nucleic acids can be determined and how often they are represented in the cell lysate. This indicates the gene promoters to which Erg binds in a given cell, and therefore, its transcriptional targets.

In another approach ENU mutagenesis modifier screen are performed i.e. a screen for mutations that can enhance or suppress the phenotypes caused by *mld2* Erg mutation in Erg. This will identify two classes of genes: 1) those that act within the Erg pathway, and 2) genes that act outside the Erg pathway, but which when mutated, produce biologic effects that can enhance or suppress that induced by mutations in Erg. The present invention therefore also contemplated modified animals comprising two or more different mutations the Erg signalling pathway or its transcriptional targets.

Further, the class 2) genes which when mutated suppresses the *mld2/+* phenotype are

further targets for the development of agents useful in the present invention.

Those skilled in the art will appreciate that the invention described herein is susceptible to variations and modifications other than those specifically described. It is to be understood that the invention includes all such variations and modifications. The invention also includes all of the steps, features, compositions and compounds referred to or indicated in this specification, individually or collectively, and any and all combinations of any two or more of said steps or features.

Table 1

Summary of sequence identifiers

SEQUENCE ID NO:	DESCRIPTION
1	nucleotide (cDNA) sequence of mouse Erg
2	amino acid sequence encoded by SEQ ID NO: 1
3	nucleotide (cDNA) sequence of human *Erg-1*
4	amino acid sequence encoded by SEQ ID NO: 3
5	nucleotide sequence (cDNA) of human *Erg-2*
6	amino acid sequence encoded by SEQ ID NO: 5

Table 2

Amino acid sub-classification

Sub-classes	Amino acids
Acidic	Aspartic acid, Glutamic acid
Basic	Noncyclic: Arginine, Lysine; Cyclic: Histidine
Charged	Aspartic acid, Glutamic acid, Arginine, Lysine, Histidine
Small	Glycine, Serine, Alanine, Threonine, Proline
Polar/neutral	Asparagine, Histidine, Glutamine, Cysteine, Serine, Threonine
Polar/large	*Asparagine, Glutamine*
Hydrophobic	*Tyrosine, Valine, Isoleucine, Leucine, Methionine, Phenylalanine, Tryptophan*
Aromatic	*Tryptophan, Tyrosine, Phenylalanine*
Residues that influence chain orientation	*Glycine and Proline*

Table 3

Exemplary and Preferred Amino Acid Substitutions

Original Residue	EXEMPLARY SUBSTITUTIONS	PREFERRED SUBSTITUTIONS
Ala	Val, Leu, Ile	Val
Arg	Lys, Gln, Asn	Lys
Asn	Gln, His, Lys, Arg	Gln
Asp	Glu	Glu
Cys	Ser	Ser
Gln	Asn, His, Lys,	Asn
Glu	Asp, Lys	Asp
Gly	Pro	Pro
His	Asn, Gln, Lys, Arg	Arg
Ile	Leu, Val, Met, Ala, Phe, Norleu	Leu
Leu	Norleu, Ile, Val, Met, Ala, Phe	Ile
Lys	Arg, Gln, Asn	Arg
Met	Leu, Ile, Phe	Leu
Phe	Leu, Val, Ile, Ala	Leu
Pro	Gly	Gly
Ser	Thr	Thr
Thr	Ser	Ser
Trp	Tyr	Tyr
Tyr	Trp, Phe, Thr, Ser	Phe
Val	Ile, Leu, Met, Phe, Ala, Norleu	Leu

Table 4

Codes for non-conventional amino acids

Non-conventional amino acid	Code	Non-conventional amino acid	Code
α-aminobutyric acid	Abu	L-N-methylalanine	Nmala
α-amino-α-methylbutyrate	Mgabu	L-N-methylarginine	Nmarg
aminocyclopropane-carboxylate	Cpro	L-N-methylasparagine	Nmasn
		L-N-methylaspartic acid	Nmasp
aminoisobutyric acid	Aib	L-N-methylcysteine	Nmcys
aminonorbornyl-carboxylate	Norb	L-N-methylglutamine	Nmgln
		L-N-methylglutamic acid	Nmglu
cyclohexylalanine	Chexa	L-Nmethylhistidine	Nmhis
cyclopentylalanine	Cpen	L-N-methylisolleucine	Nmile
D-alanine	Dal	L-N-methylleucine	Nmleu
D-arginine	Darg	L-N-methyllysine	Nmlys
D-aspartic acid	Dasp	L-N-methylmethionine	Nmmet
D-cysteine	Dcys	L-N-methylnorleucine	Nmnle
D-glutamine	Dgln	L-N-methylnorvaline	Nmnva
D-glutamic acid	Dglu	L-N-methylornithine	Nmorn
D-histidine	Dhis	L-N-methylphenylalanine	Nmphe
D-isoleucine	Dile	L-N-methylproline	Nmpro
D-leucine	Dleu	L-N-methylserine	Nmser
D-lysine	Dlys	L-N-methylthreonine	Nmthr
D-methionine	Dmet	L-N-methyltryptophan	Nmtrp
D-ornithine	Dorn	L-N-methyltyrosine	Nmtyr
D-phenylalanine	Dphe	L-N-methylvaline	Nmval
D-proline	Dpro	L-N-methylethylglycine	Nmetg
D-serine	Dser	L-N-methyl-t-butylglycine	Nmtbug
D-threonine	Dthr	L-norleucine	Nle

	D-tryptophan	Dtrp	L-norvaline	Nva
	D-tyrosine	Dtyr	α-methyl-aminoisobutyrate	Maib
	D-valine	Dval	α-methyl-γ-aminobutyrate	Mgabu
	D-α-methylalanine	Dmala	α-methylcyclohexylalanine	Mchexa
5	D-α-methylarginine	Dmarg	α-methylcylcopentylalanine	Mcpen
	D-α-methylasparagine	Dmasn	α-methyl-α-napthylalanine	Manap
	D-α-methylaspartate	Dmasp	α-methylpenicillamine	Mpen
	D-α-methylcysteine	Dmcys	N-(4-aminobutyl)glycine	Nglu
	D-α-methylglutamine	Dmgln	N-(2-aminoethyl)glycine	Naeg
10	D-α-methylhistidine	Dmhis	N-(3-aminopropyl)glycine	Norn
	D-α-methylisoleucine	Dmile	N-amino-α-methylbutyrate	Nmaabu
	D-α-methylleucine	Dmleu	α-napthylalanine	Anap
	D-α-methyllysine	Dmlys	N-benzylglycine	Nphe
	D-α-methylmethionine	Dmmet	N-(2-carbamylethyl)glycine	Ngln
15	D-α-methylornithine	Dmorn	N-(carbamylmethyl)glycine	Nasn
	D-α-methylphenylalanine	Dmphe	N-(2-carboxyethyl)glycine	Nglu
	D-α-methylproline	Dmpro	N-(carboxymethyl)glycine	Nasp
	D-α-methylserine	Dmser	N-cyclobutylglycine	Ncbut
	D-α-methylthreonine	Dmthr	N-cycloheptylglycine	Nchep
20	D-α-methyltryptophan	Dmtrp	N-cyclohexylglycine	Nchex
	D-α-methyltyrosine	Dmty	N-cyclodecylglycine	Ncdec
	D-α-methylvaline	Dmval	N-cylcododecylglycine	Ncdod
	D-N-methylalanine	Dnmala	N-cyclooctylglycine	Ncoct
	D-N-methylarginine	Dnmarg	N-cyclopropylglycine	Ncpro
25	D-N-methylasparagine	Dnmasn	N-cycloundecylglycine	Ncund
	D-N-methylaspartate	Dnmasp	N-(2,2-diphenylethyl)glycine	Nbhm
	D-N-methylcysteine	Dnmcys	N-(3,3-diphenylpropyl)glycine	Nbhe
	D-N-methylglutamine	Dnmgln	N-(3-guanidinopropyl)glycine	Narg
	D-N-methylglutamate	Dnmglu	N-(1-hydroxyethyl)glycine	Nthr
30	D-N-methylhistidine	Dnmhis	N-(hydroxyethyl))glycine	Nser

	D-N-methylisoleucine	Dnmile	N-(imidazolylethyl))glycine	Nhis
	D-N-methylleucine	Dnmleu	N-(3-indolylyethyl)glycine	Nhtrp
	D-N-methyllysine	Dnmlys	N-methyl-γ-aminobutyrate	Nmgabu
	N-methylcyclohexylalanine	Nmchexa	D-N-methylmethionine	Dnmmet
5	D-N-methylornithine	Dnmorn	N-methylcyclopentylalanine	Nmcpen
	N-methylglycine	Nala	D-N-methylphenylalanine	Dnmphe
	N-methylaminoisobutyrate	Nmaib	D-N-methylproline	Dnmpro
	N-(1-methylpropyl)glycine	Nile	D-N-methylserine	Dnmser
	N-(2-methylpropyl)glycine	Nleu	D-N-methylthreonine	Dnmthr
10	D-N-methyltryptophan	Dnmtrp	N-(1-methylethyl)glycine	Nval
	D-N-methyltyrosine	Dnmtyr	N-methyla-napthylalanine	Nmanap
	D-N-methylvaline	Dnmval	N-methylpenicillamine	Nmpen
	γ-aminobutyric acid	Gabu	N-(*p*-hydroxyphenyl)glycine	Nhtyr
	L-*t*-butylglycine	Tbug	N-(thiomethyl)glycine	Ncys
15	L-ethylglycine	Etg	penicillamine	Pen
	L-homophenylalanine	Hphe	L-α-methylalanine	Mala
	L-α-methylarginine	Marg	L-α-methylasparagine	Masn
	L-α-methylaspartate	Masp	L-α-methyl-*t*-butylglycine	Mtbug
	L-α-methylcysteine	Mcys	L-methylethylglycine	Metg
20	L-α-methylglutamine	Mgln	L-α-methylglutamate	Mglu
	L-α-methylhistidine	Mhis	L-α-methylhomophenylalanine	Mhphe
	L-α-methylisoleucine	Mile	N-(2-methylthioethyl)glycine	Nmet
	L-α-methylleucine	Mleu	L-α-methyllysine	Mlys
	L-α-methylmethionine	Mmet	L-α-methylnorleucine	Mnle
25	L-α-methylnorvaline	Mnva	L-α-methylornithine	Morn
	L-α-methylphenylalanine	Mphe	L-α-methylproline	Mpro
	L-α-methylserine	Mser	L-α-methylthreonine	Mthr
	L-α-methyltryptophan	Mtrp	L-α-methyltyrosine	Mtyr
	L-α-methylvaline	Mval	L-N-methylhomophenylalanine	Nmhphe
30	N-(N-(2,2-diphenylethyl)	Nnbhm	N-(N-(3,3-diphenylpropyl)	Nnbhe

carbamylmethyl)glycine	carbamylmethyl)glycine
1-carboxy-1-(2,2-diphenyl-ethylamino)cyclopropane	Nmbc

Table 5

	% Contribution to mature hematopoietic cells		
Source of test cells	**Test**	Competitor	Endogenous
***mld2/+ Mpl*$^{-/-}$ mouse 1**	2.1 ± 1.3	68.9 ± 7.5	28.4 ± 6.3
***mld2/+ Mpl*$^{-/-}$ mouse 2**	1.2 ± 0.4	68.5 ± 12.8	29.2 ± 11.6
+/+ *Mpl*$^{-/-}$ mouse	22.4 ± 9.5	50.8 ± 8.6	26.3 ± 10.1
+/+ *Mpl*$^{+/+}$ mouse	46.3 ± 8.4	32.2 ± 4.6	20.8 ± 4.3

Table 6

Genotype	No. Observed	No. Predicted	No. Predicted if *mld2/mld2* is lethal
+/+	21	12.8	17.0
mld2/+	30	25.5	34.0
mld2/mld2	**0**	12.8	0.0
Total	*51*	*51.0*	*51.0*

Table 7

Age of embryos[a]		Live/Total (Expected) no. of embryos[b]			Total embryo number	P[c]
		Erg				
		+/+	+/-	-/-		
	Mpl					
Weaning	+/+	10/10 (6)	14/14 (12)	**0/0** (6)	24	
	+/-	19/19(12)	31/31 (24)	**0/0** (12)	50	
	-/-	14/14 (6)	8/8 (12)	**0/0** (6)	22	
					96	
E12.5	+/+	6/6 (6)	**10/12** (12)	**0/5** (6)	23	0.029
	+/-	10/11 (12)	**23/25** (24)	**2/13** (12)	49	0.035
	-/-	5/5 (6)	11/11 (12)	**0/7** (6)	24	0.029
					96	
E11.5	+/+	4/4 (4)	**6/8** (8)	**3/4** (4)	16	1.0
	+/-	8/8 (8)	**17/18 (16)**	**2/7** (8)	33	0.092
	-/-	1/1 (4)	10/10 (8)	2/2 (4)	13	
					62	
E10.5	+/+	9/9 (6)	13/13 (12)	5/5 (6)	27	
	+/-	13/13 (12)	21/21 (23)	**8/9** (12)	43	0.612
	-/-	10/10 (6)	12/12 (12)	1/1 (6)	23	
					93	
E9.5	+/+	9/9 (5)	8/8 (9)	3/3 (5)	20	
	+/-	13/13 (9)	22/22 (19)	8/8 (9)	43	
	-/-	2/2 (5)	10/10 (9)	2/2 (5)	14	
					77	

[a] Age of embryos given in days post conception.
[b] Number and genotype of embryos from matings of $Erg^{mld2/+}$ $Mpl^{+/-}$ parents are shown to produce *Erg* and *Mpl* wild type (+/+), heterozygous (+/-) and homozygous mutant (-/-) embryos.
[c] Fischer's exact testing, using two sided P values, demonstrates a significant difference between expected and observed surviving $Erg^{mld2/mld2}$ embryos from E12.5.

Table 8

Genotype	**Cellularity ($x10^{-5}$)**	Number of colonies per yolk sac					
		M	**Meg**	**Ery**	**Meg/Ery**	**Mac/Ery**	**Multi**
$Erg^{+/+}$ $Mpl^{+/+}$ **(n=5)**	4.1 ± 1.0	131 ± 83	14 ± 6	185 ± 32	18 ± 0	72 ± 38	14 ± 19
$Erg^{mld2/+}$ $Mpl^{+/+}$ **(n=5)**	1.5 ± 0.7	81 ± 44	14 ± 8	122 ± 75	11 ± 4	43 ± 37	2 ± 4
$Erg^{mld2/mld2}$ $Mpl^{+/+}$ **(n=1)**	0.8	0	0	9	0	0	0

Table 9

Genotype	**Cellularity ($x10^{-5}$)**	Number of colonies per fetal liver ($x10^{-3}$)											
		Total	***Blast***	**G**	**GM**	**M**	**Eo**	**Meg**	**Ery**	**Meg/ Ery**	**Meg/ Myel**	**Ery/ Myel**	**Mixed**
$Erg^{+/+}$ $Mpl^{+/+}$ **(n=2)**	25.5 ± 2.1	28.8 ± 9.0	2.6 ± 1.5	2.7 ± 0.1	4.5 ± 1.8	3.0 ± 1.2	0	5.1 ± 2.6	6.1 ± 1.4	3.0 ± 0.3	0.7 ± 0.4	0.2 ± 0.3	0.9 ± 0.4
$Erg^{mld2/+}$ $Mpl^{+/+}$ **(n=3)**	17.4 ± 2.3	17.1 ± 4.2	1.2 ± 0.3	1.4 ± 0.8	1.8 ± 0.5	1.8 ± 0.4	0	1.8 ± 0.7	5.9 ± 1.7	2.5 ± 0.8	0.1 ± 0.1	0.2 ± 0.1	0.3 ± 0.1
$Erg^{mld2/mld2}$ $Mpl^{+/+}$ **(n=2)**	0.6 ± 0.4	0.9 ± 1.0	0.07 ± 0.08	0.02 ± 0.01	0.03 ± 0.03	0.08 ± 0.08	0	0.1 ± 0.1	0.4 ± 0.4	0.2 ± 0.2	0	0.01 ± 0.02	0.004 ± 0.005

BIBLIOGRAPHY

Alexander *et al., Blood, 87*:2162-2170, 1996.

Allen *et al., Nephrol Dial Transplant, 12*:701-706, 1997.

Altschul *et al., Nucl. Acids Res. 25:* 3389, 1997.

Anderson *et al., Development, 126*:3131-3148, 1998.

Arvand *et al., Oncogene, 20*:5747-5754, 2001.

Ausubel *et al., Current Protocols in Molecular Biology* John Wiley & Sons Inc, 1994-1998, Chapter 15.

Ausubel (Ed) *Current Protocols in Molecular Biology*, 5th Edition, John Wiley & Sons, Inc, NY, 2002.

Baldus *et al., Proc. Natl. Acad. Sci. U.S.A., 101*:3915-3920, 2004.

Berger, *Biochim Biophys Acta, 1455*:255-268, 1999.

Bode, *Genetics, 108*:457-470, 1984.

Bonner *et al., Eur. J. Biochem. 46:* 83, 1974.

Carpinelli *et al., Proc. Natl. Acad. Sci. U. S. A., 101*:6553-6558, 2004.

Carrere *et al., Oncogene, 16*:3261-3268, 1998.

Cartron *et al., Nature, 282*:621-623, 1979.

Ceballos *et al., Oncogene, 19*:2194-2204, 2000.

Chee *et al., Science, 274:*610-614, 1996.

Conner *et al., Proc. Natl. Acad. Sci. USA, 80*:278-282, 1983.

Copeland *et al., Science, 262*:57-66, 1993.

Culver, *Gene Therapy: A Primer for Physicians*, 2nd Ed., Mary Ann Liebert, 1996.

Dhordain *et al., Mech. Dev., 50*:17-28, 1995.

DiRisi *et al., Nature Genetics, 14:*457-460, 1996.

Douillard and Hoffman, Basic Facts about Hybridomas, in *Compendium of Immunology* Vol. II, ed. by Schwartz, 1981.

Elghanian *et al., Science 277:* 1078-1081, 1997.

Erickson *et al., Science 249:* 527-533, 1990.

Felner *et al., Biochim Biophys Acta, 1416*: 115-125, 1998.

Finkelstein *et al., Genomics, 7:*167-172, 1990.

Floege *et al.*, *J Am Soc Nephrol, 11*:2395-2403, 2000.

Friedman, In: *Therapy for Genetic Disease*, T. Friedman, Ed., Oxford University Press, pp. 105-121, 1991.

Grompe *et al.*, *Proc. Natl. Acad. Sci. USA, 86:*5855-5892, 1989.

Grompe, *Proc. Natl. Acad. Sci. USA, 86:*5855-5892, 1993.

Hacia *et al.*, *Nature Genetics, 14:*441-447, 1996.

Hart *et al.*, *Oncogene, 10*:1423-1430, 1995.

Hodgson, *Bio/Technology 9:* 19-21, 1991.

Hogan *et al.*, *Manipulating the Mouse Embryo: A Laboratory Manual*, Cold Sping Harbour Laboratory Press, CSH NY, 1986.

Horak *et al.*, *Methods Enzymol., 350*:469-483, 2002.

Ichikawa *et al.*, *Cancer Res., 54*:2865-2868, 1994.

Iwamoto *et al.*, *J. Cell Biol., 150*:27-40, 2000.

Johnson *et al "Peptide Turn Mimetics" in Biotechnology and Pharmacy, Pezzuto et al* eds Chapman and Hall, New York, 1993.

Ju *et al.*, *Proc Natl Acad Sci U S A, 99*:16613-16618, 2002.

Kim *et al.*, *EMBO J., 20*:4173-4182, 2001.

Kinszler *et al.*, *Science, 251:*1366-1370, 1991.

Kodandapani *et al.*, *Nature. 380*:456-460, 1996.

Kohler *et al.*, *Nature, 256:*495-499, 1975.

Kohler *et al.*, *European Journal of Immunology. 6:*511-519, 1976.

Kudo *et al.*, *J Biol Chem, 277*:47724-47731, 2002.

Kurrek, *Eur. J. Biochem. 270*:1628-1644, 2003.

Kyte *et al.*, *J. Mol. Biol., 157:*105-132, 1982.

Lipshutz *et al.*, *Biotechniques, 19:*442-447, 1995.

Lockhart *et al.*, *Nature Biotechnology, 14:*1675-1680, 1996.

McLaughlin *et al.*, *J. Cell Sci., 112(Pt 24)*:4695-4703, 1999.

McLaughlin *et al.*, *Blood, 98*:3332-3339, 2001.

Mansour *et al.*, *Nature 336*:348-352, 1988.

Marmur *et al.*, *J. Mol. Biol. 5:* 109, 1962.

Mizushima *et al.*, *Nucleic Acids Res, 18* 5322, 1990.

Modrich, *Ann. Rev. Genet., 25:*229-253, 1991.

Murakami *et al.*, *Oncogene, 8*:1559-1566, 1993.

Newtown *et al., Nucl. Acids. Res., 17:*2503-2516, 1989.

Orita *et al.*, *Proc. Nat. Acad. Sci. USA 86:* 2776-2770, 1989.

Pereira *et al. Proc. Natl. Acad. Sci. U. S. A., 95*:8239-8244, 1998.

Pickert, *Transgenic Animal Technology: A Laboratory Handbook*, Academic Press, San Diago, CA, 1994.

Rainis *et al.*, *Cancer Res., 65*:7596-7602, 2005.

Remington's Pharmaceutical Sciences, 20th ed. Williams and Wilkins, 2000.

Robertson *et al., Nature 322*:445-448, 1986.

Ruano *et al., Nucl. Acids. Res., 17:*8392, 1989.

Sambrook, *Molecular Cloning: A Laboratory Manual*, 3rd Edition, CSHLP, CSH, NY, 2001.

Sharrocks, *Nat. Rev. Mol. Cell Biol., 2*:827-837, 2001.

Sheffield *et al.*, *Proc. Natl. Acad. Sci. USA, 86:*232-236, 1989.

Sheffield *et al.*, *Am. J. Hum. Genet., 49:*699-706, 1991.

Shimizu *et al.*, *Proc. Natl. Acad. Sci. U. S. A., 90*:10280-10284, 1993.

Shoemaker *et al., Nature Genetics, 14:*450-456, 1996.

Sorensen *et al.*, *Nat. Genet., 6*:146-151, 1994.

Summerton *et al., Antisense and Nucleic Acid Drug Development 7*:187-195, 1997.

Thurnher *et al., J Clin Invest, 91*:2103-2110, 1993.

Vlaeminck-Guillem *et al.*, *Mech. Dev., 91*:331-335, 2000.

Wartell *et al.*, *Nucl. Acids Res., 18:*2699-2705, 1990.

Wells, *Methods Enzymol. 202:* 2699-2705, 1991.

White *et al.*, *Genomics, 12:*301-306, 1992.

Yang *et al.*, *Mol. Cell. Biol., 20*:3345-3354, 2000.

We claim:

1. A method for modulating hematopoietic cell activity, comprising administering an effective amount of a modulator which modulates the activity of an Erg polypeptide or a transcriptional target of Erg in a cell or a downstream effector of Erg polypeptide.

2. The method of claim 1 wherein the hematopoietic cell is a hematopoietic stem cell (HSC) or a multipotent or committed hematopoietic progenitor cell.

3. The method of claim 1 wherein the hematopoietic cell is one or two or more of a cell selected from the group consisting of a platelet, erythrocyte, B lymphocyte, T lymphocyte, natural killer cell, granulocyte, monocyte and macrophage.

4. The method of claim 1 wherein the hematopoietic cell is a bone marrow cell and/or a cell operably associated with junctions between bones or their associated tissues.

5. The method of claim 1 wherein the modulator enhances the activity of Erg polypeptide or a transcriptional target of Erg in a cell or a downstream effector of Erg activity.

6. The method of claim 5 wherein the modulator comprises a sequence of nucleotides encoding all or part of an Erg polypeptide.

7. The method of claim 5 or 6 wherein the modulator is a cell, plasmid or viral vector.

8. The method of claim 1 wherein the modulator reduces the activity of Erg polypeptide or a transcriptional target of Erg polypeptide in a cell.

9. The method of claim 8 wherein the modulator is a genetic agent which reduces the production of Erg polypeptide in a cell.

10. The method of claim 9 wherein the modulator is a cell, plasmid or viral vector.

11. The method of claim 5 or 8 wherein the modulator is small-molecule, peptide, peptidomimetic, a constrained peptide or a gene silencing molecule or the like.

12. The method of claim 1 when used *in vitro* or *ex vivo*.

13. The method of claim 1 when used for *in vivo* or *ex vivo* treatment or prophylaxis of a condition in a mammalian subject.

14. The method of claim 13 wherein the condition involves a HSC or a hematopoietic progenitor cell defect.

15. The method of claim 13 wherein the condition involves a defect in a hematopoietic cell operably associated with the skin and/or junctions between bones.

16. The method of claim 13 wherein the condition is a multi-lineage defect in hematopoiesis.

17. The method of claim 13 wherein the condition is a haematopoietic disorder selected from the group consisting of: thrombocytopenias, thrombocytosis, anaemias including drug induced anaemia, leukopenia, immunological disorders, autoimmune disorders, myeloproliferative disorders, cancer, conditions associated with stem cell defects or dysregulated proliferation or differentiation and bone marrow defects; including: leukemia, Hodgkin's disease, non-Hogkin's lymphoma, acute lymphocytic anaemia, plasmacytomas, multiple mycloma Burkett's lymphoma, arthritis, asthma, AIDS, rheumatoid arthritis, granulomatous disease, immune deficiency, inflammatory bowel disease, sepsis, neutropenia, neutrophilia, psoriasis, immune reaction to transplanted organ, SLE, hemophilia, hypercoagulation, diabetes, meningitis, lyme disease and allergies; autosomal

recessive traits include: Fanconi Syndrome, Thrombocytopenia with absent radius (TAR) syndrome, Bernard-Soulier syndrome, Gray platelet syndrome, and Isolated thrombocytopenia; and including autosomal dominant traits including: May-Hegglin anomaly, Alport syndrome variants, and Isolated thrombocytopenia; X-linked traits including: Wiskott-Adrich syndrome, and Isolated thrombocytopenia; aplastic pancytopenias (traditionally known as aplastic anaemia), which result from aplasia or suppression of hematopoietic stem cells, including: Fanconi's anaemia, Shwachman-Diamond syndrome, Dyskeratosis congenita, Amegakaryocytic thrombocytopenia; Down's syndrome, Dubowitz's syndrome, Seckel's syndrome, Reticular dysgenesis, and Familial aplastic anaemia (non-Fanconi's).

18. The method of claim 13 wherein the condition is a clonal hemopathy selected from the group consisting of: preleukaemias (myelodysplasias) such as Idiopathic refractory sideroblastic anaemia, Idiopathic refractory nonosideroblastic anaemia, Pancytopenia with hyperplastic marrow, Monoclonal aplastic anaemia, and Paroxysmal nocturnal hemoglobinuria; myeloproliferative diseases such as Chronic myeloproliferative diseases including Polycythemia vera, Chronic myelogenous leukemia (CML), Primary thrombocythemia, Idiopathic myelofibrosis, Subacute myeloproliferative disorders, Oligoblastic (smoldering) leukemias, Refractory leukemias with excess myeloblasts, Myelomnocytic leukemia, and Atypical myeloproliferative syndromes; acute myeloproliferative disorders such as Acute myelogenous leukemia, Myeloblastic (granuloblastic), Promyelocytic: associated with intravascular coagulation, Myelomonocytic (granlomonoblastic), Monocytic: associated with tissue infiltration, Erythroid, Megakaryocytic: associated with marrow fibrosis, Eosinophilic: associated with heart and lung fibrosis and Basophilic and Mast Cell; acute biphenotypic (myeloid and lymphoid markers) leukemia and acute leukemia with lymphoid markers evolving from a prior clonal hemopathy.

19. The method of claim 6 wherein the Erg polypeptide comprises all or part of the amino acid sequence set out in SEQ ID NO: 4 or 6 or a sequence having at least

60% sequence identity to the amino acid sequence set out in SEQ ID NO: 4 or 6.

20. The method of claim 6 wherein the nucleotide sequence encodes all or part of the amino acid sequence of SEQ ID NO: 4 or 6 or a sequence having at least 60% sequence identity to SEQ ID NO: 4 or 6.

21. The method of claim 6 wherein the nucleotide sequence is as set forth in SEQ ID NO: 3 or 5 or a sequence having at least 60% sequence identity to SEQ ID NO: 3 or 5, or a sequence that hybridises to the sequence of SEQ ID NO: 3 or 5 or to a complementary form of either of these under at least medium stringency conditions.

22. An isolated cellular agent or vector comprising a sequence of nucleotides encoding an Erg polypeptide for use in the treatment or prophylaxis of a hematopoietic lineage including a HSC defect in a mammalian subject.

23. A method for diagnosing a hematopoietic lineage deficit including a HSC defect in a mammalian subject said method comprising screening a sample from a subject for a loss of function mutation in *Erg*.

24. An isolated cell or non-human animal comprising such cells, wherein the *Erg* gene is modified to effectively modulate the activity of Erg polypeptide in the cell or animal compared to a non-modified animal of the same species.

25. The cell or non-human animal of claim 24 wherein the modification is in one allele of the *Erg* gene.

26. A method for screening for a modulator which modulates the functional activity of an HSC or a hematopoietic progenitor cell, said method comprising screening for a modulator which modulates the functional activity of Erg polypeptide or a transcriptional target of an Erg polypeptide in a cell or a downstream effector of

Erg polypeptide activity, said method comprising combining the modulator and the cell and identifying a change in the activity of the cell relative to controls indicating that Erg polypeptide activity including the activity of a transcriptional target or a downstream effector molecule has been modified.

27. The method of claim 26 wherein the identifying step includes identifying the presence of a complex between Erg or *Erg* and the modulator, or a change in the level or activity of a marker of the activity of Erg polypeptide in a cell.

28. The method of claim 26 wherein the identifying step includes assaying for one or more of the following:

i) an change in the activity of the Erg or *Erg*;

ii) a change in the level or activity of a reporter of the activity of Erg or *Erg*; and

iii) the presence of a complex between the Erg or *Erg* and the modulator.

29. The method of claim 26 wherein the method involves an initial further step of introducing random or non-random mutations or down regulating expression of *Erg* gene and/or one or more other genesin the cell.

30. The method of claim 29 wherein the mutation or down-regulating agent is introduced *via* a mutagenesis agent or vector.

31. The method of claim 26 wherein the cell is a cell of claim 24.

32. A pharmaceutical composition comprising a modulator defined in any one of claims 5 to 11 or 22 for use in the method of any one of claims 1 to 21.

33. A modulator as defined in any one of claims 5 to 11 or 22 for use in modulating hematopoietic cell activity.

34. A modulator for use according to claim 33 wherein the hematopoietic cell is a hematopoietic stem cell (HSC) or a multipotent or committed hematopoietic progenitor cell.

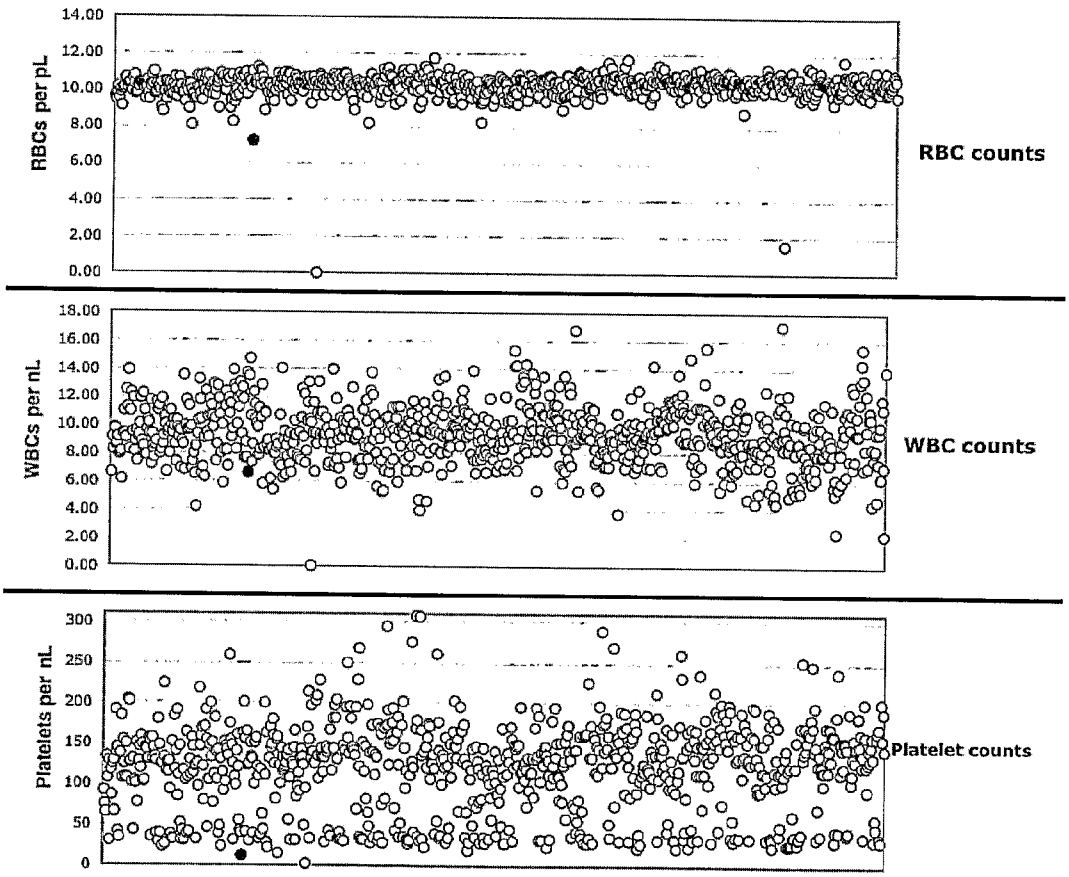

RBCs per pL
14.00
12.00
10.00
8.00
6.00
4.00
2.00
0.00
RBC counts
WBCs per nL
18.00
16.00
14.00
12.00
10.00
8.00
6.00
4.00
2.00
0.00
WBC counts
Platelets per nL
300
250
200
150
100
50
0
Platelet counts

Figure 1

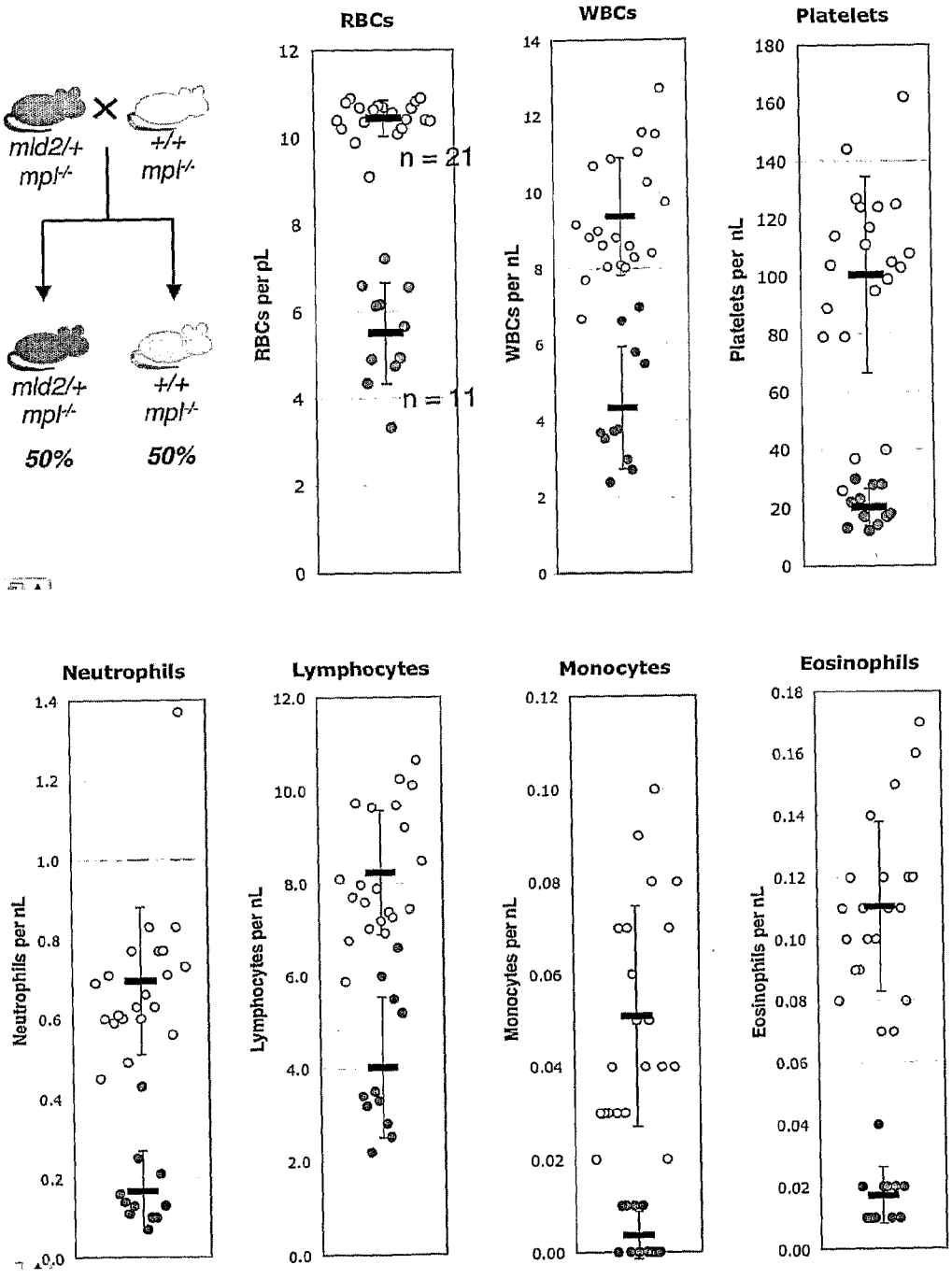

mld2/+ mpl-/-
+/+ mpl-/-
mld2/+ mpl-/-
+/+ mpl-/-
50%
50%
RBCs
RBCs per pL
n = 21
n = 11
WBCs
WBCs per nL
Platelets
Platelets per nL
Neutrophils
Neutrophils per nL
Lymphocytes
Lymphocytes per nL
Monocytes
Monocytes per nL
Eosinophils
Eosinophils per nL

Figure 2

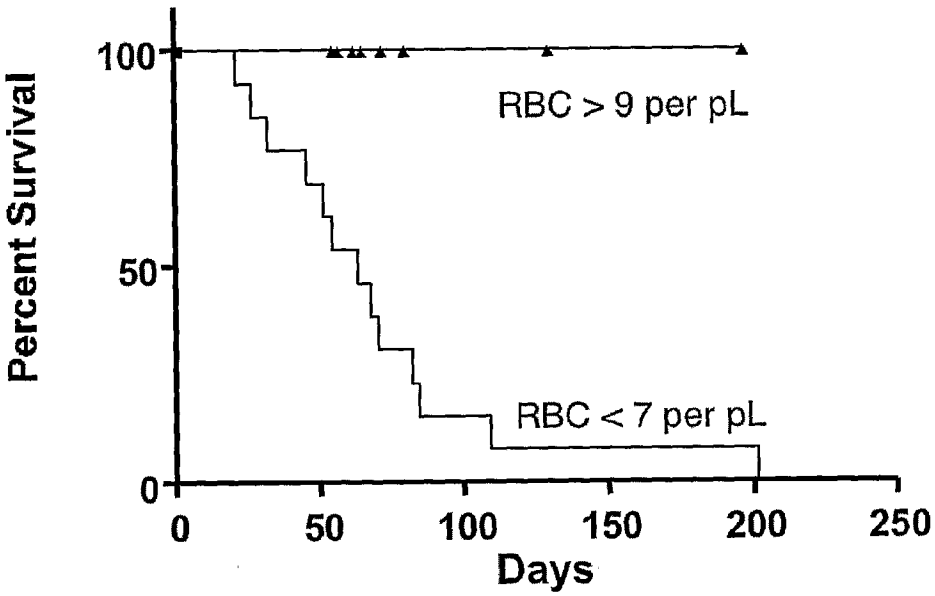

Percent Survival
100
50
0
RBC > 9 per pL
RBC < 7 per pL
0
50
100
150
200
250
Days

Figure 3

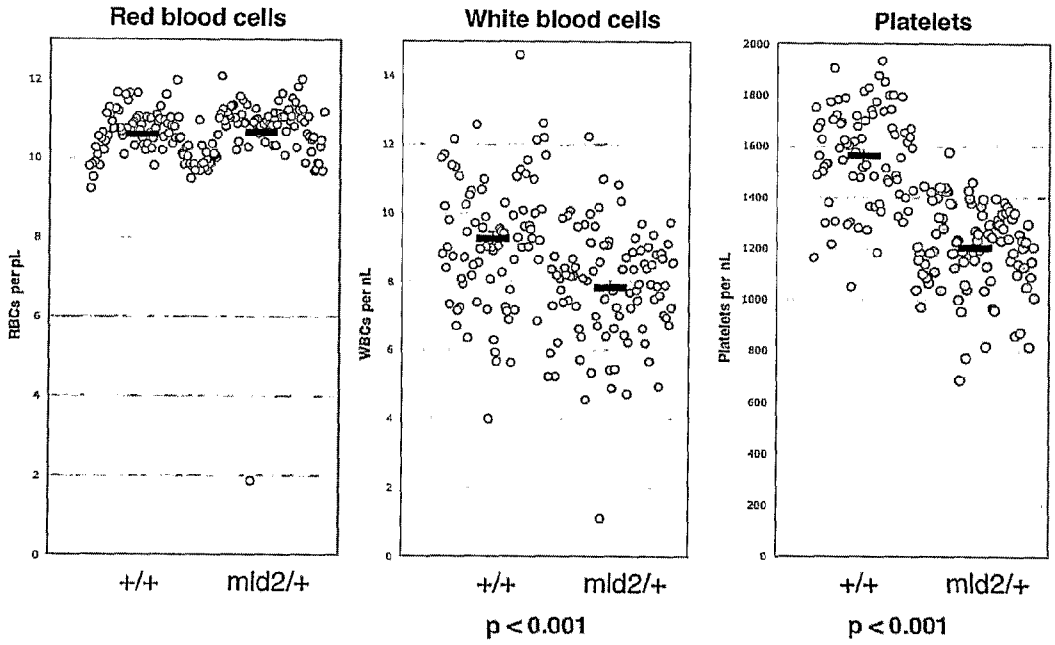

Red blood cells
White blood cells
Platelets
RBCs per pL
WBCs per nL
Platelets per nL
+/+
mld2/+
p < 0.001
p < 0.001

Figure 4

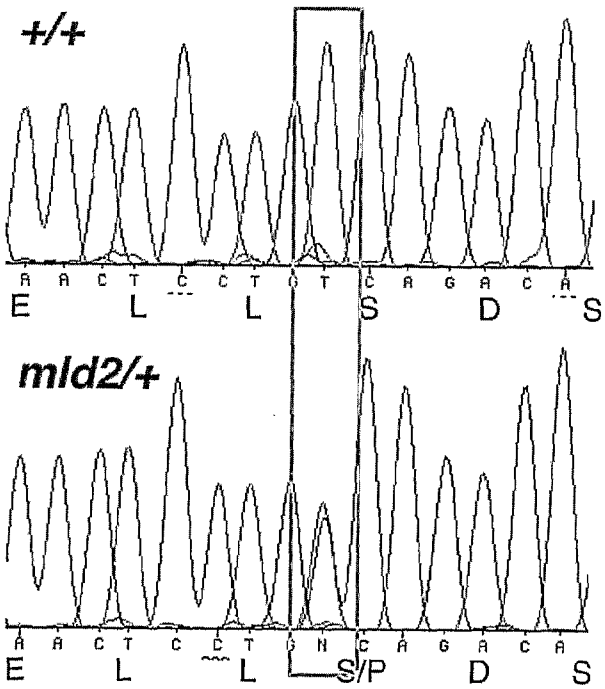


Figure 5

Alignments of Erg protein and mRNA from a range of metazoan species
Steve Loughran 10.5.06

Details of Erg sequences used in alignments

All sequences were taken from the NCBI reference sequence database (RefSeq), except for Xenopus laevis erg, whose reference is listed below.

Homo sapiens Erg-1

```
LOCUS       NP_891548               479 aa            linear   PRI 12-FEB-2006
DEFINITION  v-ets erythroblastosis virus E26 oncogene like isoform 1 [Homo
            sapiens].
ACCESSION   NP_891548
VERSION     NP_891548.1  GI:33667107
DBSOURCE    REFSEQ: accession NM_182918.2
```

Homo sapiens Erg-2

```
LOCUS       NP_004440               462 aa            linear   PRI 12-FEB-2006
DEFINITION  v-ets erythroblastosis virus E26 oncogene like isoform 2 [Homo
            sapiens].
ACCESSION   NP_004440
VERSION     NP_004440.1  GI:4758300
DBSOURCE    REFSEQ: accession NM_004449.3
```

Rattus norvegicus Erg

```
LOCUS       NP_596888               455 aa            linear   ROD 20-APR-2005
DEFINITION  v-ets erythroblastosis virus E26 oncogene like [Rattus norvegicus].
ACCESSION   NP_596888
VERSION     NP_596888.1  GI:19173756
DBSOURCE    REFSEQ: accession NM_133397.1
```

Danio rerio Erg

```
LOCUS       NP_001008616            427 aa            linear   VRT 27-NOV-2005
DEFINITION  v-ets erythroblastosis virus E26 oncogene like [Danio rerio].
ACCESSION   NP_001008616
VERSION     NP_001008616.1  GI:56693304
DBSOURCE    REFSEQ: accession NM_001008616.1
```

Gallus gallus Erg

```
LOCUS       NP_989611               478 aa            linear   VRT 16-APR-2005
DEFINITION  v-ets erythroblastosis virus E26 oncogene like [Gallus gallus].
ACCESSION   NP_989611
VERSION     NP_989611.1  GI:45383578
DBSOURCE    REFSEQ: accession NM_204280.1
KEYWORDS    .
SOURCE      Gallus gallus (chicken)
```

Figure 6

S. purpuratus Erg

```
LOCUS       NP_999833                498 aa            linear   INV 16-APR-2005
DEFINITION  transcription factor Erg [Strongylocentrotus purpuratus].
ACCESSION   NP_999833
VERSION     NP_999833.1  GI:47551301
DBSOURCE    REFSEQ: accession NM_214668.1
KEYWORDS    .
SOURCE      Strongylocentrotus purpuratus
```

Xenopus laevis Erg

```
LOCUS       CAB46566                 485 aa            linear   VRT 15-APR-2005
DEFINITION  erg [Xenopus laevis].
ACCESSION   CAB46566
VERSION     CAB46566.1  GI:5420046
DBSOURCE    embl locus XLAJ4125, accession AJ224125.1
KEYWORDS    .
SOURCE      Xenopus laevis (African clawed frog)
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Xenopus.
REFERENCE   1
  AUTHORS   Baltzinger,M., Mager-Heckel,A.M. and Remy,P.
  TITLE     Xl erg: expression pattern and overexpression during development
            plead for a role in endothelial cell differentiation
  JOURNAL   Dev. Dyn. 216 (4-5), 420-433 (1999)
   PUBMED   10633861
REFERENCE   2  (residues 1 to 485)
  AUTHORS   Baltzinger,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-FEB-1998) Baltzinger M., UPR9005, MMDCD, Centre
            National de la Recherche Scientifique, 15, RUE Rene Descartes,
            67084, FRANCE
```

Mus musculus Erg

```
LOCUS       NP_598420                486 aa            linear   ROD 25-DEC-2005
DEFINITION  avian erythroblastosis virus E-26 (v-ets) oncogene related [Mus
            musculus].
ACCESSION   NP_598420 XP_489769
VERSION     NP_598420.1  GI:19526802
DBSOURCE    REFSEQ: accession NM_133659.1
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
```

Figure 6 continued

Erg protein alignment.
Aligned by Jotun Hein method

```
     M - - - - - - - - - - - - - - - - - A S T I K E A L S V V S E  Majority
     ------------------+-------------------+-------------------+-
                       10                  20                  30
     ------------------+-------------------+-------------------+-
  1  M - - - - - - - I Q T - - V P D P A A H I K E A L S V V S E  Mus musculus Erg.pro
  1  M - - - - - - - - - - - - - - - - - A S T I K E A L S V V S E  Homo sapiens Erg-1.pro
  1  M - - - - - - - I Q T - - V P D P A A H I K E A L S V V S E  Homo sapiens Erg-2.pro
  1  M - - - - - - - - - - - - - - - - - A S T I K E A L S V V S K  Rattus novegicus Erg.pro
  1  M - - - - - - - - - - - - - - - - - A S T I K E A L S V V S E  Gallus gallus erg.pro
  1  M - - - - - - - Q H M A F I P T Y E D Y S N E A L S V V S E  Xenopus laevis Erg.pro
  1  M K Q E I E Q P - - - - - - - - - - T S G A H D G L P - - - -  S. purpuratus Erg.pro
  1  M - - - - - - - - - - - - - - - - - T A S A A A E L P - - - -  Danio rerio erg.pro

     D Q S L F E C A Y G T P H L A K T E M T A S S S S D Y G Q T  Majority
     ------------------+-------------------+-------------------+-
                       40                  50                  60
     ------------------+-------------------+-------------------+-
 22  D Q S L F E C A Y G T P H L A K T E M T A S S S S D Y G Q T  Mus musculus Erg.pro
 15  D Q S L F E C A Y G T P H L A K T E M T A S S S S D Y G Q T  Homo sapiens Erg-1.pro
 22  D Q S L F E C A Y G T P H L A K T E M T A S S S S D Y G Q T  Homo sapiens Erg-2.pro
 15  D Q S L F E C A Y G T P H L A K T E M T A S S S S D Y G Q T  Rattus novegicus Erg.pro
 15  D Q S L F E C A Y G S P H L A K T E M T A S S S S E Y G Q T  Gallus gallus erg.pro
 24  D Q S L F E C T Y G T P H L T K T E M T A S S S S D Y G Q T  Xenopus laevis Erg.pro
 18  - - - - - - - - - - G P Q V S R G G - - - - - - - - - - - -  S. purpuratus Erg.pro
 11  - - - - - - - - - - H S K L L L S A - - - - - - - - - - - -  Danio rerio erg.pro

     S K M S P R V P Q Q D W L S Q P P A R V T I K M E C N P S Q  Majority
     ------------------+-------------------+-------------------+-
                       70                  80                  90
     ------------------+-------------------+-------------------+-
 52  S K M S P R V P Q Q D W L S Q A P A R V T I K M E C N P S Q  Mus musculus Erg.pro
 45  S K M S P R V P Q Q D W L S Q P P A R V T I K M E C N P S Q  Homo sapiens Erg-1.pro
 52  S K M S P R V P Q Q D W L S Q P P A R V T I K M E C N P S Q  Homo sapiens Erg-2.pro
 45  S K M S P R V P Q Q D W L S Q P P A R V T I K M E C N P S Q  Rattus novegicus Erg.pro
 45  S K M S P R V P Q Q D W L S Q P P A R V T I K M E C N P N Q  Gallus gallus erg.pro
 54  S K M S P R V P Q Q D W L S Q P P S R V T I K M E C N P S Q  Xenopus laevis Erg.pro
 26  - - - - - - - - - - - - - - - - - - - - P E S P L D C S V A K  S. purpuratus Erg.pro
 19  - - - - - - - - - - - - - - - - - - - - V S V K L E A S A G A  Danio rerio erg.pro

     V N G S R - - N S P D E C S V A K G G K M V G S P D T V G M  Majority
     ------------------+-------------------+-------------------+-
                       100                 110                 120
     ------------------+-------------------+-------------------+-
 82  V N G S R - - N S P D E C S V N K G G K M V G S P D T V G M  Mus musculus Erg.pro
 75  V N G S R - - N S P D E C S V A K G G K M V G S P D T V G M  Homo sapiens Erg-1.pro
 82  V N G S R - - N S P D E C S V A K G G K M V G S P D T V G M  Homo sapiens Erg-2.pro
 75  V N G S R - - N S P D E C S V T K G G K M V G S P D T V G M  Rattus novegicus Erg.pro
 75  V N G S R - - N S P D D C S V A K G G K M V S S S D N V G M  Gallus gallus erg.pro
 84  V N G S R - - S S P D D C S I G K G S K M G G G S D N V G M  Xenopus laevis Erg.pro
 37  P R - Q Q - - P P P H G A P Q G T I N - - A E P T Q A A P P  S. purpuratus Erg.pro
 30  H R - S H C R S P P D D C P G V K G K - - I S P A G N A P V  Danio rerio erg.pro
```

Figure 7

```
     N Y G S - - - - - Y M E E K H M P P P N M T - - - - - - - - -  Majority
     ------------------+------------------+------------------+-
                       130                140                150
     ------------------+------------------+------------------+-
110  S Y G S - - - - - Y M E E K H V P P P N M T - - - - - - - - -  Mus musculus Erg.pro
103  N Y G S - - - - - Y M E E K H M P P P N M T - - - - - - - - -  Homo sapiens Erg-1.pro
110  N Y G S - - - - - Y M E E K H M P P P N M T - - - - - - - - -  Homo sapiens Erg-2.pro
103  S Y G S - - - - - Y M E E K H M P P P N M T - - - - - - - - -  Rattus novegicus Erg.pro
103  N Y G S - - - - - Y M E E K H I P P P N M T - - - - - - - - -  Gallus gallus erg.pro
112  N Y G S - - - - - Y I E E K H I S P P N M T - - - - - - - - -  Xenopus laevis Erg.pro
 62  - Y P S C P P Q P Y L G T T G S T P - - - - E E R S A S A G  S. purpuratus Erg.pro
 57  M Y S N - - - - - Y M E E K H T P P - - - - - - - - - - - -  Danio rerio erg.pro

     - - - - - - - - - - - T N E R R V I V P A D P T L W S T D H  Majority
     ------------------+------------------+------------------+-
                       160                170                180
     ------------------+------------------+------------------+-
127  - - - - - - - - - - - T N E R R V I V P A D P T L W S T D H  Mus musculus Erg.pro
120  - - - - - - - - - - - T N E R R V I V P A D P T L W S T D H  Homo sapiens Erg-1.pro
127  - - - - - - - - - - - T N E R R V I V P A D P T L W S T D H  Homo sapiens Erg-2.pro
120  - - - - - - - - - - - T N E R R V I V P A D P T L W S T D H  Rattus novegicus Erg.pro
120  - - - - - - - - - - - T N E R R V I V P A D P T L W S T D H  Gallus gallus erg.pro
129  - - - - - - - - - - - T N E R R V I V P A D P T L W S T D H  Xenopus laevis Erg.pro
 87  T T G R S P P P N V T T N E K R V I V P A D P N M W T A E H  S. purpuratus Erg.pro
 70  - - - - - - - - - - - S S E R R V I V P A D P G Q W S A G H  Danio rerio erg.pro

     V R Q W L E W A V K E Y G L P D V D V L L F Q N I D G K E L  Majority
     ------------------+------------------+------------------+-
                       190                200                210
     ------------------+------------------+------------------+-
146  V R Q W L E W A V K E Y G L L D V D V L L F Q N I D G K E L  Mus musculus Erg.pro
139  V R Q W L E W A V K E Y G L P D V N I L L F Q N I D G K E L  Homo sapiens Erg-1.pro
146  V R Q W L E W A V K E Y G L P D V N I L L F Q N I D G K E L  Homo sapiens Erg-2.pro
139  V R Q W L E W A V K E Y G L L D V D V L L F Q N I D G K E L  Rattus novegicus Erg.pro
139  V R Q W L E W A V K E Y G L P D V D I L L F Q N I D G K E L  Gallus gallus erg.pro
148  V R Q W L E W A I K E Y G L P D V D V L L F Q N I D G K E L  Xenopus laevis Erg.pro
117  V Q Q W V Q W A V R E Y S L V D V Q V S R F N - M D G K H L  S. purpuratus Erg.pro
 89  V R Q W L D W A V Q E Y G L P P L D V S L F Q G V D G K Q L  Danio rerio erg.pro

     C K M T K D D F Q R L T P S Y N A D I L L S H L H Y L R E -  Majority
     ------------------+------------------+------------------+-
                       220                230                240
     ------------------+------------------+------------------+-
176  C K M T K D D F Q R L T P S Y N A D I L L S H L H Y L R E -  Mus musculus Erg.pro
169  C K M T K D D F Q R L T P S Y N A D I L L S H L H Y L R E -  Homo sapiens Erg-1.pro
176  C K M T K D D F Q R L T P S Y N A D I L L S H L H Y L R E -  Homo sapiens Erg-2.pro
169  C K M T K D D F Q R L T P S Y N A D I L L S H L H Y L R E -  Rattus novegicus Erg.pro
169  C K M T K D D F Q R L T P S Y N A D I L L S H L H Y L R E -  Gallus gallus erg.pro
178  C K M T K E D F H R L T P S Y N A D I L L S H L H Y L R E -  Xenopus laevis Erg.pro
146  C K M T R D D F S R L T N N L N V D V L I S H L T F L K Q G  S. purpuratus Erg.pro
119  C R M S R E E L Q R L T S S Y S A D V L L S H L H Y L R E -  Danio rerio erg.pro

     - T P L P H L T S D D V D K A L Q N S P R L M H A R N T G G  Majority
     ------------------+------------------+------------------+-
                       250                260                270
     ------------------+------------------+------------------+-
205  - T P L P H L T S D D V D K A L Q N S P R L M H A R N T G G  Mus musculus Erg.pro
198  - T P L P H L T S D D V D K A L Q N S P R L M H A R N T G G  Homo sapiens Erg-1.pro
205  - T P L P H L T S D D V D K A L Q N S P R L M H A R N T - -  Homo sapiens Erg-2.pro
198  - T P L P H L T S D D V D K A L Q N S P R L M H A R N T - -  Rattus novegicus Erg.pro
198  - T P L P H L T S D D V D K A L Q N S P R L M H A R N T G G  Gallus gallus erg.pro
207  - T P L P H L T S D D V D K A L Q N S P R L M H A R N T G G  Xenopus laevis Erg.pro
176  Y T P G A K S N L D D N K Y N T A D H Q T T L D T N P S G G  S. purpuratus Erg.pro
148  - T P L P H L T S D D V D K A L Q N S P R L M H A R N T G G  Danio rerio erg.pro
```

Figure 7 continued

```
    A A F I F P N T S V Y P E A T Q R I T T R P D L P Y E P P -  Majority
    ------------------+------------------+------------------+-
                      280                290                300
    ------------------+------------------+------------------+-
234 A A F I F P N T S V Y P E A T Q R I T T R P D L P Y E P P -  Mus musculus Erg.pro
227 A A F I F P N T S V Y P E A T Q R I T T R P D L P Y E P P -  Homo sapiens Erg-1.pro
232 - - - - - - - - - - - - - - - - - - - - D L P Y E P P -  Homo sapiens Erg-2.pro
225 - - - - - - - - - - - - - - - - - - - - D L P Y E P P -  Rattus novegicus Erg.pro
227 A T F I F P N T S V Y P E A T Q R I T T R P D L P Y E Q A -  Gallus gallus erg.pro
236 A S F I F P N S S V Y Q D A N Q R I P S R Q D L S Y E P S -  Xenopus laevis Erg.pro
206 S A F P Y P P S T T T T V D S V H R M P R T E P S C D S L V  S. purpuratus Erg.pro
177 A N F I F P N T P V Y P P D A S R G A S R A G V G Y D A V -  Danio rerio erg.pro

    - - - R R S A W T G H S H P T P Q S K A - - - A Q P S P S -  Majority
    ------------------+------------------+------------------+-
                      310                320                330
    ------------------+------------------+------------------+-
263 - - - R R S A W T G H S H L T P Q S K A - - - A Q P S P S -  Mus musculus Erg.pro
256 - - - R R S A W T G H G H P T P Q S K A - - - A Q P S P S -  Homo sapiens Erg-1.pro
239 - - - R R S A W T G H G H P T P Q S K A - - - A Q P S P S -  Homo sapiens Erg-2.pro
232 - - - R R S T W T G H S H P T P Q S K A - - - A Q P S P S -  Rattus novegicus Erg.pro
256 - - - R R S A W T S H S H P T - Q S K A - - - T Q P S S S -  Gallus gallus erg.pro
265 - - - R R S A W T - - N H P A P P S K A - - - S Q P S - T -  Xenopus laevis Erg.pro
236 G R G R P N A W P T - - - T V P S A V S K G F T Q T S P T L  S. purpuratus Erg.pro
206 - - - R R S G W T P - - - A L P A A K V - - - S Q S S A A -  Danio rerio erg.pro

    - - - - T V P K T E D Q R - - - - - - - - - P Q L D P Y Q I  Majority
    ------------------+------------------+------------------+-
                      340                350                360
    ------------------+------------------+------------------+-
286 - - - - A V P K T E D Q R - - - - - - - - - P Q L D P Y Q I  Mus musculus Erg.pro
279 - - - - T V P K T E D Q R - - - - - - - - - P Q L D P Y Q I  Homo sapiens Erg-1.pro
262 - - - - T V P K T E D Q R - - - - - - - - - P Q L D P Y Q I  Homo sapiens Erg-2.pro
255 - - - - T V P K T E D Q R - - - - - - - - - P Q L D P Y Q I  Rattus novegicus Erg.pro
278 - - - - T V P K T E D Q R - - - - - - - - - P Q L D P Y Q I  Gallus gallus erg.pro
285 - - - - T V P K T E D P R - - - - - - - - - P Q L D P Y Q I  Xenopus laevis Erg.pro
263 P K P G I D S S A T Q I R P A G R P A Y G G S D F D P Y Q V  S. purpuratus Erg.pro
226 - - - - I M S K T E E Q R - - - - - - - - - A Q L D P Y Q I  Danio rerio erg.pro

    L G P T S S R L A N - - - - - - - - - - - - - - - - - -  Majority
    ------------------+------------------+------------------+-
                      370                380                390
    ------------------+------------------+------------------+-
303 L G P T S S R L A N - - - - - - - - - - - - - - - - - -  Mus musculus Erg.pro
296 L G P T S S R L A N - - - - - - - - - - - - - - - - - -  Homo sapiens Erg-1.pro
279 L G P T S S R L A N - - - - - - - - - - - - - - - - - -  Homo sapiens Erg-2.pro
272 L G P T S S R L A N - - - - - - - - - - - - - - - - - -  Rattus novegicus Erg.pro
295 L G P T S S R L A N - - - - - - - - - - - - - - - - - -  Gallus gallus erg.pro
302 L G P T S S R L A N - - - - - - - - - - - - - - - - - -  Xenopus laevis Erg.pro
293 F G H T S R T H A N P V I P S D W Q N L Q S V I L A L R H N  S. purpuratus Erg.pro
243 L G P T S S R L A N - - - - - - - - - - - - - - - - - -  Danio rerio erg.pro

    P G S G Q I Q L W Q F L L E L L S D S S N S N C I T W E G T  Majority
    ------------------+------------------+------------------+-
                      400                410                420
    ------------------+------------------+------------------+-
313 P G S G Q I Q L W Q F L L E L L S D S S N S N C I T W E G T  Mus musculus Erg.pro
306 P G S G Q I Q L W Q F L L E L L S D S S N S S C I T W E G T  Homo sapiens Erg-1.pro
289 P G S G Q I Q L W Q F L L E L L S D S S N S S C I T W E G T  Homo sapiens Erg-2.pro
282 P G S G Q I Q L W Q F L L E L L S D S S N S N C I T W E G T  Rattus novegicus Erg.pro
305 P G S G Q I Q L W Q F L L E L L S D S S N S N C I T W E G T  Gallus gallus erg.pro
312 P G S G Q I Q L W Q F L L E L L S D S S N S N C I T W E G T  Xenopus laevis Erg.pro
323 K G S G Q I Q L W Q F L L E L L S D S S N A N C I T W E G T  S. purpuratus Erg.pro
253 P G S G Q I Q L W Q F L L E L L S D S C N S S C I T W E G T  Danio rerio erg.pro
```

Figure 7 continued

```
    N G E F K M T D P D E V A R R W G E R K S K P N M N Y D K L  Majority
    ------------------+------------------+------------------+-
                      430                440                450
```

```
343  N G E F K M T D P D E V A R R W G E R K S K P N M N Y D K L  Mus musculus Erg.pro
336  N G E F K M T D P D E V A R R W G E R K S K P N M N Y D K L  Homo sapiens Erg-1.pro
319  N G E F K M T D P D E V A R R W G E R K S K P N M N Y D K L  Homo sapiens Erg-2.pro
312  N G E F K M T D P D E V A R R W G E R K S K P N M N Y D K L  Rattus novegicus Erg.pro
335  N G E F K M T D P D E V A R R W G E R K S K P N M N Y D K L  Gallus gallus erg.pro
342  N G E F K M T D P D E V A R R W G E R K S K P N M N Y D K L  Xenopus laevis Erg.pro
353  N G E F K M T D P D E V A R R W G E R K S K P N M N Y D K L  S. purpuratus Erg.pro
283  N G E F K M T D P D E V A R R W G E R K S K P N M N Y D K L  Danio rerio erg.pro

     S R A L R Y Y Y D K N I M T K V H G K R Y A Y K F D F H G I  Majority
     ----------+----------+----------+----------+----------+-
                         460        470        480
     ----------+----------+----------+----------+----------+-
373  S R A L R Y Y Y D K N I M T K V H G K R Y A Y K F D F H G I  Mus musculus Erg.pro
366  S R A L R Y Y Y D K N I M T K V H G K R Y A Y K F D F H G I  Homo sapiens Erg-1.pro
349  S R A L R Y Y Y D K N I M T K V H G K R Y A Y K F D F H G I  Homo sapiens Erg-2.pro
342  S R A L R Y Y Y D K N I M T K V H G K R Y A Y K F D F H G I  Rattus novegicus Erg.pro
365  S R A L R Y Y Y D K N I M T K V H G K R Y A Y K F D F H G I  Gallus gallus erg.pro
372  S R A L R Y Y Y D K N I M T K V H G K R Y A Y K F D F H G I  Xenopus laevis Erg.pro
383  S R A L R Y Y Y D K N I M T K V H G K R Y A Y K F D F A G L  S. purpuratus Erg.pro
313  S R A L R Y Y Y D K N I M T K V H G K R Y A Y K F D F H G I  Danio rerio erg.pro

     A Q A L Q P H P P E S S L Y K Y - - - P S D L P Y M G S Y H  Majority
     ----------+----------+----------+----------+----------+-
                         490        500        510
     ----------+----------+----------+----------+----------+-
403  A Q A L Q P H P P E S S L Y K Y - - - P S D L P Y M G S Y H  Mus musculus Erg.pro
396  A Q A L Q P H P P E S S L Y K Y - - - P S D L P Y M G S Y H  Homo sapiens Erg-1.pro
379  A Q A L Q P H P P E S S L Y K Y - - - P S D L P Y M G S Y H  Homo sapiens Erg-2.pro
372  A Q A L Q P H P P E S S L Y K Y - - - P S D L P Y M G S Y H  Rattus novegicus Erg.pro
395  A Q A L Q P H P P E S S M Y K Y - - - P S D L P Y M S S Y H  Gallus gallus erg.pro
402  A Q A L Q P H P P E S T M Y K Y - - - P S E L P Y M S S Y H  Xenopus laevis Erg.pro
413  A Q A M Q P V Q A D P S M Y R Y - - - Q S D L S Y L Q S Y H  S. purpuratus Erg.pro
343  A Q A L Q P H P P D S S I Y K Y P A V A A D L P Y V S S Y H  Danio rerio erg.pro

     A H P Q K M N F V A P H P P A L P V T S S S F F A A P N P Y  Majority
     ----------+----------+----------+----------+----------+-
                         520        530        540
     ----------+----------+----------+----------+----------+-
430  A H P Q K M N F V S P H P P A L P V T S S S F F A S P N P Y  Mus musculus Erg.pro
423  A H P Q K M N F V A P H P P A L P V T S S S F F A A P N P Y  Homo sapiens Erg-1.pro
406  A H P Q K M N F V A P H P P A L P V T S S S F F A A P N P Y  Homo sapiens Erg-2.pro
399  T H P Q K M N F V A P H P P A L P V T S S S F F A T P N P Y  Rattus novegicus Erg.pro
422  A H P Q K M N F V A P H P P A L P V T S S S F F A A P N P Y  Gallus gallus erg.pro
429  A H P Q K M N F V A P H P P A L P V T S S S F F A A P N A Y  Xenopus laevis Erg.pro
440  - H P T K L N F V G - - T P I N P S T N A S L F S S H S S Y  S. purpuratus Erg.pro
373  - - P Q K M S F M A P H P Q A M S V S S S G F F S G P A P Y  Danio rerio erg.pro

     W N S P T - G G I Y P N - - - T R L P A S H M P S H L G T Y  Majority
     ----------+----------+----------+----------+----------+-
                         550        560        570
     ----------+----------+----------+----------+----------+-
460  W N S P T - G G I Y P N - - - T R L P A S H M P S H L G T Y  Mus musculus Erg.pro
453  W N S P T - G G I Y P N - - - T R L P T S H M P S H L G T Y  Homo sapiens Erg-1.pro
436  W N S P T - G G I Y P N - - - T R L P T S H M P S H L G T Y  Homo sapiens Erg-2.pro
429  W N S P T - G G I Y P N - - - T R L P A S H M P S H L G T Y  Rattus novegicus Erg.pro
452  W N S P T - G G I Y P N - - - T R L P A A H M P S H L G T Y  Gallus gallus erg.pro
459  W N S P T - G S I Y P N - - - T R L P A S H M S S H L G T Y  Xenopus laevis Erg.pro
467  W S S P T G A N I Y P S G H V T H P H P G H M S S H I G T Y  S. purpuratus Erg.pro
401  W N T P A - A G I Y G G - - - P R H P P A H M P S H L G S Y  Danio rerio erg.pro
```

Figure 7 continued

```
     Y -                                                       Majority
     ----

     ----
486  Y                                                         Mus musculus Erg.pro
```

479	Y	Homo sapiens Erg-1.pro
462	Y	Homo sapiens Erg-2.pro
455	Y	Rattus novegicus Erg.pro
478	Y	Gallus gallus erg.pro
485	Y	Xenopus laevis Erg.pro
497	Y G	S. purpuratus Erg.pro
427	Y	Danio rerio erg.pro

Figure 7 continued

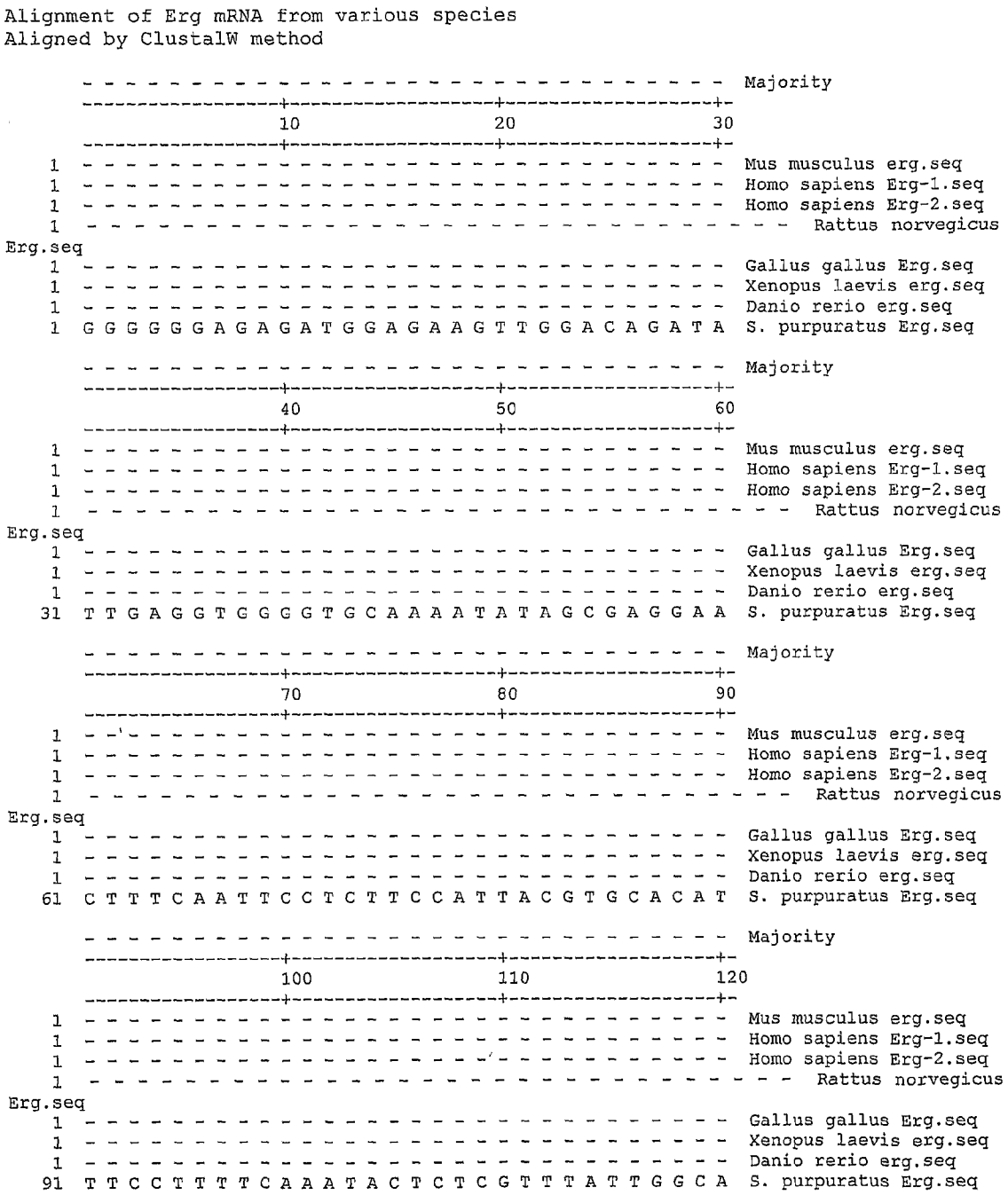

Alignment of Erg mRNA from various species
Aligned by ClustalW method
Majority
10 20 30
Mus musculus erg.seq
Homo sapiens Erg-1.seq
Homo sapiens Erg-2.seq
Rattus norvegicus Erg.seq
Gallus gallus Erg.seq
Xenopus laevis erg.seq
Danio rerio erg.seq
1 G G G G G G A G A G A T G G A G A A G T T G G A C A G A T A S. purpuratus Erg.seq
Majority
40 50 60
Mus musculus erg.seq
Homo sapiens Erg-1.seq
Homo sapiens Erg-2.seq
Rattus norvegicus Erg.seq
Gallus gallus Erg.seq
Xenopus laevis erg.seq
Danio rerio erg.seq
31 T T G A G G T G G G G T G C A A A A T A T A G C G A G G A A S. purpuratus Erg.seq
Majority
70 80 90
Mus musculus erg.seq
Homo sapiens Erg-1.seq
Homo sapiens Erg-2.seq
Rattus norvegicus Erg.seq
Gallus gallus Erg.seq
Xenopus laevis erg.seq
Danio rerio erg.seq
61 C T T T C A A T T C C T C T T C C A T T A C G T G C A C A T S. purpuratus Erg.seq
Majority
100 110 120
Mus musculus erg.seq
Homo sapiens Erg-1.seq
Homo sapiens Erg-2.seq
Rattus norvegicus Erg.seq
Gallus gallus Erg.seq
Xenopus laevis erg.seq
Danio rerio erg.seq
91 T T C C T T T T C A A A T A C T C T C G T T T A T T G G C A S. purpuratus Erg.seq

Figure 8

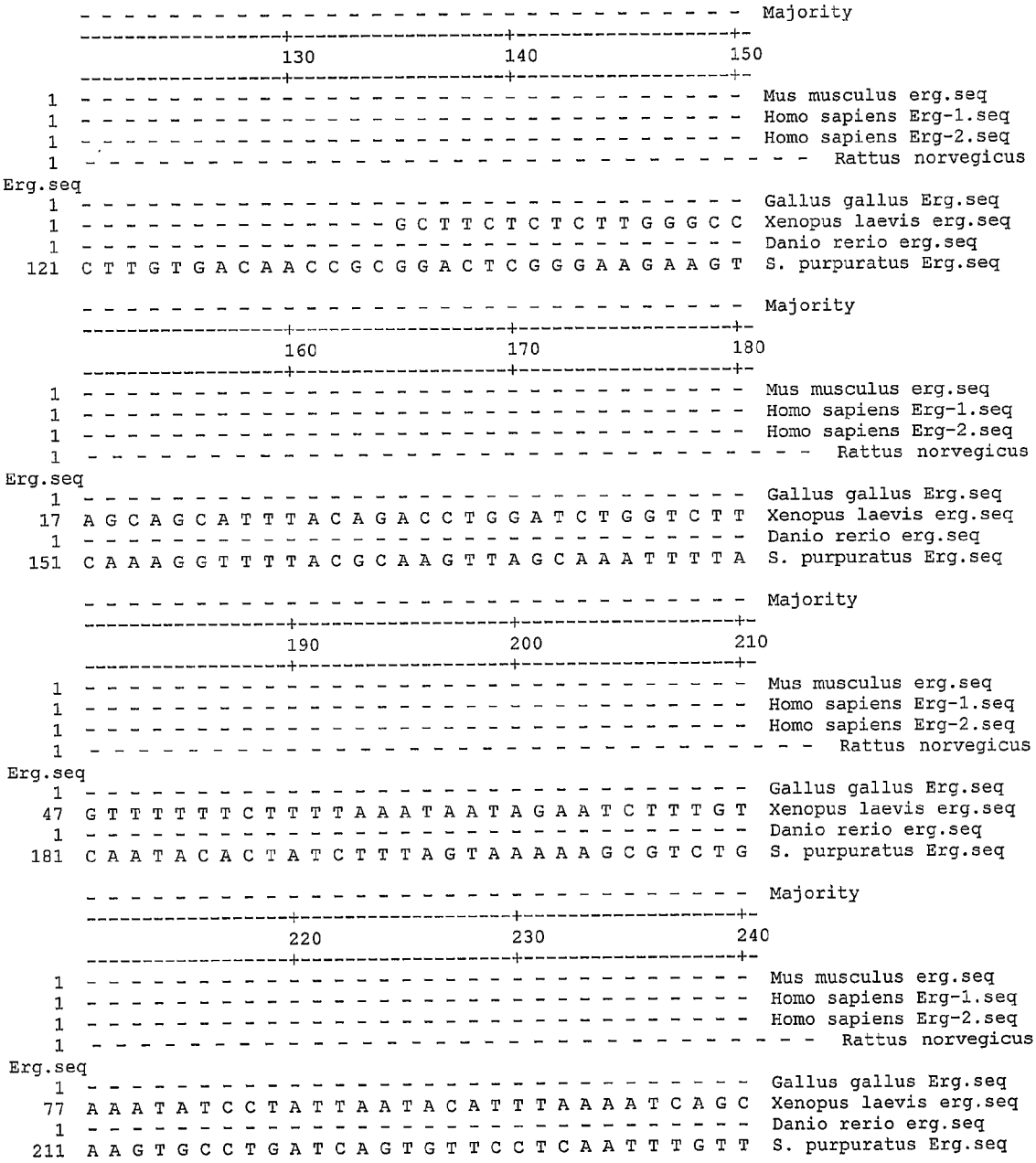


Figure 8 continued

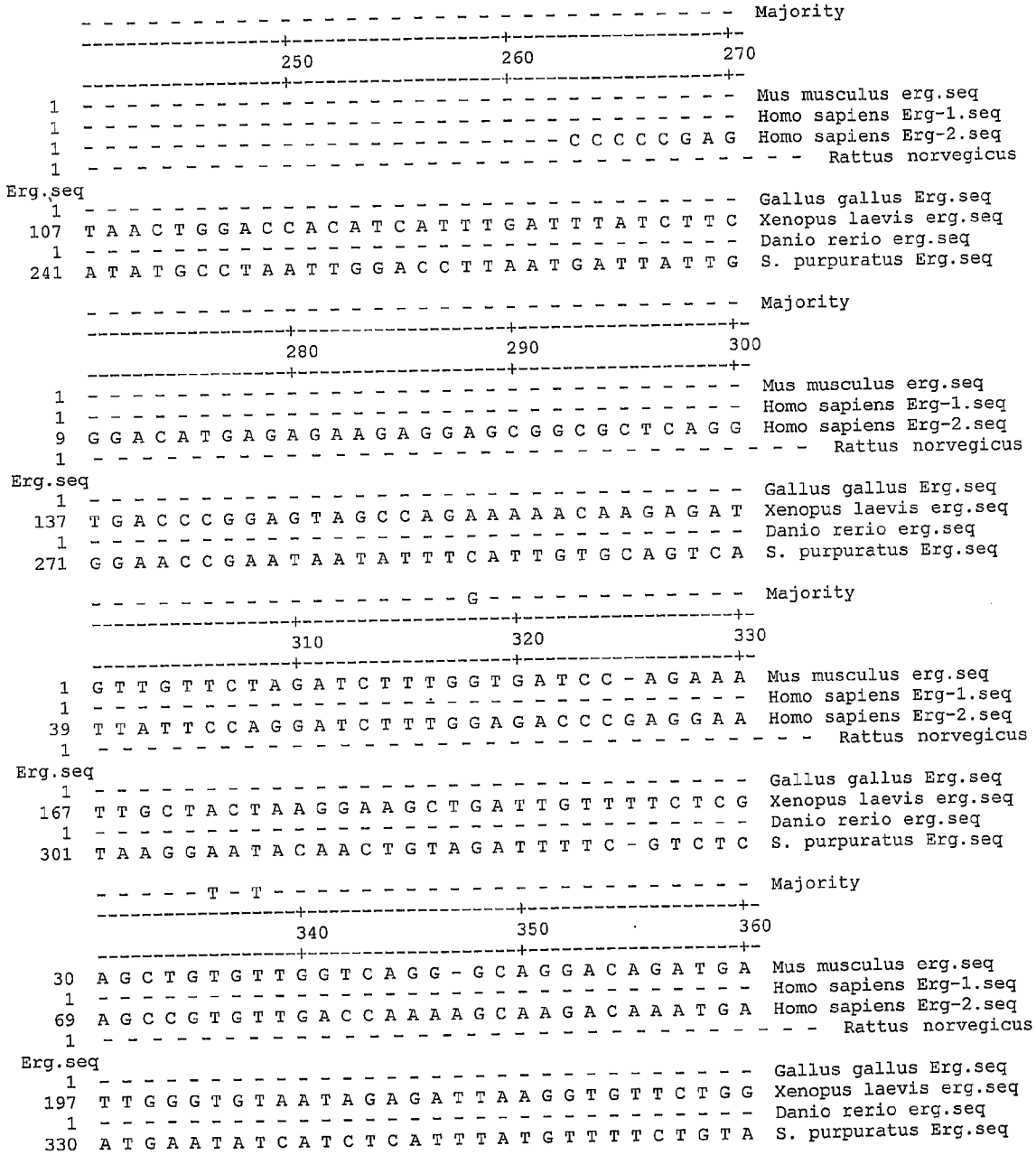


Figure 8 continued

```
      - - - - - - - - - - - - - - - - - - - - G G C - - A A - C A A C  Majority
      ----------------+-----------------+-----------------+-
                      370               380               390
      ----------------+-----------------+-----------------+-
   59 C T C C C A G A G A A A A A - G A T G G C G A A A C C A A G  Mus musculus erg.seq
    1 - - - - - - - - - - - - - - - - - - - G C T C T A A A C A A C  Homo sapiens Erg-1.seq
   99 C T C A C A G A G A A A A A A G A T G G C A G A A C C A A G  Homo sapiens Erg-2.seq
    1 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  Rattus norvegicus
Erg.seq
    1 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  Gallus gallus Erg.seq
  227 T T T A G C A C A T A T C A A T G T A G C T A T C G A T G C  Xenopus laevis erg.seq
    1 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  Danio rerio erg.seq
  360 T A G G A A G C A T T A T C G G T A G T C T C G G A C G A C  S. purpuratus Erg.seq

      C A C A - - - A A A - - - G T G C C A A G T T C T G C A A A  Majority
      ----------------+-----------------+-----------------+-
                      400               410               420
      ----------------+-----------------+-----------------+-
   88 G A C A A T C A G A G T C G T C C C G G G T T C T G C A T G  Mus musculus erg.seq
   13 C T C A T C A A A A C T A C T T T C T G G T C A G A G A G A  Homo sapiens Erg-1.seq
  129 G G C A A C T A A A - - G C C G T C A G G T T C T G A A C A  Homo sapiens Erg-2.seq
    1 - - - - - - - - - - - - - - - - G G T C A G A G A G A A G C A A T  Rattus norvegicus
Erg.seq
    1 - - - - - - - - - - - - - - - - - - - - G A A T T C C G C G - A  Gallus gallus Erg.seq
  257 C A A A C T A C A G C C A A A G C A A A A G A A A G A A A A  Xenopus laevis erg.seq
    1 - - - - - - - - - - - - - - - - - A G A T A A G T T G T G C G T G  Danio rerio erg.seq
  390 C A G A C G C T A T T T G A G A A C A C A T A C C G T G A G  S. purpuratus Erg.seq

      G C T T G T A A T T A T - - - G C T G T C A - T T A T T C A  Majority
      ----------------+-----------------+-----------------+-
                      430               440               450
      ----------------+-----------------+-----------------+-
  118 G C T G G T G A G T G A - - - G C T G G C - - C T G C T G A  Mus musculus erg.seq
   43 A G C A A T A A T T A T T - - A T T A A C A T T T A T T A A  Homo sapiens Erg-1.seq
  157 G C T G G T A G A T G G - - - G C T G G C - - T T A C T G A  Homo sapiens Erg-2.seq
   18 A A T T A T T A T T A A - - - C - - - A T - - T T A T T A A  Rattus norvegicus
Erg.seq
   12 A C G A A T A A T T A T T - - A T T A G C A A T T A T T A G  Gallus gallus Erg.seq
  287 G G A T G A G A A T T T G - - G A T T T T A A A G G A T C A  Xenopus laevis erg.seq
   17 G C G T G G C A G G A C A - - G C A G T G T G T T C T T C A  Danio rerio erg.seq
  420 G T G A A T A A G T C C A A C A C T A T C G T C T C G T C G  S. purpuratus Erg.seq

      C G A T C A A T G A T C C A G A T T G C A T T A T G C C C A  Majority
      ----------------+-----------------+-----------------+-
                      460               470               480
      ----------------+-----------------+-----------------+-
  143 A G G C C A - T G A T C C A G A C T G T A C C T G A C C C A  Mus musculus erg.seq
   71 C G A T C A A T A A A C T T G A T C G C A T T A T G G C C A  Homo sapiens Erg-1.seq
  182 A G G A C A - T G A T T C A G A C T G T C C C G G A C C C A  Homo sapiens Erg-2.seq
   40 C G A T C A A T A A A C T T G A T T G C A T T A T G G C C A  Rattus norvegicus
Erg.seq
   40 C G A T C A A T A A T C T T G A T C A C A T T A T G G C A A  Gallus gallus Erg.seq
  315 G G A A A A A T G C A A C A C A T G G C A T T C A T C C C A  Xenopus laevis erg.seq
   45 G G A C C G G T G T C T C - T G T G A G G A G C A G C T C A  Danio rerio erg.seq
  450 C C A C A C A G C A C C C A C A T C C T G A C A A C A C C T  S. purpuratus Erg.seq
```

Figure 8 continued

```
      G C - - - - - - A G C T - A T A T T A A G G A G G C C T T A  Majority
      ----------------+-----------------+-----------------+-
                      490               500               510
      ----------------+-----------------+-----------------+-
  172 G C - - - - - - A G C T C A T A T T A A G G A G G C C T T G  Mus musculus erg.seq
  101 G C - - - - - - A C - - - - T A T T A A G G A A G C C T T A  Homo sapiens Erg-1.seq
  211 G C - - - - - - A G C T C A T A T C A A G G A A G C C T T A  Homo sapiens Erg-2.seq
   70 G C - - - - - - A C - - - - T A T T A A G G A G G C C T T G  Rattus norvegicus
Erg.seq
   70 G C - - - - - - A C - - - - T A T T A A G G A A G C A T T A  Gallus gallus Erg.seq
  345 A C C T A C G A G G A T T A T A G T A A T G A A G C T C T T  Xenopus laevis erg.seq
   74 - - - - - - - - G G A T G A T C T C T C T G A G G C G C T G  Danio rerio erg.seq
  480 G T C C A C A A G C C T A A G C T A T C C G T G G T T G C A  S. purpuratus Erg.seq

      T C A G T T G T G A G T G A G G A C C A G - - T C G T T G T  Majority
      ----------------+-----------------+-----------------+-
                      520               530               540
      ----------------+-----------------+-----------------+-
  196 T C A G T T G T G A G C G A G G A C C A G - - T C A C T A T  Mus musculus erg.seq
  121 T C A G T T G T G A G T G A G G A C C A G - - T C G T T G T  Homo sapiens Erg-1.seq
  235 T C A G T T G T G A G T G A G G A C C A G - - T C G T T G T  Homo sapiens Erg-2.seq
   90 T C A G T T G T G A G C A A G G A C C A G - - T C A C T A T  Rattus norvegicus
Erg.seq
   90 T C A G T G G T G A G T G A A G A C C A G - - T C C T T G T  Gallus gallus Erg.seq
  375 T C A G T G G T G A G T G A A G A T C A G - - T C G T T A T  Xenopus laevis erg.seq
   96 T C G G T G G T G A G T G A G G A T C C G - - T G T G T A T  Danio rerio erg.seq
  510 T C G C C C A C C A A C G A G A G C C A C A T C T G G A G C  S. purpuratus Erg.seq

      T T G A G T G T G C - C T A C G G A A C G C C A C A C C T G  Majority
      ----------------+-----------------+-----------------+-
                      550               560               570
      ----------------+-----------------+-----------------+-
  224 T T G A G T G T G C - C T A C G G A A C G C C A C A C C T G  Mus musculus erg.seq
  149 T T G A G T G T G C - C T A C G G A A C G C C A C A C C T G  Homo sapiens Erg-1.seq
  263 T T G A G T G T G C - C T A C G G A A C G C C A C A C C T G  Homo sapiens Erg-2.seq
  118 T T G A G T G T G C - C T A C G G A A C G C C A C A C C T G  Rattus norvegicus
Erg.seq
  118 T T G A G T G T G C - C T A C G G A T C G C C C C A C C T T  Gallus gallus Erg.seq
  403 T C G A G T G C A C - C T A C G G A A C G C C A C A T C T T  Xenopus laevis erg.seq
  124 C A G A G T G T G T - G T A C A C A G C G G C G T C T C C T  Danio rerio erg.seq
  540 T C G T G C G T A C A G G A C C A A A G G A T G A A A C A A  S. purpuratus Erg.seq

      G C T A A G A C A G A G A T G A C C G C G T C C T C C T C C  Majority
      ----------------+-----------------+-----------------+-
                      580               590               600
      ----------------+-----------------+-----------------+-
  253 G C T A A G A C A G A G A T G A C C G C A T C C T C T T C C  Mus musculus erg.seq
  178 G C T A A G A C A G A G A T G A C C G C G T C C T C C T C C  Homo sapiens Erg-1.seq
  292 G C T A A G A C A G A G A T G A C C G C G T C C T C C T C C  Homo sapiens Erg-2.seq
  147 G C T A A G A C A G A A A T G A C C G C A T C C T C T T C C  Rattus norvegicus
Erg.seq
  147 G C A A A G A C A G A A A T G A C A G C C T C C T C T T C C  Gallus gallus Erg.seq
  432 A C T A A A A C G G A G A T G A C C G C A T C T T C C T C C  Xenopus laevis erg.seq
  153 C T G A G C G T G C A G A T G A C G G C G T C T G C A G C C  Danio rerio erg.seq
  570 G A G A T T G A G C A G C C A A C T T C C G G T G C C C A C  S. purpuratus Erg.seq
```

Figure 8 continued

```
      A G T G A C T A T - - - - G G C C A - - - G A C A T C C A A  Majority
      -----------------+-----------------+-----------------+-
                       610               620               630
      -----------------+-----------------+-----------------+-
  283 A G T G A C T A T - - - - G G C C A - - - G A C A T C C A A  Mus musculus erg.seq
  208 A G C G A C T A T - - - - G G A C A - - - G A C T T C C A A  Homo sapiens Erg-1.seq
  322 A G C G A C T A T - - - - G G A C A - - - G A C T T C C A A  Homo sapiens Erg-2.seq
  177 A G T G A C T A T - - - - G G C C A - - - G A C A T C C A A  Rattus norvegicus
Erg.seq
  177 A G T G A A T A T - - - - G G G C A - - - A A C A T C A A A  Gallus gallus Erg.seq
  462 A G T G A C T A T - - - - G G G C A - - - A A C C T C G A A  Xenopus laevis erg.seq
  183 G C T G A G C - - - - - - - - - - - - - - - T C C C T C A C  Danio rerio erg.seq
  600 G A T G G G C T T C C G G G G C C A C A G G T C A G T C G C  S. purpuratus Erg.seq

      G A T G A G C C C C C G C G T C C C T C A G C A G G A C T G  Majority
      -----------------+-----------------+-----------------+-
                       640               650               660
      -----------------+-----------------+-----------------+-
  306 G A T G A G T C C C A G A G T C C C T C A G C A G G A C T G  Mus musculus erg.seq
  231 G A T G A G C C C A C G C G T C C C T C A G C A G G A T T G  Homo sapiens Erg-1.seq
  345 G A T G A G C C C A C G C G T C C C T C A G C A G G A T T G  Homo sapiens Erg-2.seq
  200 G A T G A G T C C C A G A G T C C C A C A A C A G G A C T G  Rattus norvegicus
Erg.seq
  200 G A T G A G C C C G C G C G T T C C C C A G C A G G A C T G  Gallus gallus Erg.seq
  485 A A T G A G T C C A C G T G T G C C T C A A C A A G A C T G  Xenopus laevis erg.seq
  198 A G T A A A C T C C T G C - T G T C T - - - - - - - G C T G  Danio rerio erg.seq
  630 G G G G G A C C C G A G T - C G C C G T T G G A T T G T A G  S. purpuratus Erg.seq

      G C T G T C T C A A C C C C C A G C C A G G G T C A C C A T  Majority
      -----------------+-----------------+-----------------+-
                       670               680               690
      -----------------+-----------------+-----------------+-
  336 G C T G T C T C A A G C C C C A G C C A G G G T C A C C A T  Mus musculus erg.seq
  261 G C T G T C T C A A C C C C C A G C C A G G G T C A C C A T  Homo sapiens Erg-1.seq
  375 G C T G T C T C A A C C C C C A G C C A G G G T C A C C A T  Homo sapiens Erg-2.seq
  230 G C T G T C T C A G C C C C C A G C C A G G G T C A C C A T  Rattus norvegicus
Erg.seq
  230 G T T A T C A C A G C C C C C G G C C A G A G T T A C C A T  Gallus gallus Erg.seq
  515 G C T C T C A C A G C C T C C A T C C A G G G T G A C C A T  Xenopus laevis erg.seq
  220 T C T C T G T C A A G C T - - - - - - - G G A - G G C C A G  Danio rerio erg.seq
  659 C G T A G C G A A A C C C C - - G C C A G C A A C C T C C T  S. purpuratus Erg.seq

      C A A G A T G G A G T G C - - - - - A A C C C T A G C C A G  Majority
      -----------------+-----------------+-----------------+-
                       700               710               720
      -----------------+-----------------+-----------------+-
  366 C A A G A T G G A G T G C - - - - - A A C C C T A G T C A G  Mus musculus erg.seq
  291 C A A A A T G G A A T G T - - - - - A A C C C T A G C C A G  Homo sapiens Erg-1.seq
  405 C A A A A T G G A A T G T - - - - - A A C C C T A G C C A G  Homo sapiens Erg-2.seq
  260 C A A G A T G G A G T G C - - - - - A A C C C T A G C C A G  Rattus norvegicus
Erg.seq
  260 T A A G A T G G A G T G T - - - - - A A C C C A A A C C A G  Gallus gallus Erg.seq
  545 C A A G A T G G A A T G C - - - - - A A C C C T A G T C A A  Xenopus laevis erg.seq
  242 C - - G C T G G A G C A C - - - - - A C C G - - - - - - G A  Danio rerio erg.seq
  687 C C C C A C G G A G C G C C A C A G G G C A C C A T C A A C  S. purpuratus Erg.seq
```

Figure 8 continued

```
      G T G A A T G G A T C A A G G A A C T C T C C - T G A T G A  Majority
      ------------------+------------------+------------------+-
                        730                740                750
      ------------------+------------------+------------------+-
  391 G T G A A T G G T T C C A G G A A C T C A C C - T G A T G A  Mus musculus erg.seq
  316 G T G A A T G G C T C A A G G A A C T C T C C - T G A T G A  Homo sapiens Erg-1.seq
  430 G T G A A T G G C T C A A G G A A C T C T C C - T G A T G A  Homo sapiens Erg-2.seq
  285 G T G A A C G G T T C C A G G A A C T C A C C - T G A C G A  Rattus norvegicus
Erg.seq
  285 G T T A A T G G G T C A A G G A A T T C A C C - T G A T G A  Gallus gallus Erg.seq
  570 G T C A A T G G A T C A A G G A G C T C C C C - G G A T G A  Xenopus laevis erg.seq
  259 G C C A C T G C A - - - - G G A G T C C T C C - T G A T G A  Danio rerio erg.seq
  717 G C G G A G C C A A C A C A G G C C G C G C C A C C A T A T  S. purpuratus Erg.seq

      C T G C A G T G T G - G C C A A A G G C G G G A A G A T G G  Majority
      ------------------+------------------+------------------+-
                        760                770                780
      ------------------+------------------+------------------+-
  420 G T G C A G T G T G - A A C A A A G G T G G G A A G A T G G  Mus musculus erg.seq
  345 A T G C A G T G T G - G C C A A A G G C G G G A A G A T G G  Homo sapiens Erg-1.seq
  459 A T G C A G T G T G - G C C A A A G G C G G G A A G A T G G  Homo sapiens Erg-2.seq
  314 A T G C A G T G T G - A C C A A A G G T G G G A A G A T G G  Rattus norvegicus
Erg.seq
  314 C T G C A G C G T G - G C A A A A G G A G G G A A A A T G G  Gallus gallus Erg.seq
  599 C T G C A G T A T A - G G A A A A G G A A G C A A A A T G G  Xenopus laevis erg.seq
  284 C T G T C C G G G G - G T C A A A G G A - - - A A G A T C T  Danio rerio erg.seq
  747 C C G T C G T G T C C G C C A C A G C C G T A C C T G G G T  S. purpuratus Erg.seq

      T G G G C A G C C C A G A C A C T G T T G G G A T G A A C T  Majority
      ------------------+------------------+------------------+-
                        790                800                810
      ------------------+------------------+------------------+-
  449 T G G G C A G C C C G G A T A C T G T G G G G A T G A G C T  Mus musculus erg.seq
  374 T G G G C A G C C C A G A C A C C G T T G G G A T G A A C T  Homo sapiens Erg-1.seq
  488 T G G G C A G C C C A G A C A C C G T T G G G A T G A A C T  Homo sapiens Erg-2.seq
  343 T G G G C A G C C C T G A T A C T G T G G G A A T G A G C T  Rattus norvegicus
Erg.seq
  343 T T A G C A G T T C A G A C A A T G T T G G G A T G A A C T  Gallus gallus Erg.seq
  628 G T G G A G G G T C G G A T A A T G T T G G A A T G A A T T  Xenopus laevis erg.seq
  310 C T C C G G - - - C T G G G A A T G C T C C A G T G A T G T  Danio rerio erg.seq
  777 A C C A C G G G T T C C A C G C C C G A G G A A C G G T C G  S. purpuratus Erg.seq

      A C G G C A G C T A C A T G G A G G A G A A G C A C A T G C  Majority
      ------------------+------------------+------------------+-
                        820                830                840
      ------------------+------------------+------------------+-
  479 A C G G C A G C T A C A T G G A G G A G A A G C A T G T G C  Mus musculus erg.seq
  404 A C G G C A G C T A C A T G G A G G A G A A G C A C A T G C  Homo sapiens Erg-1.seq
  518 A C G G C A G C T A C A T G G A G G A G A A G C A C A T G C  Homo sapiens Erg-2.seq
  373 A C G G C A G C T A C A T G G A G G A G A A G C A C A T G C  Rattus norvegicus
Erg.seq
  373 A T G G A A G C T A C A T G G A A G A G A A G C A T A T T C  Gallus gallus Erg.seq
  658 A T G G C A G C T A C A T A G A A G A A A A A C A C A T T T  Xenopus laevis erg.seq
  337 A C A G C A A C T A C A T G G A G G A G A A A C A C A C A C  Danio rerio erg.seq
  807 G C T T C A G C G G G T A C C A C G G G A A G G T C A C C G  S. purpuratus Erg.seq
```

Figure 8 continued

```
      C G C C C C C A A A T A T G A C C A C C A A T G A G C G C A  Majority
      ------------------+------------------+------------------+-
                        850                860                870
      ------------------+------------------+------------------+-
  509 C G C C T C C C A A T A T G A C C A C A A A T G A G C G C A  Mus musculus erg.seq
  434 C A C C C C C A A A C A T G A C C A C G A A C G A G C G C A  Homo sapiens Erg-1.seq
  548 C A C C C C C A A A C A T G A C C A C G A A C G A G C G C A  Homo sapiens Erg-2.seq
  403 C G C C C C C C A A T A T G A C C A C G A A T G A A C G C A  Rattus norvegicus
Erg.seq
  403 C G C C T C C A A A T A T G A C A A C C A A T G A A C G A A  Gallus gallus Erg.seq
  688 C T C C C C C A A A C A T G A C A A C C A A T G A G A G G A  Xenopus laevis erg.seq
  367 C - - - G C - - - - - - - - - C G A G C A G T G A G C G C A  Danio rerio erg.seq
  837 C - - C A C C A A A T G T C A C C A C C A A T G A G A A G A  S. purpuratus Erg.seq

      G A G T T A T C G T G C C T G C A G A T C C T A C G C T G T  Majority
      ------------------+------------------+------------------+-
                        880                890                900
      ------------------+------------------+------------------+-
  539 G A G T G A T C G T C C C T G C A G A T C C T A C T C T G T  Mus musculus erg.seq
  464 G A G T T A T C G T G C C A G C A G A T C C T A C G C T A T  Homo sapiens Erg-1.seq
  578 G A G T T A T C G T G C C A G C A G A T C C T A C G C T A T  Homo sapiens Erg-2.seq
  433 G A G T G A T C G T C C C T G C A G A T C C T A C T C T G T  Rattus norvegicus
Erg.seq
  433 G A G T T A T T G T G C C A G C A G A T C C T A C G T T A T  Gallus gallus Erg.seq
  718 G A G T G A T T G T A C C T G C T G A C C C T A C T T T G T  Xenopus laevis erg.seq
  385 G A G T C A T C G T A C C T G C A G A C C C G G G G C A G T  Danio rerio erg.seq
  865 G G G T C A T C G T A C C T G C T G A C C C T A A C A T G T  S. purpuratus Erg.seq

      G G A G C A C A G A C C A T G T G C G G C A G T G G C T G G  Majority
      ------------------+------------------+------------------+-
                        910                920                930
      ------------------+------------------+------------------+-
  569 G G A G C A C A G A C C A T G T C C G A C A G T G G C T G G  Mus musculus erg.seq
  494 G G A G T A C A G A C C A T G T G C G G C A G T G G C T G G  Homo sapiens Erg-1.seq
  608 G G A G T A C A G A C C A T G T G C G G C A G T G G C T G G  Homo sapiens Erg-2.seq
  463 G G A G C A C A G A C C A T G T C C G G C A G T G G C T G G  Rattus norvegicus
Erg.seq
  463 G G A G C A C A G A C C A T G T A C G G C A G T G G C T G G  Gallus gallus Erg.seq
  748 G G A G C A C A G A T C A T G T C A G A C A G T G G C T A G  Xenopus laevis erg.seq
  415 G G T C G G C G G G T C A C G T G C G G C A G T G G C T G G  Danio rerio erg.seq
  895 G G A C T G C T G A G C A C G T A C A G C A G T G G G T C C  S. purpuratus Erg.seq

      A G T G G G C G G T G A A A G A A T A T G G C C T T C C A G  Majority
      ------------------+------------------+------------------+-
                        940                950                960
      ------------------+------------------+------------------+-
  599 A G T G G G C G G T G A A A G A A T A T G G C C T C C T C G  Mus musculus erg.seq
  524 A G T G G G C G G T G A A A G A A T A T G G C C T T C C A G  Homo sapiens Erg-1.seq
  638 A G T G G G C G G T G A A A G A A T A T G G C C T T C C A G  Homo sapiens Erg-2.seq
  493 A G T G G G C A G T G A A A G A A T A C G G C C T C C T A G  Rattus norvegicus
Erg.seq
  493 A G T G G G C A G T G A A G G A G T A T G G T C T T C C A G  Gallus gallus Erg.seq
  778 A G T G G G C A A T A A A A G A A T A C G G C C T T C C A G  Xenopus laevis erg.seq
  445 A C T G G G C G G T G C A G G A G T A C G G G C T G C C G C  Danio rerio erg.seq
  925 A G T G G G C G G T C A G G G A A T A C T C C C T G G T G G  S. purpuratus Erg.seq
```

Figure 8 continued

```
      A C G T G G A C G T C T T G T T G T T T C A G A A C A T C G  Majority
      ------------------+------------------+------------------+-
                        970                980                990
      ------------------+------------------+------------------+-
  629 A T G T G G A C G T C T T A C T A T T T C A G A A T A T C G  Mus musculus erg.seq
  554 A C G T C A A C A T C T T G T T A T T C C A G A A C A T C G  Homo sapiens Erg-1.seq
  668 A C G T C A A C A T C T T G T T A T T C C A G A A C A T C G  Homo sapiens Erg-2.seq
  523 A C G T G G A C G T C T T A T T A T T T C A G A A T A T T G  Rattus norvegicus
Erg.seq
  523 A C G T G G A C A T C T T G T T G T T C C A G A A C A T T G  Gallus gallus Erg.seq
  808 A T G T C G A T G T G C T C C T G T T C C A A A A C A T C G  Xenopus laevis erg.seq
  475 C G C T G G A C G T G T C T C T G T T T C A G G G T G T G G  Danio rerio erg.seq
  955 A T G T C C A G G T G T C C A G G T T T - - - A A C A T G G  S. purpuratus Erg.seq

      A T G G G A A G G A G C T G T G C A A G A T G A C C A A A G  Majority
      ------------------+------------------+------------------+-
                        1000               1010               1020
      ------------------+------------------+------------------+-
  659 A T G G G A A G G A G C T G T G C A A G A T G A C A A A G G  Mus musculus erg.seq
  584 A T G G G A A G G A A C T G T G C A A G A T G A C C A A G G  Homo sapiens Erg-1.seq
  698 A T G G G A A G G A A C T G T G C A A G A T G A C C A A G G  Homo sapiens Erg-2.seq
  553 A C G G G A A G G A G C T G T G C A A G A T G A C A A A A G  Rattus norvegicus
Erg.seq
  553 A T G G G A A A G A G T T G T G T A A A A T G A C C A A A G  Gallus gallus Erg.seq
  838 A C G G C A A G G A A T T A T G C A A G A T G A C C A A A G  Xenopus laevis erg.seq
  505 A T G G G A A G C A G C T C T G C A G G A T G A G C A G A G  Danio rerio erg.seq
  982 A C G G G A A A C A C C T T T G C A A G A T G A C C A G A G  S. purpuratus Erg.seq

      A T G A C T T C C A G A G G C T C A C C C C C A G C T A C A  Majority
      ------------------+------------------+------------------+-
                        1030               1040               1050
      ------------------+------------------+------------------+-
  689 A T G A C T T C C A G C G G C T C A C G C C G A G C T A C A  Mus musculus erg.seq
  614 A C G A C T T C C A G A G G C T C A C C C C C A G C T A C A  Homo sapiens Erg-1.seq
  728 A C G A C T T C C A G A G G C T C A C C C C C A G C T A C A  Homo sapiens Erg-2.seq
  583 A T G A C T T C C A G A G G C T C A C T C C G A G C T A C A  Rattus norvegicus
Erg.seq
  583 A T G A C T T C C A G A G A C T C A C G C C G A G C T A T A  Gallus gallus Erg.seq
  868 A A G A T T T C C A T C G A C T C A C G C C A A G C T A C A  Xenopus laevis erg.seq
  535 A G G A G C T C C A G C G T C T G A C C A G C A G C T A C T  Danio rerio erg.seq
 1012 A C G A C T T C A G C A G A C T G A C C A A C A A C C T C A  S. purpuratus Erg.seq

      A T G C C G A C A T C C T T C T C T C A C A T C T C C A C T  Majority
      ------------------+------------------+------------------+-
                        1060               1070               1080
      ------------------+------------------+------------------+-
  719 A T G C C G A C A T T C T T C T C T C A C A T C T C C A C T  Mus musculus erg.seq
  644 A C G C C G A C A T C C T T C T C T C A C A T C T C C A C T  Homo sapiens Erg-1.seq
  758 A C G C C G A C A T C C T T C T C T C A C A T C T C C A C T  Homo sapiens Erg-2.seq
  613 A T G C C G A C A T T C T T C T C T C A C A T C T C C A C T  Rattus norvegicus
Erg.seq
  613 A C G C A G A T A T C C T C C T G T C A C A C C T A C A C T  Gallus gallus Erg.seq
  898 A T G C T G A T A T C C T T C T C T C G C A T C T T C A C T  Xenopus laevis erg.seq
  565 C T G C A G A C G T G C T G C T C T C T C A C C T G C A C T  Danio rerio erg.seq
 1042 A C G T A G A T G T C C T C A T C T C A C A T C T A A C C T  S. purpuratus Erg.seq
```

Figure 8 continued

```
      A C C T C A G A - G A G A C T C C T C T T C C A C A T T T G  Majority
      ----------------+-----------------+-----------------+-
                      1090              1100              1110
      ----------------+-----------------+-----------------+-
  749 A C C T C A G A - G A G A C T C C C C T T C C A C A T C T G  Mus musculus erg.seq
  674 A C C T C A G A - G A G A C T C C T C T T C C A C A T T T G  Homo sapiens Erg-1.seq
  788 A C C T C A G A - G A G A C T C C T C T T C C A C A T T T G  Homo sapiens Erg-2.seq
  643 A C C T C A G A - G A G A C T C C C C T T C C A C A T C T G  Rattus norvegicus
Erg.seq
  643 A C C T C A G A - G A G A C T C C T C T T C C A C A T T T G  Gallus gallus Erg.seq
  928 A C C T C A G A - G A A A C T C C T C T T C C A C A T T T G  Xenopus laevis erg.seq
  595 A C C T C A G A - G A G A C T C C A C T T C C T C A C T T G  Danio rerio erg.seq
 1072 T C C T T A A A C A A G G T T A T A C A C C A G G G G C C A  S. purpuratus Erg.seq

      A C - - - - - T T C A G A T G A T G T T G A - T A A G G C C  Majority
      ----------------+-----------------+-----------------+-
                      1120              1130              1140
      ----------------+-----------------+-----------------+-
  778 A C - - - - - T T C C G A T G A C G T T G A - T A A G G C T  Mus musculus erg.seq
  703 A C - - - - - T T C A G A T G A T G T T G A - T A A A G C C  Homo sapiens Erg-1.seq
  817 A C - - - - - T T C A G A T G A T G T T G A - T A A A G C C  Homo sapiens Erg-2.seq
  672 A C - - - - - A T C C G A C G A C G T T G A - T A A G G C T  Rattus norvegicus
Erg.seq
  672 A C - - - - - T T C A G A T G A T G T T G A - T A A G G C C  Gallus gallus Erg.seq
  957 A C - - - - - T T C A G A T G A T G T T G A - T A A G G C C  Xenopus laevis erg.seq
  624 A C - - - - - T T C A G A T G A T G T T G A - C A A A G C C  Danio rerio erg.seq
 1102 A G A G C A A T C T G G A T G A C A A C A A A T A C A A C A  S. purpuratus Erg.seq

      T T A C A A A A C T C T - - - C C A C G G T T A A T G C A T  Majority
      ----------------+-----------------+-----------------+-
                      1150              1160              1170
      ----------------+-----------------+-----------------+-
  802 T T A C A A A A C T C T - - - C C A C G G T T A A T G C A T  Mus musculus erg.seq
  727 T T A C A A A A C T C T - - - C C A C G G T T A A T G C A T  Homo sapiens Erg-1.seq
  841 T T A C A A A A C T C T - - - C C A C G G T T A A T G C A T  Homo sapiens Erg-2.seq
  696 T T A C A A A A C T C T - - - C C A C G G T T A A T G C A T  Rattus norvegicus
Erg.seq
  696 T T A C A A A A C T C T - - - C C A C G G T T A A T G C A T  Gallus gallus Erg.seq
  981 T T A C A A A A C T C T - - - C C A C G G T T A A T G C A T  Xenopus laevis erg.seq
  648 T T G C A A A A C T C C - - - C C G C G G T T A A T G C A T  Danio rerio erg.seq
 1132 C T G C A G A C C A T C A G A C C A C T C T G G A C A C G A  S. purpuratus Erg.seq

      G C T A G A A A C A C A G G A G G T G C C T C T T T T A T T  Majority
      ----------------+-----------------+-----------------+-
                      1180              1190              1200
      ----------------+-----------------+-----------------+-
  829 G C C A G A A A C A C A G G G G G T G C A G C T T T T A T T  Mus musculus erg.seq
  754 G C T A G A A A C A C A G G G G G T G C A G C T T T T A T T  Homo sapiens Erg-1.seq
  868 G C T A G A A A C A C A - - - G - - - - - - - - - - - - - -  Homo sapiens Erg-2.seq
  723 G C T A G A A A C A C A G - - - - - - - - - - - - - - - - -  Rattus norvegicus
Erg.seq
  723 G C T A G A A A C A C A G G A G G A G C C A C T T T T A T T  Gallus gallus Erg.seq
 1008 G C T A G A A A T A C A G G A G G T G C C T C T T T T A T T  Xenopus laevis erg.seq
  675 G C T C G C A A C A C A G G A G G A G C C A A C T T C A T C  Danio rerio erg.seq
 1162 A T C C A T C A G G T G G G T C A G C C T T T C C T T A T C  S. purpuratus Erg.seq
```

Figure 8 continued

```
       T T C C - - - - - C A A A T A C A T C A G T T T A - - - T -  Majority
       ------------------+------------------+------------------+-
                         1210               1220               1230
       ------------------+------------------+------------------+-
   859 T T C C - - - - - C A A A T A C T T C A G T A T A - - - T -  Mus musculus erg.seq
   784 T T C C - - - - - C A A A T A C T T C A G T A T A - - - T -  Homo sapiens Erg-1.seq
   881 - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  Homo sapiens Erg-2.seq
   736 - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  Rattus norvegicus
Erg.seq
   753 T T T C - - - - - C A A A T A C A T C A G T T T A - - - C -  Gallus gallus Erg.seq
  1038 T T C C - - - - - C A A A C T C A T C C G T T T A - - - T -  Xenopus laevis erg.seq
   705 T T T C - - - - - C C A A C A C A C C A G T C T A C C C T -  Danio rerio erg.seq
  1192 C T C C T T C C A C T A C T A C C A C T G T T G A C A G T G  S. purpuratus Erg.seq

       - - - - C C A G A A G C T A C G C A A A G A A T - A C C A -  Majority
       ------------------+------------------+------------------+-
                         1240               1250               1260
       ------------------+------------------+------------------+-
   880 - - - - C C C G A A G C T A C G C A A A G A A T T A C A A -  Mus musculus erg.seq
   805 - - - - C C T G A A G C T A C G C A A A G A A T T A C A A -  Homo sapiens Erg-1.seq
   881 - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  Homo sapiens Erg-2.seq
   736 - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  Rattus norvegicus
Erg.seq
   774 - - - - C C A G A A G C A A C G C A A A G A A T A A C A A -  Gallus gallus Erg.seq
  1059 - - - - C A A G A C G C A A A C C A A A G A A T C C C C A -  Xenopus laevis erg.seq
   729 - - - - C C A G A C G C C A G C A G A G G A G - - - C C A -  Danio rerio erg.seq
  1222 T T C A C C G C A T G C C T C G C A C A G A A C C C T C A T  S. purpuratus Erg.seq

       - - - - - - - - - - - - - - - - - - - - G T A G G C C - -  Majority
       ------------------+------------------+------------------+-
                         1270               1280               1290
       ------------------+------------------+------------------+-
   905 - - - - - - - - - - - - - - - - - - - - C T A G G C C - -  Mus musculus erg.seq
   830 - - - - - - - - - - - - - - - - - - - - C T A G G C C - -  Homo sapiens Erg-1.seq
   881 - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  Homo sapiens Erg-2.seq
   736 - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  Rattus norvegicus
Erg.seq
   799 - - - - - - - - - - - - - - - - - - - - C A A G G C C - -  Gallus gallus Erg.seq
  1084 - - - - - - - - - - - - - - - - - - - - G C A G G C A - -  Xenopus laevis erg.seq
   751 - - - - - - - - - - - - - - - - - - - - G C A G A G C - -  Danio rerio erg.seq
  1252 G T G A T A G C C T T G T T G G A A G G G G A A G A C C C A  S. purpuratus Erg.seq

       A G A T T T A C C T T A T G A G C C T C C C A G G A G A T C  Majority
       ------------------+------------------+------------------+-
                         1300               1310               1320
       ------------------+------------------+------------------+-
   912 A G A T T T A C C T T A T G A G C C T C C C A G G A G A T C  Mus musculus erg.seq
   837 A G A T T T A C C A T A T G A G C C C C C C A G G A G A T C  Homo sapiens Erg-1.seq
   881 - - A T T T A C C A T A T G A G C C C C C C A G G A G A T C  Homo sapiens Erg-2.seq
   736 - - A T T T A C C T T A T G A G C C T C C C A G G A G A T C  Rattus norvegicus
Erg.seq
   806 A G A T T T A C C T T A T G A G C A A G C G A G G A G A T C  Gallus gallus Erg.seq
  1091 A G A T T T A T C A T A T G A G C C C T C A A G G A G A T C  Xenopus laevis erg.seq
   758 A G G T G T G G G T T A T G A C G C T G T C A G G C G C T C  Danio rerio erg.seq
  1282 A C G C T T G G C C C A C C A C C G T C C C A T C C G C T G  S. purpuratus Erg.seq
```

Figure 8 continued

```
      A G C C T G G A C C G G T C A C A G C C A C C C C A C C C C  Majority
      ----------------+-----------------+-----------------+-
                      1330              1340              1350
      ----------------+-----------------+-----------------+-
  942 A G C C T G G A C C G G C C A C A G C C A C C T C A C C C C  Mus musculus erg.seq
  867 A G C C T G G A C C G G T C A C G G C C A C C C C A C G C C  Homo sapiens Erg-1.seq
  909 A G C C T G G A C C G G T C A C G G C C A C C C C A C G C C  Homo sapiens Erg-2.seq
  764 A A C C T G G A C C G G C C A C A G C C A C C C C A C C C C  Rattus norvegicus
Erg.seq
  836 A G C G T G G A C G A G T C A C A G C C A T C C C A - - - C  Gallus gallus Erg.seq
 1121 T G C T T G G A C A A A T C A T C C T G C A C C A C - - - -  Xenopus laevis erg.seq
  788 T G G A T G G A C - - - T C C T G C T C T T C C T G - - - -  Danio rerio erg.seq
 1312 T A T C C A A G G G G T T T A C A C A A A C A T C G C C T A  S. purpuratus Erg.seq

      T C A G T C C A A A G C T G C T C A A C C A T - C T C C C T  Majority
      ----------------+-----------------+-----------------+-
                      1360              1370              1380
      ----------------+-----------------+-----------------+-
  972 T C A G T C C A A A G C T G C T C A G C C A T - C T C C C T  Mus musculus erg.seq
  897 C C A G T C G A A A G C T G C T C A A C C A T - C T C C T T  Homo sapiens Erg-1.seq
  939 C C A G T C G A A A G C T G C T C A A C C A T - C T C C T T  Homo sapiens Erg-2.seq
  794 T C A G T C C A A A G C T G C C C A G C C A T - C T C C C T  Rattus norvegicus
Erg.seq
  863 T C A G T C A A A A G C T A C C C A A C C A T - C A T C T T  Gallus gallus Erg.seq
 1147 - C A T - C A A A A G C A T C T C A A C C A T - C T A C C A  Xenopus laevis erg.seq
  811 - C A G - C T A A A G T T T C T C A G A G C T - C T G C C G  Danio rerio erg.seq
 1342 C C T T G C C C A A G C C T G G T A T C G A T T C C T C C G  S. purpuratus Erg.seq

      C C A C A G T G C C C A A A A C T G A A G A - - - - - - - -  Majority
      ----------------+-----------------+-----------------+-
                      1390              1400              1410
      ----------------+-----------------+-----------------+-
 1001 C T G C A G T G C C C A A A A C T G A A G A - - - - - - - -  Mus musculus erg.seq
  926 C C A C A G T G C C C A A A A C T G A A G A - - - - - - - -  Homo sapiens Erg-1.seq
  968 C C A C A G T G C C C A A A A C T G A A G A - - - - - - - -  Homo sapiens Erg-2.seq
  823 C C A C A G T G C C C A A A A C T G A A G A - - - - - - - -  Rattus norvegicus
Erg.seq
  892 C A A C A G T G C C C A A A A C A G A A G A - - - - - - - -  Gallus gallus Erg.seq
 1174 C A G T T C C - - - C A A A A C A G A A G A - - - - - - - -  Xenopus laevis erg.seq
  838 C C A T C A T G A G C A A A A C T G A A G A - - - - - - - -  Danio rerio erg.seq
 1372 C T A C A C A A A T T A G A C C C G C A G G G A G G C C G G  S. purpuratus Erg.seq

      C C A G C G T C C T C A G T T A - - - - G A T C C T T A T C  Majority
      ----------------+-----------------+-----------------+-
                      1420              1430              1440
      ----------------+-----------------+-----------------+-
 1023 C C A G C G T C C T C A G T T A - - - - G A T C C T T A C C  Mus musculus erg.seq
  948 C C A G C G T C C T C A G T T A - - - - G A T C C T T A T C  Homo sapiens Erg-1.seq
  990 C C A G C G T C C T C A G T T A - - - - G A T C C T T A T C  Homo sapiens Erg-2.seq
  845 C C A G C G T C C T C A G T T A - - - - G A T C C T T A C C  Rattus norvegicus
Erg.seq
  914 C C A G C G T C C T C A G T T A - - - - G A T C C T T A T C  Gallus gallus Erg.seq
 1193 C C C C A G G C C A C A G C T A - - - - G A T C C T T A T C  Xenopus laevis erg.seq
  860 G C A G A G A G C G C A G T T A - - - - G A T C C G T A T C  Danio rerio erg.seq
 1402 C C T A T G G G G G C A G T G A T T T C G A T C C G T A C C  S. purpuratus Erg.seq
```

Figure 8 continued

```
     A G A T T C T T G G A C C G A C C A G T A G C C G C C T T G  Majority
     ------------------+-------------------+-------------------+-
                      1450                1460                1470
     ------------------+-------------------+-------------------+-
1049 A G A T C C T G G G A C C G A C C A G T A G C C G C C T T G  Mus musculus erg.seq
 974 A G A T T C T T G G A C C A A C A A G T A G C C G C C T T G  Homo sapiens Erg-1.seq
1016 A G A T T C T T G G A C C A A C A A G T A G C C G C C T T G  Homo sapiens Erg-2.seq
 871 A G A T C C T G G G A C C C A C C A G T A G T C G C C T T G  Rattus norvegicus Erg.seq
 940 A G A T T C T T G G A C C G A C C A G C A G C C G T C T T G  Gallus gallus Erg.seq
1219 A G A T T C T T G G G C C A A C A A G C A G C C G A C T T G  Xenopus laevis erg.seq
 886 A G A T T T T A G G G C C G A C C A G C A G C A G A C T G G  Danio rerio erg.seq
1432 A A G T C T T C G G G C A C A C A A G C A G G A C A C A T G  S. purpuratus Erg.seq

     C A A A T C C - - - - - - - - - - - - - - - - - - - - - - -  Majority
     ------------------+-------------------+-------------------+-
                      1480                1490                1500
     ------------------+-------------------+-------------------+-
1079 C T A A T C C - - - - - - - - - - - - - - - - - - - - - - -  Mus musculus erg.seq
1004 C A A A T C C - - - - - - - - - - - - - - - - - - - - - - -  Homo sapiens Erg-1.seq
1046 C A A A T C C - - - - - - - - - - - - - - - - - - - - - - -  Homo sapiens Erg-2.seq
 901 C T A A T C C - - - - - - - - - - - - - - - - - - - - - - -  Rattus norvegicus Erg.seq
 970 C A A A T C C - - - - - - - - - - - - - - - - - - - - - - -  Gallus gallus Erg.seq
1249 C A A A T C C - - - - - - - - - - - - - - - - - - - - - - -  Xenopus laevis erg.seq
 916 C C A A C C C - - - - - - - - - - - - - - - - - - - - - - -  Danio rerio erg.seq
1462 C A A A T C C T G T A A T T C C A A G T G A T T G G C A A A  S. purpuratus Erg.seq

     - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  Majority
     ------------------+-------------------+-------------------+-
                      1510                1520                1530
     ------------------+-------------------+-------------------+-
1086 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  Mus musculus erg.seq
1011 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  Homo sapiens Erg-1.seq
1053 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  Homo sapiens Erg-2.seq
 908 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  Rattus norvegicus Erg.seq
 977 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  Gallus gallus Erg.seq
1256 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  Xenopus laevis erg.seq
 923 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  Danio rerio erg.seq
1492 A T C T A C A A T C C G T G A T A T T A G C A C T C A G A C  S. purpuratus Erg.seq

     - - - - - - - A G G C A G T G G C C A G A T C C A G C T G T  Majority
     ------------------+-------------------+-------------------+-
                      1540                1550                1560
     ------------------+-------------------+-------------------+-
1086 - - - - - - - A G G T A G T G G C C A G A T C C A G C T G T  Mus musculus erg.seq
1011 - - - - - - - A G G C A G T G G C C A G A T C C A G C T T T  Homo sapiens Erg-1.seq
1053 - - - - - - - A G G C A G T G G C C A G A T C C A G C T T T  Homo sapiens Erg-2.seq
 908 - - - - - - - A G G T A G T G G C C A G A T C C A G C T G T  Rattus norvegicus Erg.seq
 977 - - - - - - - A G G G A G T G G G C A G A T A C A G C T A T  Gallus gallus Erg.seq
1256 - - - - - - - A G G A A G T G G G C A G A T T C A G C T C T  Xenopus laevis erg.seq
 923 - - - - - - - A G G C T C T G G T C A G A T C C A G C T G T  Danio rerio erg.seq
1522 A T A A C A A A G G A A G T G G C C A G A T A C A A C T G T  S. purpuratus Erg.seq
```

Figure 8 continued

```
      G G C A G T T C C T G C T G G A G C T C C T G T C G G A C A  Majority
      ----------------+----------------+----------------+-
                      1570             1580             1590
      ----------------+----------------+----------------+-
 1109 G G C A G T T C C T G C T C G A A C T C C T G T C A G A C A  Mus musculus erg.seq
 1034 G G C A G T T C C T C C T G G A G C T C C T G T C G G A C A  Homo sapiens Erg-1.seq
 1076 G G C A G T T C C T C C T G G A G C T C C T G T C G G A C A  Homo sapiens Erg-2.seq
  931 G G C A G T T C C T G C T A G A A C T C C T G T C T G A C A  Rattus norvegicus
Erg.seq
 1000 G G C A G T T C C T A C T G G A G C T T C T G T C G G A C A  Gallus gallus Erg.seq
 1279 G G C A A T T C T T A C T G G A A C T G C T T T C G G A T A  Xenopus laevis erg.seq
  946 G G C A G T T C C T G C T G G A G C T C C T G T C C G A C A  Danio rerio erg.seq
 1552 G G C A G T T C C T T C T G G A G C T C C T A T C G G A C A  S. purpuratus Erg.seq

      G C T C C A A C T C C A A C T G C A T C A C C T G G G A A G  Majority
      ----------------+----------------+----------------+-
                      1600             1610             1620
      ----------------+----------------+----------------+-
 1139 G C T C C A A C T C C A A C T G C A T C A C C T G G G A A G  Mus musculus erg.seq
 1064 G C T C C A A C T C C A G C T G C A T C A C C T G G G A A G  Homo sapiens Erg-1.seq
 1106 G C T C C A A C T C C A G C T G C A T C A C C T G G G A A G  Homo sapiens Erg-2.seq
  961 G C T C C A A C T C C A A C T G C A T C A C C T G G G A A G  Rattus norvegicus
Erg.seq
 1030 G C T C C A A C T C C A A C T G C A T C A C C T G G G A G G  Gallus gallus Erg.seq
 1309 G T T C C A A C T C C A A C T G C A T T A C T T G G G A A G  Xenopus laevis erg.seq
  976 G C T G C A A C T C G T C C T G C A T C A C T T G G G A G G  Danio rerio erg.seq
 1582 G T T C C A A C G C C A A C T G C A T C A C T T G G G A G G  S. purpuratus Erg.seq

      G C A C C A A C G G G G A G T T C A A G A T G A C G G A C C  Majority
      ----------------+----------------+----------------+-
                      1630             1640             1650
      ----------------+----------------+----------------+-
 1169 G C A C C A A C G G G G A G T T C A A G A T G A C A G A C C  Mus musculus erg.seq
 1094 G C A C C A A C G G G G A G T T C A A G A T G A C G G A T C  Homo sapiens Erg-1.seq
 1136 G C A C C A A C G G G G A G T T C A A G A T G A C G G A T C  Homo sapiens Erg-2.seq
  991 G C A C C A A C G G G G A G T T C A A G A T G A C A G A C C  Rattus norvegicus
Erg.seq
 1060 G C A C A A A T G G G G A G T T C A A G A T G A C A G A C C  Gallus gallus Erg.seq
 1339 G A A C C A A T G G A G A A T T T A A G A T G A C C G A T C  Xenopus laevis erg.seq
 1006 G A A C C A A C G G C G A G T T C A A G A T G A C G G A C C  Danio rerio erg.seq
 1612 G A A C C A A T G G C G A G T T C A A G A T G A C C G A C C  S. purpuratus Erg.seq

      C G G A T G A G G T G G C T C G G C G C T G G G G A G A G C  Majority
      ----------------+----------------+----------------+-
                      1660             1670             1680
      ----------------+----------------+----------------+-
 1199 C G G A C G A G G T G G C T C G G C G C T G G G G G G A G A  Mus musculus erg.seq
 1124 C C G A C G A G G T G G C C C G G C G C T G G G G A G A G C  Homo sapiens Erg-1.seq
 1166 C C G A C G A G G T G G C C C G G C G C T G G G G A G A G C  Homo sapiens Erg-2.seq
 1021 C G G A T G A G G T G G C T C G G C G C T G G G G G G A G C  Rattus norvegicus
Erg.seq
 1090 C T G A T G A A G T G G C T C G G C G T T G G G G A G A G A  Gallus gallus Erg.seq
 1369 C C G A T G A A G T G G C A A G A C G T T G G G G G G A G A  Xenopus laevis erg.seq
 1036 C G G A C G A G G T G G C G C G A C G C T G G G G C C G A G C  Danio rerio erg.seq
 1642 C G G A T G A A G T C G C C A G G A G A T G G G G A G A G C  S. purpuratus Erg.seq
```

Figure 8 continued

```
      G G A A G A G C A A A C C C A A C A T G A A C T A T G A C A  Majority
      ----------------+----------------+----------------+-
                      1690             1700             1710
      ----------------+----------------+----------------+-
 1229 G G A A G A G C A A G C C C A A C A T G A A C T A T G A C A  Mus musculus erg.seq
 1154 G G A A G A G C A A A C C C A A C A T G A A C T A C G A T A  Homo sapiens Erg-1.seq
 1196 G G A A G A G C A A A C C C A A C A T G A A C T A C G A T A  Homo sapiens Erg-2.seq
 1051 G G A A G A G C A A G C C C A A C A T G A A C T A T G A C A  Rattus norvegicus
Erg.seq
 1120 G G A A A A G C A A A C C T A A C A T G A A C T A T G A C A  Gallus gallus Erg.seq
 1399 G G A A A A G C A A A C C C A A C A T G A A C T A T G A C A  Xenopus laevis erg.seq
 1066 G G A A G A G C A A G C C C A A C A T G A A C T A C G A C A  Danio rerio erg.seq
 1672 G G A A A A G C A A A C C C A A C A T G A A C T A C G A C A  S. purpuratus Erg.seq

      A G C T C A G C C G T G C C C T C C G C T A C T A C T A T G  Majority
      ----------------+----------------+----------------+-
                      1720             1730             1740
      ----------------+----------------+----------------+-
 1259 A G C T C A G C C G C G C C C T C C G C T A C T A C T A C G  Mus musculus erg.seq
 1184 A G C T C A G C C G C G C C C T C C G T T A C T A C T A T G  Homo sapiens Erg-1.seq
 1226 A G C T C A G C C G C G C C C T C C G T T A C T A C T A T G  Homo sapiens Erg-2.seq
 1081 A A C T C A G C C G T G C C C T C C G C T A C T A C T A C G  Rattus norvegicus
Erg.seq
 1150 A A C T C A G C C G T G C A C T T C G C T A C T A C T A T G  Gallus gallus Erg.seq
 1429 A G C T C A G C C G T G C A C T T C G T T A C T A C T A C G  Xenopus laevis erg.seq
 1096 A G C T G A G T C G A G C C C T G C G C T A C T A C T A T G  Danio rerio erg.seq
 1702 A G C T C A G C C G G G C T C T G C G C T A C T A C T A C G  S. purpuratus Erg.seq

      A C A A G A A C A T C A T G A C C A A G G T G C A T G G G A  Majority
      ----------------+----------------+----------------+-
                      1750             1760             1770
      ----------------+----------------+----------------+-
 1289 A C A A A A A C A T C A T G A C C A A G G T G C A C G G G A  Mus musculus erg.seq
 1214 A C A A G A A C A T C A T G A C C A A G G T C C A T G G G A  Homo sapiens Erg-1.seq
 1256 A C A A G A A C A T C A T G A C C A A G G T C C A T G G G A  Homo sapiens Erg-2.seq
 1111 A C A A A A A C A T C A T G A C C A A G G T G C A C G G G A  Rattus norvegicus
Erg.seq
 1180 A C A A A A A T A T T A T G A C T A A A G T T C A T G G T A  Gallus gallus Erg.seq
 1459 A T A A A A A T A T T A T G A C T A A A G T C C A T G G C A  Xenopus laevis erg.seq
 1126 A C A A G A A C A T C A T G A C C A A A G T G C A C G G C A  Danio rerio erg.seq
 1732 A C A A G A A C A T C A T G A C C A A G G T T C A T G G C A  S. purpuratus Erg.seq

      A G C G C T A C G C C T A C A A G T T T G A C T T C C A C G  Majority
      ----------------+----------------+----------------+-
                      1780             1790             1800
      ----------------+----------------+----------------+-
 1319 A G C G C T A C G C C T A C A A G T T T G A C T T C C A C G  Mus musculus erg.seq
 1244 A G C G C T A C G C C T A C A A G T T C G A C T T C C A C G  Homo sapiens Erg-1.seq
 1286 A G C G C T A C G C C T A C A A G T T C G A C T T C C A C G  Homo sapiens Erg-2.seq
 1141 A G C G C T A T G C C T A C A A G T T T G A C T T C C A C G  Rattus norvegicus
Erg.seq
 1210 A A C G C T A T G C C T A C A A A T T T G A T T T C C A C G  Gallus gallus Erg.seq
 1489 A G C G C T A T G C T T A C A A A T T T G A T T T T C A T G  Xenopus laevis erg.seq
 1156 A G C G C T A C G C C T A C A A G T T C G A C T T C C A T G  Danio rerio erg.seq
 1762 A G A G A T A C G C G T A C A A G T T C G A C T T C G C C G  S. purpuratus Erg.seq
```

Figure 8 continued

```
      G G A T T G C C C A G G C C C T G C A G C C C C A C C C T C  Majority
      -----------------+-----------------+-----------------+-
                       1810              1820              1830
      -----------------+-----------------+-----------------+-
 1349  G G A T T G C C C A G G C C C T G C A G C C C C A C C C T C  Mus musculus erg.seq
 1274  G G A T C G C C C A G G C C C T C C A G C C C C A C C C C C  Homo sapiens Erg-1.seq
 1316  G G A T C G C C C A G G C C C T C C A G C C C C A C C C C C  Homo sapiens Erg-2.seq
 1171   G G A T T G C C C A G G C C C T G C A G C C C C A T C C C C   Rattus norvegicus
Erg.seq
 1240  G A A T C G C T C A G G C C C T C C A G C C T C A C C C T C  Gallus gallus Erg.seq
 1519  G G A T T G C T C A G G C T C T T C A A C C T C A T C C T C  Xenopus laevis erg.seq
 1186  G C A T T G C G C A A G C C T T G C A G C C G C A C C C G C  Danio rerio erg.seq
 1792  G G T T A G C C C A A G C C A T G C A G C C G G T C C A A G  S. purpuratus Erg.seq

      C T G A G T C A T C C C T G T A C A A G T A C C C C T C A G  Majority
      -----------------+-----------------+-----------------+-
                       1840              1850              1860
      -----------------+-----------------+-----------------+-
 1379  C T G A G T C G T C C C T G T A C A A G T A C C C C T C C G  Mus musculus erg.seq
 1304  C G G A G T C A T C T C T G T A C A A G T A C C C C T C A G  Homo sapiens Erg-1.seq
 1346  C G G A G T C A T C T C T G T A C A A G T A C C C C T C A G  Homo sapiens Erg-2.seq
 1201   C C G A G T C G T C C C T G T A C A A G T A C C C C T C C G   Rattus norvegicus
Erg.seq
 1270  C A G A A T C A T C C A T G T A C A A A T A C C C A T C A G  Gallus gallus Erg.seq
 1549  C A G A A T C A A C C A T G T A C A A G T A T C C C T C A G  Xenopus laevis erg.seq
 1216  C C G A C T C C T C C A T C T A C A A G T A C C C T G C G G  Danio rerio erg.seq
 1822  C T G A T C C T A G C A T G T A C C G C T A C C A G T C A G  S. purpuratus Erg.seq

      A C C T C C C G T A C A T G G G C T - - - - C C T A T C A C  Majority
      -----------------+-----------------+-----------------+-
                       1870              1880              1890
      -----------------+-----------------+-----------------+-
 1409  A C C T G C C A T A C A T G G G C T - - - - C C T A T C A C  Mus musculus erg.seq
 1334  A C C T C C C G T A C A T G G G C T - - - - C C T A T C A C  Homo sapiens Erg-1.seq
 1376  A C C T C C C G T A C A T G G G C T - - - - C C T A T C A C  Homo sapiens Erg-2.seq
 1231   A C C T G C C G T A C A T G G G C T - - - - C C T A T C A C   Rattus norvegicus
Erg.seq
 1300  A C C T C C C C T A C A T G A G T T - - - - C C T A C C A T  Gallus gallus Erg.seq
 1579  A A C T T C C A T A T A T G A G C T - - - - C A T A C C A T  Xenopus laevis erg.seq
 1246  T G G C G G C G G A T C T G C C C T A C G T C A G C T C C T  Danio rerio erg.seq
 1852  A C C T C T C G T A C C T G C A G A - - - - - - G C T A T C  S. purpuratus Erg.seq

      G C C C A C C C C C A G A A G A T G A A C T T T G T G G C T  Majority
      -----------------+-----------------+-----------------+-
                       1900              1910              1920
      -----------------+-----------------+-----------------+-
 1435  G C C C A C C C C C A G A A G A T G A A C T T T G T G T C T  Mus musculus erg.seq
 1360  G C C C A C C C A C A G A A G A T G A A C T T T G T G G C G  Homo sapiens Erg-1.seq
 1402  G C C C A C C C A C A G A A G A T G A A C T T T G T G G C G  Homo sapiens Erg-2.seq
 1257   A C C C A C C C C C C A G A A G A T G A A C T T T G T G G C T   Rattus norvegicus
Erg.seq
 1326  G C A C A C C C C C C A G A A G A T G A A C T T T G T A G C T  Gallus gallus Erg.seq
 1605  G C A C A T C C A C A G A A G A T G A A C T T T G T T G C C  Xenopus laevis erg.seq
 1276  A C - C A C C C G C A G A A G A T G A G C T T C A T G G C T  Danio rerio erg.seq
 1876  A C - C A C C C G A C C A A G C T G A A C T T T G T G G - -  S. purpuratus Erg.seq
```

Figure 8 continued

```
      C C C C A T C C T C C G G C C C T C C C C G T C A C A T C T  Majority
      -----------------+-----------------+-----------------+-
                       1930              1940              1950
      -----------------+-----------------+-----------------+-
 1465 C C C C A C C C T C C C G C T C T C C C A G T C A C A T C T  Mus musculus erg.seq
 1390 C C C C A C C C T C C A G C C C T C C C C C G T G A C A T C T  Homo sapiens Erg-1.seq
 1432 C C C C A C C C T C C A G C C C T C C C C C G T G A C A T C T  Homo sapiens Erg-2.seq
 1287  C C C C A C C C T C C G G C C C T C C C A G T C A C A T C T  Rattus norvegicus
Erg.seq
 1356 C C C C A T C C C C C T G C T T T G C C C G T A A C C T C A  Gallus gallus Erg.seq
 1635 C C C C A T C C G C C A G C T T T A C C T G T G A C A T C G  Xenopus laevis erg.seq
 1305 C C T C A T C C T C A G G C C A T G A G C G T C T C C T C C  Danio rerio erg.seq
 1903 - - - - G T A C C C C G A T C A A C C C G T C C A C C A A T  S. purpuratus Erg.seq

      T C C A G T T T C T T T G C T G C C C C G A A T C C A T A C  Majority
      -----------------+-----------------+-----------------+-
                       1960              1970              1980
      -----------------+-----------------+-----------------+-
 1495 T C C A G T T T C T T T G C T T C C C C G A A C C C A T A C  Mus musculus erg.seq
 1420 T C C A G T T T T T T T G C T G C C C C A A A C C C A T A C  Homo sapiens Erg-1.seq
 1462 T C C A G T T T T T T T G C T G C C C C A A A C C C A T A C  Homo sapiens Erg-2.seq
 1317  T C C A G T T T C T T T G C T A C C C C G A A C C C A T A C  Rattus norvegicus
Erg.seq
 1386 T C C A G C T T T T T T G C T G C C C C T A A T C C A T A C  Gallus gallus Erg.seq
 1665 T C A A G T T T C T T T G C A G C C C C T A A T G C A T A C  Xenopus laevis erg.seq
 1335 T C C G G C T T C T T C A G C G G G C C G G C T C C G T A C  Danio rerio erg.seq
 1929 G C C A G C C T C T T T A G C T C T C A C A G T T C A T A C  S. purpuratus Erg.seq

      T G G A A T T C A C C A A C T G G G G - - - G T A T C T A C  Majority
      -----------------+-----------------+-----------------+-
                       1990              2000              2010
      -----------------+-----------------+-----------------+-
 1525 T G G A A T T C A C C G A C T G G G G - - - G C A T C T A C  Mus musculus erg.seq
 1450 T G G A A T T C A C C A A C T G G G G - - - G T A T A T A C  Homo sapiens Erg-1.seq
 1492 T G G A A T T C A C C A A C T G G G G - - - G T A T A T A C  Homo sapiens Erg-2.seq
 1347  T G G A A T T C G C C G A C T G G G G - - - G C A T C T A C  Rattus norvegicus
Erg.seq
 1416 T G G A A T T C A C C A A C T G G A G - - - G C A T C T A C  Gallus gallus Erg.seq
 1695 T G G A A T T C A C C A A C T G G A A - - - G T A T T T A T  Xenopus laevis erg.seq
 1365 T G G A A C A C A C C A G C G G C A G - - - G G A T C T A C  Danio rerio erg.seq
 1959 T G G T C A T C T C C A A C A G G C G C C A A T A T C T A C  S. purpuratus Erg.seq

      C C C A A C - - - - - - - - - - A C T A G G C T C C C C G C C  Majority
      -----------------+-----------------+-----------------+-
                       2020              2030              2040
      -----------------+-----------------+-----------------+-
 1552 C C G A A C - - - - - - - - - - A C T A G G C T C C C A G C C  Mus musculus erg.seq
 1477 C C C A A C - - - - - - - - - - A C T A G G C T C C C C A C C  Homo sapiens Erg-1.seq
 1519 C C C A A C - - - - - - - - - - A C T A G G C T C C C C A C C  Homo sapiens Erg-2.seq
 1374  C C G A A C - - - - - - - - - - A C T A G G C T C C C A G C C  Rattus norvegicus
Erg.seq
 1443 C C C A A T - - - - - - - - - - A C C A G G C T G C C A G C T  Gallus gallus Erg.seq
 1722 C C A A A T - - - - - - - - - - A C T C G G C T G C C A G C T  Xenopus laevis erg.seq
 1392 G G C G G G - - - - - - - - - - C C A C G G C A C C C C C C T  Danio rerio erg.seq
 1989 C C C A G C G G T C A C G T G A C A C A T C C T C A C C C G  S. purpuratus Erg.seq
```

Figure 8 continued

```
      A G C C A T A T G C C T T C T C A T C T G G G C A C C T A C  Majority
      ------------------+------------------+------------------+-
                        2050               2060               2070
      ------------------+------------------+------------------+-
 1573 A G C C A T A T G C C C T C T C A C C T G G G C A C C T A C  Mus musculus erg.seq
 1498 A G C C A T A T G C C T T C T C A T C T G G G C A C T T A C  Homo sapiens Erg-1.seq
 1540 A G C C A T A T G C C T T C T C A T C T G G G C A C T T A C  Homo sapiens Erg-2.seq
 1395 A G C C A T A T G C C C T C T C A C C T G G G C A C C T A C  Rattus norvegicus
Erg.seq
 1464 G C T C A T A T G C C T T C C C A T C T T G G C A C C T A C  Gallus gallus Erg.seq
 1743 A G C C A T A T G T C T T C T C A T C T T G G A A C C T A C  Xenopus laevis erg.seq
 1413 G C T C A C A T G C C C T C A C A T C T G G G C T C A T A C  Danio rerio erg.seq
 2019 G G T C A C A T G T C C T C G C A C A T A G G C A C A T A C  S. purpuratus Erg.seq

      T A C T A G A G A C C T G G C G G A G G C C T T T C C C A A  Majority
      ------------------+------------------+------------------+-
                        2080               2090               2100
      ------------------+------------------+------------------+-
 1603 T A C T A G A G A C C A G A T G G A G G C C T T T C C C A A  Mus musculus erg.seq
 1528 T A C T A A A G A C C T G G C G G A G G C T T T T C C C A T  Homo sapiens Erg-1.seq
 1570 T A C T A A A G A C C T G G C G G A G G C T T T T C C C A T  Homo sapiens Erg-2.seq
 1425 T A C T A G A G A C C A G A T A G A G G C C T T T C C C A A  Rattus norvegicus
Erg.seq
 1494 T A C T A A G T G G G G A A A G A A A G A A A               Gallus gallus Erg.seq
 1773 T A T T A A A T C A A G A T C T C A G A C A C G T A         Xenopus laevis erg.seq
 1443 T A C T A G A A G A A C A G C A G G A A C A T C T C A A G A  Danio rerio erg.seq
 2049 T A T - - G G A T A A C A C C A A G T G C C C T T T A C - -  S. purpuratus Erg.seq

      C A G X G T G C A A T C G C C X G X C X A T C G C C A C A A  Majority
      ------------------+------------------+------------------+-
                        2110               2120               2130
      ------------------+------------------+------------------+-
 1633 C A G T G T A C A A T C G C C T A G C A G T C G C C A C A A  Mus musculus erg.seq
 1558 C A G C G T G C A T T C A C C A G C C C A T C G C C A C A A  Homo sapiens Erg-1.seq
 1600 C A G C G T G C A T T C A C C A G C C C A T C G C C A C A A  Homo sapiens Erg-2.seq
 1455 C A G T G C A C A A T C G C C T A G C A A T C A C C A C A G  Rattus norvegicus
Erg.seq
 1516                                                           Gallus gallus Erg.seq
 1798                                                           Xenopus laevis erg.seq
 1473 C C A - G C A G A G A C A C - - - - - - - T C A C A T A A T  Danio rerio erg.seq
 2075 C C G - T T G C A A T T G C C G G A C T - C T A T A A C A G  S. purpuratus Erg.seq

      A C T C T - T - G G A G A A C A C X A A T C A A A A G T G C  Majority
      ------------------+------------------+------------------+-
                        2140               2150               2160
      ------------------+------------------+------------------+-
 1663 A C T C - - - - A G C G A A C A C A A A T C A A A A G T G C  Mus musculus erg.seq
 1588 A C T C T A T C G G A G A A C A T G A A T C A A A A G T G C  Homo sapiens Erg-1.seq
 1630 A C T C T A T C G G A G A A C A T G A A T C A A A A G T G C  Homo sapiens Erg-2.seq
 1485 A C T C - - - - A G A G A A C A C A A A T C G A A A G T G C  Rattus norvegicus
Erg.seq
 1516                                                           Gallus gallus Erg.seq
 1798                                                           Xenopus laevis erg.seq
 1495 A G T T - - - - G G A G A A C A G - - - - C A G G A A C A T  Danio rerio erg.seq
 2103 A C T C T - T T A T G A A G T G C - - - - C A T C A C T T T  S. purpuratus Erg.seq
```

Figure 8 continued

```
      C T C A T G A G G A A T G G A A A X X G X T T X X C X X G X  Majority
      -----------------+-----------------+-----------------+-
                       2170              2180              2190
      -----------------+-----------------+-----------------+-
 1689  C T C G T G A G G A A T G G A A A C G G - T T G G C C A G A  Mus musculus erg.seq
 1618  C T C A A G A G G A A T G A A A A A A A G C T T T A C T G G G  Homo sapiens Erg-1.seq
 1660  C T C A A G A G G A A T G A A A A A A A G C T T T A C T G G G  Homo sapiens Erg-2.seq
 1511  C T C C T G A G G A A T G G A A A C G G - T T G G C C A G A  Rattus norvegicus
Erg.seq
 1516                                                             Gallus gallus Erg.seq
 1798                                                             Xenopus laevis erg.seq
 1517  C T C A A G A C C A G C G G A G A T G C - C C A G C A G A G  Danio rerio erg.seq
 2128  T T T G T A T T G T A T G A A                               S. purpuratus Erg.seq

      G C T G G G G A A G G A A G C X G G G G X - - - X A A G A X  Majority
      -----------------+-----------------+-----------------+-
                       2200              2210              2220
      -----------------+-----------------+-----------------+-
 1718  G C T G G G G A A G G A A G C T G G G G G G G A A A A A A A  Mus musculus erg.seq
 1648  G C T G G G G A A G G A A G C C G G G G - - - - - A A G A G  Homo sapiens Erg-1.seq
 1690  G C T G G G G A A G G A A G C C G G G G - - - - - A A G A G  Homo sapiens Erg-2.seq
 1540  G C T G G G G A A G G A A G C C G G G G - - - - - A A A G A A  Rattus norvegicus
Erg.seq
 1516                                                             Gallus gallus Erg.seq
 1798                                                             Xenopus laevis erg.seq
 1546  G C A G G A C A G G G G A T G T A A A G G - - - G C A G A T  Danio rerio erg.seq
 2142                                                             S. purpuratus Erg.seq

      A X C C A A A G A C T C X X X G G X G X G X G T X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       2230              2240              2250
      -----------------+-----------------+-----------------+-
 1748  A A C C A A A G A C T C G C A G G G G A G G G T G G G T G C  Mus musculus erg.seq
 1673  A T C C A A A G A C T C T T G G G A G G G A G T - - - - - -  Homo sapiens Erg-1.seq
 1715  A T C C A A A G A C T C T T G G G A G G G A G T - - - - - -  Homo sapiens Erg-2.seq
 1566  A A C C A A A G A C T C A C A G G G G A G G G T G G G C A C  Rattus norvegicus
Erg.seq
 1516                                                             Gallus gallus Erg.seq
 1798                                                             Xenopus laevis erg.seq
 1573  A T C A G G C C G C T G T T A G A G A C C T G C T G C A - -  Danio rerio erg.seq
 2142                                                             S. purpuratus Erg.seq

      X A C X X A A G T X T T A C T A C A G A A A T X X X G A X G  Majority
      -----------------+-----------------+-----------------+-
                       2260              2270              2280
      -----------------+-----------------+-----------------+-
 1778  C A C C C A A G T A T T A C T A C A G A A A T A - - G A A G  Mus musculus erg.seq
 1697  T A C T G A A G T C T T A C T A C A G A A A T G A G G A G G  Homo sapiens Erg-1.seq
 1739  T A C T G A A G T C T T A C T A C A G A A A T G A G G A G G  Homo sapiens Erg-2.seq
 1596  C A C C C A A G T A T T A C T A C A G A A A T A - - G A A G  Rattus norvegicus
Erg.seq
 1516                                                             Gallus gallus Erg.seq
 1798                                                             Xenopus laevis erg.seq
 1601  T G A T C A A T A C T T A T T C A T C T A T A A T G T T A T  Danio rerio erg.seq
 2142                                                             S. purpuratus Erg.seq
```

Figure 8 continued

```
      A X G C T X A A A A T X T X X X G X A X A T G G A C A T A T  Majority
      ----------------+-----------------+-----------------+-
                      2290              2300              2310
      ----------------+-----------------+-----------------+-
 1806 A A G C T C A A A A T A T A C T G T A C A T G G A C A T A T  Mus musculus erg.seq
 1727 A T G C T A A A A A T G T C A C G A A T A T G G A C A T A T  Homo sapiens Erg-1.seq
 1769 A T G C T A A A A A T G T C A C G A A T A T G G A C A T A T  Homo sapiens Erg-2.seq
 1624 A A G C T C A A A A T A T C C T G T A C A T G G A C A T A T  Rattus norvegicus
Erg.seq
 1516                                                            Gallus gallus Erg.seq
 1798                                                            Xenopus laevis erg.seq
 1631 C T G T A C C T T A T T G T A C A A A C A T G C A T A T C A  Danio rerio erg.seq
 2142                                                            S. purpuratus Erg.seq

      X A X C T G T G G X C X X X C C T T G T X A A A G A C A X T  Majority
      ----------------+-----------------+-----------------+-
                      2320              2330              2340
      ----------------+-----------------+-----------------+-
 1836 A A C C T G T G G T C C A C C C T T G T C A A A G A C A T T  Mus musculus erg.seq
 1757 C A T C T G T G G A C T G A C C T T G T A A A A G A C A G T  Homo sapiens Erg-1.seq
 1799 C A T C T G T G G A C T G A C C T T G T A A A A G A C A G T  Homo sapiens Erg-2.seq
 1654 A A C C T G T G G T C C A C C C T T G T C A A A G A C A T T  Rattus norvegicus
Erg.seq
 1516                                                            Gallus gallus Erg.seq
 1798                                                            Xenopus laevis erg.seq
 1661 G A G A G A G G C G G A G C G T A G G T G A G A C G G G T C  Danio rerio erg.seq
 2142                                                            S. purpuratus Erg.seq

      G T A T G T A G A A X X X X X X X X C X X X A X G A X X A  Majority
      ----------------+-----------------+-----------------+-
                      2350              2360              2370
      ----------------+-----------------+-----------------+-
 1866 G T A T G T A G A A A A G C T T G A C C A A A A A G A A C A  Mus musculus erg.seq
 1787 G T A T G T A G A A G C A T G A A G T C T T A A G G A C A A  Homo sapiens Erg-1.seq
 1829 G T A T G T A G A A G C A T G A A G T C T T A A G G A C A A  Homo sapiens Erg-2.seq
 1684 G T A T G T A G A A A A G C G T G A C C G A G A A G A A C A  Rattus norvegicus
Erg.seq
 1516                                                            Gallus gallus Erg.seq
 1798                                                            Xenopus laevis erg.seq
 1691 A G A G G A C A G A C A G A C A G G C A G A G G G A G A C G  Danio rerio erg.seq
 2142                                                            S. purpuratus Erg.seq

      A X T G X C X A A G X A A G T G G X C T T A A G X A X T G T  Majority
      ----------------+-----------------+-----------------+-
                      2380              2390              2400
      ----------------+-----------------+-----------------+-
 1896 A C T G C C C A A G G A A G T G G C C T T A A G T A G T G T  Mus musculus erg.seq
 1817 A G T G C C A A A G A A A G T G G T C T T A A G A A A T G T  Homo sapiens Erg-1.seq
 1859 A G T G C C A A A G A A A G T G G T C T T A A G A A A T G T  Homo sapiens Erg-2.seq
 1714 A C T G T C C A A G G A A G T G G C C T T A A G T A G T G T  Rattus norvegicus
Erg.seq
 1516                                                            Gallus gallus Erg.seq
 1798                                                            Xenopus laevis erg.seq
 1721 G A T G G T C A G A G G A G G G C A G A C A G A C A G A C A  Danio rerio erg.seq
 2142                                                            S. purpuratus Erg.seq
```

Figure 8 continued

```
      A X A A A C T X X X X X G T A G A G T X T G G X A X C X X A  Majority
      -----------------+-----------------+-----------------+-
                       2410              2420              2430
      -----------------+-----------------+-----------------+-
 1926 A C A A A C T - - - - - G T A G A G T G T G G G A C C T T G  Mus musculus erg.seq
 1847 A T A A A C T T T A G A G T A G A G T T T G G A A T C C C A  Homo sapiens Erg-1.seq
 1889 A T A A A C T T T A G A G T A G A G T T T G G A A T C C C A  Homo sapiens Erg-2.seq
 1744  A C A A A C T - - - - - G T A G A G T G T G G G A A A T A A  Rattus norvegicus
Erg.seq
 1516                                                            Gallus gallus Erg.seq
 1798                                                            Xenopus laevis erg.seq
 1751 A T G G A G G T C A A A G A A G A T C A G A A G A T C A C A  Danio rerio erg.seq
 2142                                                            S. purpuratus Erg.seq

      C T X X X X X X X A C X X G G A T X X A X X X A A A - - - G  Majority
      -----------------+-----------------+-----------------+-
                       2440              2450              2460
      -----------------+-----------------+-----------------+-
 1951 C T T G A C C - T A C C A G G A T T G A G T G A A A - - - G  Mus musculus erg.seq
 1877 C T A A T G C A A A C T G G G A T G A A A C T A A A - - - G  Homo sapiens Erg-1.seq
 1919 C T A A T G C A A A C T G G G A T G A A A C T A A A - - - G  Homo sapiens Erg-2.seq
 1769  C T T G - - - - T A C C A G G A T T C T G T G A G A - - - G  Rattus norvegicus
Erg.seq
 1516                                                            Gallus gallus Erg.seq
 1798                                                            Xenopus laevis erg.seq
 1781 G A C A G A T G G A C G G A C A G A C A A T G A A A T A G A  Danio rerio erg.seq
 2142                                                            S. purpuratus Erg.seq

      C A A T A G A A X C A X X A X A X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       2470              2480              2490
      -----------------+-----------------+-----------------+-
 1977 C A A T A G A A C A G G A A C A G A C T T G G C C T A A C A  Mus musculus erg.seq
 1904 C A A T A G A A A C A A C A C A G T T T T G A C C T A A C A  Homo sapiens Erg-1.seq
 1946 C A A T A G A A A C A A C A C A G T T T T G A C C T A A C A  Homo sapiens Erg-2.seq
 1792 C A A T A G A A C C A T A A A T C                            Rattus norvegicus
Erg.seq
 1516                                                            Gallus gallus Erg.seq
 1798                                                            Xenopus laevis erg.seq
 1811 C A G A A G A G G C A A G A G A C A G A C G G T C A G A C A  Danio rerio erg.seq
 2142                                                            S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       2500              2510              2520
      -----------------+-----------------+-----------------+-
 2007 C A C C A T G T A T G A T G T C A T T T G A A G G G A A A C  Mus musculus erg.seq
 1934 T A C C G T T T A T A A T G C C A T T T T A A G G A A A A C  Homo sapiens Erg-1.seq
 1976 T A C C G T T T A T A A T G C C A T T T T A A G G A A A A C  Homo sapiens Erg-2.seq
 1808                                                            Rattus norvegicus
Erg.seq
 1516                                                            Gallus gallus Erg.seq
 1798                                                            Xenopus laevis erg.seq
 1841 G A A A T A C A G T C A G A T G G A C A G A C A G T C A A A  Danio rerio erg.seq
 2142                                                            S. purpuratus Erg.seq
```

Figure 8 continued

```
      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       2530              2540              2550
      -----------------+-----------------+-----------------+-
 2037  T A C C T G T G T T T A A A A A T A G A A A C C A T G T C A  Mus musculus erg.seq
 1964  T A C C T G T A T T T A A A A A T A G A A A C - A T A T C A  Homo sapiens Erg-1.seq
 2006  T A C C T G T A T T T A A A A A T A G A A A C - A T A T C A  Homo sapiens Erg-2.seq
 1808                                                             Rattus norvegicus
Erg.seq
 1516                                                             Gallus gallus Erg.seq
 1798                                                             Xenopus laevis erg.seq
 1871  C A G G C A A T C A C A T A G A G G A G G A C A G A C G G T  Danio rerio erg.seq
 2142                                                             S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       2560              2570              2580
      -----------------+-----------------+-----------------+-
 2067  A A A C C A A A G A A G A A G C A G A C A A A A A C A G A G  Mus musculus erg.seq
 1993  A A A A C A A G A G A A A A G A C A C G A G A G A G A C T G  Homo sapiens Erg-1.seq
 2035  A A A A C A A G A G A A A A G A C A C G A G A G A G A C T G  Homo sapiens Erg-2.seq
 1808                                                             Rattus norvegicus
Erg.seq
 1516                                                             Gallus gallus Erg.seq
 1798                                                             Xenopus laevis erg.seq
 1901  C A G A T G G T C A G A C A G A C C A A G G A C A G A G A G  Danio rerio erg.seq
 2142                                                             S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       2590              2600              2610
      -----------------+-----------------+-----------------+-
 2097  A T G G C G G G C C C C T C C A C A G A C C A C A A T G G C  Mus musculus erg.seq
 2023  - T - - - - G G C C C A T C A A C A G A C G T T G A T - - -  Homo sapiens Erg-1.seq
 2065  - T - - - - G G C C C A T C A A C A G A C G T T G A T - - -  Homo sapiens Erg-2.seq
 1808                                                             Rattus norvegicus
Erg.seq
 1516                                                             Gallus gallus Erg.seq
 1798                                                             Xenopus laevis erg.seq
 1931  A T A G A T G G A C A A T C A G A C C A A A A C A C A G T C  Danio rerio erg.seq
 2142                                                             S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       2620              2630              2640
      -----------------+-----------------+-----------------+-
 2127  T G T G G A C T T T C T T G G C T G G C A T T T G C T T G T  Mus musculus erg.seq
 2045  - - - - - A T G C A A C T G C A T G G C A T G T G C T - G T  Homo sapiens Erg-1.seq
 2087  - - - - - A T G C A A C T G C A T G G C A T G T G C T - G T  Homo sapiens Erg-2.seq
 1808                                                             Rattus norvegicus
Erg.seq
 1516                                                             Gallus gallus Erg.seq
 1798                                                             Xenopus laevis erg.seq
 1961  A C A C C G A T A A A C A G G A G G A C A G A T A G A T G G  Danio rerio erg.seq
 2142                                                             S. purpuratus Erg.seq
```

Figure 8 continued

```
      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       2650              2660              2670
      -----------------+-----------------+-----------------+-
 2157 T T T C T T T G A A A T C A G A T T C G T T C C A T T T G A  Mus musculus erg.seq
 2069 T T T G G T T G A A A T C A A A T A C A T T C C G T T T G A  Homo sapiens Erg-1.seq
 2111 T T T G G T T G A A A T C A A A T A C A T T C C G T T T G A  Homo sapiens Erg-2.seq
 1808                                                            Rattus norvegicus
Erg.seq
 1516                                                            Gallus gallus Erg.seq
 1798                                                            Xenopus laevis erg.seq
 1991 A C A G A T C A T A - - C A G A C A G A T A A A G G A T G A  Danio rerio erg.seq
 2142                                                            S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       2680              2690              2700
      -----------------+-----------------+-----------------+-
 2187 T A G A G A G C T G C C A G C T C T G T C A A              Mus musculus erg.seq
 2099 T G G A C A G C T G T C A G C T T T C T C A A A C T G T G A  Homo sapiens Erg-1.seq
 2141 T G G A C A G C T G T C A G C T T T C T C A A A C T G T G A  Homo sapiens Erg-2.seq
 1808                                                            Rattus norvegicus
Erg.seq
 1516                                                            Gallus gallus Erg.seq
 1798                                                            Xenopus laevis erg.seq
 2019 C A G A C A A T C A G A G G G A C A A T C A G A A G A G G A  Danio rerio erg.seq
 2142                                                            S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       2710              2720              2730
      -----------------+-----------------+-----------------+-
 2209                                                            Mus musculus erg.seq
 2129 A G A T G A C C C A A A G T T T C C A A C T C C T T T A C A  Homo sapiens Erg-1.seq
 2171 A G A T G A C C C A A A G T T T C C A A C T C C T T T A C A  Homo sapiens Erg-2.seq
 1808                                                            Rattus norvegicus
Erg.seq
 1516                                                            Gallus gallus Erg.seq
 1798                                                            Xenopus laevis erg.seq
 2049 A A G A T A G T C A A A C G G G C A A A C A - - - - - G - -  Danio rerio erg.seq
 2142                                                            S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       2740              2750              2760
      -----------------+-----------------+-----------------+-
 2209                                                            Mus musculus erg.seq
 2159 G T A T T A C C G G G A C T A T G A A C T A A A A G G T G G  Homo sapiens Erg-1.seq
 2201 G T A T T A C C G G G A C T A T G A A C T A A A A G G T G G  Homo sapiens Erg-2.seq
 1808                                                            Rattus norvegicus
Erg.seq
 1516                                                            Gallus gallus Erg.seq
 1798                                                            Xenopus laevis erg.seq
 2072 A C A G A G C A G G G A A G A C C A A T G G A C A A G C A G  Danio rerio erg.seq
 2142                                                            S. purpuratus Erg.seq
```

Figure 8 continued

```
      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      ----------------+-----------------+-----------------+-
                      2770              2780              2790
      ----------------+-----------------+-----------------+-
 2209                                                           Mus musculus erg.seq
 2189  G A C T G A G G A T G T G T A T A G A G T G A G C G T G T G  Homo sapiens Erg-1.seq
 2231  G A C T G A G G A T G T G T A T A G A G T G A G C G T G T G  Homo sapiens Erg-2.seq
 1808                                                                 Rattus norvegicus
Erg.seq
 1516                                                           Gallus gallus Erg.seq
 1798                                                           Xenopus laevis erg.seq
 2102  A C A C A C A G A C A A T C A G A C A G A G G A C A G A C G  Danio rerio erg.seq
 2142                                                           S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      ----------------+-----------------+-----------------+-
                      2800              2810              2820
      ----------------+-----------------+-----------------+-
 2209                                                           Mus musculus erg.seq
 2219  A T T G T A G A C A G A G G G G T G A A G A A G G A G G A G  Homo sapiens Erg-1.seq
 2261  A T T G T A G A C A G A G G G G T G A A G A A G G A G G A G  Homo sapiens Erg-2.seq
 1808                                                                 Rattus norvegicus
Erg.seq
 1516                                                           Gallus gallus Erg.seq
 1798                                                           Xenopus laevis erg.seq
 2132  G T A A A G C A G A C A G T C A C A G T C A G A C A G G C A  Danio rerio erg.seq
 2142                                                           S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      ----------------+-----------------+-----------------+-
                      2830              2840              2850
      ----------------+-----------------+-----------------+-
 2209                                                           Mus musculus erg.seq
 2249  G A A G A G G C A G A G A A G G A G G A G A C C A G G C T G  Homo sapiens Erg-1.seq
 2291  G A A G A G G C A G A G A A G G A G G A G A C C A G G C T G  Homo sapiens Erg-2.seq
 1808                                                                 Rattus norvegicus
Erg.seq
 1516                                                           Gallus gallus Erg.seq
 1798                                                           Xenopus laevis erg.seq
 2162  G A C T G C A G G T T G G T G A T G T G G G T C T G C A G T  Danio rerio erg.seq
 2142                                                           S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      ----------------+-----------------+-----------------+-
                      2860              2870              2880
      ----------------+-----------------+-----------------+-
 2209                                                           Mus musculus erg.seq
 2279  G G A A A G A A A C T T C T C A A G C A A T G A A G A C T G  Homo sapiens Erg-1.seq
 2321  G G A A A G A A A C T T C T C A A G C A A T G A A G A C T G  Homo sapiens Erg-2.seq
 1808                                                                 Rattus norvegicus
Erg.seq
 1516                                                           Gallus gallus Erg.seq
 1798                                                           Xenopus laevis erg.seq
 2192  G T G T G T T A A A G C A T G A A G A A T G A A G C T C C T  Danio rerio erg.seq
 2142                                                           S. purpuratus Erg.seq
```

Figure 8 continued

```
      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       2890              2900              2910
      -----------------+-----------------+-----------------+-
 2209                                                            Mus musculus erg.seq
 2309  G A C T C A G G A C A T T T G G G G A C T G T G T A C A A T  Homo sapiens Erg-1.seq
 2351  G A C T C A G G A C A T T T G G G G A C T G T G T A C A A T  Homo sapiens Erg-2.seq
 1808                                                                  Rattus norvegicus
Erg.seq
 1516                                                            Gallus gallus Erg.seq
 1798                                                            Xenopus laevis erg.seq
 2222  G C C G T G T T T C T C A G T A T T A C T C C G T C C T C G  Danio rerio erg.seq
 2142                                                            S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       2920              2930              2940
      -----------------+-----------------+-----------------+-
 2209                                                            Mus musculus erg.seq
 2339  G A G T T A T G G A G A C T C G A G G G T T C A T G C A G T  Homo sapiens Erg-1.seq
 2381  G A G T T A T G G A G A C T C G A G G G T T C A T G C A G T  Homo sapiens Erg-2.seq
 1808                                                                  Rattus norvegicus
Erg.seq
 1516                                                            Gallus gallus Erg.seq
 1798                                                            Xenopus laevis erg.seq
 2252  G G G C A T C A G A G A G T G T T C T G T G C C T C T A A A  Danio rerio erg.seq
 2142                                                            S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       2950              2960              2970
      -----------------+-----------------+-----------------+-
 2209                                                            Mus musculus erg.seq
 2369  C A G T G T T A T A C C A A A C C C A G T G T T A G G A G A  Homo sapiens Erg-1.seq
 2411  C A G T G T T A T A C C A A A C C C A G T G T T A G G A G A  Homo sapiens Erg-2.seq
 1808                                                                  Rattus norvegicus
Erg.seq
 1516                                                            Gallus gallus Erg.seq
 1798                                                            Xenopus laevis erg.seq
 2282  C T A T G C A A T - - C A T G T A G A A A A C T G G A A G C  Danio rerio erg.seq
 2142                                                            S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       2980              2990              3000
      -----------------+-----------------+-----------------+-
 2209                                                            Mus musculus erg.seq
 2399  A A G G A C A C A G C G T A A T G G A G A A A G G G A A G T  Homo sapiens Erg-1.seq
 2441  A A G G A C A C A G C G T A A T G G A G A A A G G G A A G T  Homo sapiens Erg-2.seq
 1808                                                                  Rattus norvegicus
Erg.seq
 1516                                                            Gallus gallus Erg.seq
 1798                                                            Xenopus laevis erg.seq
 2310  T T T C A G A T G T T T T T C T T G A C C T C A G A T C T T  Danio rerio erg.seq
 2142                                                            S. purpuratus Erg.seq
```

Figure 8 continued

```
      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       3010              3020              3030
      -----------------+-----------------+-----------------+-
 2209                                                           Mus musculus erg.seq
 2429  A G T A G A A T T C A G A A A C A A A A A T G C G C A T C T  Homo sapiens Erg-1.seq
 2471  A G T A G A A T T C A G A A A C A A A A A T G C G C A T C T  Homo sapiens Erg-2.seq
 1808                                                                 Rattus norvegicus
Erg.seq
 1516                                                           Gallus gallus Erg.seq
 1798                                                           Xenopus laevis erg.seq
 2340  T C C A G A C T T A A G G T G C G T T T A T A T G A A G C C  Danio rerio erg.seq
 2142                                                           S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       3040              3050              3060
      -----------------+-----------------+-----------------+-
 2209                                                           Mus musculus erg.seq
 2459  C T T T C T T T G T T T G - - - - T C A A A T G A A A A T T  Homo sapiens Erg-1.seq
 2501  C T T T C T T T G T T T G - - - - T C A A A T G A A A A T T  Homo sapiens Erg-2.seq
 1808                                                                 Rattus norvegicus
Erg.seq
 1516                                                           Gallus gallus Erg.seq
 1798                                                           Xenopus laevis erg.seq
 2370  G T G C C T T A C A G A C A G A C T C A G C G G A A G A C A  Danio rerio erg.seq
 2142                                                           S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       3070              3080              3090
      -----------------+-----------------+-----------------+-
 2209                                                           Mus musculus erg.seq
 2485  T T A A C T G G A A T T G T C T G A T A T T T A A G A G A A  Homo sapiens Erg-1.seq
 2527  T T A A C T G G A A T T G T C T G A T A T T T A A G A G A A  Homo sapiens Erg-2.seq
 1808                                                                 Rattus norvegicus
Erg.seq
 1516                                                           Gallus gallus Erg.seq
 1798                                                           Xenopus laevis erg.seq
 2400  T T T A C C A A A G A C A G C A G T A T T T C T A T T A T T  Danio rerio erg.seq
 2142                                                           S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       3100              3110              3120
      -----------------+-----------------+-----------------+-
 2209                                                           Mus musculus erg.seq
 2515  A C A T T C A G G A C C T C A T C A T T A T G T G G G G G C  Homo sapiens Erg-1.seq
 2557  A C A T T C A G G A C C T C A T C A T T A T G T G G G G G C  Homo sapiens Erg-2.seq
 1808                                                                 Rattus norvegicus
Erg.seq
 1516                                                           Gallus gallus Erg.seq
 1798                                                           Xenopus laevis erg.seq
 2430  A T T A T T A T A A T G A C A A T G T T C C G A G T C G C G  Danio rerio erg.seq
 2142                                                           S. purpuratus Erg.seq
```

Figure 8 continued

```
      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      ------------------+------------------+------------------+-
                        3130               3140               3150
      ------------------+------------------+------------------+-
 2209                                                            Mus musculus erg.seq
 2545  T T T G T T C T C C A C A G G G T C A G G T A A G A G A T G  Homo sapiens Erg-1.seq
 2587  T T T G T T C T C C A C A G G G T C A G G T A A G A G A T G  Homo sapiens Erg-2.seq
 1808                                                                  Rattus norvegicus
Erg.seq
 1516                                                            Gallus gallus Erg.seq
 1798                                                            Xenopus laevis erg.seq
 2460  A A C G C T G T G C G T T T G T T A G A T T G A A G T A T A  Danio rerio erg.seq
 2142                                                            S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      ------------------+------------------+------------------+-
                        3160               3170               3180
      ------------------+------------------+------------------+-
 2209                                                            Mus musculus erg.seq
 2575  G C C T T C T T G G C T G C C A C A A T C A G A A A T C A C  Homo sapiens Erg-1.seq
 2617  G C C T T C T T G G C T G C C A C A A T C A G A A A T C A C  Homo sapiens Erg-2.seq
 1808                                                                  Rattus norvegicus
Erg.seq
 1516                                                            Gallus gallus Erg.seq
 1798                                                            Xenopus laevis erg.seq
 2490  C T A T T C A G T G T T A T A A T T C T T - - - - - - - - T  Danio rerio erg.seq
 2142                                                            S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      ------------------+------------------+------------------+-
                        3190               3200               3210
      ------------------+------------------+------------------+-
 2209                                                            Mus musculus erg.seq
 2605  G C A G G C A T T T T G G G T A G G C G G C C T C C A G T T  Homo sapiens Erg-1.seq
 2647  G C A G G C A T T T T G G G T A G G C G G C C T C C A G T T  Homo sapiens Erg-2.seq
 1808                                                                  Rattus norvegicus
Erg.seq
 1516                                                            Gallus gallus Erg.seq
 1798                                                            Xenopus laevis erg.seq
 2512  T T C T A T C A T A G G T A T A T A A C T T A T G C A T T T  Danio rerio erg.seq
 2142                                                            S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      ------------------+------------------+------------------+-
                        3220               3230               3240
      ------------------+------------------+------------------+-
 2209                                                            Mus musculus erg.seq
 2635  T T C C T T T G A G T C G C G A A C G C T G T G C G T T T G  Homo sapiens Erg-1.seq
 2677  T T C C T T T G A G T C G C G A A C G C T G T G C G T T T G  Homo sapiens Erg-2.seq
 1808                                                                  Rattus norvegicus
Erg.seq
 1516                                                            Gallus gallus Erg.seq
 1798                                                            Xenopus laevis erg.seq
 2542  A A A C A C T A T G A G T T G A T C T C A G C C A G C C A A  Danio rerio erg.seq
 2142                                                            S. purpuratus Erg.seq
```

Figure 8 continued

```
      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       3250              3260              3270
      -----------------+-----------------+-----------------+-
 2209                                                           Mus musculus erg.seq
 2665  T C A G A A T G A A G T A T A C A A G T C A A T G T T T T T  Homo sapiens Erg-1.seq
 2707  T C A G A A T G A A G T A T A C A A G T C A A T G T T T T T  Homo sapiens Erg-2.seq
 1808                                                                 Rattus norvegicus
Erg.seq
 1516                                                           Gallus gallus Erg.seq
 1798                                                           Xenopus laevis erg.seq
 2572  T C A G G C G T T A G T C C T A G A C T C T G T C T G A C G  Danio rerio erg.seq
 2142                                                           S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       3280              3290              3300
      -----------------+-----------------+-----------------+-
 2209                                                           Mus musculus erg.seq
 2695  C C C C C T T T T T A T A T A A T A A T T A T A T A A C T T  Homo sapiens Erg-1.seq
 2737  C C C C C T T T T T A T A T A A T A A T T A T A T A A C T T  Homo sapiens Erg-2.seq
 1808                                                                 Rattus norvegicus
Erg.seq
 1516                                                           Gallus gallus Erg.seq
 1798                                                           Xenopus laevis erg.seq
 2602  G C G G C T T T A C A G T G A T G C T G G A C T G C A G A A  Danio rerio erg.seq
 2142                                                           S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       3310              3320              3330
      -----------------+-----------------+-----------------+-
 2209                                                           Mus musculus erg.seq
 2725  A T G C A T T T A T A C A C T A C G A G T T G A T C T C G G  Homo sapiens Erg-1.seq
 2767  A T G C A T T T A T A C A C T A C G A G T T G A T C T C G G  Homo sapiens Erg-2.seq
 1808                                                                 Rattus norvegicus
Erg.seq
 1516                                                           Gallus gallus Erg.seq
 1798                                                           Xenopus laevis erg.seq
 2632  G C T C A C G G C C T C A G C C C G A C T G A C C C A C A G  Danio rerio erg.seq
 2142                                                           S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       3340              3350              3360
      -----------------+-----------------+-----------------+-
 2209                                                           Mus musculus erg.seq
 2755  C C A G C C A A A G A C A C A C G A C A A A A G A G A C A A  Homo sapiens Erg-1.seq
 2797  C C A G C C A A A G A C A C A C G A C A A A A G A G A C A A  Homo sapiens Erg-2.seq
 1808                                                                 Rattus norvegicus
Erg.seq
 1516                                                           Gallus gallus Erg.seq
 1798                                                           Xenopus laevis erg.seq
 2662  C A C G G C A C A C A C A C G G C C T G A G A C T G C T G A  Danio rerio erg.seq
 2142                                                           S. purpuratus Erg.seq
```

Figure 8 continued

```
      X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       3370              3380              3390
      -----------------+-----------------+-----------------+-
 2209                                                          Mus musculus erg.seq
 2785  T C G A T A T A A T G T G G C C T T G A A T T T T A A C T C  Homo sapiens Erg-1.seq
 2827  T C G A T A T A A T G T G G C C T T G A A T T T T A A C T C  Homo sapiens Erg-2.seq
 1808                                                                Rattus norvegicus
Erg.seq
 1516                                                          Gallus gallus Erg.seq
 1798                                                          Xenopus laevis erg.seq
 2692  C A A A G G G A A A A C A C A T A T A A A T T A T A T G A T  Danio rerio erg.seq
 2142                                                          S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       3400              3410              3420
      -----------------+-----------------+-----------------+-
 2209                                                          Mus musculus erg.seq
 2815  T G T A T G C T T A A T G T T T A C A A T A T G A A G T T A  Homo sapiens Erg-1.seq
 2857  T G T A T G C T T A A T G T T T A C A A T A T G A A G T T A  Homo sapiens Erg-2.seq
 1808                                                                Rattus norvegicus
Erg.seq
 1516                                                          Gallus gallus Erg.seq
 1798                                                          Xenopus laevis erg.seq
 2722  T T T C T A C T G A A T C A A T T T A A T A A A G T A T T T  Danio rerio erg.seq
 2142                                                          S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       3430              3440              3450
      -----------------+-----------------+-----------------+-
 2209                                                          Mus musculus erg.seq
 2845  T T A G T T C T T A G A A T G C A G A A T G T A T G T A A T  Homo sapiens Erg-1.seq
 2887  T T A G T T C T T A G A A T G C A G A A T G T A T G T A A T  Homo sapiens Erg-2.seq
 1808                                                                Rattus norvegicus
Erg.seq
 1516                                                          Gallus gallus Erg.seq
 1798                                                          Xenopus laevis erg.seq
 2752  T C T C C T G G A T A A A G A A A A A A A A A A A A A A A A  Danio rerio erg.seq
 2142                                                          S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       3460              3470              3480
      -----------------+-----------------+-----------------+-
 2209                                                          Mus musculus erg.seq
 2875  A A A A T A A G C T T G G C C T A G C A T G G C A A A T C A  Homo sapiens Erg-1.seq
 2917  A A A A T A A G C T T G G C C T A G C A T G G C A A A T C A  Homo sapiens Erg-2.seq
 1808                                                                Rattus norvegicus
Erg.seq
 1516                                                          Gallus gallus Erg.seq
 1798                                                          Xenopus laevis erg.seq
 2782  A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A  Danio rerio erg.seq
 2142                                                          S. purpuratus Erg.seq
```

Figure 8 continued

```
      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       3490              3500              3510
      -----------------+-----------------+-----------------+-
 2209                                                           Mus musculus erg.seq
 2905  G A T T T A T A C A G G A G T C T G C A T T T G C A C T T T  Homo sapiens Erg-1.seq
 2947  G A T T T A T A C A G G A G T C T G C A T T T G C A C T T T  Homo sapiens Erg-2.seq
 1808                                                                 Rattus norvegicus
Erg.seq
 1516                                                           Gallus gallus Erg.seq
 1798                                                           Xenopus laevis erg.seq
 2811                                                           Danio rerio erg.seq
 2142                                                           S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       3520              3530              3540
      -----------------+-----------------+-----------------+-
 2209                                                           Mus musculus erg.seq
 2935  T T T T A G T G A C T A A A G T T G C T T A A T G A A A A C  Homo sapiens Erg-1.seq
 2977  T T T T A G T G A C T A A A G T T G C T T A A T G A A A A C  Homo sapiens Erg-2.seq
 1808                                                                 Rattus norvegicus
Erg.seq
 1516                                                           Gallus gallus Erg.seq
 1798                                                           Xenopus laevis erg.seq
 2811                                                           Danio rerio erg.seq
 2142                                                           S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       3550              3560              3570
      -----------------+-----------------+-----------------+-
 2209                                                           Mus musculus erg.seq
 2965  A T G T G C T G A A T G T T G T G G A T T T T G T G T T A T  Homo sapiens Erg-1.seq
 3007  A T G T G C T G A A T G T T G T G G A T T T T G T G T T A T  Homo sapiens Erg-2.seq
 1808                                                                 Rattus norvegicus
Erg.seq
 1516                                                           Gallus gallus Erg.seq
 1798                                                           Xenopus laevis erg.seq
 2811                                                           Danio rerio erg.seq
 2142                                                           S. purpuratus Erg.seq

      X X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                       3580              3590              3600
      -----------------+-----------------+-----------------+-
 2209                                                           Mus musculus erg.seq
 2995  A A T T T A C T T T G T C C A G G A A C T T G T G C A A G G  Homo sapiens Erg-1.seq
 3037  A A T T T A C T T T G T C C A G G A A C T T G T G C A A G G  Homo sapiens Erg-2.seq
 1808                                                                 Rattus norvegicus
Erg.seq
 1516                                                           Gallus gallus Erg.seq
 1798                                                           Xenopus laevis erg.seq
 2811                                                           Danio rerio erg.seq
 2142                                                           S. purpuratus Erg.seq
```

Figure 8 continued

```
      X X X X X X X X X X X X X X X X X X X X X X X X X X X X  Majority
      -----------------+-----------------+-----------------+-
                      3610              3620              3630
      -----------------+-----------------+-----------------+-
 2209                                                          Mus musculus erg.seq
 3025  G A G A G C C A A G G A A A T A G G A T G T T T G G C A C C  Homo sapiens Erg-1.seq
 3067  G A G A G C C A A G G A A A T A G G A T G T T T G G C A C C  Homo sapiens Erg-2.seq
 1808                                                                Rattus norvegicus
Erg.seq
 1516                                                          Gallus gallus Erg.seq
 1798                                                          Xenopus laevis erg.seq
 2811                                                          Danio rerio erg.seq
 2142                                                          S. purpuratus Erg.seq

      X                                                        Majority
      --

      --
 2209                                                          Mus musculus erg.seq
 3055  C                                                       Homo sapiens Erg-1.seq
 3097  C                                                       Homo sapiens Erg-2.seq
 1808                                                                Rattus norvegicus
Erg.seq
 1516                                                          Gallus gallus Erg.seq
 1798                                                          Xenopus laevis erg.seq
 2811                                                          Danio rerio erg.seq
 2142                                                          S. purpuratus Erg.seq
```

Figure 8 continued

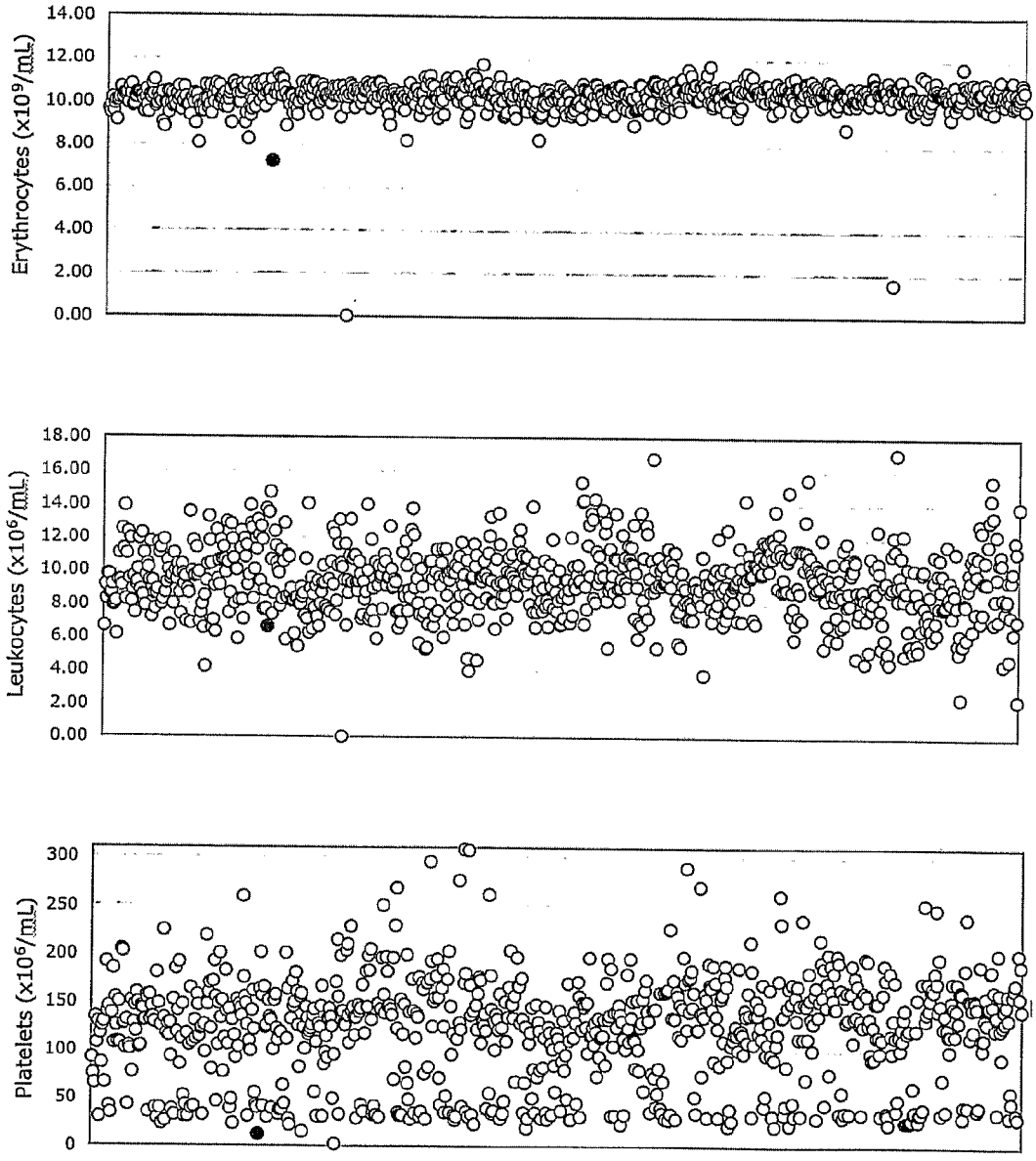

Erythrocytes (x10⁹/mL)
14.00
12.00
10.00
8.00
6.00
4.00
2.00
0.00
Leukocytes (x10⁶/mL)
18.00
16.00
14.00
12.00
10.00
8.00
6.00
4.00
2.00
0.00
Platelets (x10⁶/mL)
300
250
200
150
100
50
0

Figure 9

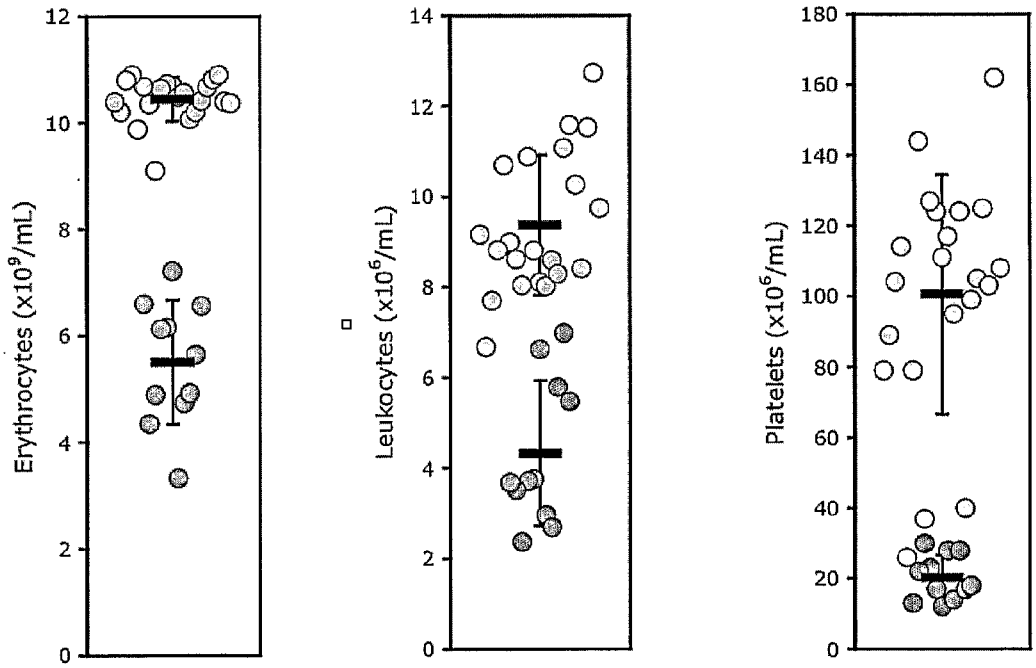

Erythrocytes (x10⁹/mL)
0
2
4
6
8
10
12
Leukocytes (x10⁶/mL)
0
2
4
6
8
10
12
14
Platelets (x10⁶/mL)
0
20
40
60
80
100
120
140
160
180

Figure 10

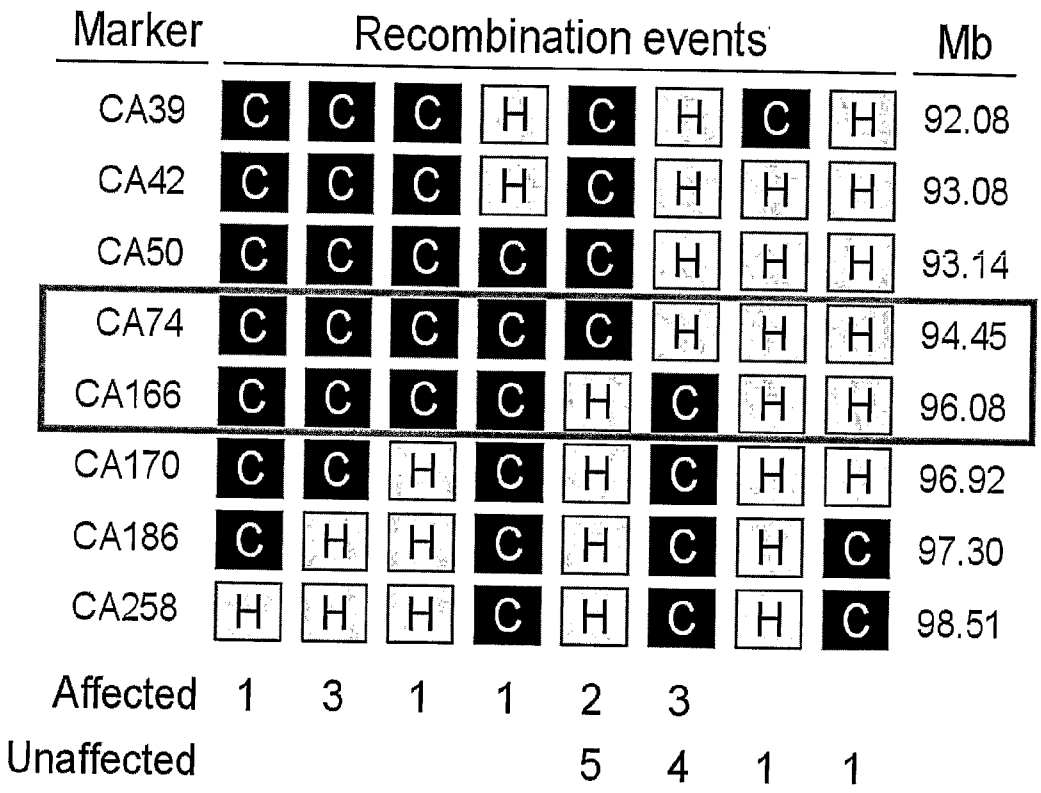

Marker
Recombination events
Mb
CA39 C C C H C H C H 92.08
CA42 C C C H C H H H 93.08
CA50 C C C C C H H H 93.14
CA74 C C C C C H H H 94.45
CA166 C C C C H C H H 96.08
CA170 C C H C H C H H 96.92
CA186 C H H C H C H C 97.30
CA258 H H H C H C H C 98.51
Affected 1 3 1 1 2 3
Unaffected 5 4 1 1

Figure 11

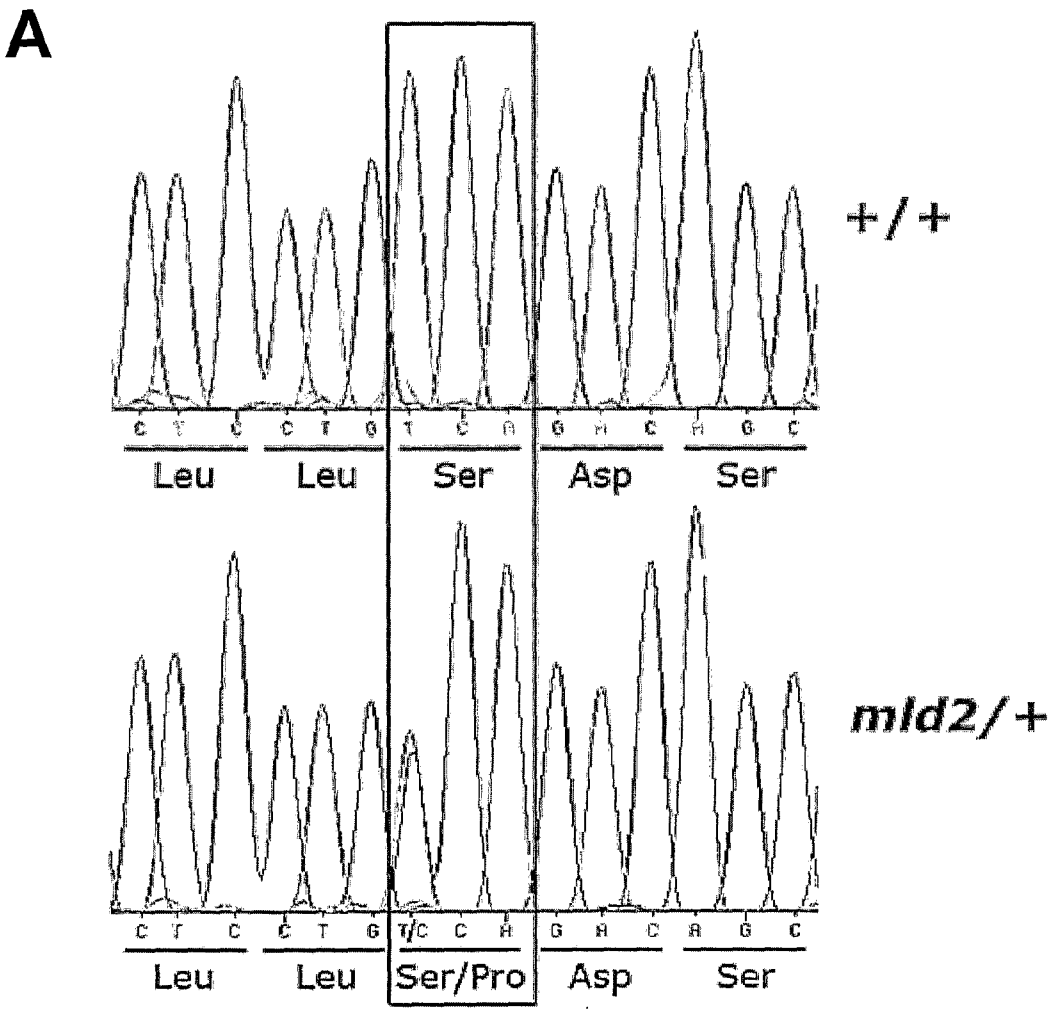


Figure 12

B

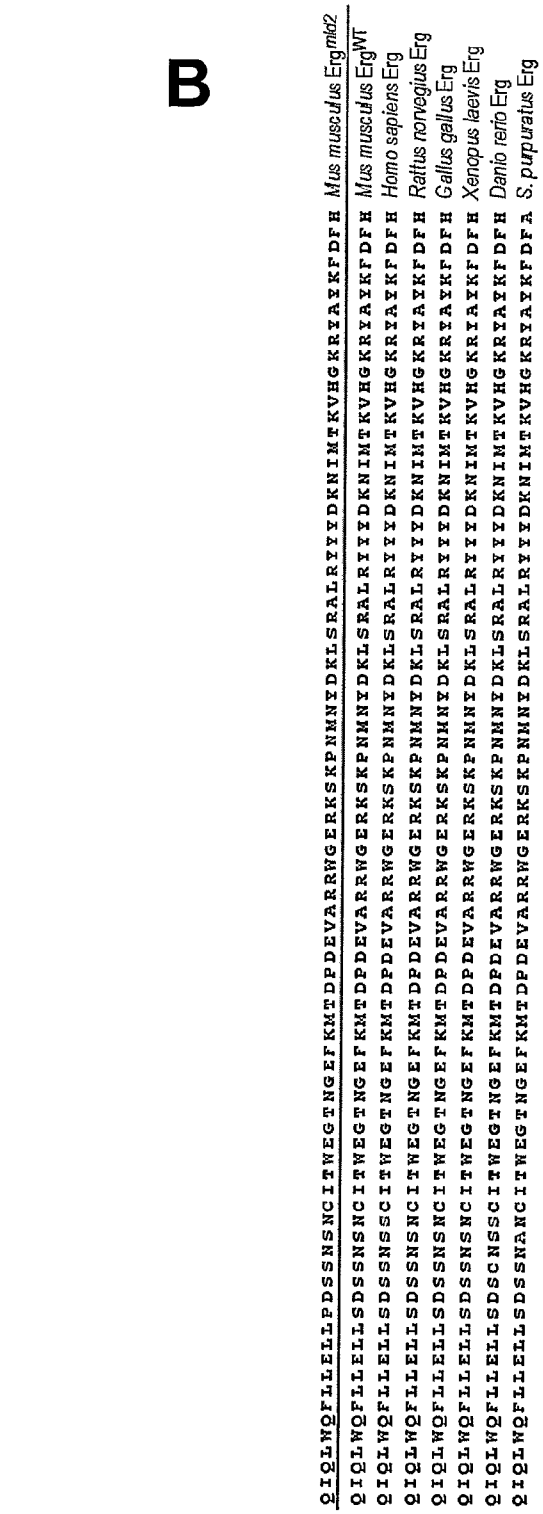


Figure 12 continued

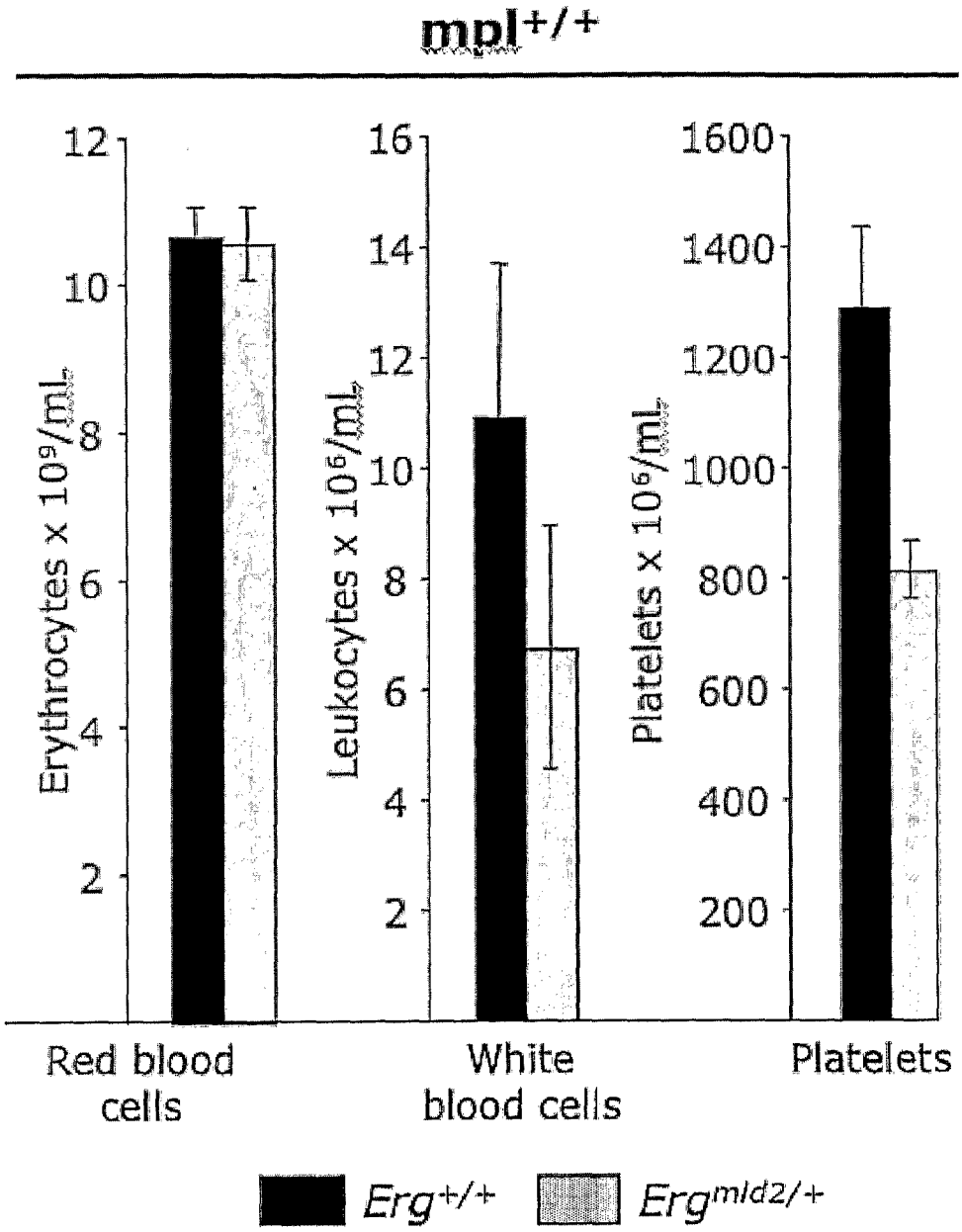

mpl+/+
Erythrocytes x 10⁹/mL
12
10
8
6
4
2
Leukocytes x 10⁶/mL
16
14
12
10
8
6
4
2
Platelets x 10⁶/mL
1600
1400
1200
1000
800
600
400
200
Red blood cells
White blood cells
Platelets
Erg+/+
Erg mld2/+

Figure 13

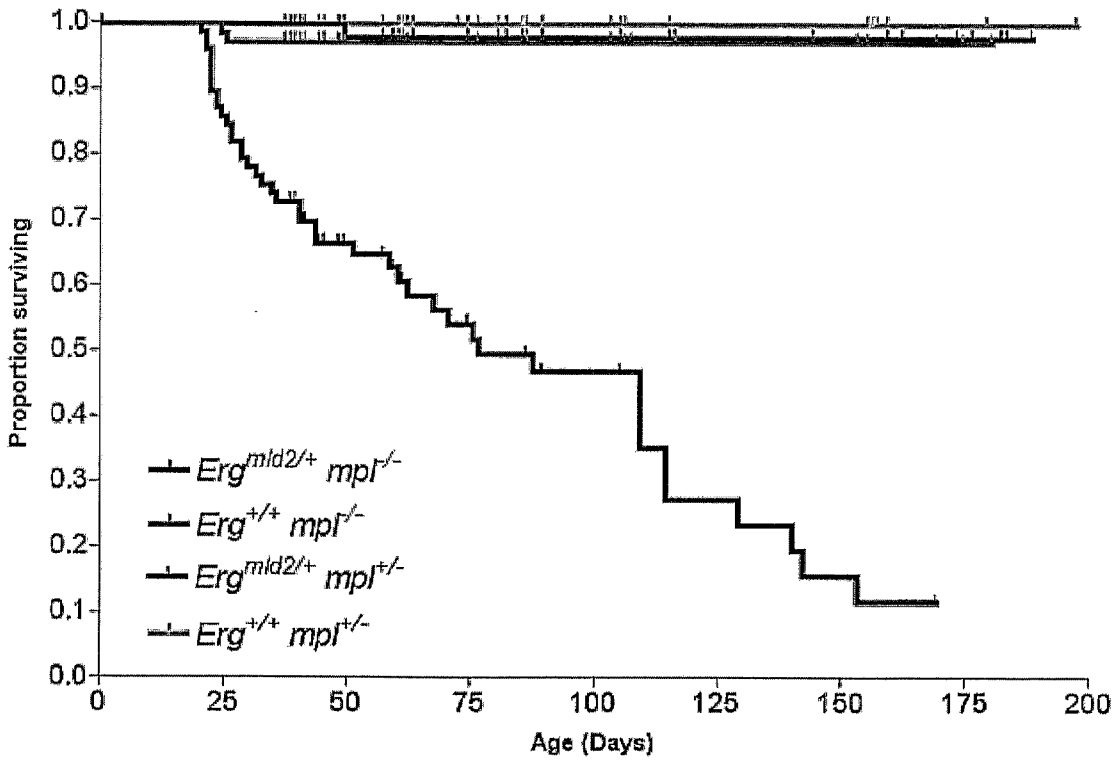

1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Proportion surviving
0
25
50
75
100
125
150
175
200
Age (Days)
Erg^mld2/+ mpl^-/-
Erg^+/+ mpl^-/-
Erg^mld2/+ mpl^+/-
Erg^+/+ mpl^+/-

Figure 14

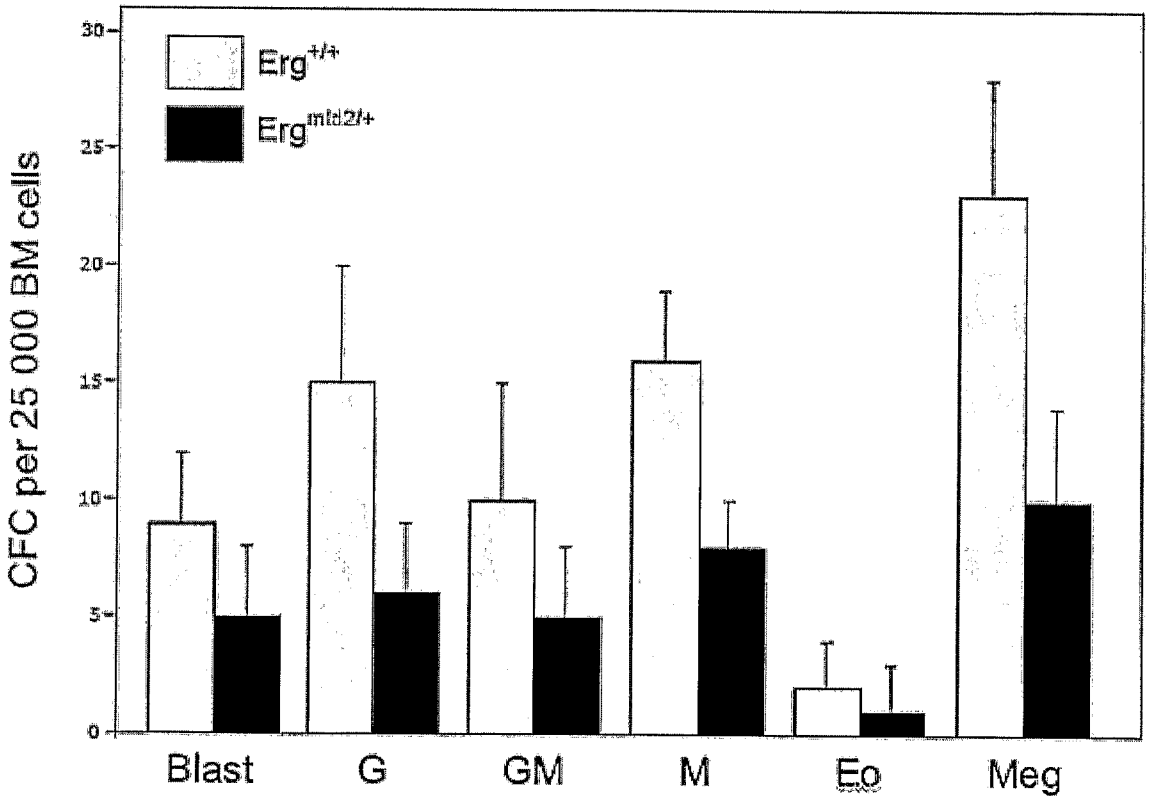

Erg+/+
Erg mld2/+
CFC per 25 000 BM cells
30
25
20
15
10
5
0
Blast
G
GM
M
Eo
Meg

Figure 15

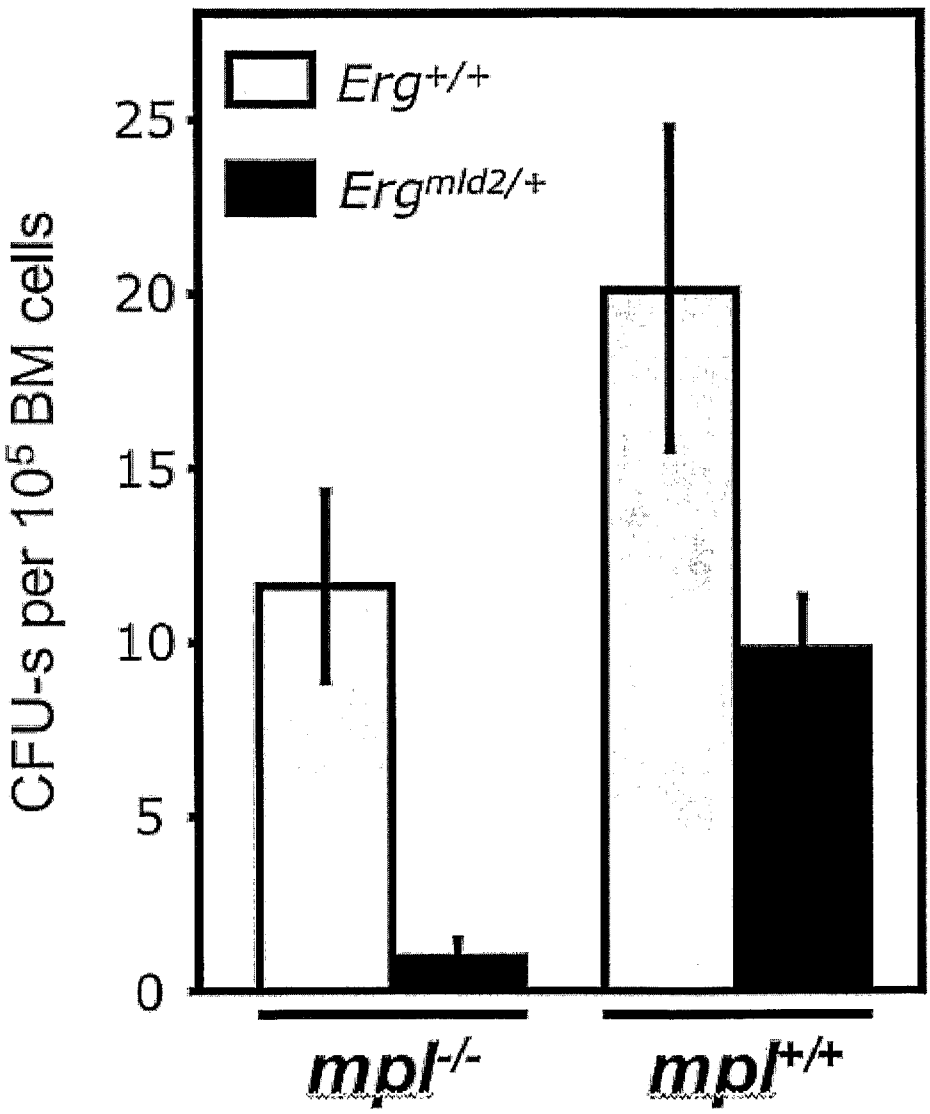


Figure 16

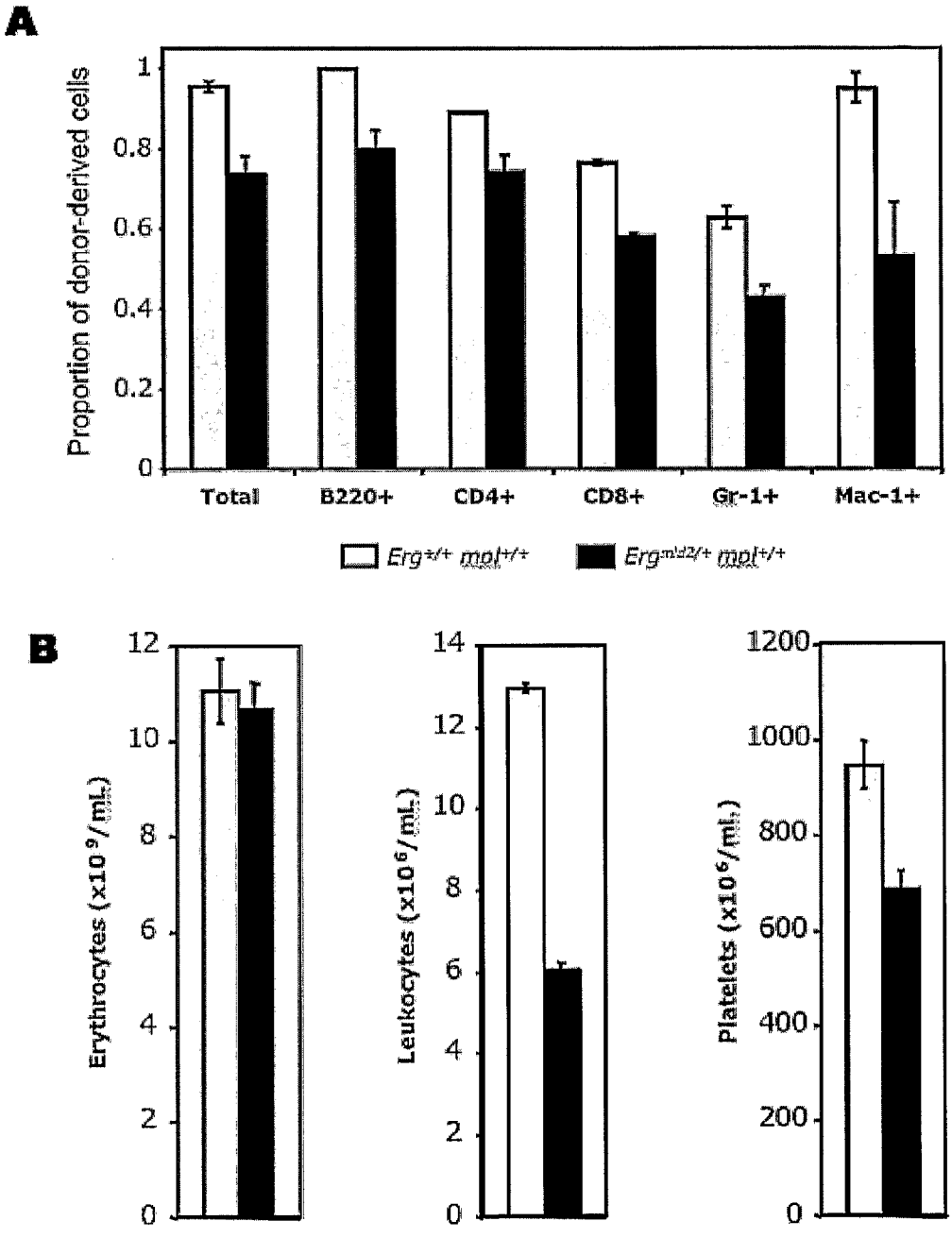

A
Proportion of donor-derived cells
1
0.8
0.6
0.4
0.2
0
Total
B220+
CD4+
CD8+
Gr-1+
Mac-1+
Erg+/+ mpl+/+
Erg+d2/+ mpl+/+
B
Erythrocytes (x10⁹/mL)
12
10
8
6
4
2
0
Leukocytes (x10⁶/mL)
14
12
10
8
6
4
2
0
Platelets (x10⁶/mL)
1200
1000
800
600
400
200
0

Figure 17

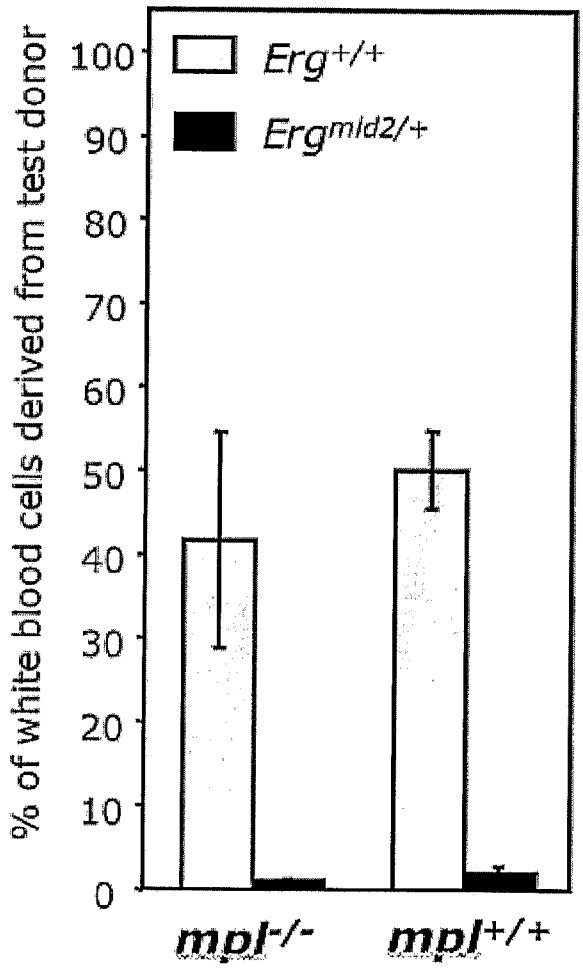

% of white blood cells derived from test donor
100
90
80
70
60
50
40
30
20
10
0
Erg+/+
Ergmld2/+
mpl-/-
mpl+/+

Figure 18

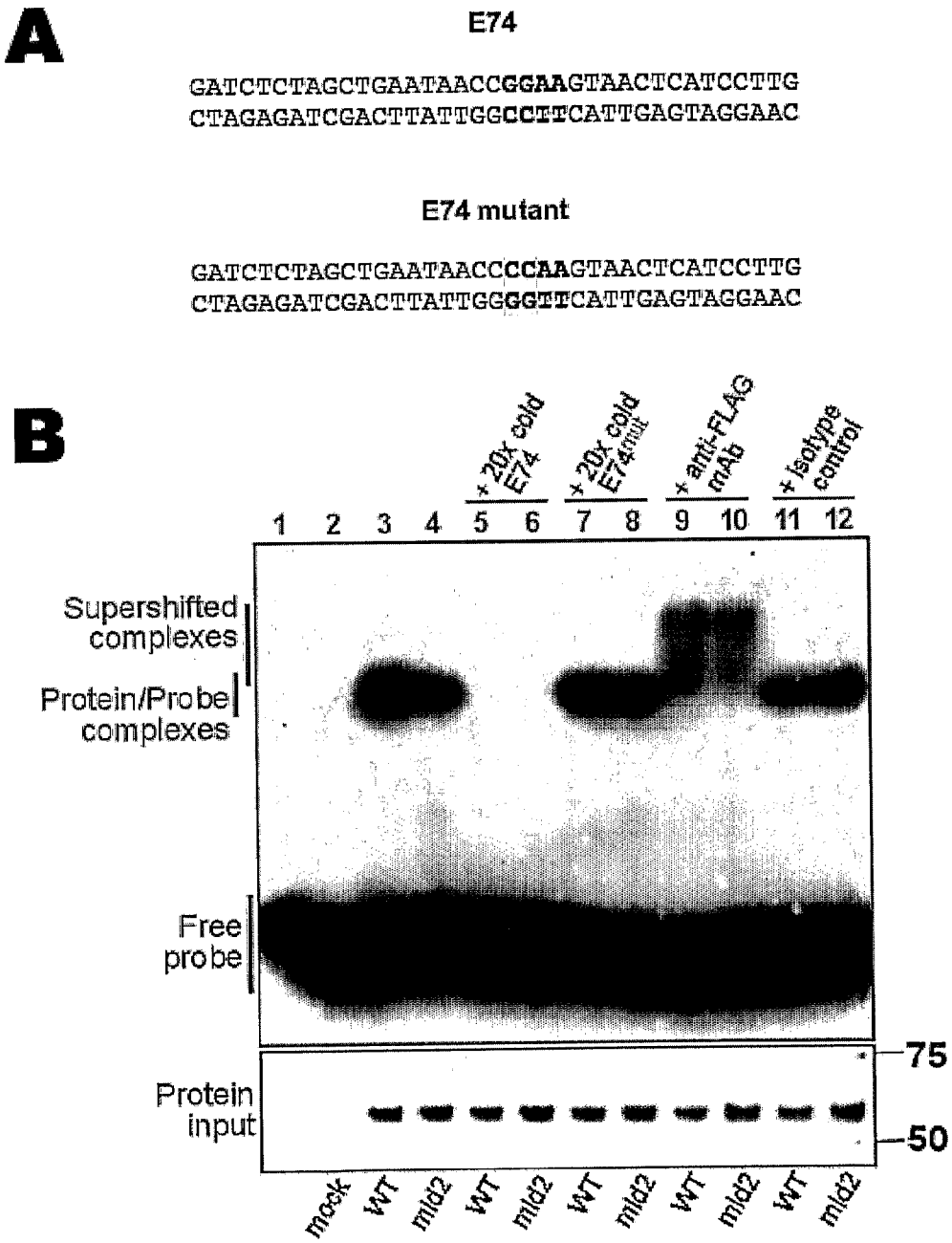

A
E74
GATCTCTAGCTGAATAACCGGAAGTAACTCATCCTTG
CTAGAGATCGACTTATTGGCCTTCATTGAGTAGGAAC
E74 mutant
GATCTCTAGCTGAATAACCCCAAGTAACTCATCCTTG
CTAGAGATCGACTTATTGGGGTTCATTGAGTAGGAAC
B
+ 20x cold E74
+ 20x cold E74mut
+ anti-FLAG mAb
+ isotype control
1 2 3 4 5 6 7 8 9 10 11 12
Supershifted complexes
Protein/Probe complexes
Free probe
Protein input
75
50
mock WT mld2 WT mld2 WT mld2 WT mld2 WT mld2

Figure 19

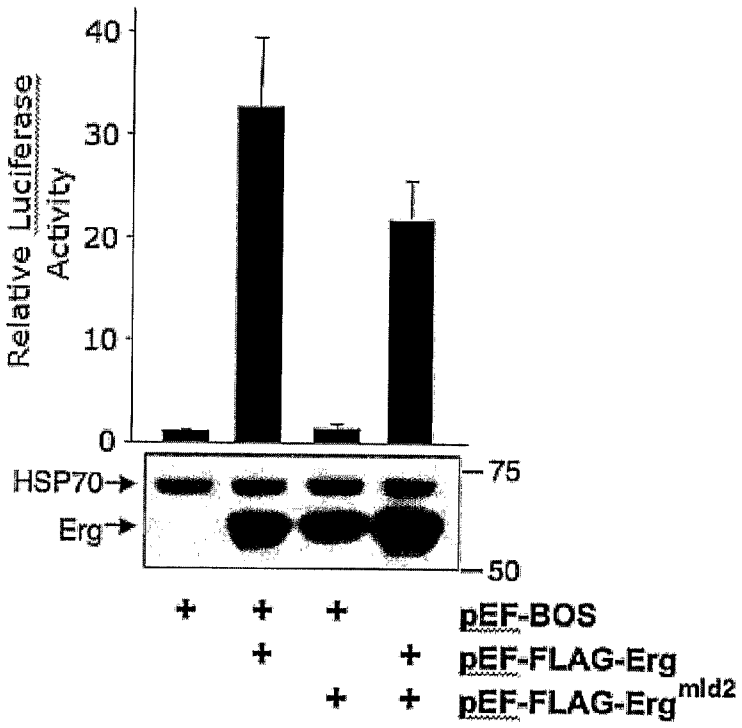

Relative Luciferase Activity
40
30
20
10
0
HSP70
Erg
75
50
+ + + pEF-BOS
+ + pEF-FLAG-Erg
+ + pEF-FLAG-Erg^mld2

Figure 20

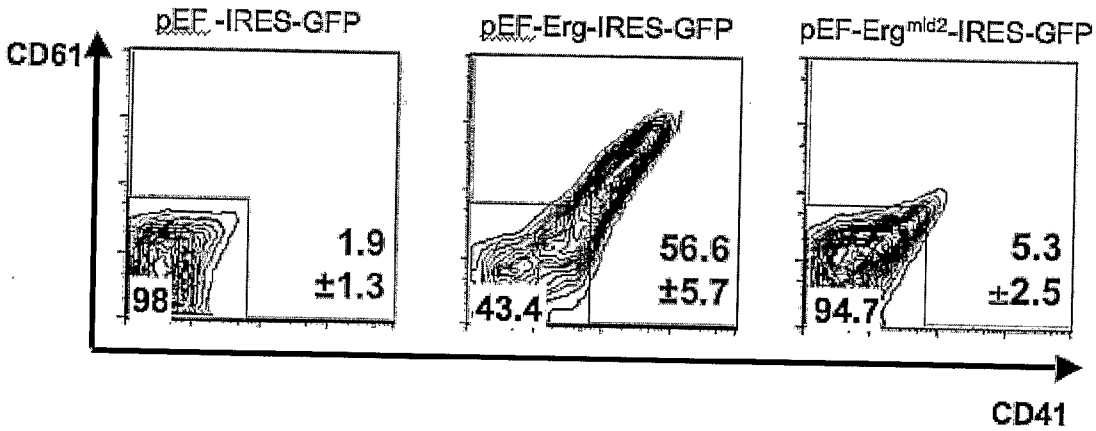

pEF-IRES-GFP
pEF-Erg-IRES-GFP
pEF-Erg^{mid2}-IRES-GFP
CD61
98
1.9
±1.3
43.4
56.6
±5.7
94.7
5.3
±2.5
CD41

Figure 21

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU2007/000645

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl.

C12N 15/11 (2006.01) *A61P 7/00* (2006.01)
A61K 38/17 (2006.01) *C12N 5/10* (2006.01)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
CAPLUS, MEDLINE, WPIDS: ETS related gene OR Erg, transcription factors, hematopoiesis, leukemia

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Hart AH, Corrick CM, Tymms MJ, Hertzog PJ, Kola I. Human ERG is a proto-oncogene with mitogenic and transforming activity. Oncogene. 1995 Apr 6;10(7):1423-30. See the whole document, particularly the abstract, Figures 1a, 2, and associated text, and the discussion.	22, 24, 32-34
X Y	Deramaudt TB, Remy P, Stiegler P. Identification of interaction partners for two closely-related members of the ETS protein family, FLI and ERG. Gene. 2001 Aug 22;274(1-2):169-77. Erratum in: Gene 2001 Oct 31;278(1-2):265. See the whole document, particularly Figure 1 and associated text, Materials and Methods parts 2.1 and 2.3, results, and the discussion.	22, 26-28, 32-34 1-21
X	Rainis L, Toki T, Pimanda JE, Rosenthal E, Machol K, Strehl S, Gottgens B, Ito E, Izraeli S. The proto-oncogene ERG in megakaryoblastic leukemias. Cancer Res. 2005 Sep 1;65(17):7596-602. See the whole document, particularly the abstract, Figures 3, 5A, 5B and associated text, and the discussion.	22, 24, 32-34

[X] Further documents are listed in the continuation of Box C [X] See patent family annex

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
09 July 2007

Date of mailing of the international search report
23 JUL 2007

Name and mailing address of the ISA/AU

AUSTRALIAN PATENT OFFICE
PO BOX 200, WODEN ACT 2606, AUSTRALIA
E-mail address: pct@ipaustralia.gov.au
Facsimile No. (02) 6285 3929

Authorized officer
Jason MacKenzie
AUSTRALIAN PATENT OFFICE
(ISO 9001 Quality Certified Service)
Telephone No : (02) 6283 7934

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU2007/000645

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: **1 and 26 (in part)**
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

Please see extra sheet.

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU2007/000645

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	McLaughlin F, Ludbrook VJ, Cox J, von Carlowitz I, Brown S, Randi AM. Combined genomic and antisense analysis reveals that the transcription factor Erg is implicated in endothelial cell differentiation. Blood. 2001 Dec 1;98(12):3332-9.	
X	See the whole document, particularly the abstract, "Delivery of GeneBlocs to	24
Y	HUVECS" on page 3333, Figures 1, 2, 4, 6, and associated text, Table 1 and associated text, and the discussion.	1-21
	WO2001/088124 A2 (RIBOZYME PHARMACEUTICALS INC) 22 November 2001.	
X	See the whole document, particularly the claims.	24
Y		1-21
	Yang L, Xia L, Wu DY, Wang H, Chansky HA, Schubach WH, Hickstein DD, Zhang Y. Molecular cloning of ESET, a novel histone H3-specific methyltransferase that interacts with ERG transcription factor. Oncogene. 2002 Jan 3;21(1):148-52.	
X	See the whole document, particularly the abstract, Figure 3 and the associated text.	22, 26-28, 32-34
Y		1-21
	Anderson MK, Hernandez-Hoyos G, Diamond RA, Rothenberg EV. Precise developmental regulation of Ets family transcription factors during specification and commitment to the T cell lineage. Development. 1999 Jun;126(14):3131-48.	
Y	See the whole document, particularly the abstract, Figures 4-5 and the associated text.	1-21
	Pereira DS, Dorrell C, Ito CY, Gan OI, Murdoch B, Rao VN, Zou JP, Reddy ES, Dick JE. Retroviral transduction of TLS-ERG initiates a leukemogenic program in normal human hematopoietic cells. Proc Natl Acad Sci U S A. 1998 Jul	
X	7;95(14):8239-44. See the whole document.	22, 24, 32-34
	Marcucci G, Baldus CD, Ruppert AS, Radmacher MD, Mrozek K, Whitman SP, Kolitz JE, Edwards CG, Vardiman JW, Powell BL, Baer MR, Moore JO, Perrotti D, Caligiuri MA, Carroll AJ, Larson RA, de la Chapelle A, Bloomfield CD. Overexpression of the ETS-related gene, ERG, predicts a worse outcome in acute myeloid leukemia with normal karyotype: a Cancer and Leukemia Group B study. J Clin Oncol. 2005 Dec 20;23(36):9234-42. Epub 2005 Nov 7.	
Y	See the whole document.	1-21
	Yi H, Fujimura Y, Ouchida M, Prasad DD, Rao VN, Reddy ES. Inhibition of apoptosis by normal and aberrant Fli-1 and erg proteins involved in human solid tumors and leukemias. Oncogene. 1997 Mar 20;14(11):1259-68.	
X	See the whole document, particularly the abstract, Figures 2, 4, 5, 7d, 8b and the associated text.	1-22, 24 and 32-34

Supplemental Box
(To be used when the space in any of Boxes I to VIII is not sufficient)

Continuation of Box No: II (2)

Claims 1 and 26 define methods directed to transcriptional targets and downstream effectors of Erg polypeptide. The description only provides support for Erg modulators, and is wholly silent to both the modulation of, and identification of inhibitors to Erg transcriptional targets and downstream effectors. As such, the search for claims 1, 26, and the appended claims thereto is necessarily directed to that which is fully supported, ie both the modulation of, and identification of inhibitors against the Erg polypeptide *per se*.

Claim 26 is not clear. The claimed method comprises the step of identifying change in activity of cells incubated with a potential Erg modulator, and attributing altered activity to modulated Erg function. However many tested modulators may alter cell activity in the total absence of any specific modulation of Erg. For the purpose of this search, the method of claim 26 has been construed as comprising the step of identifying altered Erg function in cells.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.
PCT/AU2007/000645

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report	Patent Family Member
WO 2001/088124	AU 61676/01

Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.

END OF ANNEX