

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
9 October 2008 (09.10.2008)

PCT

(10) International Publication Number
WO 2008/121007 A1

(51) International Patent Classification:

C12N 15/12 (2006.01) *A61K 39/395* (2006.01)
C07H 21/02 (2006.01) *A61K 48/00* (2006.01)
C12N 15/63 (2006.01) *A61K 35/00* (2006.01)
C07K 14/82 (2006.01) *A61K 35/04* (2006.01)
C07K 16/32 (2006.01) *A61K 38/10* (2006.01)
A61K 38/18 (2006.01)

(74) Agent: BALDWIN; PO Box 852, Wellington, 6001 (NZ).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, 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, ME, 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.

(21) International Application Number:

PCT/NZ2008/000069

(22) International Filing Date: 29 March 2008 (29.03.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/908,967 29 March 2007 (29.03.2007) US

(71) Applicant (for all designated States except US): **AUCKLAND UNISERVICES LIMITED** [NZ/NZ]; Level 10, 70 Symonds Street, Auckland 1010 (NZ).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **DONGXU, Liu** [CN/NZ]; Liggins Institute, University of Auckland, 2-6 Park Avenue, Grafton, Auckland (NZ). **LOBIE, Peter, Edward** [AU/NZ]; 593A Riddell Road, Glendowie, Auckland (NZ).

(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, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, 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

(54) Title: NOVEL SARATAN POLYPEPTIDES AND POLYNUCLEOTIDES AND METHODS OF USE THEREOF

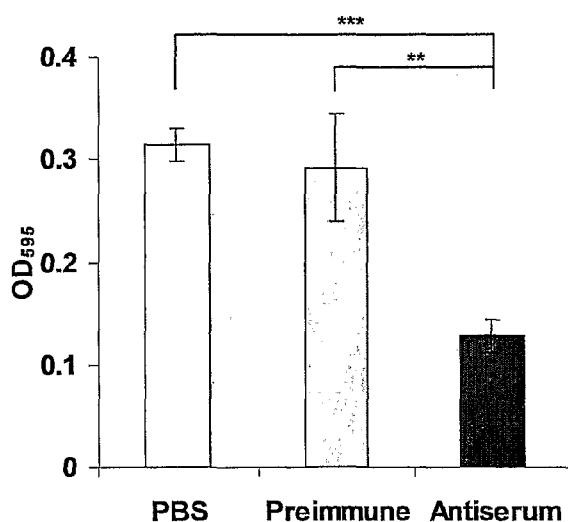


FIGURE 14

(57) Abstract: The invention relates to Saratan polypeptides, polynucleotides, and antibodies. The invention also relates to expression vectors and host cells for producing these polypeptides, polynucleotides, or antibodies. The invention further relates to diagnostics and therapeutics, especially for cancer, and particularly breast cancer, comprising one or more of the disclosed polypeptides, polynucleotides, antibodies, expression vectors, host cells, or compositions thereof.

WO 2008/121007 A1

NOVEL SARATAN POLYPEPTIDES AND POLYNUCLEOTIDES AND METHODS OF USE THEREOF

FIELD OF THE INVENTION

5 The invention relates to Saratan polypeptides, as well as Saratan polynucleotides and antibodies to Saratan. The invention also relates to expression vectors and host cells for producing these polypeptides, polynucleotides, or antibodies. The invention further relates to methods for diagnosis and treatment, especially for cancer, and particularly breast cancer, colon cancer, lung cancer, prostate cancer, endometrial cancer, and
10 stomach cancer, using one or more of the disclosed polypeptides, polynucleotides, antibodies, expression vectors, host cells, or compositions thereof.

BACKGROUND OF THE INVENTION

Traditional therapies used to treat cancer patients include surgery, radiation,
15 chemotherapy, and sex hormone therapy. However, there are always a large number of cancer patients whose condition does not respond to these traditional therapies. Because uncontrolled cell proliferation, cell survival, and/or cell motility are the hallmarks of cancer, for many decades, the most direct and effective strategy in the field of cancer therapy has been to attempt to stop cancer growth by blocking cell division. Traditional
20 chemotherapy agents are therefore toxic to healthy cells as well as cancer cells. Moreover, the inherent genetic instability of neoplastic cells eventually results in the selection of drug-resistant clones after prolonged exposure to chemotherapy.

Targeting specific pathways to stop cancer growth has emerged as an alternative
25 strategy for new drug development because of less toxicity and high tolerability. For example, an increasing number of compounds directed against the EGFR and HER2 receptors have entered clinical development and are currently in clinical trials. Thus, there is a continuing need for new anti-cancer agents. In particular, there is a need to identify specific genes and proteins that can be targeted to inhibit cancer cells.

30

SUMMARY OF INVENTION

The present invention stems from the identification of a novel open reading frame which encodes a new polypeptide, termed Saratan, which among other things, acts to promote cancer cell proliferation, migration, and invasion.

5

The invention disclosure features an isolated Saratan polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NO:1, 3, 8-10, 49-64, 102-104. In a particular aspect, the polypeptide comprises the amino acid sequence of SEQ ID NO:1, 8-10. In a further aspect, the polypeptide comprises the amino acid sequence of SEQ ID NO:1. In another aspect, the polypeptide is a fragment, for example, comprising at least one amino acid sequence extending from amino acids 59-81 of SEQ ID NO:1, or from amino acids 1-61 of SEQ ID NO:1.

10

The invention disclosure also features an isolated Saratan polynucleotide comprising a coding sequence for at least one polypeptide. In one aspect, the polynucleotide comprises a coding sequence for at least one amino acid sequence selected from the group consisting of SEQ ID NO:1, 3, 8-10, 49-64, 102-104. In a particular aspect, the polynucleotide comprises a coding sequence for SEQ ID NO:1, 8-10. In a further aspect, the polynucleotide comprises a coding sequence for SEQ ID NO:1. In another aspect, the polynucleotide comprises a fragment of a coding sequence, for example, at least one amino acid sequence extending from amino acids 59-81 of SEQ ID NO:1, or from amino acids 1-61 of SEQ ID NO:1.

15

20

In an additional aspect, the invention disclosure features an isolated Saratan polynucleotide comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO:14-16, 19, 22, 25-27, 28-30, 37, 38, 41, 42, 45-48, 65, 66, 68-73, 74-101, 120-122. In a particular aspect, the polynucleotide comprises the nucleic acid sequence of SEQ ID NO:14. In another aspect, the polynucleotide is a fragment or an oligonucleotide comprising, for example, the nucleic acid sequence extending from nucleotides 175-249 of SEQ ID NO:14. In addition, the invention encompasses an isolated polynucleotide, or fragment thereof, which hybridizes to any one of the nucleic acid sequences of SEQ ID NO:14-16, 19, 22, 25-27, 28-30, 37, 38, 41, 42, 45-48, 65, 66, 68-73, 74-101, 120-122. The invention further encompasses an isolated

25

30

polynucleotide comprising the complement, reverse complement, reverse sequence, or fragments thereof, of any one of the nucleic acid sequences.

Further featured as part of the invention disclosure are inhibitory agents, such as Saratan
5 antisense polynucleotides and Saratan interfering RNAs (iRNAs) and small interfering
RNAs (siRNAs). Preferably, the agent is a nucleic acid adapted to inhibit Saratan in use.
More preferably, the agent is selected from the group consisting of: an antisense
nucleic acid directed to a Saratan transcript; a nucleic acid adapted to express such
antisense in use; iRNA directed to a Saratan transcript; and a nucleic acid adapted to
10 express such iRNA in use. In a particular aspect, the invention disclosure features an
iRNA directed to the Saratan transcript, or a nucleic acid adapted in use to express an
iRNA directed to the Saratan transcript, wherein the iRNA comprises any one of the
nucleotide sequences of SEQ ID NO:16-24, 25-36, 45-48, 68-73, 75-89, 96-100. Such
antisense polynucleotides and iRNAs, in particular, siRNAs, can inhibit expression of a
15 Saratan polynucleotide, e.g., SEQ ID NO:14-16, 19, 22, 75-77, 101, 120-122, or
modified sequences thereof. The invention disclosure features expression vectors, as
well as host cells, for producing these antisense polynucleotides or iRNAs.

The invention disclosure features an expression vector comprising a Saratan
20 polynucleotide. In one aspect, the expression vector comprises a coding sequence for at
least one amino acid sequence selected from the group consisting of SEQ ID NO:1, 3,
8-10, 49-64, 102-104. In a particular aspect, the expression vector comprises a coding
sequence for at least one amino acid sequence of SEQ ID NO:1, 8-10. In a further
aspect, the expression vector comprises a coding sequence for at least one amino acid
25 sequence of SEQ ID NO:1. In another aspect, the expression vector comprises a coding
sequence for at least one amino acid sequence extending from amino acids 59-81 of
SEQ ID NO:1, or from amino acids 1-61 of SEQ ID NO:1.

The invention disclosure also features a host cell, for example, a microbial host cell,
30 comprising at least one expression vector.

The invention disclosure additionally features modified Saratan polypeptides, e.g., for at
least one of SEQ ID NO:1, 3, 8-10, 49-64, 102-104, including biologically active

alterations, fragments, variants, and derivatives, described herein. Also featured are polynucleotides encoding these modified polypeptides, as well as alterations, fragments, variants, and derivatives of the disclosed polynucleotides, expression vectors comprising these polynucleotides, and host cells comprising these vectors. In specific
5 aspects, the compositions and methods of the invention employ these modified polynucleotides or polypeptides, or corresponding expression vectors or host cells.

In addition, the invention disclosure features Saratan polypeptides, e.g., at least one of SEQ ID NO:1, 3, 8-10, 49-64, 102-104, or modified sequences thereof, which include
10 fusions or conjugates with a heterologous agent as described in detail herein.

As another feature, the invention disclosure includes Saratan antibodies. The antibodies can be polyclonal or monoclonal antibodies, and any modified antibodies. Such antibodies can bind to a Saratan polypeptide, e.g., SEQ ID NO:1, 3, 8-10, 49-64, 102-
15 104, or a modified sequence thereof. In particular aspects, the antibodies can be directed to the native polypeptide, any peptides derived from this polypeptide, any modifications of these polypeptides or peptides (e.g., where primary structure is based on the sequence of Saratan), or any polypeptides or peptides which mimic the 3-D conformation of Saratan.

20

The invention disclosure features a composition comprising an isolated Saratan polypeptide, e.g., at least one of SEQ ID NO:1, 3, 8-10, 49-64, 102-104. Also featured is a composition comprising an isolated polynucleotide, e.g., at least one of SEQ ID NO:14-48, 65, 66, 68-73, 74-101, 120-122. Further featured is a composition that
25 includes an antibody, expression vector, or host cell comprising an expression vector, in accordance with the invention. The composition can include any one of the biologically active alterations, fragments, variants, and derivatives described herein. The composition can include at least one fusion or conjugate. The composition can be formulated, for example, as pharmaceutical compositions described in detail herein.

30

The invention disclosure also features a composition of the invention as part of a kit for diagnosis or treatment, especially for cancer, and particularly breast cancer, colon cancer, lung cancer, prostate cancer, endometrial cancer, or stomach cancer, in

accordance with the disclosed methods. The kits can comprise: a) at least one Saratan component (e.g., polypeptide, polynucleotide, or antibody) as set out herein; and b) optionally, instructions for use, for example, in diagnosing or treating cancer.

- 5 The invention disclosure also features a method for producing a Saratan polypeptide, the method comprising the steps of: a) culturing an expression vector or host cell comprising an expression vector, which comprises at least part of a coding sequence for at least one Saratan polypeptide under conditions suitable for the expression of the polypeptide; and b) recovering the polypeptide from the culture. In particular aspects,
- 10 the polypeptide comprises at least one amino acid sequence selected from the group consisting of SEQ ID NO:1, 3, 8-10, 49-64, 102-104, or modified sequences thereof.

- The invention disclosure additionally features a method for producing a Saratan polypeptide, e.g., for at least one of SEQ ID NO:1, 3, 8-10, 49-64, 102-104, which
- 15 comprises a fusion or conjugate as described in detail herein. Such method comprises: a) culturing an expression vector or host cell comprising an expression vector, which comprises a coding sequence for at least one Saratan polypeptide under conditions suitable for the expression of the polypeptide; b) forming the Saratan fusion or conjugate (e.g., by expression of the fused sequence or by chemical conjugation); and c)
- 20 recovering the fusion or conjugate. In particular aspects, the polypeptide comprises at least one amino acid sequence selected from the group consisting of SEQ ID NO:1, 3, 8-10, 49-64, 102-104, or modified sequences thereof.

- In addition, the invention disclosure features a method of inhibiting a cancer cell (e.g.,
- 25 inhibiting cell proliferation, cell survival, or cell motility), in particular, a cell from breast cancer, colon cancer, lung cancer, prostate cancer, endometrial cancer, or stomach cancer. The method comprises use of at least one inhibitor of Saratan (e.g., polynucleotides or antibodies) as described herein, comprising contacting the cell with the inhibitor. In particular aspects, the inhibitor is selected from chemical compounds
- 30 (e.g., small molecules), antagonists, antibodies, antisense polynucleotides, and iRNAs. Exemplary cancers and cancer cells are described in detail herein.

As an added feature, the invention disclosure encompasses a method of treating cancer in a subject, in particular, breast cancer, colon cancer, lung cancer, prostate cancer, endometrial cancer, or stomach cancer. In certain aspects, the method comprises administering at least one inhibitor of Saratan (e.g., polynucleotides or antibodies) to the
5 subject, or a composition, for example, pharmaceutical composition, thereof. In particular aspects, the inhibitor is selected from chemical compounds (e.g., small molecules), antagonists, antibodies, antisense polynucleotides, and iRNAs. Applicable cancers and cancer cells are described in detail herein.

10 In various aspects, the methods of the invention disclosure utilize *in vivo* or *in vitro* expression systems. In other aspects, the methods employ polynucleotides or polypeptides produced by recombinant, synthetic, or semi-synthetic means, or polynucleotides or polypeptides produced by endogenous means (e.g., naturally occurring components).

15

In one specific aspect, the invention comprises an isolated Saratan RNA molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO:101 and 120-122, or a sequence complementary thereto.

20 In another aspect, the invention comprises an isolated Saratan RNA molecule comprising the nucleic acid sequence of SEQ ID NO:74, or a sequence complementary thereto.

In yet another aspect, the invention comprises an isolated Saratan RNA molecule
25 comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO:75-77, or a sequence complementary thereto.

In still another aspect, the invention comprises an isolated Saratan RNA molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID
30 NO:68-73, 78-89, and 96-98, or a sequence complementary thereto.

In an additional aspect, the invention comprises an isolated Saratan RNA molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO:90-95, and 99-100, or a sequence complementary thereto.

- 5 In a further aspect, the invention comprises an isolated Saratan RNA molecule comprising a nucleic acid sequence having at least 95% identity to any one of SEQ ID NO:101 and 120-122, or a sequence complementary thereto.

- 10 In yet a further aspect, the invention comprises an isolated Saratan RNA molecule comprising a nucleic acid sequence having at least 95% identity to SEQ ID NO:74, or a sequence complementary thereto.

- The invention also comprises an antisense RNA molecule which binds to the nucleic acid sequence of any one of the preceding aspects.

15

The invention additionally comprises an antisense RNA molecule comprising the nucleic acid sequence of any one of the preceding aspects.

- 20 The invention further comprises a small interfering RNA molecule which inhibits translation of the nucleic acid sequence of any one of the preceding aspects.

The invention yet further comprises a small interfering RNA molecule comprising the nucleic acid sequence of any one of the preceding aspects.

- 25 The invention still further comprises an RNA oligonucleotide comprising the nucleic acid sequence of any one of the preceding aspects.

Also included in the invention is a vector which expresses the nucleic acid sequence of any one of the preceding aspects.

30

Additionally included in the invention is a host cell which comprises the vector of any preceding aspect.

In another aspect, the invention comprises an isolated Saratan polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1 and 102-104, or a functional fragment thereof.

- 5 In yet another aspect, the invention comprises an isolated Saratan polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:8-10.

- 10 In still another aspect, the invention comprises an isolated Saratan polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:49-64, and 115.

- 15 In an additional aspect, the invention comprises an isolated Saratan polypeptide comprising an amino acid sequence having at least 97% identity to SEQ ID NO:1.

The invention also comprises an isolated Saratan RNA molecule which encodes the amino acid sequence of any one of the preceding aspects.

- 20 The invention additionally comprises a Saratan antibody or antibody fragment which binds to the amino acid sequence of any one of the preceding aspects.

The invention further comprises a Saratan antibody or antibody fragment which binds to the amino acid sequence of any one of the preceding aspects.

- 25 The invention yet further comprises a Saratan antibody of any one of the preceding aspects which is monoclonal.

The invention still further comprises a composition comprising the isolated Saratan RNA molecule of any one of the preceding aspects.

- 30 Also included in the invention is a composition comprising the antisense RNA molecule of any one of the preceding aspects, or the small interfering RNA molecule of any one of the preceding aspects.

Additionally included in the invention is a composition comprising the vector of any preceding aspect or the host cell of any preceding aspect.

- 5 Further included in the invention is a composition comprising the isolated Saratan polypeptide of any one of the preceding aspects.

Yet further included in the invention is a composition comprising the Saratan antibody of any one of the preceding aspects.

10

Still further included in the invention is pharmaceutical composition comprising the composition of any one of the preceding aspects, and a physiologically acceptable carrier.

- 15 In another aspect, the invention comprises the use of the composition of any one of the preceding aspects, in the preparation of a medicament for the treatment of cancer in an individual, wherein the cancer is selected from the group consisting of breast cancer, colon cancer, lung cancer, prostate cancer, endometrial cancer, and stomach cancer. In particular aspects of the invention, the cancer may be metastatic cancer.

20

In yet another aspect, the invention comprises a method for treating cancer in an individual comprising administering the pharmaceutical composition of any preceding aspect to the individual, thereby treating the cancer, wherein the cancer is selected from the group consisting of breast cancer, colon cancer, lung cancer, prostate cancer,

- 25 endometrial cancer, and stomach cancer. In particular aspects, the cancer may be metastatic cancer.

The invention disclosure may also be said broadly to consist in the parts, elements and features referred to or indicated in the specification of the application, individually or
30 collectively, in any or all combinations of two or more of said parts, elements or features, and where specific integers are mentioned herein which have known equivalents in the art to which the invention relates, such known equivalents are deemed to be incorporated herein as if individually set forth.

Other aspects and embodiments of the invention disclosure are described herein below.

BRIEF DESCRIPTION OF THE DRAWINGS

5 These and other aspects of the present invention, which should be considered in all its aspects, will become apparent from the following description, which is given by way of example only, with reference to the accompanying figures:

Fig. 1. Expression of Saratan in MCF-7 mammary carcinoma cells. Total RNA was
10 isolated from the MCF-7 cells. The expression of Saratan was detected by RT-PCR using One-step RT-PCR kit (Qiagen). M, 1 kb plus DNA ladder; hST, human Saratan.

Fig. 2A-B. mRNA expression pattern of human Saratan in human normal tissues and cancer cell lines. (Fig. 2A) Saratan was amplified by PCR with Saratan specific
15 primers from cDNAs made from a panel of different human tissues (Primgen). The tissue of origin is marked above each lane. β 2M, the β 2-microglobulin was used as the cDNA input control. (Fig. 2B) The expression of Saratan in human cancer cell lines was examined by RT-PCR with the same Saratan specific primer pairs using One-step RT-PCR kit (Qiagen). β -actin was included as the RNA input control.

20

Figs. 3A-B. Cytoplasmic localization of Saratan-EGFP in MCF-7 cells. MCF-7 cells were transiently transfected with pEGFP-N1-hST (Fig. 3A) or the pEGFP-N1 vector (Fig. 3B). 24 h after transfection, cells were fixed and nuclei were counterstained with DAPI (blue). The expression and cellular localization of the green fluorescent Saratan-
25 EGFP fusion protein (hST-EGFP) or EGFP protein in MCF-7 was monitored using a confocal laser scanning microscope. Bar, 10 μ M.

Figs. 4A-B. The N-terminal of Saratan is located outside of the cell membrane. (Fig. 4A) Western blot of MCF-7 cells transiently transfected with plasmids phCMVn-hST, phCMVc-hST or the empty phCMV1 vector plasmid as a control. The expression
30 of HA-tagged Saratan was detected by the rabbit polyclonal HA-Probe (Y-11) (Santa Cruz Biotech). (Fig. 4B) MCF-7 cells were transiently transfected with plasmids

phCMVn-hST or phCMVc-hST expressing N-terminal HA-tagged Saratan (HA-hST, panels a-b, **Fig. 4B**) and a C-terminal HA-tagged Saratan (hST-HA, panels c-d, **Fig. 4B**), respectively, or with the empty phCMV1 vector plasmid as a control (panels e-f, **Fig. 4B**). After fixation, cells were permeabilized with 0.5% Triton[®] X-100 for 30 min (panels a, c and e, **Fig. 4B**) or were not permeabilized (panels b, d and f, **Fig. 4B**). Then, the expression of HA-tagged Saratan was examined with the rabbit polyclonal HA-Probe (Y-11) (Santa Cruz Biotech) followed by incubation with a FITC-labeled anti-rabbit second antibody (Sigma). Nuclei were counterstained with DAPI (blue). Bar, 10 μ M.

Figs. 5A-B. Forced Saratan overexpression in MCF-7 cells. Cells were stably transfected with plasmids expressing Saratan (MCF7-hST, **Fig. 5A**) or the empty vector plasmid (MCF7-Vec, **Fig. 5B**) as control as indicated. The expression of Saratan was examined by RT-PCR with Saratan specific primers. β -actin was used as input control for RT-PCR. M, 1 kb plus DNA ladder.

Fig. 6. Forced overexpression of Saratan significantly increases cell proliferation. MCF-7 cells that stably express hST were produced by introducing an hST expression plasmid (MCF7-hST). As a control, cells were transfected with empty vector (MCF7-Vec). Cells were seeded at a density of 50,000 cells per a well in six-well plates in triplicates in 3 ml of RPMI 1640 with 10% serum. Cells were cultured up to 10 days with media changed every other day. Cell number was determined after trypsinization every 2 days.

Figs. 7A-B. Efficiency of siRNA construct in reducing Saratan expression. MCF-7 cells were stably transfected with plasmids expressing Saratan specific siRNA constructs pSilencer 2.1-siRNA-A, B, and C, as well as empty vector as a control (siRNA-vec). The expression of Saratan (hST) was determined by RT-PCR (**Fig. 7A**) with β -actin as an internal control. The quantification was done with a UVP-system and normalised to β -actin with the density of siRNA-vec as 1.00 (**Fig. 7B**).

Fig. 8. Inhibition of cell proliferation by siRNA of targeting Saratan. MCF-7 cells that stably express siRNA specific for Saratan (siRNA-A), and control siRNA with

limited homology to any known sequences in the human genomes (Ambion) (siRNA-CK), were seeded at a density of 50,000 cells per a well in six-well plates in triplicate in 3 ml of RPMI 1640 with 10% serum. Cells were cultured up to 10 days with media changed every other day. Cell number was determined after trypsinization every 2 days.

5

Fig. 9. Over-expression of Saratan increased BrdU incorporation while siRNA specific for Saratan decreased BrdU incorporation. MCF-7 cells that stably express Saratan were produced by transfecting a Saratan expression plasmid (MCF7-hST). Alternatively, cells were transfected with empty vector (MCF7-Vec), or by vector for expressing siRNA specific for hST (siRNA-A), or control siRNA (siRNA-CK). Cells were seeded on glass coverslip in six-well plates in full media and incubated overnight before 18 hours of serum starvation in serum-free media. Cells were then pulse labelled with BrdU. Proliferating cells were detected by anti-BrdU mouse monoclonal antibody using a BrdU staining kit (Zymed). The cell nuclei were counterstained with hematoxylin. Cells with dark brown nuclear staining were counted as BrdU-positive cells. The percentage of BrdU-positive nuclei relative to the total number of nuclei was calculated from 10 random fields per coverslip.

10

15

Figs. 10A-D. Forced expression of Saratan significantly promotes the migration and invasion of mammary carcinoma MCF-7 cells. (Fig. 10A) Migration assay. MCF7-hST cells over-expressing Saratan or MCF7-Vec control cells in serum-free media were loaded into the upper side of the 8- μ m-pore-size Transwell inserts. These were placed into lower chambers with full media containing 10% FBS. The loaded chambers were incubated for 48h. Cells in the upper chamber were removed with cotton swabs and those that had migrated to the lower surface of the inner chamber were identified microscopically under a UV-visible fluorescence microscope after staining with Hoechst 33258. Representative fields of cells which had migrated to the bottom side of the Transwell filters are presented on the right. **(Fig. 10B)** Invasion assay. The same as the cell migration assay except that the upper side of the filters was coated with a layer of Matrigel. Representative fields of cells which had invaded to the bottom side of the Transwell filters are presented on the right. **(Fig. 10C)** Additional migration assay. The same as in Fig. 10A except that transiently transfected MCF-7 cells were used. Cells were transiently transfected by vector producing siRNA to specifically

20

25

30

target Saratan mRNA (siRNA-A) or by vector producing siRNA negative control (siRNA-CK). (**Fig. 10D**) Invasion assay. The same as in Fig. 10C except that the upper sides of the filters were coated with a layer of Matrigel. All numerical data presented are the means $\pm$ standard deviation (SD) of triplicates. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

Fig. 11. Production of recombinant human Saratan protein in bacteria. The N-terminal peptide including 61 amino acids of Saratan was expressed in bacteria as a GST fusion protein (GST-hSTn) and purified with glutathione Sepharose 4B matrix (Amersham Biosciences). Purified human recombinant N-terminal peptide of Saratan (hSTn) protein was analysed by SDS-PAGE and visualised by Coomassie blue staining and was shown as a 12 kDa band as indicated by arrow. GST-hSTn fusion protein is also indicated. Molecular mass markers are indicated in kDa by bars on the left.

Fig. 12. Characterization of antiserum against synthesized N-terminal Saratan peptide. Synthesized N-terminal Saratan peptide and the recombinant protein were separated by SDS-PAGE and immunoblotted by polyclonal antibody against the synthesized N-terminal Saratan peptide by Western blot. Molecular mass markers are indicated by bars in kDa on the left.

Figs. 13A-H. Immunofluorescence staining of Saratan expression in MCF-7 cells. Cells were transfected with EGFP-tagged Saratan expressing plasmid pEGFP-N1-hST (right column) or empty pEGFP-N1 vector plasmid (left column). 24 h after transfection, cells were fixed and Saratan expression was detected with the rabbit polyclonal antibody against Saratan by indirect immunofluorescence staining with tetramethyl Rhodamine isothiocyanate (TRITC) conjugated secondary antibody. Nuclei of cells were visualized by DAP staining. Images were taken using a UV-visible fluorescence microscope with proper filters. Merged, combined images of green, red and blue fluorescence.

Fig. 14. Rabbit polyclonal antibodies against Saratan significantly inhibit the growth of MCF-7 cells. MCF-7 cells were seeded into 96 wells microplates without (PBS) or with affinity purified rabbit anti-Saratan antibodies (Antiserum) or the

preimmune serum as the control (Preimmune) at a concentration of 600 µg/ml. After incubation of 48 h, the cell proliferation was determined by MTT assay. All numerical data presented are the means $\pm$ standard deviation (SD) of triplicates. **, $p < 0.01$; ***, $p < 0.001$.

5

Figs. 15A-B. Rabbit anti-Saratan polyclonal antibodies significantly impair migration and invasion of mammalian carcinoma MCF-7 cells. (Fig. 15A) Cell migration assay. Cells in serum-free media containing affinity purified rabbit anti-Saratan antibodies (Antiserum) or the preimmune serum as the control (Preimmune) at a concentration of 600 µg/ml were loaded into the upper chambers of non-coated 8 µm pore size Transwell inserts. These were placed into lower chambers with full media containing 10% FBS. The loaded chambers were incubated for 24h. Cells in the upper chamber were removed with cotton swabs and those that had migrated to the lower surface of the inner chamber were identified under a UV-visible fluorescence microscope after staining with Hoechst 33258. Representative fields of cells which had migrated to the bottom side of the Transwell filters are presented on the right. (Fig. 15B) Cell invasion assay. Same as the cell migration assay except that the upper side of the filters was coated with a layer of Matrigel. Representative fields of cells which had invaded to the bottom side of the Transwell filters are presented on the right. All numerical data presented are the means $\pm$ standard deviation (SD) of triplicates. **, $p < 0.01$.

10
15
20

Fig. 16. Expression of Saratan in MDA-231 mammary carcinoma cells. Total RNA was isolated from the MDA-231 cells. The expression of Saratan was detected by RT-PCR with hST-F4 and hST-R4 primers using One-step RT-PCR kit (Qiagen). The specific band amplified is indicated by arrow. M, 1 kb plus DNA ladder; hST, human Saratan.

25

Figs. 17A-H. Immunofluorescence staining of Saratan expression in MDA-231 cells. Indirect immunofluorescence staining was performed on fixed cells with either the affinity-purified rabbit anti-Saratan polyclonal antibodies (Antiserum, right panels) or the preimmune serum (Preimmune, left panels) with fluorescein isothiocyanate (FITC) conjugated secondary antibodies. Nuclei of cells were visualized by propidium iodide

30

(PI) staining. Images were taken using a UV-visible fluorescence microscope with proper filters. Merged, combined images of green and red fluorescence; Visible, images of cell morphology under a normal visible light.

5 **Figs. 18A-B. Rabbit anti-Saratan polyclonal antibody significantly impaired migration and invasion of mammalian carcinoma MDA-231 cells. (Fig. 18A)** Cell migration assay. Cells in serum-free media containing affinity purified rabbit anti-Saratan antibodies (Antiserum) or the preimmune serum as the control (Preimmune) at a concentration of 600 µg/ml were loaded into the upper chambers of non-coated 8 µm
10 pore size Transwell inserts. These were placed into lower chambers with full media containing 10% FBS. The loaded chambers were incubated for 24h. Cells in the upper chamber were removed with cotton swabs and those that had migrated to the lower surface of the inner chamber were identified under a UV-visible fluorescence microscope after staining with Hoechst 33258. Representative fields of cells migrated to
15 the bottom side of the Transwell filters were presented on the right. **(Fig. 18B)** Cell invasion assay. The same as the cell migration assay except that the upper side of the filters was coated with a layer of Matrigel. Representative fields of cells invaded to the bottom side of the Transwell filters were presented on the right. All numerical data presented are the means ± standard deviation (SD) of triplicates. **, p<0.01; ***,
20 p<0.001.

Figs. 19A-F. Saratan amino acid and nucleotide sequence information. (Fig. 19A) The derived amino acid sequence of Saratan protein is presented in single letter code. The predicted transmembrane domain for Saratan is depicted in bold, with underlining.
25 **(Fig. 19B)** The nucleotide sequence of Saratan cDNA is shown (GenBank Acc. No. NM_020215). The Saratan coding region (underlined) extends from nucleotides 161 to 412, with both the start (ATG) and stop (TGA) codons shown in bold. **(Fig. 19C)** Schematic diagram of alternatively spliced mRNA isoforms of human Saratan. Exons are shown in solid bars with coding regions representing by bars shaded in red and non-coding regions in yellow. Introns are shown by lines and their sizes are indicated above.
30 The GenBank accession numbers of representative EST clones of each transcript are shown in brackets. (Adapted from NCBI AceView for Homo sapiens gene C14orf132 (Thierry-Mieg D, Thierry-Mieg J. AceView: a comprehensive cDNA-supported gene

and transcripts annotation. Genome Biol. 2006;7 Suppl 1:S12.1-14. Epub 2006 Aug 7).
 (Fig. 19D) The amino acid sequences and their alignment of Saratan isoforms α , β , and
 γ encoded by alternatively spliced transcripts a and b, c, and d, respectively. (Fig. 19 E)
 Amino acid sequence alignments of Saratan with XP_875290 (*Bos taurus*),
 5 XP_001147850 (*Pan troglodytes*), XP_001476208 (*Mus musculus*), EDL81821 (*Rattus*
norvegicus) and XP_001334647 (*Danio rerio*) obtained from BLAST (hypertext
 transfer protocol://world wide web.ncbi.nlm.nih.gov/BLAST/) and CLUSTAL W
 (hypertext transfer protocol://world wide web.ebi.ac.uk/clustalw/) analysis. (Fig. 19F).
 Multiple sequence alignments. (Fig. 19G) Saratan fusion proteins and fragments
 10 described in the Examples, below. (Fig. 19H) Saratan RNA sequences. (Fig. 19I)
 Nucleotide sequence of human Saratan transcript a mRNA deduced from the cloned
 cDNA of Saratan (GenBank Accession number NM_020215). The initiation codon
 AUG and the termination codon UGA are underlined and in bold. Polyadenylation
 signals, AAUAAA and AUUAAA are also underlined and in italics. * indicates the
 15 possible position of poly(A) tails. (Fig. 19J) Nucleotide sequence of human Saratan
 transcript b mRNA. The initiation codon AUG and the termination codon UGA are
 underlined and in bold. Polyadenylation signals, AAUAAA and AUUAAA are also
 underlined and in italics. * indicates the possible position of poly(A) tails. (Fig. 19K)
 Nucleotide sequence of human Saratan Transcript c mRNA. The initiation codon AUG
 20 and the termination codon UGA are underlined and in bold. Polyadenylation signals,
 AAUAAA and AUUAAA are also underlined and in italics. * indicates the possible
 position of poly(A) tails. (Fig. 19L) Nucleotide sequence of human Saratan transcript d
 mRNA. The initiation codon AUG and the termination codon UGA are underlined and
 in bold. Polyadenylation signals, AAUAAA and AUUAAA are also underlined and in
 25 italics. * indicates the possible position of poly(A) tails.

Figs. 20A-D. The mouse homolog of human Saratan. (Fig. 20A) The cDNA
 sequence of mouse Saratan as predicted (GenBank accession number XM_001476158).
 The initiation codon ATG and the termination codon TGA are underlined and in bold.
 30 (Fig. 20B) cDNA sequence alignment of human (NM_020215) and mouse
 (XM_001476158) Saratan by Blast 2 sequence program. (Fig. 20C) The amino acid

sequence of mouse Saratan. (Fig. 20D) The protein sequence alignment of human and mouse Saratan.

DETAILED DESCRIPTION OF THE INVENTION

5 The following is a description of the present invention, including preferred embodiments thereof, given in general terms. The invention is further elucidated from the disclosure given under the section "Examples" which provides experimental data supporting the invention and specific examples thereof.

10 Definitions

The term "antibody" should be understood in the broadest possible sense and is intended to include intact monoclonal antibodies and polyclonal antibodies. It is also intended to cover fragments and derivatives of antibodies so long as they exhibit the desired biological activity. Antibodies encompass immunoglobulin molecules and
15 immunologically active portions of immunoglobulin (Ig) molecules, i.e., molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen. These include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, Fc, Fab, Fab', and Fab₂ fragments, and a Fab expression library. Antibody molecules relate to any of the classes IgG, IgM, IgA, IgE, and IgD, which differ from one another
20 by the nature of heavy chain present in the molecule. These include subclasses as well, such as IgG1, IgG2, and others. The light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all classes, subclasses, and types. Also included are chimeric antibodies, for example, monoclonal antibodies or fragments thereof that are specific to more than one source, e.g., a mouse or human
25 sequence. Further included are camelid antibodies or nanobodies. It will be understood that each reference to "antibodies" or any like term, herein includes intact antibodies, as well as any fragments, derivatives, or variants thereof.

"Altered" polynucleotides, as used herein, include those with deletions, insertions, or
30 substitutions of different nucleotides resulting in polynucleotides that encode the same or functionally equivalent. The encoded polypeptides may also be "altered" and contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in functionally equivalent sequences. Deliberate 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 as long as at least one biological activity (e.g., stimulation of cell proliferation, cell survival, or cell motility) or immunogenic activity of the polypeptide is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid; positively charged amino acids may include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values may include leucine, isoleucine, and valine, glycine and alanine, asparagine and glutamine, serine and threonine, and phenylalanine and tyrosine. Guidance in making substitutions and/or deletions or additions can be obtained, for example, by sequence comparisons to homologues, orthologues, or paralogues, as shown in the figures, herein. Of particular note are the homologous sequences shown herein as SEQ ID NO:4, 7, 107, 110, 114, 116, 117, and 119.

The terms “cancer” and “cancerous” refer to a physiological condition in mammals that is typically characterized by abnormal or unregulated cell proliferation, cell survival, and/or cell motility. Cancer and cancer pathology can be associated, for example, with metastasis, interference with the normal functioning of neighbouring cells, release of cytokines or other secretory products at abnormal levels, suppression or aggravation of inflammatory or immunological response, neoplasia, premalignancy, malignancy, invasion of surrounding or distant tissues or organs, such as lymph nodes, etc. Specifically included are breast cancers, which can include epithelial tumours, nonepithelial tumours, carcinomas, for example, carcinomas *in situ*, as well as invasive breast cancers. Also included are colon cancers, lung cancers, prostate cancers, endometrial cancers, and stomach cancers, among others.

The terms “complementary” or “complementarity,” as used herein, refer to the natural binding of polynucleotides under permissive salt and temperature conditions by base-pairing. For the sequence A-G-T, the complementary sequence is T-C-A, the reverse complement is A-C-T, and the reverse sequence is T-G-A. Complementarity between two single stranded molecules may be partial, in which only some of the nucleic acids bind, or it may be complete when total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has

significant effects on the efficiency and strength of hybridization between nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acids strands and in the design and use of iRNAs and PNAs.

- 5 The term “derivative,” as used herein, refers to the chemical modification of a polynucleotide, or a polynucleotide complementary thereto. Such modifications include, for example, replacement of hydrogen by an alkyl, acyl, or amino group. In preferred aspects, a polynucleotide derivative encodes a polypeptide which retains the biological or immunological function of the natural molecule. A derivative polypeptide is one
10 which is modified by glycosylation, pegylation, or any similar process which retains one or more biological functions (e.g., stimulation of cell proliferation, cell survival, or cell motility) or immunogenic function of the sequence from which it was derived. In reference to antibodies, the term “derivatives” includes, for example, hybrid and recombinant antibodies. “Hybrid” and “recombinant” versions of an antibody include,
15 for example, humanised antibodies, diabodies, triabodies, and single chain antibodies.

- The term “expression” includes production of polynucleotides and polypeptides, in particular, the production of RNA (e.g., mRNA) from a gene or portion of a gene, and includes the production of a polypeptide encoded by an RNA or gene or portion of a
20 gene, and the appearance of a detectable material associated with expression. For example, the formation of a complex, for example, from a polypeptide-polypeptide interaction, polypeptide-nucleotide interaction, or the like, is included within the scope of the term “expression”. Another example is the binding of a binding ligand, such as a hybridization probe or antibody, to a gene or other polynucleotide or oligonucleotide, a
25 polypeptide or a protein fragment, and the visualization of the binding ligand. Thus, increased intensity of a spot on a microarray, on a hybridization blot such as a Northern blot, or on an immunoblot such as a Western blot, or on a bead array, or by PCR analysis, is included within the term “expression” of the underlying biological molecule.

- 30 “Inhibition” or “inhibiting” Saratan is intended to refer to blocking or reducing biological activity and/or levels of Saratan. While it may be desirable to completely inhibit the activity of Saratan, this need not be essential. “Inhibition” of Saratan may occur at the level of expression and production of Saratan (e.g., transcriptional or

translational levels) or by targeting Saratan function, for example. The terms “inhibit” or “inhibition” of Saratan, as used herein, refer to a decrease, for example, in DNA levels (e.g., decreased DNA synthesis, increased turnover, and/or decreased stability), RNA levels (e.g., decreased transcription, increased turnover, and/or decreased stability), or polypeptide levels (e.g., decreased translation, increased turnover, and/or decreased stability) or activity, or post-translational modification. An inhibitor can also decrease or block the activities or expression levels of downstream or upstream agents in the Saratan pathway. While the invention is generally directed to Saratan inhibitors, in some circumstances, there may be beneficial aspects in maintaining or increasing Saratan levels. The disclosed polynucleotides and polypeptides can also be used for such purposes, in accordance with well-known methods.

The term “modified” refers to altered sequences and to sequence fragments, variants, and derivatives, as described herein. The term includes polypeptides, polynucleotides, antibodies, and like agents according to the present disclosure. Preferably, the modified sequences of the invention retain at least one biological activity (e.g., effect on cell proliferation, cell survival, or cell motility) or binding or immunogenic function. Such modifications can be termed “functional” as noted herein.

The term “oligonucleotide” refers to a polynucleotide, typically a probe or primer, including, without limitation, single stranded DNAs, single or double stranded RNAs, RNA:DNA hybrids, and double stranded DNAs. Oligonucleotides, such as single stranded DNA probe oligonucleotides, are often synthesized by chemical methods, for example using automated oligonucleotide synthesizers that are commercially available, or by a variety of other methods, including *in vitro* expression systems, recombinant techniques, and expression in cells and organisms.

The term “patient” or “subject” includes human and non-human animals. Non-human animals include, but are not limited to, birds and mammals, in particular, mice, rabbits, cats, dogs, pigs, sheep, goats, cows, and horses.

As used herein, the phrase “pharmaceutically acceptable diluents, carriers, and/or excipients” is intended to include substances that are useful in preparing a

pharmaceutical composition, and may be co-administered with an agent in accordance with the invention while allowing same to perform its intended function. These are generally safe, non-toxic, and neither biologically nor otherwise undesirable. Examples of pharmaceutically acceptable diluents, carriers, and/or excipients include solutions, solvents, dispersion media, delay agents, emulsions, and the like. Diluents, carriers, and/or excipients may contain minor amounts of additives such as substances that enhance isotonicity and chemical stability.

“Polynucleotide” (e.g., “Saratan,” which can be used to discuss a polynucleotide) when used in the singular or plural, generally refers to any nucleic acid sequence, e.g., any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. This includes, without limitation, single and double stranded DNA, DNA including single and double-stranded regions, single and double stranded RNA, and RNA including single and double stranded regions, hybrid molecules comprising DNA and RNA that may be single stranded or, more typically, double stranded or include single and double stranded regions. Also included are triple-stranded regions comprising RNA or DNA or both RNA and DNA. Specifically included are mRNAs, cDNAs, and genomic DNAs, and any fragments thereof. The term includes DNAs and RNAs that contain one or more modified bases, such as tritiated bases, or unusual bases, such as inosine. The polynucleotides of the invention can encompass coding or non-coding sequences, or sense or antisense sequences, or iRNAs such as siRNAs. It will be understood that each reference to a “polynucleotide” or like term, herein, will include the full length sequences as well as any fragments, derivatives, or variants thereof.

It will be further understood that any reference to a “polynucleotide” (or “oligonucleotide,” or “probe,” or “primer,” etc.) having a particular SEQ ID NO. will encompass both the DNA and the counterpart RNA sequences. Regarding the particular RNA/DNA sequences disclosed herein, it is noted that SEQ ID NO: 74 corresponds to SEQ ID NO: 15; SEQ ID NO: 75 corresponds to SEQ ID NO: 16; SEQ ID NO: 76 corresponds to SEQ ID NO: 19; SEQ ID NO: 77 corresponds to SEQ ID NO: 22; SEQ ID NO: 78 corresponds to SEQ ID NO: 25; SEQ ID NO: 79 corresponds to SEQ ID NO: 26; SEQ ID NO: 80 corresponds to SEQ ID NO: 27;

SEQ ID NO: 81 corresponds to SEQ ID NO: 28; SEQ ID NO: 82 corresponds to
SEQ ID NO: 29; SEQ ID NO: 83 corresponds to SEQ ID NO: 30; SEQ ID NO: 84
corresponds to SEQ ID NO: 31; SEQ ID NO: 85 corresponds to SEQ ID NO: 32;
SEQ ID NO: 86 corresponds to SEQ ID NO: 33; SEQ ID NO: 87 corresponds to
5 SEQ ID NO: 34; SEQ ID NO: 88 corresponds to SEQ ID NO: 35; SEQ ID NO: 89
corresponds to SEQ ID NO: 36; SEQ ID NO: 90 corresponds to SEQ ID NO: 37;
SEQ ID NO: 91 corresponds to SEQ ID NO: 38; SEQ ID NO: 92 corresponds to
SEQ ID NO: 39; SEQ ID NO: 93 corresponds to SEQ ID NO: 40; SEQ ID NO: 94
corresponds to SEQ ID NO: 41; SEQ ID NO: 95 corresponds to SEQ ID NO: 42;
10 SEQ ID NO: 96 corresponds to SEQ ID NO: 45; SEQ ID NO: 97 corresponds to
SEQ ID NO: 46; SEQ ID NO: 98 corresponds to SEQ ID NO: 47; SEQ ID NO: 99
corresponds to SEQ ID NO: 65; SEQ ID NO: 100 corresponds to SEQ ID NO: 66.

“Polypeptide” (e.g., “Saratan,” which can be used to discuss a polypeptide) as used
15 herein, refers to an oligopeptide, peptide, or protein, or fragment thereof, and to
naturally occurring, recombinant, synthetic, or semi-synthetic molecules. Where
“polypeptide” is recited herein to refer to an amino acid sequence of a naturally
occurring protein molecule, “polypeptide” and like terms, are not meant to limit the
amino acid sequence to the complete, native amino acid sequence for the full length
20 molecule. It will be understood that each reference to a “polypeptide” or like term,
herein, will include the full length sequence, as well as any fragments, derivatives, or
variants thereof.

“SEQ ID NO:” as referred to herein, can indicate each sequence identifier individually,
25 or any combination thereof, or all such sequence identifiers.

The terms “substantially purified” or “isolated” as used herein, refer to nucleic or amino
acid sequences that are removed from their cellular, recombinant, or synthetic
environment, and are at least 60% free, preferably 75% free, and most preferably at
30 least 90% free or at least 99% free from other components with which they are
associated in their environment. “Isolated” polynucleotides and polypeptides have been
identified and separated from at least one contaminant nucleic acid molecule with which
they are associated in their natural state. Accordingly, it will be understood that isolated

polynucleotides and polypeptides are in a form which differs from the form or setting in which they are found in nature. It will further be appreciated that "isolated" does not necessarily reflect the exact extent (e.g., a specific percentage) to which the sequence has been purified.

5

"Treatment" and like terms refer to methods and compositions to prevent, cure, or ameliorate a medical disorder (e.g., medical disease, condition, or syndrome), or reduce at least a symptom of such disorder. In particular, this includes methods and compositions to prevent or delay onset of a medical disorder; to cure, correct, reduce, 10 slow, or ameliorate the physical or developmental effects of a disorder; and/or to prevent, end, reduce, or ameliorate the pain or suffering caused the disorder. The term "treatment" is to be considered in its broadest context. The term does not necessarily imply that the subject is treated until total recovery. Accordingly, "treatment" as used herein broadly includes inhibiting, reducing or preventing cell proliferation, cell 15 survival, and/or cell motility; ameliorating the symptoms or severity of cell proliferation, cell survival, and/or cell motility; or preventing or otherwise reducing the risk of developing cell proliferation, cell survival, and/or cell motility, for example cancer, and in particular, breast cancer, colon cancer, lung cancer, prostate cancer, endometrial cancer, and stomach cancer, among others.

20

A "variant" of polypeptide, as used herein, refers to an amino acid sequence that is altered by one or more amino acids. Similarly, a variant antibody is altered by one or more amino acids. A variant polynucleotide is altered by one or more nucleotides. A variant may result in "conservative" changes, wherein a substituted amino acid has 25 similar structural or chemical properties, e.g., replacement of leucine with isoleucine. More rarely, a variant may result in "nonconservative" changes, e.g., replacement of a glycine with a tryptophan. Analogous minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without abolishing biological or immunogenic 30 activity may be found using computer programs well known in the art, for example, LASERGENE software (DNASTAR).

The invention also encompasses variants which retain at least one biological activity (e.g., effect on cell proliferation, cell survival, or cell motility) or immunogenic function. A preferred variant is one having at least 80%, and more preferably at least 90%, sequence identity to a disclosed sequence. A most preferred variant is one having
5 at least 95%, at least 97%, at least 98%, or at least 99% sequence identity to a sequence disclosed herein. The percentage identity is determined by aligning the two sequences to be compared as described below, determining the number of identical residues in the aligned portion, dividing that number by the total number of residues in the inventive (queried) sequence, and multiplying the result by 100. A useful alignment program is
10 AlignX (Vector NTI).

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology (including recombinant techniques), microbiology, cell biology, and biochemistry, which are within the skill of the art.
15 Such techniques are explained fully in the literature, such as, Molecular Cloning: A Laboratory Manual, 2nd edition, Sambrook et al., 1989; also, Molecular Cloning: A Laboratory Manual, 3rd Edition, Sambrook et al., 2000; Oligonucleotide Synthesis, MY Gait, ed., 1984; Animal Cell Culture, R.I. Freshney, ed., 1987; Methods in Enzymology, Academic Press, Inc.; Handbook of Experimental Immunology, 4th
20 edition, D .M. Weir & CC. Blackwell, eds., Blackwell Science Inc., 1987; Gene Transfer Vectors for Mammalian Cells, JAM. Miller & MAP. Calos, eds., 1987; Current Protocols in Molecular Biology, FEM. Ausubel et al., eds., 1987; and PCR: The Polymerase Chain Reaction, Mullis et al., eds., 1994.

25 Saratan

Saratan comprises a novel amino acid sequence that, to the inventors' knowledge, has not been previously published. The previously predicted sequence for Saratan (GenBank Acc. No. NM_020215) has been suppressed from the database as erroneous. The coding sequence for Saratan was previously calculated by a computer programme
30 to include a 5' untranslated region of 4,219 bp. This prior coding sequence was found to be unsupported by experimental data, and the protein sequence was suppressed by GenBank. In contrast, as determined by the inventors, the correct open reading frame for Saratan codes for a novel protein of 83 amino acids (Fig. 19A). With the corrected

data, the Saratan transcript is predicted as 7.5 kb (Fig. 19B; also designated as transcript a, see below). It is the inventors' understanding that there has been no prior public record of the corrected sequence information. As shown herein, Saratan is predicted to comprise a small protein with theoretical molecular mass of 8.8 kDa. According to the disclosed sequence analysis, the Saratan protein includes a large extracellular region at its N-terminus, followed by a single transmembrane region of 22 amino acids and a relatively short (4 amino acids) cellular region at C-terminus (Fig. 19A).

Database searching has revealed that, in addition to transcript a, there are at least three other alternatively spliced mRNAs. Transcript a is shown in Fig. 19B, and is represented by IMAGE clone 5493356 (GenBank accession number BC043593). The protein it encodes is designated herein as isoform α . Transcript b is shown in Fig. 19C, and is represented by EST clones BM967279, BI791809, BM352639, BI713575, BQ189729, and BM893938. Transcript b includes an extra intron as compared with transcript a. However, its coding region is the same as transcript a, and therefore the protein it encodes is the same as isoform α . Transcript c is also shown in Fig. 19C, and is represented by an EST clone DA249054. Transcript c includes an extra exon, in-frame, located in the 4.6 kb intron of transcript a. This exon encodes an additional sequence of 31 amino acids (NKDKSRISPSVCDPREDPELAPGKGCVWPGR; SEQ ID NO: 115). Thus, the full-length protein (designated as isoform β) is 114 amino acids (Fig. 19D). Transcript d is shown in Fig. 19C, and is represented by EST clone DA114412. Transcript d includes an extra exon situated in the intron of transcript a. However, this exon is not in-frame with the coding region of transcript a, and a downstream start codon is used. Transcript d is therefore predicted to encode a distinct protein (designated isoform γ) comprising 92 amino acids, with a divergent N-terminal sequence.

A database search using BlastP in NCBI revealed that Saratan protein shares significant homology with two predicted hypothetical proteins XP_875290 (*Bos taurus*) XP_001147850 (*Pan troglodytes*), XP_001476208 (*Mus musculus*), EDL81821 (*Rattus norvegicus*) and XP_001334647 (*Danio rerio*) at the corresponding N-terminal region (Fig. 19 E). The cDNA sequence of predicted mouse Saratan (GenBank accession

number XM_001476158) based on an EST clone AK052429 is shown in Fig. 20A. It is supported by numerous other EST clones such as CB521084, CD351912, CD350794, CF182588, CB245973, CB520382, CB246030, BU704198, BB354535, CA752314, CF532626, CF537731, BB659851, BQ921856, CJ162628, CJ176112, CJ164386, CJ169326, AA450691, CJ167344, AI481362, AI613651, CJ179881, BY004487, BY727205, BB644402, BB639576, CJ173864, CJ170896, and CJ159608. There is nominal sequence similarity shared by the human and mouse sequences within the 5' and 3'-UTR. However, we found 87% sequence identity (229/261) in the ORF regions, which indicates that the coding regions are highly conserved (Fig. 20B). The predicted mouse homolog shows the same length (83 amino acids) as human Saratan (Fig. 20C). At the protein level, the similarity is as high as 92%, and 74 out of 83 amino acids are identical (Fig. 20D).

In accordance with the findings disclosed herein, Saratan is highly or moderately expressed in several cancer cell lines, including cell lines for breast cancer, colon cancer, lung cancer, prostate cancer, endometrial cancer, and stomach cancer (see, Examples). Moreover, MCF-7 breast cancer cells overexpressing Saratan grow significant faster than control cells stably transfected by the empty vector phCMV1, as shown by total cell number assays over a period of time of 10 days. BrdU incorporation assays showed that the increase in total cell number was the result of increased cell proliferation. Further, this increase in cell proliferation could be effectively reversed by limiting the endogenous expression levels of Saratan using siRNA technology. siRNA reduced the expression of endogenous Saratan, resulting in a reduction in growth rate and inhibition of cell proliferation (see, Examples).

Thus, the data shown herein demonstrate that Saratan expression promotes the proliferation of MCF-7 breast carcinoma cells, and that this proliferation can be effectively inhibited by reducing the expression levels, of Saratan via siRNA. As such, Saratan represents an ideal novel target for cancer treatment and diagnosis, particularly for breast cancer. Any reagents that inhibit the biological activity of Saratan can be used to inhibit the proliferation, survival, and/or motility of cancer cells. These reagents can include, for example, chemical compounds (e.g., small molecules), antagonists, antibodies, and iRNAs. Similarly, diagnostic agents (e.g., polynucleotides and

antibodies) can be used to determine Saratan levels and detect a cancerous condition, e.g., cancer onset, progression, or recurrence.

Saratan polynucleotides and polypeptides

5 The invention encompasses Saratan polypeptides, including those comprising at least one of SEQ ID NO: 1, 3, 8-10, 49-64, 102-104, and fragments, variants, and derivatives thereof. The invention also encompasses the use of these polypeptides in the diagnosis of cancer, especially breast cancer. The invention further encompasses the use of the polypeptides for preparing antibodies to inhibit the cell proliferation, cell survival, or
10 cell motility of such cells.

The polypeptides of the present invention comprise at least one sequence selected from the group consisting of: (a) polypeptides comprising at least one amino acid sequence selected from the group consisting of SEQ ID NO:1, 3, 8-10, 49-64, 102-104, or
15 fragments, variants, or derivatives thereof; (b) polypeptides comprising a functional domain of at least one amino acid sequence selected from the group consisting of SEQ ID NO:1 3, 8-10, 49-64, 102-104, and fragments and variants thereof; and (c) polypeptides comprising at least a specified number of contiguous residues of at least one amino acid sequence selected from the group consisting of SEQ ID NO:1, 3, 8-10,
20 49-64, 102-104, or variants or derivatives thereof. In one particular embodiment, the invention encompasses an isolated Saratan polypeptide comprising the amino acid sequence of at least one of SEQ ID NO:1, 3, 8-10, 49-64, 102-104. All of these sequences are collectively referred to herein as polypeptides of the invention.

25 The polypeptides of the present invention may be expressed and used in various assays to determine their biological activity. The polypeptides may be used for large-scale synthesis and isolation protocols, for example, for commercial production. Such polypeptides may be used to raise antibodies, to isolate corresponding amino acid sequences, and to quantitatively determine levels of the amino acid sequences.

30

The invention encompasses Saratan polynucleotides including those of SEQ ID NO:14-24, 25-27, 28-30, 37, 38, 41, 42, 45-48, 65, 66, 68-73, 74-101, 120-122, and fragments, variants, and derivatives thereof. The invention also encompasses the use of these

polynucleotides in the diagnosis of cancer, especially breast cancer. The invention further encompasses the use of Saratan polynucleotides for the inhibition of cell proliferation, cell survival, or cell motility of such cells. Accordingly, the invention encompasses the use of these polynucleotides for preparing expression vectors and host
5 cells, and for preparing antisense polynucleotides and iRNAs.

The polynucleotides of the present invention comprise at least one sequence selected from the group consisting of: (a) sequences comprising a coding sequence for at least one amino acid sequence selected from the group consisting of SEQ ID NO:1, 3, 8-10,
10 49-64, 102-104, or fragments, variants, or derivatives thereof; (b) complements, reverse sequences, and reverse complements of a coding sequence for at least one amino acid sequence selected from the group consisting of SEQ ID NO:1, 3, 8-10, 49-64, 102-104, or fragments, variants, or derivatives thereof; (c) open reading frames contained in the coding sequence for at least one amino acid sequence selected from the group consisting
15 of SEQ ID NO:1, 3, 8-10, 49-64, 102-104, or their fragments, variants; or derivatives (d) functional domains of a coding sequence for at least one amino acid sequence selected from the group consisting of SEQ ID NO:1, 3, 8-10, 49-64, 102-104, or fragments, variants, or derivatives thereof; (e) sequences comprising at least a specified number of contiguous residues of a coding sequence for at least one amino acid
20 sequence selected from the group consisting of SEQ ID NO:1, 3, 8-10, 49-64, 102-104, or variants or derivatives thereof; and (f) sequences comprising at least a specified number of contiguous nucleotides of SEQ ID NO:14-24, 25-27, 28-30, 37, 38, 41, 42, 45-48, 65, 66, 68-73, 74-101, 120-122, or variants or derivatives thereof. In one particular embodiment, the invention encompasses an isolated polynucleotide
25 comprising a coding sequence for at least one amino acid sequence selected from the group consisting of SEQ ID NO:1-3, 5, 6, 8-10, 49-64, 102-104. In another particular embodiment, the invention encompasses an isolated polynucleotide comprising a sequence selected from the group consisting of SEQ ID NO:14-21, 23, 24-48, 49-64, 102-104. Oligonucleotide probes and primers and their variants are also provided. All of
30 these polynucleotides and oligonucleotide probes and primers are collectively referred to herein, as polynucleotides of the invention.

The isolated polynucleotides of the present invention also have utility in genome mapping, in physical mapping, and in cloning of genes of more or less related species. Probes designed using the polynucleotides of the present invention may be used to detect the presence and examine the expression patterns of genes in any organism
5 having sufficiently homologous DNA and RNA sequences in their cells, using techniques that are well known in the art, such as slot blot techniques or microarray analysis. Primers designed using the polynucleotides of the present invention may be used for sequencing and PCR amplifications. The polynucleotides of the present invention may also be used as compositions, for example, pharmaceutical compositions.
10 The polynucleotides of the present invention can also be used to provide health benefits. For such benefits, the polynucleotides can be presented as expression vectors or host cells comprising expression vectors.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of nucleotide sequences encoding the polypeptides of the
15 invention, some bearing minimal homology to the nucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of nucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in
20 accordance with the standard triplet genetic code as applied to naturally occurring amino acid sequences, and all such variations are to be considered as being specifically disclosed.

Nucleotide sequences which encode the Saratan polypeptides, or their fragments or
25 variants, are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring sequence under appropriately selected conditions of stringency. However, it may be advantageous to produce nucleotide sequences encoding a polypeptide, or its fragment or derivative, possessing a substantially different codon usage. Codons may be selected to increase the rate at which expression of the
30 polypeptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding polypeptides and its derivatives without altering the encoded amino acid sequences include the production of RNA

transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of Saratan polynucleotides, or fragments
5 variants, or derivatives thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents that are well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding a Saratan polypeptide, or any variants, fragments, or derivatives thereof. Also encompassed by the
10 invention are polynucleotide sequences that are capable of hybridizing to the disclosed nucleotide sequences, and in particular, those shown in SEQ ID NO:14-16, 19, 22, under various conditions of stringency as taught in Wahl, G. M. and S. L. Berger (1987; Methods Enzymol. 152:399-407) and Kimmel, A. R. (1987; Methods Enzymol. 152:507-511).

15 Methods for DNA sequencing which are well known and generally available in the art may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (U.S. Biochemical Corp, Cleveland, OH), Taq polymerase (Perkin Elmer), thermostable
20 T7 polymerase (Amersham Pharmacia Biotech, Piscataway, NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE Amplification System (Life Technologies, Gaithersburg, MD). Preferably, the process is automated with machines such as the Hamilton Micro Lab 2200 (Hamilton, Reno, NV), Peltier Thermal Cycler (PTC200; MJ Research, Watertown, MA) the ABI Catalyst and
25 373 and 377 DNA Sequencers (Perkin Elmer), or the Genome Sequencer 20TM (Roche Diagnostics).

The nucleic acid sequences encoding the Saratan polypeptides may be extended utilizing a partial nucleotide sequence and employing various methods known in the art
30 to detect upstream sequences such as promoters and regulatory elements. For example, one method which may be employed, "restriction-site" PCR, uses universal primers to retrieve unknown sequence adjacent to a known locus (Sarkar, G. (1993) PCR Methods Applic. 2:318-322). In particular, genomic DNA is first amplified in the presence of

primer to a linker sequence and a primer specific to the known region. The amplified sequences are then subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different fluorescent dyes (one for each nucleotide), which are laser activated, and detection of the emitted wavelengths by a charge coupled device camera. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and Sequence NAVIGATOR, Perkin Elmer) and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for the sequencing of small pieces of DNA which might be present in limited amounts in a particular sample.

In another embodiment of the invention, Saratan polynucleotides or fragments thereof which encode polypeptides may be used in recombinant DNA molecules to direct expression of the polypeptides, or fragments or variants thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced, and these sequences may be used to clone and express Saratan polypeptides. The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter amino acid encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, site-directed mutagenesis may be used to insert new restriction sites, alter glycosylation patterns, change codon preference, introduce mutations, and so forth.

In another embodiment of the invention, a natural, modified, or recombinant nucleic acid sequence encoding a polypeptide may be ligated to a heterologous sequence to encode a fusion protein. For example, it may be useful to encode a chimeric sequence that can be recognized by a commercially available antibody. A fusion protein may also
5 be engineered to contain a cleavage site located between the polypeptide of the invention and the heterologous protein sequence, so that the polypeptide may be cleaved and purified away from the heterologous moiety.

In another embodiment, sequences encoding polypeptides may be synthesized, in whole
10 or in part, using chemical methods well known in the art (see Caruthers, M. H. et al. (1980) Nucl. Acids Res. Symp. Ser. 215-223, Horn, T. et al. (1980) Nucl. Acids Res. Symp. Ser. 225-232). Alternatively, the polypeptide itself may be produced using chemical methods to synthesize the amino acid sequence, or a fragment thereof. For example, polypeptide synthesis can be performed using various solid-phase techniques
15 (Roberge, J. Y. et al. (1995) Science 269:202-204; Merrifield J. (1963) J. Am. Chem. Soc. 85:2149-2154) and automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer). Various fragments of polypeptides may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

20

The newly synthesized polypeptide may be isolated by preparative high performance liquid chromatography (e.g., Creighton, T. (1983) Proteins Structures and Molecular Principles, WH Freeman and Co., New York, NY). The composition of the synthetic polypeptides may be confirmed by amino acid analysis or sequencing (e.g., the Edman
25 degradation procedure; Creighton, supra). Additionally, the amino acid sequence of the polypeptide, or any part thereof, may be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins, or any part thereof, to produce a variant molecule.

30 In order to express a biologically active polypeptides, the nucleotide sequences encoding the polypeptide or functional equivalents, may be inserted into appropriate expression vector, e.g., a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods which are well

known to those skilled in the art may be used to construct expression vectors containing sequences encoding the polypeptide and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. Such techniques are described
5 in Sambrook, J. et al. (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Plainview, NY; also, Sambrook, J. et al. (2000) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Plainview, NY; and Ausubel, F. M. et al. (1989) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, NY.

10

A variety of expression vector/host systems may be utilized to contain and express sequences encoding the polypeptides of the invention. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant phage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors;
15 insect cell systems infected with virus expression vectors (e.g., baculovirus); plant cell systems transformed with virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. For bacteria, useful plasmids include pET, pRSET, pTrcHis2, and pBAD plasmids from Invitrogen, pET and pCDF plasmids from
20 Novagen, and DirectorTM plasmids from Sigma-Aldrich. In particular, *E. coli* can be used with the expression vector pET. The invention is not limited by the expression vector or host cell employed.

The "control elements" or "regulatory sequences" are those non-translated regions (e.g.,
25 enhancers, promoters, 5' and 3' untranslated regions) which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible
30 promoters such as the hybrid lacZ promoter of the BLUESCRIPT phagemid (Stratagene, LaJolla, CA) or pSPORT1 plasmid (Life Technologies) and the like may be used. The baculovirus polyhedrin promoter may be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (e.g., heat shock, RUBISCO, and

storage protein genes) or from plant viruses (e.g., viral promoters or leader sequences) may be cloned into the vector.

5 In bacterial systems, a number of expression vectors may be selected depending upon the use intended for the polypeptide. For example, when large quantities of polypeptide are needed, vectors which direct high level expression of fusion proteins that are readily purified may be used. Such vectors include, but are not limited to, the multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene), in which the sequence encoding a polypeptide may be ligated into the vector in frame with sequences
10 for the amino-terminal Met and the subsequent 7 residues of β -galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke, G. and S. M. Schuster (1989) J. Biol. Chem. 264:5503-5509); and the like.

pGEX vectors (Promega, Madison, WI) may also be used to express the polypeptides as
15 fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems may be designed to include heparin, thrombin, or factor Xa protease cleavage sites so that the cloned polypeptide of interest can be released from the GST
20 moiety at will. In the yeast, *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. For reviews, see Ausubel et al. (supra) and Grant et al. (1987) Methods Enzymol. 153:516-544.

25 Specific initiation signals may also be used to achieve more efficient translation of sequences encoding the polypeptides of the invention. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding a polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control
30 signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including the ATG initiation codon should be provided. Furthermore, the initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements

and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used, such as those described in the literature (Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162).

5

In addition, a host cell may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed polypeptide in the desired fashion. Such modifications of the sequence include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing
10 which cleaves a "prepro" form of the polypeptide may also be used to facilitate correct insertion, folding, and/or function. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities are available from the American Type Culture Collection (ATCC; Bethesda, MD) and may be chosen to ensure the correct modification and processing of the sequence. Specific host cells
15 include, but are not limited to, *Rhodotorula*, *Aureobasidium*, *Saccharomyces*, *Sporobolomyces*, *Pseudomonas*, *Erwinia* and *Flavobacterium*; or such other organisms as *Escherichia*, *Lactobacillus*, *Bacillus*, *Streptomyces*, and the like. Particular host cells include *Escherichia coli*, which is particularly suited for use with the present invention, *Saccharomyces cerevisiae*, *Bacillus thuringiensis*, *Bacillus subtilis*, *Streptomyces*
20 *lividans*, and the like.

There are several techniques for introducing nucleic acids into eukaryotic cells cultured *in vitro*. These include chemical methods (Felgner et al., Proc. Natl. Acad. Sci., USA, 84:7413 7417 (1987); Bothwell et al., Methods for Cloning and Analysis of Eukaryotic
25 Genes, Eds., Jones and Bartlett Publishers Inc., Boston, Mass. (1990), Ausubel et al., Short Protocols in Molecular Biology, John Wiley and Sons, New York, NY (1992); and Farhood, Annal. NY Acad. Sci., 716:23 34 (1994)), use of protoplasts (Bothwell, supra) or electrical pulses (Vatteroni et al., Mutn. Res., 291:163 169 (1993); Sabelnikov, Prog. Biophys. Mol. Biol., 62: 119 152 (1994); Bothwell et al., supra; and
30 Ausubel et al., supra), use of attenuated viruses (Davis et al., J. Virol. 1996, 70(6), 3781 3787; Brinster et al. J. Gen. Virol. 2002, 83(Pt 2), 369 381; Moss, Dev. Biol. Stan., 82:55 63 (1994); and Bothwell et al., supra), as well as physical methods (Fynan et al.,

supra; Johnston et al., *Meth. Cell Biol.*, 43(Pt A):353 365 (1994); Bothwell et al., supra; and Ausubel et al., supra).

Successful delivery of nucleic acids to animal tissue can be achieved by cationic liposomes (Watanabe et al., *Mol. Reprod. Dev.*, 38:268 274 (1994)), direct injection of naked DNA or RNA into animal muscle tissue (Robinson et al., *Vacc.*, 11:957 960 (1993); Hoffman et al., *Vacc.* 12:1529 1533; (1994); Xiang et al., *Virol.*, 199:132 140 (1994); Webster et al., *Vacc.*, 12:1495 1498 (1994); Davis et al., *Vacc.*, 12:1503 1509 (1994); Davis et al., *Hum. Molec. Gen.*, 2:1847 1851 (1993); Dalemans et al. *Ann NY Acad. Sci.* 1995, 772, 255 256. Conry, et al. *Cancer Res.* 1995, 55(7), 1397-1400), and embryos (Naito et al., *Mol. Reprod. Dev.*, 39:153 161 (1994); and Burdon et al., *Mol. Reprod. Dev.*, 33:436 442 (1992)), intramuscular injection of self replicating RNA vaccines (Davis et al., *J. Virol.* 1996, 70(6), 3781 3787; Balasuriya et al. *Vaccine* 2002, 20(11 12), 1609 1617) or intradermal injection of DNA using "gene gun" technology (Johnston et al., supra).

A variety of protocols for detecting and measuring the expression of the polypeptides of the invention, using either polyclonal or monoclonal antibodies specific for the protein are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay can be used with monoclonal antibodies reactive to two non-interfering epitopes on the polypeptide, but a competitive binding assay can also be used. These and other assays are described, among other places, in Hampton, R. et al. (1990; *Serological Methods, a laboratory Manual*, APS Press, St Paul, MN) and Maddox, D. E. et al. (1983; *J. Exp. Med.* 158:1211-1216).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding the polypeptides, or any fragments or variants thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may

be used to synthesize RNA probes *in vitro* by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits Amersham Pharmacia Biotech, Promega, and US Biochemical. Suitable reporter molecules or labels, which may be
5 used for ease of detection, include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Expression vectors or host cells transformed with expression vectors may be cultured
10 under conditions suitable for the expression and recovery of the polypeptide from culture. The culture can comprise components for *in vitro* or *in vivo* expression. *In vitro* expression components include those for rabbit reticulocyte lysates, *E. coli* lysates, and wheat germ extracts, for example, Expressway™ or RiPs systems from Invitrogen, Genelator™ systems from iNtRON Biotechnology, EcoPro™ or STP3™ systems from
15 Novagen, TNT® Quick Coupled systems from Promega, and EasyXpress systems from QIAGEN. The polypeptide produced from culture may be secreted or contained intracellularly depending on the sequence and/or the vector used. In particular aspects, expression vectors which encode a phage polypeptide can be designed to contain signal sequences which direct secretion of the polypeptide through a prokaryotic or eukaryotic
20 cell membrane.

Other constructions may include an amino acid domain which will facilitate purification of the polypeptide. Such domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan (e.g., 6×-HIS) modules that allow purification on
25 immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAG® extension/affinity purification system (Immunex Corp., Seattle, WA). Useful epitope tags include 3X-FLAG®, HA, VSV-G, V5, HSV, GST, GFP, MBP, GAL4, and β -galactosidase. Useful plasmids include those comprising a biotin tag (e.g., PinPoint™ plasmids from Promega),
30 calmodulin binding protein (e.g., pCAL plasmids from Stratagene), streptavidin binding peptide (e.g., InterPlay™ plasmids from Stratagene), a c-myc or FLAG® tag (e.g., Immunoprecipitation plasmids from Sigma-Aldrich), or a histidine tag (e.g., QIAExpress plasmids from QIAGEN).

To facilitate purification, expression vectors can include a cleavable linker sequences such as those specific for Factor Xa or enterokinase (Invitrogen, San Diego, CA). For example, the vector can include one or more linkers between the purification domain
5 and the polypeptide. One such expression vector provides for expression of a fusion protein comprising a polypeptide of the invention and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMAC (immobilized metal ion affinity chromatography as described in Porath, J. et al. (1992) Prot. Exp. Purif. 3: 263-281)
10 while the enterokinase cleavage site provides a means for purifying the polypeptide from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D. J. et al. (1993; DNA Cell Biol. 12:441-453).

Polynucleotides for inhibition of Saratan

15 As mentioned herein before, in one embodiment, polynucleotides may be utilised to inhibit Saratan in accordance with the invention. Such polynucleotides may be DNA, RNA, single stranded, or double stranded. Polynucleotides for use with the invention may be referred to herein as "isolated" polynucleotides. Isolated Saratan polynucleotides may be obtained using a number of techniques known in the art. For
20 example, recombinant DNA technology may be used as described for example in Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual (2nd edition), Cold Spring Harbor Laboratory Press, Plainview, NY and Sambrook and Russell (2000) Molecular Cloning: A Laboratory Manual (3rd Edition), Cold Spring Harbor Laboratory Press, Plainview, NY. Similarly, chemical synthesis, for example, using
25 phosphoramidite and solid phase chemistry, may be used. Polynucleotides may be designed on the basis of the disclosed Saratan nucleic acid sequence data, the known relative interactions between nucleotide bases, known sequence homology, and the particular nucleic acid technology to be employed, as may be exemplified herein after.

30 In one embodiment, interfering RNAs (iRNAs or siRNAs) may be utilised to inhibit Saratan. Polynucleotides of use in iRNA techniques will typically have 100% complementarity to their target. However, it should be appreciated that this need not be the case, provided the iRNA retains specificity for its target and the ability to block

translation. Exemplary iRNA molecules may be in the form of ~18 to 21 bp double stranded RNAs with 3' dinucleotide overhangs, although shorter or longer molecules may be appropriate. In cases where the iRNA is produced *in vivo* by an appropriate nucleic acid vector, it will typically take the form of an RNA molecule having a stem-loop structure, for example, an approximately 19 nucleotide stem and a 9 nucleotide loop with 2-3 Us at the 3' end. Algorithms of use in designing siRNA are available from Cenix (Dresden, Germany, via Ambion, TX).

Exemplary siRNA molecules can include the following sequences:

10

Target sequence 1:	5'- AAATAATCACTCGTGTGACTT -3'	SEQ ID NO:16
Sense siRNA strand:	5'- AUAAUCACUCGUGUGACUUt -	SEQ ID NO:17
Antisense strand:	3' -	SEQ ID NO:18
	3'- ttUAUUAGUGAGCACACUGAA -5'	
Sense sequence		SEQ ID NO:68
Antisense sequence	AUAAUCACUCGUGUGACUU	SEQ ID NO:69
	UAUUAGUGAGCACACUGAA	
Target sequence 2:	5'- AATTACAAGAAAGCTGGACTT -3'	SEQ ID NO:19
Sense siRNA strand:	5'- UUACAAGAAAGCUGGACUUt -	SEQ ID NO:20
Antisense strand:	3' -	SEQ ID NO:21
	3'- ttAAUGUUCUUUCGACCUGAA -5'	
Sense sequence		SEQ ID NO:70
Antisense sequence	UUACAAGAAAGCUGGACUU	SEQ ID NO:71
	AAUGUUCUUUCGACCUGAA	
Target sequence 3:	5'- AAGAGCAACTTACTGCTTTGT -3'	SEQ ID NO:22
Sense siRNA strand:	5'- GAGCAACUUACUGCUUUGUt -	SEQ ID NO:23
Antisense strand:	3' -	SEQ ID NO:24
	3'- ttCUCGUUGAAUGACGAAACA -5'	
Sense sequence		SEQ ID NO:72
Antisense sequence	GAGCAACUUACUGCUUUGU	SEQ ID NO:73
	CUCGUUGAAUGACGAAACA	

Exemplary iRNA molecules can include the structures described as follows.

15

In one aspect, the iRNA comprises a nucleotide sequence selected from the group consisting of:

- 1) 5'-ATAATCACTCGTGTGACTT-3' (SEQ ID NO:25)
- 2) 5'-TTACAAGAAAGCTGGACTT-3' (SEQ ID NO:26)
- 3) 5'-GAGCAACTTACTGCTTTGT-3' (SEQ ID NO:27)

20

In one preferred aspect, the iRNA comprises a nucleotide sequence selected from the group consisting of:

5' -ATAATCACTCGTGTGACTT**CTCAAGAGA**AAAGTCACACGAGTGATTAT-3'
 |---Sense Strand---|--Loop--|--Antisense Strand-|
 (SEQ ID NO:28)

5' -TTACAAGAAAGCTGGACTT**CTCAAGAGA**AAAGTCCAGCTTTCTTGTA-3'
 |---Sense Strand---|--Loop--|--Antisense Strand-|
 (SEQ ID NO:29)

5' -GAGCAACTTACTGCTTTGTT**CTCAAGAGA**ACAAAGCAGTAAGTTGCTC-3'
 |---Sense Strand---|--Loop--|--Antisense Strand-|
 (SEQ ID NO:30)

In another preferred aspect, the iRNA comprises a nucleotide sequence selected from the group consisting of:

5' -XXXXATAATCACTCGTGTGACTT**CTCAAGAGA**AAAGTCACACGAGTGATTATXXXX-3'
 |---Sense Strand---|--Loop--|--Antisense Strand-|
 (SEQ ID NO:31)

5' -XXXXGTTACAAGAAAGCTGGACTT**CTCAAGAGA**AAAGTCCAGCTTTCTTGTAAXXXX-3'
 |---Sense Strand---|--Loop--|--Antisense Strand-|
 (SEQ ID NO:32)

5' -XXXXGAGCAACTTACTGCTTTGTT**CTCAAGAGA**ACAAAGCAGTAAGTTGCTCXXXX-3'
 |---Sense Strand---|--Loop--|--Antisense Strand-|
 (SEQ ID NO:33)

"X" indicates any number of additional nucleotides which may be present; for example termination signals and restriction sites which may be of use in cloning and expressing the iRNA.

By way of non-limiting examples, the following nucleic acids may be used to clone and express, in desired vectors, the iRNAs of use in the invention:

40 *Bam*HI *Hind*
 III
 5' -ggatccATAATCACTCGTGTGACTT**CTCAAGAGA**AAAGTCACACGAGTGATTATTTTTT**GGAA**aagtcc-
 3'
 | Sense Strand | Loop | Antisense Strand | Terminator
 45 (SEQ ID NO:34)
*Bam*HI *Hind*
 III

5

15

20

35

completely complementary to their targets may be utilised provided they retain specificity for their target and the ability to inhibit expression.

Persons of ordinary skill in the art to which the invention relates will appreciate
5 antisense molecules of use in the invention having regard to the description provided
herein, and available Saratan sequence data. Further information regarding antisense
technology can be gained, for example, from: Kandimalla ER, Manning A, Lathan C,
Byrn RA, Agrawal S. Design, biochemical, biophysical and biological properties of
10 cooperative antisense oligonucleotides; *Nucleic Acids Res.* 1995 Sep 11;23(17):3578-
84; Tseng BY, Brown KD. Antisense oligonucleotide technology in the development of
cancer therapeutics; *Cancer Gene Ther.* 1994 Mar;1(1):65-71; Brysch W,
Schlingensiepen KH. Design and application of antisense oligonucleotides in cell
culture, *in vivo*, and as therapeutic agents; *Cell Mol Neurobiol.* 1994 Oct;14(5):557-68;
15 Han J, Zhu Z, Hsu C, Finley WH. Selection of antisense oligonucleotides on the basis of
genomic frequency of the target sequence; *Antisense Res Dev.* 1994 Spring;4(1):53-65.

It should be appreciated that DNAzymes, single stranded DNA, ribozymes, and triple
helix DNA may also be of use in inhibiting Saratan in accordance with the invention.
Ribozymes, DNAzymes, triple helix, and single stranded DNA may be readily
20 appreciated by persons of general skill in the art to which the invention relates having
regard to the description provided herein, available Saratan sequence data and current
methodologies. However, by way of example methodology associated with these
technologies is described in Joseph Sambrook and David W. Russell. *Molecular
Cloning: A Laboratory Manual (Third Edition)*, Cold Spring Harbor Laboratory Press,
25 NY.

Polynucleotides of use in the invention, including antisense, iRNA, ribozymes and
DNAzymes may be chemically modified to increase stability or prevent degradation or
otherwise. For example, the nucleic acid molecules may include analogs with unnatural
30 bases, modified sugars, especially at the 2' position of the ribose, or altered phosphate
backbones. Polynucleotides of use in the invention may also include sequences which
allow for targeted degradation of any transcript to which they bind. For example, a
sequence specific for RNase H, may be included. Another example is the use of

External Guide Sequences (EGSs), which may recruit a ribozyme (RNase P) to digest the transcript to which an antisense molecule is bound.

5 One can help ensure specificity of the likes of antisense oligonucleotides, iRNAs, ribozymes, DNazymes, and cDNAs by screening candidate sequences for homology with other sequences in the transcriptome, the full complement of activated genes, mRNAs, or transcripts in a particular cell. Also, skilled persons may appreciate appropriate algorithms of use in designing and ensuring specificity of such polynucleotides.

10

Polynucleotides of use in the invention may be used in the form of nucleic acid molecules produced *in vitro*, for example, single stranded DNA, iRNA, antisense RNA, or DNazymes. Alternatively, where appropriate, they may be used in the form of a vector adapted to produce appropriate nucleic acids, for example, antisense molecules, 15 iRNA, or ribozymes. The inventors contemplate the use of any vectors as may be known in the art. For example, naked plasmids that employ CMV promoters may be used. Viral vectors may also be suitable, such as adeno-associated virus (AAV) and lentiviruses. Other examples of suitable promoters and viral vectors are provided herein after. One advantage of using such viral vectors is that they may allow for systemic 20 administration, as opposed to localised administration to a tissue or tumour.

The vectors or constructs of use in the invention may include appropriate genetic elements, such as promoters, enhancers, origins of replication as are known in the art, including inducible, constitutive, or tissue-specific promoters. In a specific embodiment, 25 a vector comprises an inducible promoter operably linked to the region coding a nucleic acid of the invention (for example, antisense Saratan or suitable siRNA), such that expression of the nucleic acid is controllable by controlling the presence or absence of the appropriate inducer of transcription. In another embodiment, nucleic acid molecules of the invention are flanked by regions that promote homologous recombination at a 30 desired site in the genome, thus providing for intrachromosomal integration of the desired nucleic acids (Koller and Smithies, 1989, Proc. Natl. Acad. Sci. USA 86:8932-8935; Zijlstra et al., 1989, Nature 342:435-438). Of course, the vectors may remain extrachromosomal.

In another embodiment of the invention, PNAs are used. PNAs are peptide-nucleic acid hybrids in which the phosphate backbone has been replaced by an achiral and neutral backbone made from N-(2-aminoethyl)-glycine units (see, e.g., Eureka Bioscience
5 Collection. PNA and Oligonucleotide Inhibitors of Human Telomerase. G. Gavory and S. Balasubramanian, Landes Bioscience, 2003). The bases A, G, T, C are attached to the amino nitrogen on the backbone via methylenecarbonyl linkages (P.E. Nielsen et al., Science 1991. 254: 1497-1500; M. Egholm et al., Nature 1993. 365: 566-568). PNAs
10 bind complementary sequences with high specificity, and higher affinity relative to analogous DNA or RNA (M. Egholm et al., supra). PNA/DNA or PNA/RNA hybrids also exhibit higher thermal stability compared to the corresponding DNA/DNA or DNA/RNA duplexes (M. Egholm et al., supra). PNAs also possess high chemical and biological stability, due to the unnatural amide backbone that is not recognized by
15 nucleases or proteases (V. Demidov et al., Biochem Pharmacol 1994. 48: 1310-1313). Typically, PNAs are at least 5 bases in length, and include a terminal lysine. PNAs may be pegylated to further extend their lifespan (Nielsen, P. E. et al. (1993) Anticancer Drug Des. 8:53-63).

As non-limiting examples, the following antigene PNAs are provided as complementary
20 to unique sequences in the coding DNA strand of the first exon of the Saratan gene, and are designed to inhibit mRNA synthesis.

1) H-AGTGCGCCGTGCGGTCT-NH₂ (SEQ ID NO:37)

2) H-TCTCCGGACGCTCGCTG-NH₂ (SEQ ID NO:38)

25 **Antibodies for inhibition of Saratan**

The invention encompasses Saratan antibodies, for example, antibodies that bind to at least a portion of or a modified sequence thereof. In certain aspects of the invention, antibodies may be used to inhibit Saratan. It will be understood that, for the purposes of the invention, a fragment or derivative of an antibody need not act fully as an antibody;
30 that is to say, the fragment or derivative need not be capable of recruiting immune system cells to the site of binding to Saratan *in vivo*. It is not necessary to produce neutralising antibodies. Those of ordinary skill in the art to which the invention relates

will recognise methods to generate antibody fragments. However, by way of general example proteolytic digestion of intact antibodies may be used, or the fragments may be directly produced via recombinant nucleic acid technology.

- 5 Humanisation of antibodies may be used to reduce the immunogenicity of antibodies generated in other animals. Production of humanised antibodies or humanization of antibodies can be achieved using techniques known in the art, for example in the case of humanisation of murine antibodies by epitope-guided selection. The most frequently used strategies for the humanization of rodent monoclonal antibodies are CDR grafting
10 (Reichmann, L., M. Clark, H. Waldman, and G. Winter. 1998. Reshaping human antibodies for therapy. *Nature* 332L323-327) and resurfacing (Pedersen, J. T., A. H. Henry, S. J. Searle, B. C. Guild, M. Roguska, and A. R. Rees. 1994. Comparison of surface accessible residues in human and murine immunoglobulin Fv domains. Implication for humanization of murine antibodies. *J. Mol. Biol.* 235:959-973). These
15 are based on a rational design approach and iterative optimization, i.e., site-directed mutagenesis of framework residues aided by computer modelling.

- Other selective humanization strategies using phage display may be used (Baca, M., L. G. Presta, S. J. O'Connor, and J. A. Wells. 1997. Antibody humanization using
20 monovalent phage display. *J. Biol. Chem.* 272:10678-10684; Hoogenboom, H. R., A. P. de Bruine, S. E. Hufton, R. M. Hoet, J. W. Arends, and R. C. Roovers. 1998. Antibody phage display technology and its applications. *Immunotechnology*. 4:1-20; Jespers, L. S., A. Roberts, S. M. Mahler, G. Winter, and H. R. Hoogenboom. 1994. Guiding the selection of human antibodies from phage display repertoires to a single epitope of an
25 antigen. *Biotechnology (NY)* 12:899-903; Rader, C. and C. F. Barbas, III. 1997. Phage display of combinatorial antibody libraries. *Curr. Opin. Biotechnol.* 8:503-508; Rader, C., D. A. Cheresch, and C. F. Barbas, III. 1998. A phage display approach for rapid antibody humanization: designed combinatorial V gene libraries. *Proc. Natl. Acad. Sci. U. S. A* 95:8910-8915; Rosok, M. J., D. E. Yelton, L. J. Harris, J. Bajorath, K. E.
30 Hellstrom, I. Hellstrom, G. A. Cruz, K. Kristensson, H. Lin, W. D. Huse, and S. M. Glaser. 1996. A combinatorial library strategy for the rapid humanization of anticarcinoma BR96 Fab. *J. Biol. Chem.* 271:22611-22618).

Those of skill in the art to which the invention relates will appreciate the terms “diabodies” and “triabodies”. These are molecules which comprise a heavy chain variable domain (VH) connected to a light chain variable domain (VL) by a short peptide linker that is too short to allow pairing between the two domains on the same chain. This promotes pairing with the complementary domains of one or more other chains and encourages the formation of dimeric or trimeric molecules with two or more functional antigen binding sites. The resulting antibody molecules may be monospecific or multispecific (e.g., bispecific in the case of diabodies). Such antibody molecules may be created from two or more antibodies using methodology standard in the art to which the invention relates; for example, as described by Todorovska et al. (Design and application of diabodies, triabodies and tetrabodies for cancer targeting. *J. Immunol. Methods.* 2001 Feb 1;248(1-2):47-66).

The production of antibodies may be carried out according to standard methodology in the art. For example, in the case of the production of polyclonal antibodies the methodology described by Bean (Eric S. Bean (2001) *Polyclonal Antibodies*. In: *Basic Methods in Antibody Production and Characterization* antibodies. Howard, G, and Bethel D. (ed.), CRC Press, 5:21-50, 2000) may be used. Monoclonal antibodies may be prepared, for example, in accordance with the methodology of Stewart (Sandy J. Stewart (2001) *Monoclonal Antibody Production*. In: *Basic Methods in Antibody Production and Characterization* antibodies. Howard, G. and Bethel D. (ed.), CRC Press, 6:51-68, 2000) or in “*Monoclonal Antibody Production Techniques and Applications*,” Lawrence B Schook eds., Marcel Dekker Inc., New York, 1987.

Standard methods can be used to identify amino acid sequences useful for antibody production. Antigenic segments of a polypeptide can be predicted, for example, by Abie Pro 3.0: Peptide Antibody Design (hypertext transfer protocol://world wide web.changbioscience.com/abie/abie.html). More particularly, antigenic segments can be predicted according to the Hopp-Woods scale (Hopp TP, Woods KR. Prediction of protein antigenic determinants from amino acid sequences. *Proc Natl Acad Sci U S A.* 1981 Jun;78(6):3824-8.) and/or the Kyte and Doolittle scale (Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J Mol Biol.* 1982 May 5;157(1):105-32.). Such epitopes may be conformational specific, in that they may

include non-contiguous residues, and may constitute various portions of the predicted antigenic sequences (e.g., portions of any one of SEQ ID NO:49-64, see below), or a combination of one or more portions of the predicted antigenic sequences (e.g., combinations of one or more of SEQ ID NO:49-64), or one or more of the full length sequences (e.g., one or more of SEQ ID NO:49-64). Antigenic sequences can comprise any combination of amino acids or their derivatives that would form a similar 3-D structure (e.g., surface residues) as would be encountered in the native polypeptide.

In Saratan, sequence analysis shows two highly antigenic sequences, including NVHAAANGQGQPEDPPRSSNDAVL (residues 37-60, SEQ ID NO:49) and FMDSPNEDFSTEYSLF (residues 17-32, SEQ ID NO:50). Fourteen antigenic sequences of 14 amino acids are also included: FMDSPNEDFSTEYS (residues 17-30, SEQ ID NO:51), MDSPNEDFSTEYSL (residues 18-31, SEQ ID NO:52), DSPNEDFSTEYSLF (residues 19-32, SEQ ID NO:53), NVHAAANGQGQPED (residues 37-50, SEQ ID NO:54), VHAAANGQGQPEDP (residues 38-51, SEQ ID NO:55), HAAANGQGQPEDPP (residues 39-52, SEQ ID NO:56), AAANGQGQPEDPPR (residues 40-53, SEQ ID NO:57), AANGQGQPEDPPRS (residues 41-54, SEQ ID NO:58), ANGQGQPEDPPRSS (42-55, SEQ ID NO:59), NGQGQPEDPPRSSN (residues 43-56, SEQ ID NO:60), GQGQPEDPPRSSND (residues 44-57, SEQ ID NO:61), QGQPEDPPRSSNDA (residues 45-58, SEQ ID NO:62), GQPEDPPRSSNDAV (residues 46-59, SEQ ID NO:63), and QPEDPPRSSNDAVL (residues 47-60, SEQ ID NO:64).

Antibodies of use in the invention may also be produced via standard recombinant techniques and as described for example by Welschof et al. (Welschof, M., C. Christ, I. Hermes, A. Keller, C. Kleist, and M. Braunagel. 2003. Generation and screening of a modular human scFv expression library from multiple donors. *Methods Mol. Biol.* 207:103-121). The inventors consider recombinant techniques to be a preferable means of producing antibodies on a commercial scale. Nucleic acids encoding an antibody may be readily identified on the basis of the amino acid sequence of the antibody, the genetic code, and the understood degeneracy therein. Nucleic acids encoding antibodies may be isolated from hybridoma cells, for example, and subsequently characterised using procedures standard in the art. For example, a nucleic acid probe may be designed based

on the amino acid sequence of a portion of an antibody and then used to isolate genes encoding the heavy and/or light chains of the antibody. Alternatively, nucleic acids may be generated by standard chemical synthesis methodology, for example, using phosphoramidite and solid phase chemistry. The amino acid sequence of an antibody of the invention may be determined using standard methodology; for example, Edman degradation and HPLC or mass spectroscopy analysis, may be used.

In addition to therapeutic use of antibodies directed against Saratan, they may find use in purification of Saratan or in diagnostic applications. For example, the antibodies may be immobilised on a solid phase. This would aid in purification and/or quantitation of the level of Saratan in a sample. In the case of diagnostic procedures and purification of Saratan it is not necessary for the antibody to have to inhibit Saratan activity. Additionally, as may be useful for certain applications, antibodies may be modified by labelling with a compound which provides a detectable signal; for example, enzymes, fluorescent agents, and radioisotopes can be used. Those of general skill in the art to which the invention relates will readily identify such suitable labelling systems. Additionally, antibodies may be used as carriers, for example to carry toxins, radionucleotides, isotopes, genes, or other therapeutic molecules to cells or tissues to aid in therapy.

Persons of ordinary skill in the art will readily appreciate methods for determining the efficacy of an antibody in preventing, decreasing, or inhibiting cell proliferation, cell survival, or cell motility. However, by way of example, the methodology described elsewhere herein, including one or more of the assays referred to in the "Examples" section, may be used.

Compositions for inhibition of Saratan

The agents of use in inhibiting Saratan may be used on their own, or in the form of compositions in combination with one or more pharmaceutically acceptable diluents, carriers, and/or excipients. Those skilled in the art to which the invention relates will readily appreciate a variety of pharmaceutically acceptable diluents, carriers, and/or excipients which may be employed in compositions of the invention. As will be appreciated, the choice of such diluents, carriers, and/or excipients will be dictated to

some extent by the nature of the agent to be used, the intended dosage form of the composition, and the mode of administration. By way of example, in the case of administration of polynucleotides, such as vectors adapted to express antisense or iRNA, suitable carriers include isotonic solutions, water, aqueous saline solution, aqueous dextrose solution, and the like.

In addition to standard diluents, carriers, and/or excipients, a pharmaceutical composition of the invention may be formulated with additional constituents, or in such a manner, so as to enhance the activity of the agent or help protect the integrity of the agent. For example, the composition may further comprise adjuvants or constituents which provide protection against degradation, or decrease antigenicity of an agent, upon administration to a subject. Alternatively, the agent may be modified so as to allow for targeting to specific cells, tissues, or tumours.

Agents may be formulated to incorporate a sustained-release system. Inasmuch as this is the case, compositions may include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or microcapsules. Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919; EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman et al., 1983, Biopolymers: 22: 547-56), poly(2-hydroxyethyl methacrylate) (Langer et al., 1981, J. Biomed. Mater. Res.: 15: 267), ethylene vinyl acetate (Langer et al., 1981, J. Biomed. Mater. Res.: 15: 267), or poly-D-(-)-3-hydroxybutyric acid (EP 133,988).

Agents of the invention may also be formulated into liposomes. Liposomes comprising the compound may be prepared using techniques known in the art to which the invention relates. By way of example see: DE 3,218,121, EP 52,322, EP 36,676, EP 88,046, EP 143,949, EP 142,641, Japanese Pat. Appln. 83-118008, U.S. Pat. Nos. 4,485,045 and 4,544,545, and EP 102,324. Ordinarily, the liposomes are of the small (from or about 200 to 800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol percent cholesterol, the selected proportion being adjusted for the most efficacious therapy. Agents of use in the invention may also be pegylated to increase their lifetime.

Additionally, it is contemplated that a composition in accordance with the invention may be formulated with other ingredients which may be of benefit to a subject in particular instances. For example, there may be benefit in incorporating, where appropriate, one or more anti-neoplastic agents. Examples of such agents include:

5 alkylating agents, for example, chlorambucil (e.g., LeukeranTM), cyclophosphamide (e.g., EndoxanTM, CycloblastinTM, NeosarTM, CyclophosphamideTM), ifosfamide (e.g., HoloxanTM, IfexTM, MesnexTM), thiotepa (e.g., ThioplexTM, ThiotepaTM); and antimetabolites/S-phase inhibitors, for example, methotrexate sodium (e.g., FolexTM, AbitrexateTM, EdertrexateTM), 5-fluorouracil (e.g., EfudixTM, EfudexTM), hydroxyurea

10 (e.g., DroxiaTM, Hydroxyurea, HydreaTM), amsacrine, gemcitabine (e.g., GemzarTM), dacarbazine, thioguanine (e.g., LanvisTM).

Also included are antimetabolites/mitotic poisons, for example, etoposide (EtopophosTM, Etoposide, ToposarTM), vinblastine (e.g., VelbeTM, VelbanTM), vindesine

15 (e.g., EldesineTM), vinorelbine (e.g., NavelbineTM), paclitaxel (e.g., TaxolTM); antibiotic-type agents, for example, doxorubicin (e.g., RubexTM), bleomycin (e.g., BlenoxaneTM), dactinomycin (e.g., CosmegenTM), daunorubicin (e.g., CerubidinTM), mitomycin (e.g., MutamycinTM); hormonal agents, for example, aminoglutethimide (e.g., CytadrenTM), anastrozole (e.g., ArimidexTM), estramustine (e.g., EstracytTM, EmcytTM), goserelin (e.g.,

20 ZoladexTM), hexamethylmelanine (e.g., HexametTM), letrozole (e.g., FemaraTM), anastrozole (e.g., ArimidexTM), and tamoxifen (e.g., EstroxynTM, GenoxTM, NovaldexTM, SoltamoxTM, TamofenTM).

Further included are combinations of any two or more anti-neoplastic agents (for

25 example, Adriamycin/5-fluorouracil/cyclophosphamide (FAC); and cyclophosphamide/methotrexate/5-fluorouracil (CMF)). Particularly useful are combinations that include, for example, at least two or more agents such as cyclophosphamide (e.g., CYTOXAN), methotrexate (e.g., RHEUMATREX), 5-fluorouracil (e.g., ADRUCIL), doxorubicin (e.g., ADRIAMYCIN), and

30 cyclophosphamide (e.g., CYTOXAN). Useful for metastatic disease are agents such as capecitabine (e.g., XELODA), doxorubicin (e.g., ADRIAMYCIN), including its liposomal formulation, gemcitabine (e.g., GEMZAR), the taxanes, including paclitaxel (e.g., TAXOL) and docetaxel (e.g., TAXOTERE), vinorelbine (e.g., NAVELBINE),

and trastuzumab (e.g., HERCEPTIN). Persons of ordinary skill in the art to which the invention relates will readily appreciate examples of other agents which may be of benefit. Agents of the invention may also be formulated with compounds and agents, other than those specifically mentioned herein, in accordance with accepted pharmaceutical practice.

As will be appreciated, in the case of administration of polynucleotides, they may be packaged into viral delivery systems, which viral systems may themselves be formulated into compositions as herein described. Persons of skill in the art to which the invention relates may appreciate a variety of suitable viral vectors having regard to the nature of the invention described herein. However, by way of example, retroviral vectors, adenoviral vectors, and adeno-associated virus (AAV) can be used. Persons of skill in the art to which the invention relates will readily appreciate methods which may be employed to implement such vectors in the present invention. However, by way of example only: the use of retroviral vectors is reported in Miller et al., 1993, Meth. Enzymol. 217:581-599, and Boesen et al., 1994, Biotherapy 6:291-302; the use of adenoviral vectors is reported for example in Kozarsky and Wilson, 1993, Current Opinion in Genetics and Development 3:499-503; Rosenfeld et al., 1991, Science 252:431-434; Rosenfeld et al., 1992, Cell 68:143-155; Mastrangeli et al., 1993, J. Clin. Invest. 91:225-234; PCT Publication WO 94/12649; and Wang, et al., 1995, Gene Therapy 2:775-783; and, the use of AAV has been reported in Walsh et al., 1993, Proc. Soc. Exp. Biol. Med. 204:289-300; U.S. Patent No. 5,436,146.

In accordance with the mode of administration to be used, and the suitable pharmaceutical excipients, diluents and/or carriers employed, compositions of the invention may be adapted into customary dosage forms such as solutions, orally administrable liquids, injectable liquids, tablets, coated tablets, capsules, pills, granules, suppositories, transdermal patches, suspensions, emulsions, sustained release formulations, gels, aerosols, liposomes, powders and immunoliposomes. The dosage form chosen will reflect the mode of administration desired to be used, the disorder to be treated, and the nature of the agent to be used. Particularly preferred dosage forms include orally administrable tablets, gels, pills, capsules, semisolids, powders, sustained release formulations, suspensions, elixirs, aerosols, ointments, or solutions for topical

administration, and injectable liquids. Skilled persons will readily recognise appropriate dosage forms and formulation methods. Generally, compositions are prepared by contacting or mixing specific agents and ingredients with one another. Then, if necessary, the product is shaped into the desired formulation. By way of example, certain methods of formulating compositions may be found in references such as Gennaro AR: Remington: The Science and Practice of Pharmacy, 20th ed., Lippincott, Williams & Wilkins, 2000.

The amount of an agent of the invention in a composition can vary widely depending on the type of composition, size of a unit dosage, kind of carriers, diluents and/or excipients, and other factors well known to those of ordinary skill in the art. In general, the final composition can comprise from 0.0001 percent by weight (% w) to 100% w of the actives of this invention, preferably 0.001% w to 10% w, with the remainder being any other active agents present and/or carrier(s), diluent(s) and/or excipient(s).

Methods of treatment

While the inventors' primary studies have involved breast cancer cells, Saratan is predicted to also act in the small intestine, spleen, liver, fetal liver and kidney, and also heart, prostate, uterus, colon, stomach, skin, lung, trachea, brain, cerebellum, fetal brain, spinal cord, placenta, adrenal gland, adipose, cartilage, hematopoietic and immune systems, pancreas, and also skeletal muscle, thymus, salivary gland, thyroid, umbilical cord, and ovaries. Specifically, Saratan is predicted to act in breast cancer, colon cancer, lung cancer, prostate cancer, endometrial cancer, and stomach cancer, among others. Accordingly, the inventors contemplate the inhibition of Saratan being applicable to the treatment of a variety of disorders characterised by altered cell proliferation, cell survival, or cell motility. In addition, based on homology to the mouse Saratan sequence, we predict a role for human Saratan in treating disorders of the brain, for example, brain cancers, and neurodegenerative diseases such as multiple sclerosis, Parkinson's disease, Alzheimer's disease, or Amyotrophic Lateral Sclerosis (ALS) (see, e.g., [hypertext transfer protocol://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/av.cgi?c=geneid&org=9606&l=5696](http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/av.cgi?c=geneid&org=9606&l=5696)). For such disorders, it is

In one embodiment, the invention relates to a method of preventing, reducing, or inhibiting cell proliferation, cell survival, or cell motility by inhibiting Saratan. Preferably, the method is for the treatment of a disorder characterised by aberrant cell in subject. This aberrant cell proliferation, cell survival, or cell motility may occur in one or more cell type within a subject and can include metastatic disorders. Specific disorders include, for example, cancer (breast, colon, lung prostate, endometrial, or stomach cancer, for example) and endometriosis. Examples of disorders include cancers such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and cancers of the adrenal gland, bladder, bone, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, bone marrow, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. In various embodiments, the disorder is an epithelial tumor of the breast, colon, lung, prostate, pancreas, stomach, endometrium, or ovary, or squamous cell carcinoma, or a melanoma, or a renal cancer or tumour.

Regarding breast cancers, these can include epithelial tumours (e.g., from cells lining ducts or lobules) or nonepithelial tumours (e.g., from the supporting stroma), such as angiosarcomas, primary stromal sarcomas, and phyllodes tumor. Breast cancers can also include carcinomas, for example, carcinomas *in situ*, as well as invasive cancers. Carcinoma *in situ* includes proliferation of cancer cells within ducts or lobules and without invasion of stromal tissue. Lobular carcinoma *in situ* (LCIS) includes nonpalpable lesions which can indicate increased risk of subsequent invasive carcinoma in either breast. In breast cancer, invasive carcinoma generally comprises adenocarcinoma, with most comprising infiltrating ductal type carcinoma and the remainder comprising infiltrating lobular carcinoma. Rare forms of breast cancer include medullary, mucinous, and tubular carcinomas. Breast cancer disorders also include Paget's disease of the nipple, and metastatic breast cancer.

Regarding colon cancer, this can generally include cancer of the colon, rectum, and/or anus, and especially, adenocarcinomas, and also carcinomas (e.g., squamous cloacogenic carcinomas), melanomas, lymphomas, and sarcomas. Epidermoid (nonkeratinizing squamous cell or basaloid) carcinomas are also included. The colon cancer may be associated with particular types of polyps or other lesions, for example,

tubular adenomas, tubulovillous adenomas (e.g., villoglandular polyps), villous (e.g., papillary) adenomas (with or without adenocarcinoma), hyperplastic polyps, hamartomas, juvenile polyps, polypoid carcinomas, pseudopolyps, lipomas, or leiomyomas. The cancer may be associated with familial polyposis and related conditions such as Gardner's syndrome or Peutz-Jeghers syndrome. The cancer may be associated, for example, with chronic fistulas, irradiated anal skin, leukoplakia, lymphogranuloma venereum, Bowen's disease (intraepithelial carcinoma), condyloma acuminatum, or human papillomavirus. In other aspects, the cancer may be associated with basal cell carcinoma, extramammary Paget's disease, cloacogenic carcinoma, or malignant melanoma.

Persons of general skill in the art to which the invention relates may readily appreciate alternative types of disorder which the invention may be applicable, especially having regard to the expression of Saratan provided herein. In addition, it will be appreciated by those of general skill in the art to which the invention relates, having regard to the nature of the invention and the results reported herein, that the present invention is applicable to a variety of different animals. Accordingly, the diagnostics and treatments can apply to any animal of interest. In particular, the invention is applicable to mammals, more particularly humans.

Persons of ordinary skill in the art to which the invention relates will appreciate various means and agents of use to inhibit Saratan. By way of example, nucleic acid technology including iRNA, antisense, and triple helix DNA may be employed to inhibit expression. Alternatively antibodies directed against Saratan or functional derivatives of such antibodies may be used. Exemplary agents are described in detail herein. Those agents of use in the invention will preferably exhibit one or more of the following characteristics: 1) the ability to prevent, reduce or inhibit cell proliferation; 2) the ability to prevent, reduce or inhibit cell survival; 3) the ability to prevent, reduce or inhibit cell motility; 3) the ability to prevent, reduce or inhibit expression or activity of Saratan; 4) the ability to prevent, decrease, reduce or control metastasis of tumours. Preferably, suitable agents will exhibit two or more of these characteristics.

As shown herein, Saratan is encoded as a cellular factor that is expressed in certain cancer cells, and also by at least one subset of normal adult cells. Therefore, Saratan can be considered a tumor-associated antigen. Several approaches can be used to target Saratan based on differences in expression and access in normal and cancer cells (reviewed, generally, in Paul, *Fundamental Immunology*, 1999, Lippincott-Raven Publishers, Philadelphia, PA, Chapter 37). Cancer cells are likely to express Saratan at much higher levels and such differences in expression levels between normal and cancer cells can be exploited therapeutically (see, e.g., Brown JP, et al., Quantitative analysis of melanoma-associated antigen p97 in normal and neoplastic tissues. *Proc Natl Acad Sci USA* 1981;78:539–543). Targeting may also be attained because of better access of Saratan-specific effector cells to cancer cells than to normal cells. For example, Saratan expressed on cancer cells may be more available for binding due to incomplete glycosylation (e.g., as in the case of epithelial mucins). For cancer cells, increased expression of MHC molecules may make tumours a direct target for T cells (see, e.g., Uyttenhove C, et al. The expression of mouse gene P1A in testis does not prevent safe induction of cytolytic T cells against a P1A-encoded tumor antigen. *Int J Cancer* 1997;70:349–356).

Multiple immunotherapeutic strategies involving innate or acquired immunity can be used to control Saratan-associated cancer, including (a) local application of a live bacterial vaccine, e.g., BCG; (b) use of cytokines; (c) active immunization against Saratan; (d) passive therapy with Saratan antibodies; and (e) adoptive transfer of effector cells (e.g., T cells). Active immunization against Saratan can be used to induce immune responses or passive immunization with a murine monoclonal antibody directed against Saratan. For general guidance, see, e.g., Riethmüller G, et al., Monoclonal antibody therapy for resected Dukes' C colorectal cancer: seven-year outcome of a multicenter randomized trial. *J Clin Oncol* 1998;16:1788–1794; Riethmüller G, et al., Randomized trial of monoclonal antibody for adjuvant therapy of resected Dukes' C colorectal carcinoma. German Cancer Aid 17-1A Study Group. *Lancet* 1994;343:1177–1183; Herlyn DM, et al., Inhibition of growth of colorectal carcinoma in nude mice by monoclonal antibody. *Cancer Res* 1980;40:717–721.

For immunization, pure antigens are often ineffective in inducing an acquired (i.e., antigen-specific) immune response unless certain adjuvants are used to stimulate innate immunity. Therefore, numerous approaches are designed to stimulate innate immunity at the site of vaccinations by the use of chemical and/or bacterial agents. Synthetic peptides used in vaccines can be designed for particular MHC haplotypes (see, e.g., Toes RE et al. Peptide vaccination can lead to enhanced tumor growth through specific T-cell tolerance induction. *Proc Natl Acad Sci USA* 1996;93:7855–7860). Antigenic peptides can be loaded onto heat-shock protein (or as recombinant virus-like particles) to increase the efficacy of immunization. Generally, effective induction of an immune response requires antigen presentation in an environment that provides appropriate help or secondary signals.

Several experimental designs use dendritic cells pulsed with virus-specific or tumor-associated peptides to induce tumor-reactive T cells and rejection of transplanted tumor cells. Dendritic cells can be loaded with synthetic antigenic peptides or recombinant proteins. Dendritic cells can also be loaded with one or more of: native peptides stripped from tumor cell surfaces; tumor-derived, peptide-loaded heat-shock proteins; tumor-derived mRNA; or fused tumor cells (for review, see Shurin MR. Dendritic cells presenting tumor antigen. *Cancer Immunol Immunother* 1996;43:158–164). One advantage of these strategies is that powerful immunity can be induced to (unique) individually distinct tumor antigens, as well as tumor-associated antigens.

For Saratan antigens, recombinant vaccines can be developed using vaccinia, *Listeria*, or virus-like particles. In other approaches, genetic vaccination can be used, for example, by injecting naked DNA plasmid constructs, intramuscularly encoding the tumor antigen (see e.g., Donnelly JJ, Ulmer JB, Liu MA. DNA vaccines. *Life Sci* 1997;60:163–172). GM-CSF may also be used to improve the presentation of the antigen by dendritic cells at the site of injection (see, e.g., Syrengelas AD, Chen TT, Levy R. DNA immunization induces protective immunity against B-cell lymphoma. *Nat Med* 1996;2:1038–1041). In other approaches, vaccination with anti-idiotypic antibodies can be used.

As further approaches, passive antibody therapy or adoptive transfer of tumor-specific T cells can be used. For example, passive immunization with Saratan antibody can protect against challenge with tumor cells and can be therapeutic when given soon after challenge with the cancer cells (e.g., see Riethmüller G, et al. Monoclonal antibody therapy for resected Dukes' C colorectal cancer: seven-year outcome of a multicenter randomized trial. J Clin Oncol 1998;16:1788–1794; Riethmüller G, et al. Randomized trial of monoclonal antibody for adjuvant therapy of resected Dukes' C colorectal carcinoma. German Cancer Aid 17-1A Study Group. Lancet 1994;343:1177–1183; Herlyn DM, et al. Inhibition of growth of colorectal carcinoma in nude mice by monoclonal antibody. Cancer Res 1980;40:717–721).

As alternate approaches, anti-idiotypic antibody treatment can be used to induce cancer cells to go into a long-lasting dormant state (see, e.g., Miller RA, Maloney DG, Warnke R, Levy R. Treatment of B-cell lymphoma with monoclonal anti-idiotypic antibody. N Engl J Med 1982;306:517–522). In addition, antibodies to tumour cells can be used as carriers for cytokines or cytotoxic agents, such as radiochemicals or natural toxins (see, e.g., Ghetie V, Vitetta E. Immunotoxins in the therapy of cancer: from bench to clinic. Pharmacol Ther 1994;63:209–234; Reisfeld RA, Gillies SD. Recombinant antibody fusion proteins for cancer immunotherapy. Curr Top Microbiol Immunol 1996;213:27–53). The recombinant antibody-cytokine or antibody-toxin fusion proteins may be used to concentrate these agents in the stroma surrounding the tumor cells. As alternatives, bispecific monoclonal antibodies can be engineered to bind effector cells as well as tumor antigens on the cancer cells. Monoclonal antibodies can also be humanized to reduce the stimulation of neutralizing anti-murine antibodies by patients. As still further approaches, adoptive transfer of T cells can be used with longer established tumor loads. T cells that have been isolated from patients can be expanded *in vitro* with IL-2 and then infused into patients who receive IL-2 as well (see, e.g., Smith CA, et al. Adoptive immunotherapy for Epstein-Barr virus-related lymphoma. Leuk Lymphoma 1996;23:213–220).

The efficacy of an agent in inhibiting Saratan may be determined having regard to the description of the invention herein and known methodology. For example, efficacy of agents may be determined by observing their ability to prevent, reduce, or inhibit

expression of Saratan or one or more of the functional effects of Saratan. By way of particular example, the affect of the agent on one or more of cellular invasion, cellular migration, the level of gene transcription of Saratan and Saratan responsive genes may be studied. Such studies may be conducted in *vitro* or in *vivo*.

5

The assays described herein after under the heading "Examples" may be used to determine the suitability of an agent in accordance with the invention. Specifically, RT-PCR and Northern blot analysis can be used to detect Saratan expression at the mRNA level, and Western blotting and direct or indirect immunostaining can be used to detect
10 the expression at the protein level. To detect Saratan activity, cell-based assays for cell proliferation, cell survival, or cell motility can be used.

In respect of the effects of Saratan inhibition on metastasis, an *in vivo* assay may be used, as described, for example, in Fidler, I. J. (1973) Nat. New Biol. 242, 148-149; and
15 Price J. E. The biology of cancer metastasis. Prog. Clin. Biol. Res., 354A: 237-255, 1990, or Kerbel R. S. What is the optimal rodent model for anti-tumor drug testing? Cancer Metastasis Rev., 17: 301-304, 1998; Killion J. J., Radinsky R., Fidler I. J. Orthotopic models are necessary to predict therapy of transplantable tumours in mice. Cancer Metastasis Rev., 17: 279-284, 1998; and Price J. E. Analyzing the metastatic
20 phenotype. J. Cell. Biochem., 56: 16-22, 1994.

Administration routes/regimes

The inventors contemplate administration of any of the agents or compositions of the invention by any means capable of delivering the desired activity (e.g., inhibition of
25 Saratan) to a target site within the body of a subject. A target site may be any site within the body which may have or be susceptible to a disorder, and may include one or more cells, tissues or a specific tumor. For example, administration may include parenteral administration routes, systemic administration routes, oral and topical administration. Administration may be by way of injection, subcutaneous, intraorbital, ophthalmic,
30 intraspinal, intracisternal, topical, infusion (using, e.g., slow release devices or minipumps such as osmotic pumps or skin patches), implant, aerosol, inhalation, scarification, intraperitoneal, intracapsular, intramuscular, intratumoral, intranasal, oral, buccal, transdermal, pulmonary, rectal or vaginal delivery. As will be appreciated, the

administration route chosen may be dependent on the position of the target site within the body of a subject, as well as the nature of the agent or composition being used.

5 In a specific embodiment, in the case of polynucleotides, they may be administered for example by infection using defective or attenuated retroviral or other viral vectors (see e.g., U.S. Patent No. 4,980,286), or by direct injection of naked DNA, or by use of microparticle bombardment (e.g., a gene gun; Biolistic, DuPont); by coating with lipids or cell-surface receptors or transfecting agents; encapsulation in liposomes, microparticles, or microcapsules; by linkage to a peptide which is known to enter the
10 nucleus; or by administering it in linkage to a ligand subject to receptor-mediated endocytosis (see, e.g., Wu and Wu, 1987, J. Biol. Chem. 262:4429-4432), which can be used to target cell types specifically expressing the receptors, and the like. In addition, nucleic acid-ligand complexes can be formed in which the ligand comprises a fusogenic viral peptide to disrupt endosomes, allowing the nucleic acid molecules to avoid
15 lysosomal degradation.

In yet another embodiment, the polynucleotides can be targeted *in vivo* for cell specific uptake and expression, by targeting a specific receptor, as described for example in WO 92/06180 dated April 16, 1992 (Wu et al.); WO 92/22635 dated December 23, 1992
20 (Wilson et al.); WO 92/20316 dated November 26, 1992 (Findeis et al.); WO 93/14188 dated July 22, 1993 (Clarke et al.); and, WO 93/20221 dated October 14, 1993 (Young). Alternatively, the polynucleotides can be introduced intracellularly and incorporated within host cell DNA for expression, by homologous recombination Koller and Smithies, 1989, Proc. Natl. Acad. Sci. USA 86:8932-8935; Zijlstra et al., 1989, Nature
25 342:435-438. Cells into which polynucleotides can be introduced for purposes of the present invention encompass any desired, available cell type. The appropriate cell type will depend on the nature of the disorder to be treated. However, by way of example, the polynucleotide can be introduced to a cancer cell.

30 As will be appreciated, the dose of an agent or composition administered, the period of administration, and the general administration regime may differ between subjects depending on such variables as the nature of the condition to be treated, severity of symptoms of a subject, the size of any tumour to be treated, the target site to be treated,

the mode of administration chosen, and the age, sex and/or general health of a subject. Persons of average skill in the art to which the invention relates will readily appreciate or be able to determine appropriate administration regimes having regard to such factors. It should be appreciated that administration may include a single daily dose or
5 administration of a number of discrete divided doses as may be appropriate. The inventors also contemplate the administration regimes which combine different modes or routes of administration. For example, intratumoural injection and systemic administration could be combined.

10 It should be appreciated that a method of the invention may comprise further steps such as the administration of additional agents or compositions which may be beneficial to a subject having regard to the condition to be treated. For example, other agents of use in treating proliferative disorders (such as the anti-neoplastic agents mentioned above) could be administered. It should be appreciated that such additional agents and
15 compositions may be administered concurrently with the agents and compositions of the invention, or in a sequential manner. For example, the additional agents or compositions could be administered before or after administration of the agents or compositions of the invention. It should be appreciated in relation to sequential delivery of agents or compositions, that sequential administration of one agent or composition after the other
20 need not occur immediately, although this may be preferable. There may be a time delay between delivery of the agents or compositions. The period of the delay will depend on factors such as the condition to be treated and the nature of the compositions or agents to be delivered. However, by way of example, the inventors contemplate periods of between hours to several days or months.

25 **Diagnostic methods and compositions**

In one embodiment, the invention relates to use of one or more reagents of the invention in a method of diagnosing a disorder associated with cell proliferation, cell survival, or cell motility. Preferably, the method is for the diagnosis of a disorder characterised by
30 aberrant cell proliferation, cell survival, or cell motility in subject. This aberrant cell proliferation, cell survival, or cell motility may occur in one or more cell type within a subject and can include metastatic disorders. Specific disorders include, for example, cancer (breast, lung, colon, prostate, endometrial, or stomach cancer, for example) and

endometriosis. Examples of disorders include cancers such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and cancers of the adrenal gland, bladder, bone, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, bone marrow, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. In various embodiments, the disorder is an epithelial tumor of the breast, lung, prostate, colon, pancreas, endometrium, stomach, or ovary, or a squamous cell carcinoma, or a melanoma, or a renal cancer or tumour. Specific breast and colon cancers are described in detail herein.

In accordance with the invention, antibodies which specifically bind Saratan may be used for the diagnosis of conditions or disorders characterized by expression of Saratan, or in assays to monitor patients being treated with Saratan inhibitors. The antibodies useful for diagnostic purposes may be prepared in the same manner as those described above for therapeutics. Diagnostic assays for Saratan include methods which utilize the antibody and a label to detect Saratan in human body fluids or extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by joining them, either covalently or non-covalently, with a reporter molecule. A wide variety of reporter molecules which are known in the art may be used, several of which are described above.

A variety of protocols, including ELISA, RIA, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of Saratan expression. Normal or standard values for Saratan expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibody to Saratan under conditions suitable for complex formation. The amount of standard complex formation may be quantified by various methods, but preferably by photometric means. Quantities of Saratan expressed in subject, control, and disease, samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disorders.

In another embodiment of the invention, the polynucleotides encoding Saratan may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues
5 in which expression of Saratan may be correlated with disorders. The diagnostic assay may be used to distinguish between absence, presence, and excess expression of Saratan, and to monitor regulation of Saratan levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting
10 polynucleotide sequences, including genomic sequences, encoding Saratan or closely related molecules, may be used to identify nucleic acid sequences which encode Saratan. The specificity of the probe, whether it is made from a highly specific region, e.g., 10 unique nucleotides in the 5' regulatory region, or a less specific region, e.g., especially in the 3' coding region, and the stringency of the hybridization or
15 amplification (maximal, high, intermediate, or low) will determine whether the probe identifies only naturally occurring sequences encoding Saratan, alleles, or related sequences.

Probes may also be used for the detection of related sequences, and should preferably
20 contain at least 50% of the nucleotides from any of the Saratan encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and derived from the nucleotide sequence of SEQ ID NO:14-16, 19, 22, or fragments encompassing a nucleic acid sequence of SEQ ID NO:17, 18, 20, 21, 23, 24-48, or from genomic sequences including promoter, enhancer elements, and introns of the naturally occurring
25 Saratan.

Means for producing specific hybridization probes for DNAs encoding Saratan include the cloning of nucleic acid sequences encoding Saratan or modified sequences into vectors for the production of mRNA probes. Such vectors are known in the art,
30 commercially available, and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for

example, radionuclides such as ^{32}P or ^{35}S , or enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding Saratan may be used for the diagnosis of disorders which are associated with either increased or decreased expression of Saratan. The polynucleotide sequences encoding Saratan may be used in Southern or northern analysis; dot blot or other membrane-based technologies; in PCR technologies; or in dipstick, pin, ELISA assays; or microarrays utilizing fluids or tissues from patient biopsies to detect altered Saratan expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding Saratan may be useful in assays that detect activation or induction of various cancers, particularly those mentioned above. The nucleotide sequences encoding Saratan may be labeled by standard methods, and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. If the amount of signal in the biopsied or extracted sample is significantly altered from that of a comparable control sample, the nucleotide sequences have hybridized with nucleotide sequences in the sample, and the presence of altered levels of nucleotide sequences encoding Saratan in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or in monitoring the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of Saratan, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, which encodes Saratan, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with those from an experiment where a known amount of a substantially purified polynucleotide is used. Standard values obtained from normal samples may be compared with values obtained

from samples from patients who are symptomatic for a disorder. Deviation between standard and subject values is used to establish the presence of the disorder.

Once the disorder is diagnosed and a treatment protocol is initiated, hybridization assays
5 may be repeated on a regular basis to evaluate whether the level of expression in the patient begins to approximate that which is observed in the normal patient. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months. With respect to cancer, the presence of a relatively high amount of transcript in biopsied tissue from an individual may indicate a
10 predisposition for the development of the disorder, or may provide a means for detecting the disorder prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

15 Additional diagnostic uses for oligonucleotides designed from the sequences encoding Saratan may involve the use of PCR. Such oligomers may be chemically synthesized, generated enzymatically, or produced *in vitro*. Oligomers will preferably consist of two nucleotide sequences, one with sense orientation (5'.fwdarw.3') and another with
20 antisense (3'.fwdarw.5'), employed under optimized conditions for identification of a specific gene or condition. The same two oligomers, nested sets of oligomers, or even a degenerate pool of oligomers may be employed under less stringent conditions for detection and/or quantitation of closely related DNA or RNA sequences.

25 Methods which may also be used to quantitate the expression of Saratan include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and standard curves onto which the experimental results are interpolated (Melby, P. C. et al. (1993) J. Immunol. Methods, 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 229-236). The speed of quantitation of multiple samples may be accelerated by running
30 the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously (to produce a transcript image), and to identify genetic variants, mutations and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of the disorder, to diagnose the disorder, and to develop and monitor the activities of therapeutic agents. In one embodiment, the microarray is prepared and used according to the methods known in the art such as those described in PCT application WO 95/11995 (Chee et al.), Lockhart, D. J. et al. (1996; Nat. Biotech. 14: 1675-1680) and Schena, M. et al. (1996; Proc. Natl. Acad. Sci. 93: 10614-10619).

The microarray is preferably composed of a large number of unique, single stranded nucleic acid sequences, usually either synthetic antisense oligonucleotides or fragments of cDNAs, fixed to a solid support. The oligonucleotides are preferably about 6 to 60 nucleotides in length, more preferably about 15 to 30 nucleotides in length, and most preferably about 20 to 25 nucleotides in length. For a certain type of microarray, it may be preferable to use oligonucleotides which are 7 to 10 nucleotides in length. The microarray may contain oligonucleotides which cover the known 5' or 3' sequence, or may contain sequential oligonucleotides which cover the full length sequence; or unique oligonucleotides selected from particular areas along the length of the sequence.

Polynucleotides used in the microarray may be oligonucleotides that are specific to a gene or genes of interest in which at least a fragment of the sequence is known or that are specific to one or more unidentified cDNAs which are common to a particular cell or tissue type or to a normal, developmental, or disease state. In certain situations, it may be appropriate to use pairs of oligonucleotides on a microarray. The pairs will be identical, except for one nucleotide preferably located in the centre of the sequence. The second oligonucleotide in the pair (mismatched by one) serves as a control. The number of oligonucleotide pairs may range from 1 to 1,000,000.

In order to produce oligonucleotides to a known sequence for a microarray, the gene of interest is examined using a computer algorithm which starts at the 5' or more

preferably at the 3' end of the nucleotide sequence. The algorithm identifies oligomers of defined length that are unique to the gene, have a GC content within a range suitable for hybridization, and lack predicted secondary structure that may interfere with hybridization. In one aspect, the oligomers are synthesized at designated areas on a substrate using a light-directed chemical process. The substrate may be paper, nylon or any other type of membrane, filter, chip, glass slide, or any other suitable solid support.

In one aspect, the oligonucleotides may be synthesized on the surface of the substrate by using a chemical coupling procedure and an ink jet application apparatus, such as that described in PCT application WO 95/251116 (Baldeschweiler et al.). In another aspect, a gridded array analogous to a dot or slot blot (HYBRIDOT apparatus, Life Technologies) may be used to arrange and link cDNA fragments or oligonucleotides to the surface of a substrate using a vacuum system, thermal, UV, mechanical or chemical bonding procedures. In yet another aspect, an array may be produced by hand or by using available devices, materials, and machines (including multichannel pipettors or robotic instruments; Brinkmann, Westbury, NY) and may include about 8, 24, 96, 384, 1536 or 6144 oligonucleotides, or any other multiple from 2 to 1,000,000, which lends itself to the efficient use of commercially available instrumentation.

In order to conduct sample analysis using the microarrays, polynucleotides are extracted from a biological sample. The biological samples may be obtained from any bodily fluid (e.g., blood, urine, saliva, phlegm, gastric juices, etc.), cultured cells, biopsies, or other tissue preparations. To produce probes, the polynucleotides extracted from the sample are used to produce nucleic acid sequences which are complementary to the oligonucleotides on the microarray. If the microarray consists of cDNAs, antisense RNAs are appropriate probes. Therefore, in one aspect, mRNA is used to produce cDNA which, in turn and in the presence of fluorescent nucleotides, is used to produce fragment or oligonucleotide antisense RNA probes. These fluorescently labeled probes are incubated with the microarray so that the probe sequences hybridize to the cDNA oligonucleotides of the microarray. In another aspect, nucleic acid sequences used as probes can include polynucleotides, fragments, and complementary or antisense sequences produced using restriction enzymes, PCR technologies, and oligolabeling kits

(Amersham Pharmacia Biotech), which are well known in the area of hybridization technology.

Incubation conditions are adjusted so that hybridization occurs with precise
5 complementary matches or with various degrees of less complementarity. After removal
of non-hybridized probes, a scanner is used to determine the levels and patterns of
fluorescence. The scanned images are examined to determine degree of
complementarity and the relative abundance of each oligonucleotide sequence on the
microarray. A detection system may be used to measure the absence, presence, and
10 amount of hybridization for all of the distinct sequences simultaneously. This data may
be used for large scale correlation studies or functional analysis of the sequences,
mutations, variants, or polymorphisms among samples (Heller, R. A. et al., (1997) Proc.
Natl. Acad. Sci. 94:2150-55).

15 In another embodiment of the invention, the nucleic acid sequences which encode
Saratan may be used to generate hybridization probes which are useful for mapping the
naturally occurring genomic sequence. The sequences may be mapped to a particular
chromosome, to a specific region of a chromosome, or to artificial chromosome
constructions, such as human artificial chromosomes (HACs), yeast artificial
20 chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1
constructions or single chromosome cDNA libraries (cf. Price, C. M. (1993) Blood Rev.
7:127-134; Trask, B. J. (1991) Trends Genet. 7:149-154).

Fluorescent *in situ* hybridization (FISH as described in Verma et al. (1988) Human
25 Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York, NY) may
be correlated with other physical chromosome mapping techniques and genetic map
data. Examples of genetic map data can be found in various scientific journals or at the
Online Mendelian Inheritance in Man (OMIM) site. Correlation between the location of
the gene encoding Saratan on a physical chromosomal map and a specific disorder, or
30 predisposition to a specific disorder, may help delimit the region of DNA associated
with that disorder. The nucleotide sequences of the subject invention may be used to
detect differences in gene sequences between normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, linkage analysis using established chromosomal markers, may be used to extend genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms, or parts thereof, by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disorder has been crudely localized by genetic linkage to a particular genomic region, for example, AT to 11q22-23 (Gatti, R. A. et al. (1988) Nature 336:577-580), any sequences mapping to that area may represent associated or regulatory genes for further investigation. The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, and affected individuals.

In another embodiment of the invention, Saratan, or its functional or immunogenic fragments or oligopeptides thereof, can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes, between Saratan and the agent being tested, may be measured. In a further embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding Saratan specifically compete with a test compound for binding Saratan. In this same manner, the antibodies can be used to detect the presence of any amino acid sequence which shares one or more antigenic determinants with Saratan.

Another technique for drug screening which may be used provides for high throughput screening of compounds having suitable binding affinity to the protein of interest as described in WO 84/03564. In this method, as applied to Saratan, large numbers of different small test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The test compounds are reacted with Saratan, or fragments thereof, and washed. Bound Saratan is then detected by methods well known in the art. Purified Saratan can also be coated directly onto plates for use in the aforementioned

drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

5 In additional embodiments, the nucleotide sequences which encode Saratan may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

10 **Kits for treatment or diagnosis**

The agents and compositions of the invention may be used in kits suitable for controlling or inhibiting Saratan or for the treatment of a disorder as defined herein. In other aspects, the agents and compositions may be used in diagnostic kits. Kits can comprise at least one agent of the invention in a suitable container. The agents may be
15 formulated suitable for direct administration to a subject for example, as agents or pharmaceutical compositions. Alternatively, the kit may comprise one or more agents in one container and pharmaceutical diluents, carriers and/or excipients in another; the contents of each container being mixed together prior to administration. The kit may also comprise additional agents and compositions in further separate containers as may
20 be necessary for a particular application. Any container suitable for storing and/or administering an agent or composition may be used in a kit of the invention. Suitable containers will be appreciated by persons skilled in the art. By way of example, such containers include vials and syringes. The containers may be suitably sterilised and hermetically sealed. Further, kits of the invention can also comprise instructions for the
25 use and administration of the components of the kit. The invention is further elucidated with reference to the examples below.

EXAMPLES

The examples described herein are for purposes of illustrating embodiments of the
30 invention. Other embodiments, methods, and types of analyses are within the scope of persons of ordinary skill in the molecular diagnostic arts and need not be described in detail herein. Other embodiments within the scope of the art are considered to be part of this invention.

EXAMPLE 1: MATERIALS AND METHODS*Cell culture*

The human breast cancer MCF-7 and MDA-MB-231 (referred to as MDA-231) cell lines were obtained from the American Type Culture Collection (Manassas, VA). MCF-7 cells were cultured in RPMI 1640 medium supplemented with 10% heat-inactivated foetal bovine serum (FBS), 100 units/ml penicillin, 100 µg/ml streptomycin, and 2 mM L-glutamine. MDA-231 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% FBS and 100 units/ml penicillin, 100 µg/ml streptomycin. All cell cultures were incubated in a humidified incubator with 5% CO₂ at 37°C.

Plasmid constructs

The insert (~7.5 kb) of IMAGE cDNA clone (5493356) was excised by *EcoRV* and *NotI* digestion and subcloned into phCMV1 vector (Genlantis, CA). The resulting plasmid was designated phCMV-hST.

A sequence encoding the HA (hemagglutinin YPYDVPDYA, SEQ ID NO:11) tag was added in frame at the 5' or 3' end of the coding sequence for Saratan. This was cloned into phCMV1 vector to obtain expression plasmids phCMVn-hST and phCMVc-hST for N- or C-terminal HA-tagged Saratan fusion proteins, respectively. The N-terminal HA-tagged Saratan included the sequence MYPYDVPDYASLEF (SEQ ID NO:12) prior to the first amino acid of Saratan. The C-terminal HA-tagged Saratan included the sequence EFYPYDVPDYA (SEQ ID NO:13) immediately after the last amino acid of Saratan.

The coding sequence of Saratan was also cloned into the pEGFP-N1 vector in frame with the 5' EGFP tag (pEGFP-N1-hST) to express the C-terminal EGFP tagged Saratan.

The Glutathione S-transferase (GST) Gene Fusion System from Amersham Biosciences was used to produce recombinant Saratan protein in *E. coli*. The human Saratan cDNA fragment coding for the N-terminal peptide of the first 61 amino acids was amplified by PCR with the following primer pair: 5'-ctaggatccATGGATCTCTCCTTTATGGCCG-3' (SEQ ID NO:39, forward, cloning linker in lower case) and 5'-

tgactcgagtcTAGCAAGACGGCGTCGTTG-3' (SEQ ID NO:40, reverse). The PCR product was cloned into the pGEX 4T1 vector (Amersham Biosciences) upon 5' *Bam* HI and 3' *Xho*I digestion to generate pGEX 4T1-hSTn plasmid for the expression of GST fusion proteins in *E. coli*.

5

For each construct, the insert was verified by DNA sequencing.

Establishing MCF-7 cells stably expressing Saratan

The human breast cancer cell line MCF-7 was stably transfected with phCMV-hST using Saint-Mix (Synvolux Therapeutics B.V., the Netherlands) transfection reagent. The resulting cells were designated as MCF7-hST. As a control, MCF-7 cells were stably transfected with the empty vector phCMV1 to produce cells designated as MCF7-Vec. Cell clones were selected by addition of G418 (Bio-Rad Laboratories, CA) to a concentration of 0.8 mg/ml in the medium. Transfected cell lines were generated as pools of positive cell clones. The overexpression of Saratan in MCF7-Saratan was confirmed by Western blotting.

15

Preparation of total RNA

Total RNA was isolated from MCF-7 cells with Trizol reagent (Invitrogen) at 1 ml/10 cm² according to the manufacturer's instructions and resuspended in diethyl pyrocarbonate-treated nuclease-free water. RNA samples were further treated with DNase I for 30 min at 37°C. The reaction was stopped by addition of 25 mM EDTA and incubation at 65°C for 15 min. RNA samples were then purified by extraction in phenol/chloroform (pH 5.2, phenol:chloroform:isoamyl alcohol at 25:24:1) followed by an additional chloroform extraction and ethanol precipitation. Quantification and purity of the RNA was assessed by A₂₆₀/A₂₈₀ absorption, and RNA quality was assessed by agarose gel electrophoresis. RNA samples with ratios of A₂₆₀/A₂₈₀ greater than 1.6 were stored at - 80°C for further analysis.

25

Reverse transcription-PCR

One-step reverse transcription (RT)-PCR kit (Qiagen) was used to determine the presence of Saratan using specific primers hST5 (sense, 5'-ATGGCGCAGCTCTTCTACTC-3', SEQ ID NO:41) and hST3 (antisense, 5'-

CGTGCATCTCTCTGCATCCTC-3'; SEQ ID NO:42) or hST-F4 (sense, 5'-ACACCATGGATCTCTCCTTTATG-3', SEQ ID NO:65) and hST-R4 (antisense, 5'-TGACTCGAGTCATAGCAAGACGGCGTCGTTG-3', SEQ ID NO:66). For RT-PCR, the following procedure was employed. To start, 1 µg of total RNA was diluted to 0.1 µg/µl to minimize the variation of sample handling. This dilution was treated by DNase I for 15 min, followed by inactivation of DNase by adding EDTA to 5 mM and heating to 70°C for 15 min.

The DNase-treated RNA was then mixed with a master cocktail containing RT-PCR buffer, sense and antisense primers, dNTPs, RNase inhibitor, and an enzyme mixture containing reverse transcriptase (Omniscript and Sensiscript) and HotStart Taq DNA polymerase at the concentrations recommended by the manufacturer to a final volume of 50 µl. The temperature-cycle protocol included: 60 min at 50°C for RT reaction, followed by denaturation and activation of HotStart DNA polymerase for 15 min at 95°C, and PCR amplification for 10 sec at 95°C, 30 sec at 60°C, and 1 min at 72°C for 30 cycles. A final extension for 5 min at 72°C was performed at the end of the cycles. β-actin was similarly amplified by RT-PCR using 0.2 µg of total RNA as an internal control. Ten microlitres of each of the RT-PCR product was fractionated on 1% agarose gels. The identity of RT-PCR product was confirmed by the size, restriction enzyme digestion, and DNA sequencing.

mRNA expression assays

For mRNA expression studies, a cDNA panel (Primgen) consisting of 10 ng of first-strand cDNA from various human tissues was screened by PCR with the Saratan specific hST5 and hST3 primers. As recommended by the manufacturer, human beta-2-microglobulin gene was used as the cDNA input control. The primers used to amplified beta-2-microglobulin gene included 5'-TCTCGCTCCGTGGCCTTAGCTG-3' (sense primer, SEQ ID NO:43) and 5'-AACATGTCTCGATCCCACTTA ACTATC-3' (anti-sense primer, (SEQ ID NO:44). The expression of Saratan in a variety of cancer cell lines was determined by RT-PCR. The cell lines included A549 human lung carcinoma cells; AGS human stomach cancer cells; Colo320DM and DLD-1 neuroendocrine colon carcinoma cells; DU145 and PC3 human prostate cancer cells; BxPC3 pancreatic

cancer; RL95-2 endometrial carcinoma cells; HepG2 human liver cancer cells; and T47D, BT459 and MCF-7 human breast cancer cells.

Confocal Laser Scanning Microscopy

- 5 To investigate the subcellular localization of Saratan, MCF-7 mammary carcinoma cells grown on glass coverslips were transiently transfected with pEGFP-N1-hST or the empty pEGFP-N1 vector. Cells were grown in RPMI 1640 containing 10% serum for 24 h. Cells were fixed in ice-cold 4% paraformaldehyde, washed with PBS, permeabilized for 10 min with 0.1% Triton[®] X-100, and nuclei were counterstained with
- 10 DAPI. In another experiment, MCF-7 cells grown on glass coverslips were transiently transfected with phCMVn-hST, phCMVc-hST, or the empty phCMV1 vector as a control.

Cells were fixed in ice-cold 4% paraformaldehyde and washed with PBS. A subset of

15 each group of transfectants was permeabilized with 0.5% Triton[®] X-100 for 30 min. The remaining subset was left untreated. Cells were incubated with primary rabbit polyclonal HA-Probe (Y-11) (Cat # sc-805, Santa Cruz Biotech) followed by incubation with a FITC-labeled goat anti-rabbit second antibody (Cat # F-0511, Sigma). Coverslips were mounted on slides with the nuclei counterstained by DAPI. The fluorescent images

20 were collected from equatorial planes of cells with optical thickness of 2 μ m using an inverted Zeiss LSM 510 confocal microscope (40 $\times$ oil immersion objective, NA=1.3). Images were converted to the tagged-information-file format and processed with Adobe Photoshop.

25 *Immunoblotting*

- Cells were washed twice with ice-cold phosphate-buffered saline (PBS) and lysed at 4°C in lysis buffer (20 mM Tris•HCl, pH 7.4, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1% Triton[®] X-100, 1% Nonidet P-40, 1 μ g/ml protease inhibitor cocktail (GE Healthcare) and 0.1 mM PMSF). The lysates were next sonicated and then cleared by
- 30 centrifugation at 15,000 $\times$ g for 15 min at 4°C. SDS-polyacrylamide gel electrophoresis (PAGE) sample buffer (50 mM Tris•HCl, pH 6.8; 2% SDS; 2% β -mercaptoethanol, and bromophenol blue) was added to each sample and the samples were boiled for 5 min. Samples were subjected to discontinuous SDS-PAGE with a 12% resolving gel and

transferred to nitrocellulose membranes (Hybond™ C-extra) using standard electroblotting procedures.

5 Membranes were blocked with 5% non-fat dry milk in PBS with 0.1% Tween® 20 (PBST) for 1 h at room temperature. The blots were then immunolabelled with monoclonal antibodies 9E10 against c-Myc (1:1000 dilution) in PBST containing 1% non-fat dry milk at 4°C overnight. After three washes with PBST at room temperature, immunolabelling was detected by ECL plus™ chemiluminescence as described by the manufacturer (GE Healthcare). Blots were stripped and reprobed with monoclonal
10 antibody against β -actin to ensure equal loading of the cell lysate proteins. Blots were stripped by incubation for 30 min at 50°C in a solution containing 62.5 mM Tris•HCl, pH 6.7; 2% SDS; and 0.7% β -mercaptoethanol. Blots were then washed for 30 min with several changes of PBST at room temperature. Efficacy of stripping was determined by re-exposure of the membranes to ECL plus™. Thereafter, blots were re-blocked and
15 immunolabelled as described above.

Cell growth and proliferation assay

Cells were seeded onto 6-well plates at a density of 5×10^4 cells per well in triplicate in 3 ml of complete RPMI medium containing 10% FBS. Cells were cultured under this
20 condition for up to 10 days with medium change every other day until being harvested. To compare the cell proliferation rate, the total cell number in each well was quantified with use of a haemocytometer every 2 days. Cell viability was assessed by using trypan blue. Cell cycle kinetics were measured by Bromodeoxyuridine (BrdU) incorporation using a BrdU staining kit (Zymed) following the manufacturer's protocol. Briefly, cells
25 were seeded on glass coverslips in 6-well plates in full media and incubated overnight before 18 hours of serum starvation in serum-free media. Then cells were pulse labelled with BrdU at a concentration of 10 μ M in serum-free media for 45 min.

After fixing, denaturing, and blocking, cells were incubated with biotinylated anti-BrdU
30 mouse monoclonal antibody. Cells were then incubated with streptavidin-peroxidase followed by incubation with Diaminobenzidine (DAB) chromogen solution for colour development in the presence of hydrogen peroxide. The cell nuclei were counterstained with hematoxylin. Cells with dark brown nuclear staining were counted as BrdU-

labelled cells. For quantification, 10 random fields per coverslip (magnification of 40×) were documented by photomicroscopy. The percentage of BrdU-positive cell nuclei relative to the total number of cell nuclei was calculated. A minimum of 500 cells per coverslip were counted.

5

Apoptosis assay

Apoptotic cell death was measured by fluorescent microscopic analysis of cell DNA staining patterns with karyophilic Hoechst 33258 (Del, B. G., Z. Darzynkiewicz, C. Degraef, R. Mosselmans, D. Fokan, and P. Galand. 1999. Comparison of methods based on annexin-V binding, DNA content or TUNEL for evaluating cell death in HL-60 and adherent MCF-7 cells. *Cell Prolif.* 32:25-37). Cells were seeded onto 6-well plates. Cells were serum depleted for 24 hours. Cells were then fixed with 4% paraformaldehyde, permeabilized with 0.6% Tween[®] 20, and stained with Hoechst 33258 (1:10,000; Molecular Probes, Eugene, OR) for 5 min at room temperature. Following washing with PBS, nuclear morphology was examined under a UV-visible fluorescence microscope with a 40× objective. Cells with apoptotic nuclear characteristics such as nuclear condensation and fragmentation were scored as apoptotic. For each treatment, 200-300 cells were examined on each treatment.

20 *Cell migration and invasion*

Cell migration was assayed using modified Boyden chambers containing polycarbonate membranes (tissue culture treated, 6.5-mm diameter, 10-μm thickness, 8-μm pores, Transwell[®]; Costar, Cambridge, MA) as previously described (Doerr, M. E. and J. I. Jones. 1996. The roles of integrins and extracellular matrix proteins in the insulin-like growth factor I-stimulated chemotaxis of human breast cancer cells. *J. Biol. Chem.* 271:2443-2447) with some modification. Cell invasion assays were carried using the same protocol, except that a layer of Matrigel was polymerized onto the membranes according to the manufacturer's specifications. This was done prior to the addition of cells to the upper wells.

30

To ensure single cell suspension for migration/invasion assay, cells at near confluence were trypsinized the day before the assay, passed through an 18G needle several times and incubated at least 12 hours prior to assays. The next day, cells were harvested with

5 mM EDTA in PBS, pH 7.4, washed twice, and then suspended in serum-free RPMI/BSA. A 100- μ l volume of cells (1.0×10^5) was loaded into Transwell inserts. The lower chambers were filled with 600 μ l of RPMI containing 10% FBS. The loaded chamber was placed in a 37°C incubator in a humidified atmosphere with 5% CO₂.
5 After 24 h, the inserts were removed and the cells were then fixed with 4% paraformaldehyde. Cells (and Matrigel) on the upper surface were removed by scraping with cotton swabs. Those that migrated to the lower surface of the membrane were identified microscopically using a UV-visible fluorescence microscope, following staining with Hoechst 33258.

10

Cell counts were obtained by counting all cells in six grids on one row across the centre of the round filter. This was done using a Lovins Microslide field finder (Gurley, Troy, NY) under a microscope at a total magnification of 100 $\times$. Pictures were taken from representative fields of each membrane. All data presented are the means $\pm$ standard
15 deviation (SD) of triplicate wells.

Soft agar assay for colony formation

For soft agar colony formation assay, cells were cultured in 6-well plates first covered with a layer of agar (0.5%). Cells were grown in RPMI 1640 media containing 10%
20 FBS. The middle layer contained 5×10^3 cells in 0.35% agar in RPMI 1640 media containing 10% FBS. Medium was added as the top layer to prevent drying of the agarose gels. The plates are incubated at 37°C in a humidified incubator for 14 days, after which the cells were stained with 0.5 ml of 0.005% crystal violet in distilled water for 1 h and destained for 2 h in water. Colonies were inspected and photographed.

25

Construction of Saratan siRNA vector

The pSilencer 2.1-U6 vector (Ambion) was used to express Saratan siRNA using the RNA polymerase III U6 promoter. Briefly, oligonucleotides with 3' single stranded overhangs were designed to contain a 19mer hairpin sequence specific for Saratan
30 mRNA target. The oligonucleotides also included a loop sequence separating the two complementary domains and a transcription termination sequence. The oligonucleotides were annealed to form the hairpin, and ligated into pSilencer digested with *Bam* HI and *Hind* III using standard methods. Three different siRNA constructs were generated to

target the following sequence specific for Saratan mRNA:
ATAATCACTCGTGTGACTT (SEQ ID NO:45) for siRNA-A;
TTACAAGAAAGCTGGACTT (SEQ ID NO:46) for siRNA-B and
GAGCAACTTACTGCTTTGT (SEQ ID NO:47) for siRNA-C. The sequence
5 ACTACCGTTGTTATAGGTG (SEQ ID NO:48) was designed to include limited
homology to any known sequences in the human genomes (Ambion) and used as the
negative control.

Production of recombinant human Saratan protein in bacteria

10 The recombinant N-terminal peptide from Saratan protein was GST tagged, expressed,
and purified from *E. coli*. In particular, the pGEX 4T1-hSTn plasmid was used to
transform BL21-Gold cells (Stratagene). A single recombinant *E. coli* colony was
inoculated into LB medium containing carbenicillin (50 µg/ml). The overnight culture
was diluted 1:200 in LB medium with carbenicillin, pH 7.4, and cultured at 37°C to
15 optical density at 600 nm of ~0.5. Protein expression was induced by addition of
isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.2 mM and the
cultures were incubated for an additional 5-6 hours at room temperature.

Cells were harvested and lysed by sonication in buffer A (50mM Tris•HCl, pH 8.0,
20 150mM NaCl, 1mM EDTA and 5mM DTT) containing lysozyme. The lysate was
cleared by centrifugation and Triton® X-100 was added to a final concentration of 1%.
GST-hSTn was purified from the lysate with columns containing glutathione
Sephadex® 4B matrix (Amersham Biosciences). In-column bound protein was
incubated for 10 min in buffer A containing 5 mM ATP, 5 mM MgSO4 and 0.1 mg/ml
25 denatured *E. coli* proteins. This was washed with a buffer of 50mM Tris•HCl (pH 8.0),
150mM NaCl and 1mM EDTA with/without 1% Triton® X-100. The GST tag was
cleaved by in-column incubation with Thrombin for 2 h at room temperature. Purified
protein was eluted with PBS. This was analysed by SDS-PAGE on a 4-12% NuPAGE
Bis-tris gel (Invitrogen) and visualised by Coomassie blue staining. The yield of
30 purified protein was estimated by Bradford's assay.

Protein Synthesis and production of polyclonal antibodies

While recombinant Saratan was produced in *E. coli*, the N-terminal peptide including 61

amino acids of Saratan (MDLSFMAAQLPMMGGAFMDSPNEDFSTEYSLFNSSA NVHAAANGQGQPEDPPRSSNDAVLL; SEQ ID NO:10) was synthesized by Open Biosystems, AL, USA. Polyclonal antisera against the synthesized peptide were also produced by Open Biosystems. This was done using subcutaneous and intramuscular injections of the immunogen into rabbits as described by Bean (Eric S. Bean (2001) Polyclonal Antibodies. In: Basic Methods in Antibody Production and Characterization antibodies. Howard, G. and Bethel D. (ed.), CRC Press, 5:31-50, 2000). The antibodies were affinity purified from the antisera using standard methodology.

10 *Sequence analysis and transmembrane domain prediction*

Sequence analysis of Saratan protein was performed with the following programs: HMMTOP 2.0 (world wide web enzim.hu/hmmtop/) (Tusnady GE, Simon I. Principles governing amino acid composition of integral membrane proteins: application to topology prediction. J Mol Biol. 1998 Oct 23;283(2):489-506, and Tusnady GE, Simon I. The HMMTOP transmembrane topology prediction server. Bioinformatics. 2001 Sep;17(9):849-50), TMHMM 2.0 (world wide web cbs.dtu.dk/services/TMHMM-2.0/) (Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol. 2001 Jan 19;305(3):567-80.), Tmpred (hypertext transfer protocol://world wide web ch.embnet.org/software/TMPRED_form.html), and PRODIV-TMHMM (hypertext transfer protocol://world wide web pdc.kth.se/~hakanv/prodiv-tmhmm/) (Viklund H, Elofsson A. Best alpha-helical transmembrane protein topology predictions are achieved using hidden Markov models and evolutionary information. Protein Sci. 2004 Jul;13(7):1908-17.).

25

MTT cell proliferation assay

MCF-7 cells were seeded at a concentration of 5×10^3 cells/well in 100 μ l RPMI-1640 culture medium containing 10% FBS into 96 well flat bottom tissue culture microplates with affinity purified rabbit anti-Saratan antibodies or the preimmune serum as the control at a concentration of 600 μ g/ml, or with PBS as the untreated control. The cells were incubated for 48 h in a humidified incubator with 5% CO₂ at 37°C. After the incubation period, 10 μ l of the MTT (3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyl tetrazolium bromide) labelling reagent was added at a final concentration 0.5 mg/ml to

each well and the cells were labelled for 5 h in the incubator. At the end of the incubation period, the medium was removed and the converted dye was solubilized with 100 µl solubilizing solution (0.01N HCl in 10 % SDS in water). The plates were then sealed and placed at 37°C overnight to completely solubilize the purple formazan crystals. The absorbance of the converted dye was measured using Spectra Max 250 multiplate reader at a wavelength of 595 nm. Each treatment was performed in triplicate.

EXAMPLE 2: RESULTS

10 *Expression of Saratan in MCF-7 mammary carcinoma cells.*

RT-PCR was used to detect Saratan expression in mammary carcinoma MCF-7 cells. As shown in Fig. 1, RT-PCR produced a specifically amplified fragment of the predicted size (515 bp). The identity of the PCR product was confirmed by restriction enzyme digestion and DNA sequencing.

15

Tissue-specific mRNA expression

A panel of cDNAs from various human tissues was screened for mRNA expression with primers that are specific for human Saratan. Saratan transcript was detected in most tissues tested (Fig. 2A). The highest expression was found in fetal brain, cerebellum, brain, and spinal cord: Modest expression was seen in skin, kidney, fetal liver, adrenal gland, cartilage, ovary, uterus, and pancreas. Some expression or undetected expression was noted in prostate, skeletal muscle, normal colon, small intestine, stomach, thymus, liver, lung, trachea, placenta, adipose, stimulated PBL, thyroid, umbilical cord and PBL, salivary gland, heart, and spleen.

25

We also examined the expression of Saratan in a number of human cancer cell lines using RT-PCR. As shown in Fig. 2B, Saratan was differentially expressed in the cancer cells tested. Higher expression was detected in DLD-1 colon carcinoma cells and T47D (estrogen receptor positive), BT549 (estrogen receptor negative) and MCF-7 (estrogen receptor positive) breast cancer cells. Moderate expression was observed in A549 lung carcinoma cells, DU145 prostate cancer cells, and RL95-2 endometrial carcinoma cells. Some expression was observed in AGS stomach cancer cells.

30

Saratan is a cytoplasmic and membrane bound protein.

To investigate the subcellular localization of Saratan, MCF-7 cells were transiently transfected with an expression construct that expresses Saratan with a C-terminal EGFP-tag, or with the pEGFP-N1 vector as a negative control. The localization of the fluorescent fusion protein was determined by using laser-scanning confocal microscopy. As shown in Fig. 3, Saratan-EGFP fusion protein was localized in discrete cytoplasmic structures and also at the cell membrane whereas the EGFP alone was diffusely present in both the cytoplasm and nucleus.

10 *Saratan is exposed to the outside of the cell membrane.*

On the basis of hydropathy analysis using the web-based prediction programs HMMTOP, TMHMM, Tmpred and PRODIV-TMHMM, Saratan was predicted to contain a single transmembrane domain (residues 59–81) at the C-terminus. However, the topographical analysis by those programmes differs. The large portion of the N-terminal (residues 1–58) was predicted outside of the cell membrane by HMMTOP, 15 TMHMM, and Tmpred.

To investigate the membrane topology of Saratan, we constructed two Saratan expression plasmids that express Saratan with the HA epitope tag attached to either the N or C terminus, respectively. As shown by Western blot analysis (Fig. 4A), both N- and C-terminal HA-tagged Saratan could be easily detected by the rabbit polyclonal HA-Probe (Y-11). We thus used a membrane permeabilisation strategy coupled with confocal immunofluorescence microscopy to directly examine the membrane orientation of Saratan (panels a-f, Figs. 4B). Triton® X-100 is the most commonly used permeabilisation agent for immunofluorescent staining and it efficiently solvates cellular membranes without disturbing protein-protein interactions. MCF-7 cells transfected with phCMVc-hST or phCMVn-hST were fixed with formaldehyde, permeabilized with 0.5% Triton® X-100, and then immunostained with antibodies against HA. 25

30

When we permeabilized MCF-7 cells containing phCMVn-hST (panel a, Fig. 4B) or phCMVc-hST (panel c, Fig. 4B) with Triton® X-100, we observed the cytoplasmic distribution of HA-tagged Saratan immunofluorescence, similar to the pattern observed

with EGFP-tagged Saratan (Fig. 3). Thus, the HA epitope tags did not alter the localization of Saratan, and both the N- and the C-termini of Saratan were accessible when the membranes were permeabilized. Without permeabilisation with Triton[®] X-100, HA-tagged Saratan immunofluorescence was still observed in cells containing
5 phCMVn-hST with clumped distribution around cell membrane (panel b, Fig. 4B), indicating that the outer membrane remained permeable to HA-specific antibodies. However, the HA-tagged Saratan immunofluorescence was not detected in cells containing phCMVc-hST without permeabilisation (panel d, Fig. 4B), indicating that the inner membrane remained impermeable to the antibodies.

10

As a control, MCF-7 cells were also transfected with the empty control vector phCMV1, however, no immunofluorescence was observed in cells with treatment (panel e, Fig. 4B) or without treatment (panel f, Fig. 4B) with permeabilisation agent. Therefore, we concluded that human Saratan contained a C-proximal transmembrane
15 domain with its N-terminal portion exposed to the outside of the cell membrane which was accessible by antibodies without detergent treatment.

Establishing MCF-7 cells stably expressing Saratan.

To elucidate the role of Saratan in breast cancer, Saratan was cloned and Saratan-
20 overexpressing cells were generated. An insert containing the coding sequence of Saratan in an IMAGE cDNA clone was subcloned into phCMV1 expression vector. The human breast cancer cell line MCF-7 was stably transfected with the Saratan expression plasmid construct phCMV-hST (designated MCF7-hST) or the empty vector phCMV1 (designated MCF7-Vec). As shown in Fig. 5, overexpression of Saratan was detected in
25 MCF7-hST cells compared with the control MCF7-Vec cells.

Forced expression of Saratan significantly increases mammary carcinoma cell proliferation.

To see whether Saratan expression affects cell proliferation, total cell number was
30 assessed for MCF-7 cells overexpressing Saratan. As shown in Fig. 6, forced overexpression of Saratan in MCF-7 cells significantly increased cell proliferation. MCF7-hST cells replicated much faster than the control MCF7-Vec cells. After 10 days, the cell number of MCF7-hST transfectants increased by 85% (calculated as 1.85

fold) as compared with that of MCF7-Vec transfectants.

Inhibition of Saratan inhibits cancer cell proliferation.

Three different target sites in Saratan mRNA were used to knock down Saratan
5 expression. MCF-7 cells stably expressing each of these siRNAs were established. The
effectiveness of each siRNA to knock down Saratan expression was determined by RT-
PCR with Saratan specific primers. As shown in Figs. 7A-B, all three siRNA constructs
tested were able to reduce the mRNA level of Saratan. The knock down efficiencies
were measured as 88%, 74%, and 12% for siRNA-A, -B and -C, respectively. MCF7-
10 siRNA-A cells were used in subsequent assays.

To investigate the functional consequences of reduced Saratan expression by siRNA,
total cell number was assessed. As shown in Fig. 8, siRNA depletion of Saratan in
MCF-7 cells significantly reduced cell count. The total cell number of MCF-7 cells
15 stably expressing siRNA-A was decreased by 40% compared with MCF-7 cells
expressing control siRNA-CK. This indicated that the stimulatory effect of Saratan on
cell proliferation could be reversed by reducing the expression of endogenous Saratan
via siRNA.

20 *Forced expression of Saratan increases BrdU incorporation in MCF-7 cells while
Saratan-specific siRNA specific decreases BrdU incorporation.*

BrdU incorporation assays were used to examine the mechanism by which Saratan
stimulates proliferation of MCF-7 cells. As shown in Fig. 9, forced overexpression of
Saratan in stable transfectant MCF-7 cells increased BrdU incorporation by 27%
25 compared with the stably transfected control cells with vector alone. Reduction of
Saratan expression in MCF-7 siRNA-A cells resulted in decreased BrdU incorporation
by 82% compared to control MCF-7 siRNA-CK cells. From these experiments, we
deduced that the stimulatory effect of Saratan on cell proliferation was due to increased
DNA synthesis. Conversely, the inhibitory effect of effect of siRNA-A was due to
30 decreased DNA synthesis.

*Forced expression of Saratan significantly promotes the migration and invasion of
mammary carcinoma MCF-7 cells, while Saratan-specific siRNA specifically decreases*

such migration and invasion.

The effect of Saratan on cell mobility was evaluated by migration and invasion assays using the standard Transwell inserts. As shown in Figs. 10A and B, forced expression of Saratan significantly increased MCF-7 cell migration and invasion through Matrigel.

- 5 Compared with MCF7-Vec control cells, MCF7-hST cells overexpressing Saratan showed a 1.6 fold increase in migration (Fig. 10A) and a 5.5 fold increase in invasion (Fig. 10B).

- 10 To examine the effect of Saratan on cell mobility, we knocked down the endogenous expression of Saratan using siRNA specific to Saratan mRNA. As shown in Figs. 10C and D, in contrast with forced expression, inhibition of Saratan expression by siRNA significantly inhibited both MCF-7 cell migration and invasion. Compared with the siRNA-CK control, Saratan-specific siRNA-A reduced migration by 40.5% (Fig. 10C) and invasion by 57.9% (Fig. 10D). As noted for forced expression, the knock down of
- 15 endogenous Saratan had more effect on invasion than migration.

Production of recombinant human Saratan protein in bacteria.

- The cDNA coding for the first 61 amino acids of human Saratan was cloned into the GST fusion protein expression vector pGEX 4T1. After transforming the pGEX 4T1-hSTn plasmid into BL21-Gold cells, the production of GST-hSTn was successfully
- 20 achieved upon induction with IPTG. GST-hSTn fusion protein was purified with glutathione Sepharose 4B matrix. After cleavage of GST tag, recombinant Saratan protein was produced. The integrity and purity of the recombinant GST-hSTn protein was determined by SDS-PAGE (Fig. 11). The recombinant N-terminal peptide included
- 25 a sequence of gsMDLSFMAAQLPMMGGAFMDSPNEDFSTEYSLFNSSANVHAAA NGQGQPEDPPRSSNDAVLL (SEQ ID NO:67; the linker sequence is shown in lower case) with a predicted MW of 6.6 kDa, and an apparent MW of ~12 kDa on SDS-PAGE.

- 30 *Characterization of antiserum against synthesized Saratan N-terminal peptide.*

While recombinant Saratan was being produced in bacteria, rabbit polyclonal antibodies were generated against the synthesized N-terminal Saratan peptide. As shown in Fig. 12, the antiserum was able to recognise not only the immunogen but also recombinant

Saratan in Western blot analysis. The antiserum recognised the recombinant N-terminal Saratan as a band of ~12 kDa, as shown by Coomassie blue stained SDS-PAGE (Fig. 11), and the synthesized immunogen as a band of 6.5 kDa, as predicted. The antibodies recognised both synthesized and bacterially produced Saratan in Western blots, in amounts as low as 5 ng.

Endogenous expression of Saratan in MCF-7 cells.

The rabbit polyclonal antibodies raised against Saratan were used to examine the endogenous expression of Saratan in MCF-7 cells. MCF-7 cells were transiently transfected with pEGFP-N1-hST expressing EGFP-tagged Saratan or pEGFP-N1 empty vector. EGFP-tagged Saratan or EGFP proteins were viewed as green fluorescence using UV-visible fluorescence microscopy. Consistent with the results shown in Fig. 3, Saratan-EGFP fusion protein was localized to discrete cytoplasmic structures and also at the cell membrane (Fig. 13B), while EGFP alone was diffusely localized to both the cytoplasm and nucleus (Fig. 13A). As expected, a similar pattern of subcellular distribution as Saratan-EGFP fusion protein was observed using the rabbit polyclonal antibodies against Saratan in both transfected and non-transfected MCF-7 cells (Figs. 13C and D). Moreover, stronger Saratan signals were observed in pEGFP-N1-hST transfected cells (Fig. 13D), but not in pEGFP-N1 transfected control cells (Fig. 13C). These results clearly demonstrated that the antibodies were able to recognise EGFP-tagged Saratan as well as endogenous Saratan.

Rabbit anti-Saratan polyclonal antibodies significantly inhibit MCF-7 cell proliferation.

The MTT assay is based on the observation that only metabolically active cells are able to cleave MTT (yellow tetrazolium salt) to produce formazan (purple crystals), which can be solubilized and quantified by spectrophotometric means. The MTT assay has been accepted as one of the most sensitive and reliable cell biological approaches to quantitatively measure cellular proliferation, viability, and activation of a population in response to external factors.

As shown by RT-PCR (Figs. 1 and 2B) and by immunofluorescence staining (Figs. 13C and D), MCF-7 cells express Saratan at relative high levels. We used MTT assay to test the effect of rabbit anti-Saratan polyclonal antibodies on the proliferation of MCF-7

cells. As shown in Fig. 14, while the inclusion of the preimmune serum had little effect on MCF-7 cell proliferation as compared to the untreated PBS control, the Saratan antiserum strongly inhibited the MCF-7 cell proliferation. After 48 h of incubation, Saratan antiserum reduced the viability of MCF-7 cells by 56.1% compared with the preimmune control.

Rabbit anti-Saratan polyclonal antibodies significantly impair migration and invasion of mammalian carcinoma MCF-7 cells.

To demonstrate a physiological role of Saratan in cell migration/invasion, we assessed whether using antibodies to block Saratan function would impair motility/invasiveness. As shown in Figs. 15A and B, the number of cells that migrated through the non-coated (migration) or Matrigel-coated (invasion) porous filters of Transwell inserts was significantly reduced by addition of rabbit anti-Saratan polyclonal antibodies, as compared to the preimmune serum control. This was in marked contrast to forced overexpression of Saratan in MCF-7 cells. Using anti-Saratan antibodies, we observed a 54.9% decrease in migration and 35.0% decrease in invasion.

Rabbit anti-Saratan polyclonal antibodies significantly impair migration and invasion of mammalian carcinoma MDA-231 cells.

Further experiments were performed to confirm the inhibitory effects of Saratan antibodies on cell mobility. In an initial assay, mammalian carcinoma MDA-231 cells were screened for the expression of Saratan. As shown in Fig. 16, Saratan transcript was easily detected by RT-PCR. This expression was further demonstrated immunofluorescence staining with the rabbit polyclonal antibodies against Saratan. Only weak signals were observed by the preimmune serum (Fig. 17A) whereas enhanced signals from anti-Saratan antibodies were detected (Fig. 17B). Like in MCF-7 cells, Saratan was distributed in discrete cytoplasmic structures and at the cell membrane (Fig. 17DB).

To determine whether Saratan antibodies would inhibit MDA-231 cell migration/invasion as in MCF-7 cells, MDA-231 cells were treated with the purified rabbit anti-Saratan antibodies or the preimmune serum as the control at a concentration of 600 µg/ml in migration and invasion assays. As shown in Figs. 18A and B, both migration

and invasion of MDA-231 cells were markedly reduced by inclusion of the rabbit anti-Saratan polyclonal antibodies compared with the preimmune serum control. We observed a 41.0% decrease in migration and a 87.4% decrease in invasion.

- 5 Unlike MCF-7, which is estrogen-dependent and poorly invasive, MDA-231 is estrogen-independent and highly invasive (Hiraguri S, Godfrey T, Nakamura H, Graff J, Collins C, Shayesteh L, Doggett N, Johnson K, Wheelock M, Herman J, Baylin S, Pinkel D, Gray J. Mechanisms of inactivation of E-cadherin in breast cancer cell lines. Cancer Res. 1998 May 1;58(9):1972-7; and Larkins TL, Nowell M, Singh S, Sanford
- 10 GL. Inhibition of cyclooxygenase-2 decreases breast cancer cell motility, invasion and matrix metalloproteinase expression. BMC Cancer. 2006 Jul 10;6:181). Therefore, it was not surprising that a more dramatic decrease in cell invasion was observed (Fig. 18B).
- 15 From the sum of these observations, we concluded that inhibition of Saratan expression/function can be utilized to decrease cell proliferation, cell survival, and/or cell motility, and treat disorders that are associated with these cellular activities.

The invention has been described herein, with reference to certain preferred

20 embodiments, in order to assist the reader in practising the invention without undue experimentation. However, a person having ordinary skill in the art will readily recognise that many of the components and parameters may be varied or modified to a certain extent without departing from the scope of the invention. Furthermore, titles, headings, or the like are provided to enhance the reader's comprehension of this

25 document, and should not be read as limiting the scope of the present invention.

The entire disclosures of all patent applications, patents, and publications, cited above and below, if any, are hereby incorporated by reference in their entirety. Throughout this specification, and any description which follows, unless the context requires

30 otherwise, the words "comprise," "comprising" and the like, are to be construed in an inclusive sense as opposed to an exclusive sense, that is to say, in the sense of "including, but not limited to".

WHAT IS CLAIMED IS:

1. An isolated Saratan RNA molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO:101 and 120-122, or a sequence complementary thereto.
5
2. An isolated Saratan RNA molecule comprising the nucleic acid sequence of SEQ ID NO:74, or a sequence complementary thereto.
- 10 3. An isolated Saratan RNA molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO:75-77, or a sequence complementary thereto.
4. An isolated Saratan RNA molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO:68-73, 78-89, and 96-98, or a sequence
15 complementary thereto.
5. An isolated Saratan RNA molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO:90-95, and 99-100, or a sequence complementary thereto.
20
6. An isolated Saratan RNA molecule comprising a nucleic acid sequence having at least 95% identity to any one of SEQ ID NO:101 and 120-122, or a sequence complementary thereto.
- 25 7. An isolated Saratan RNA molecule comprising a nucleic acid sequence having at least 95% identity to SEQ ID NO:74, or a sequence complementary thereto.
8. An antisense RNA molecule which binds to the nucleic acid sequence of any one of claims 1 to 3.
30
9. An antisense RNA molecule comprising the nucleic acid sequence of claim 4 or claim 5.

10. A small interfering RNA molecule which inhibits translation of the nucleic acid sequence of any one of claims 1 to 3.
11. A small interfering RNA molecule comprising the nucleic acid sequence of claim 4 or claim 5.
12. An RNA oligonucleotide comprising the nucleic acid sequence of any one of claims 3 to 5.
13. A vector which expresses the nucleic acid sequence of any one of claims 1 to 7.
14. A host cell which comprises the vector of claim 13.
15. An isolated Saratan polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1 and 102-104, or a functional fragment thereof.
16. An isolated Saratan polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:8-10.
17. An isolated Saratan polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:49-64, and 115.
18. An isolated Saratan polypeptide comprising an amino acid sequence having at least 97% identity to SEQ ID NO:1.
19. An isolated Saratan RNA molecule which encodes the amino acid sequence of any one of claims 15 to 18.
19. A Saratan antibody or antibody fragment which binds to the amino acid sequence of claim 15.
20. A Saratan antibody or antibody fragment which binds to the amino acid sequence of claim 16 or claim 17.

21. The Saratan antibody of claim 19 or claim 20 which is monoclonal.
22. A composition comprising the isolated Saratan RNA molecule of any one of claims
5 1 to 7.
23. A composition comprising the antisense RNA molecule of claim 8 or claim 9, or
the small interfering RNA molecule of claim 10 or claim 11.
- 10 24. A composition comprising the vector of claim 13 or the host cell of claim 14.
25. A composition comprising the isolated Saratan polypeptide of any one of claims 15
to 18.
- 15 26. A composition comprising the Saratan antibody of any one of claims 19 to 21.
27. A pharmaceutical composition comprising the composition of any one of claims 23,
24, or 26, and a physiologically acceptable carrier.
- 20 28. Use of the composition of any one of claims 23, 24, or 26, in the preparation of a
medicament for the treatment of cancer in an individual, wherein the cancer is selected
from the group consisting of breast cancer, colon cancer, lung cancer, prostate cancer,
endometrial cancer, and stomach cancer.
- 25 29. The use of claim 28, wherein the cancer is metastatic cancer.
30. A method for treating cancer in an individual comprising administering the
pharmaceutical composition of claim 27 to the individual, thereby treating the cancer,
wherein the cancer is selected from the group consisting of breast cancer, colon cancer,
30 lung cancer, prostate cancer, endometrial cancer, and stomach cancer.
31. The method of claim 30, wherein the cancer is metastatic cancer.

FIGURE 1

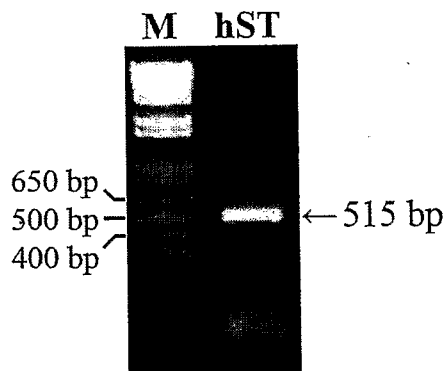


FIGURE 2A

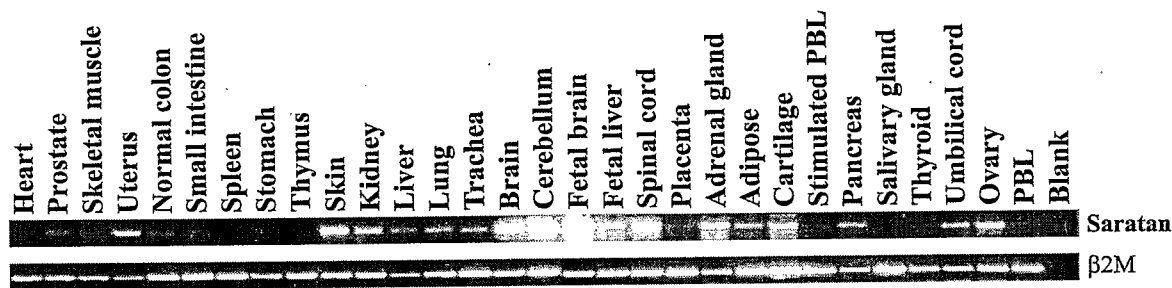


FIGURE 2B

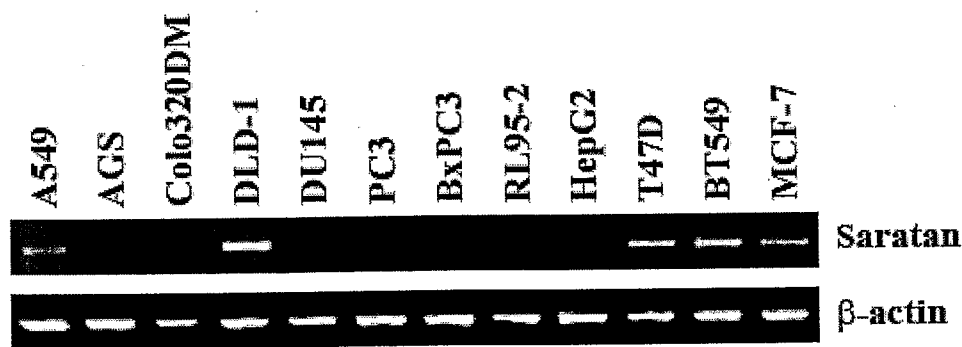


FIGURE 3A

EGFP

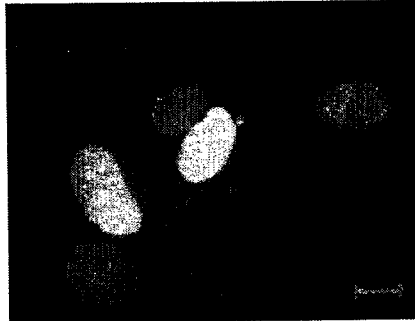


FIGURE 3B

hST-EGFP



FIGURE 4A

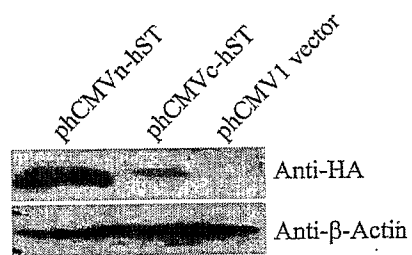


FIGURE 4B

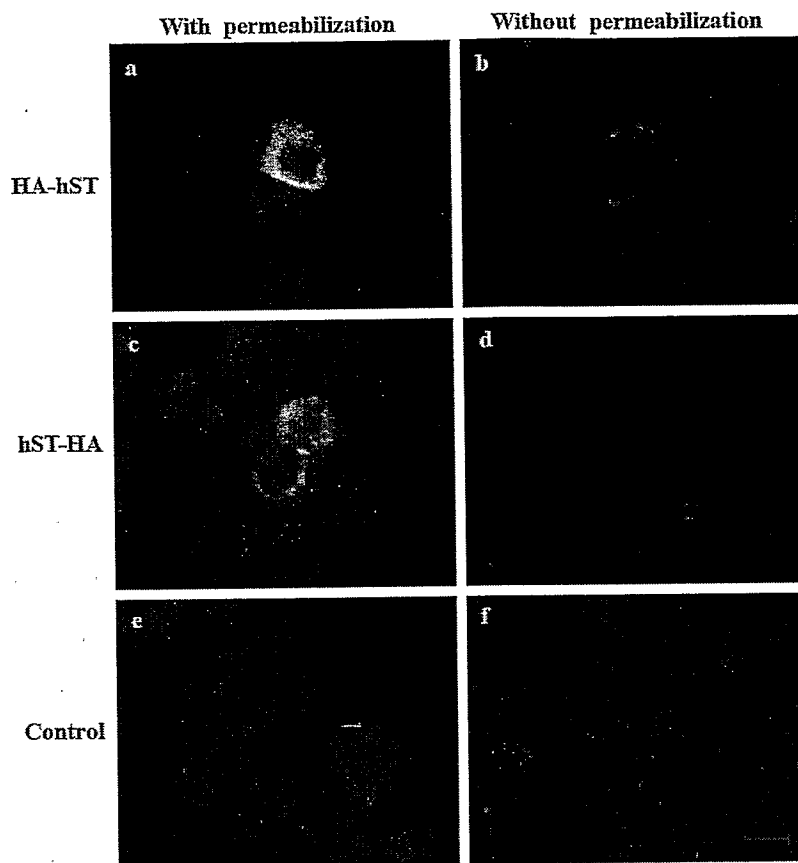


FIGURE 5A

FIGURE 5B

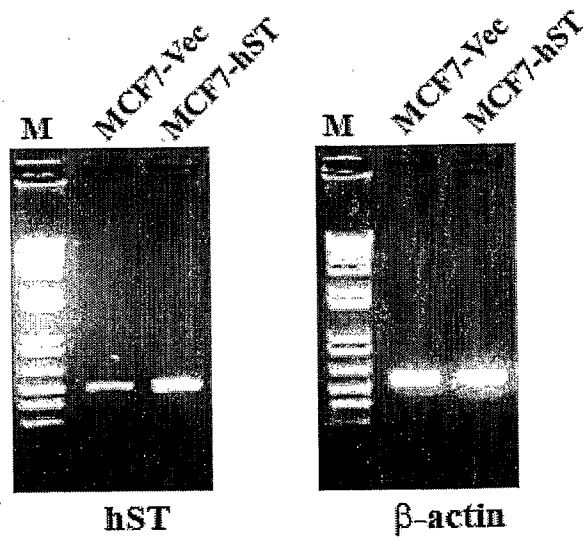


FIGURE 6

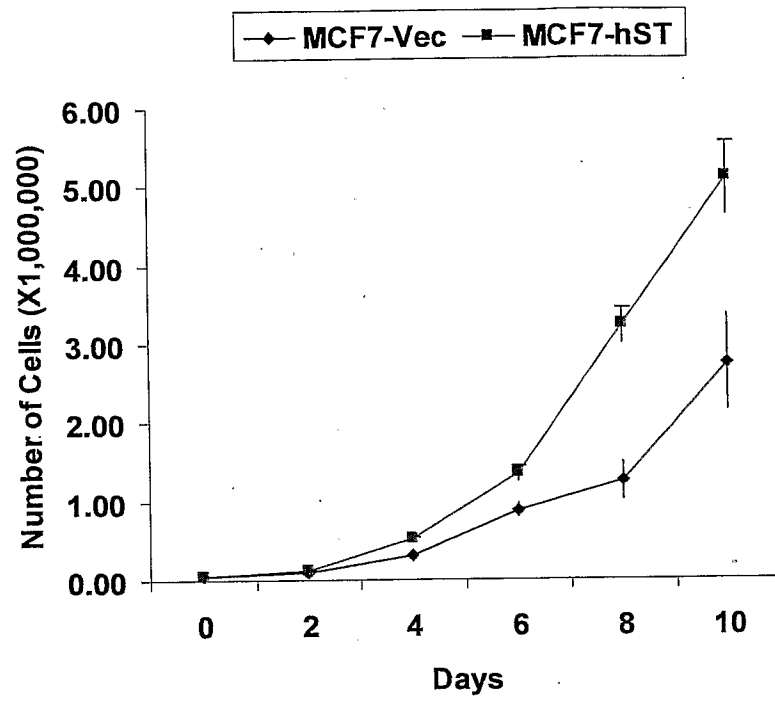


FIGURE 7A

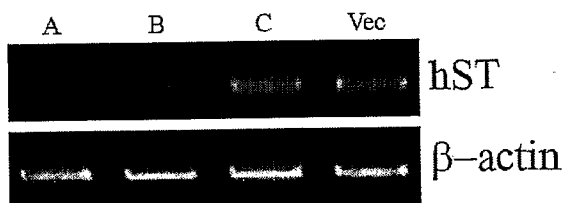


FIGURE 7B

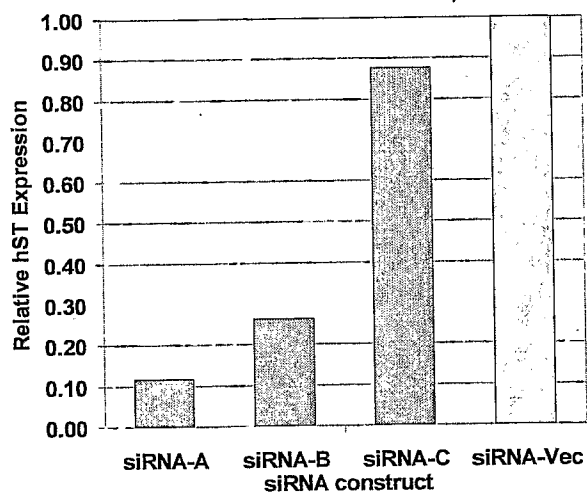


FIGURE 8

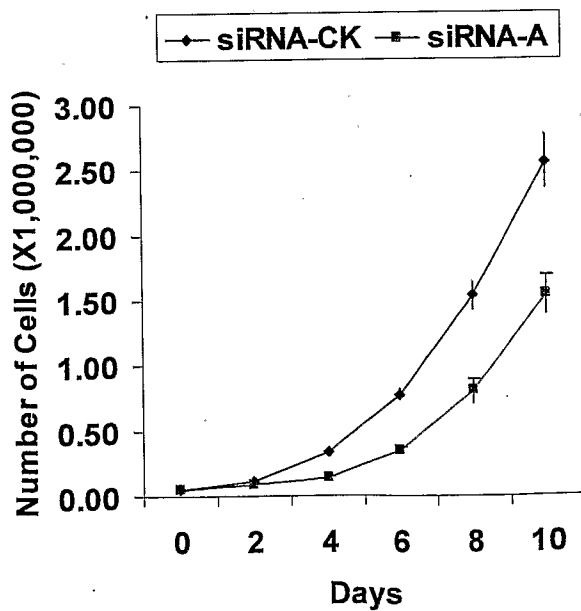


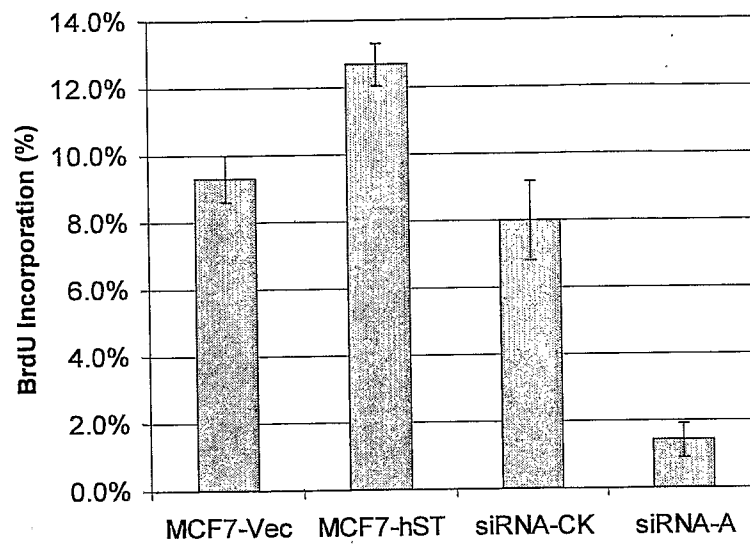
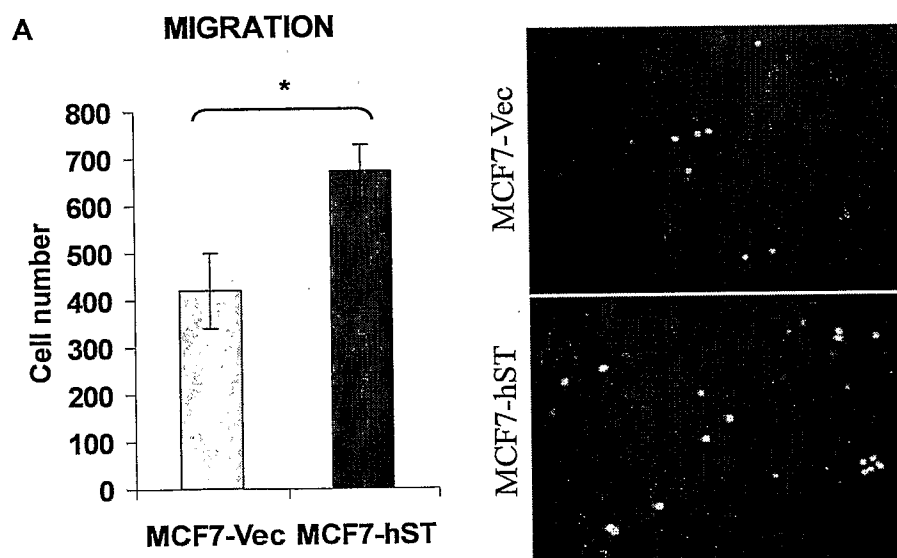
FIGURE 9**FIGURE 10A**

FIGURE 10B

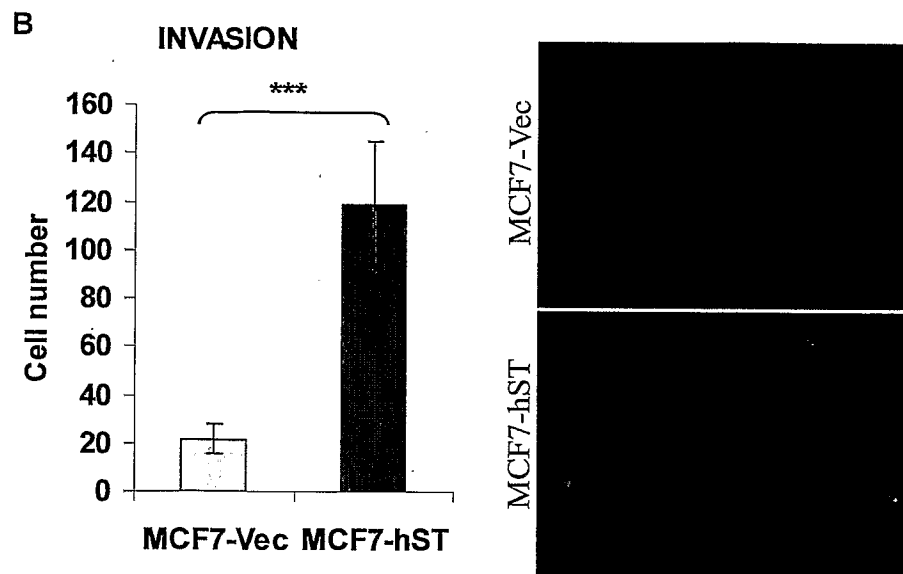


FIGURE 10C

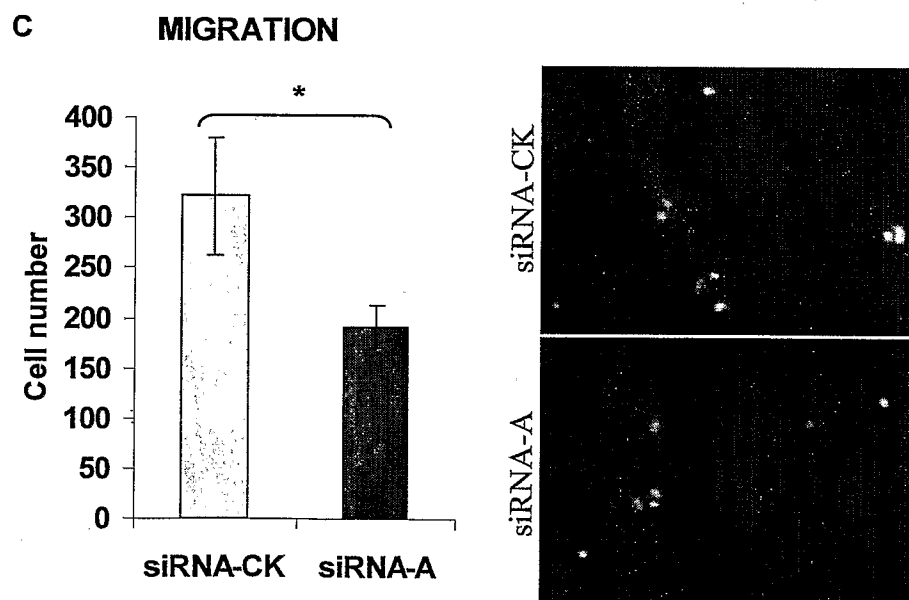


FIGURE 10D

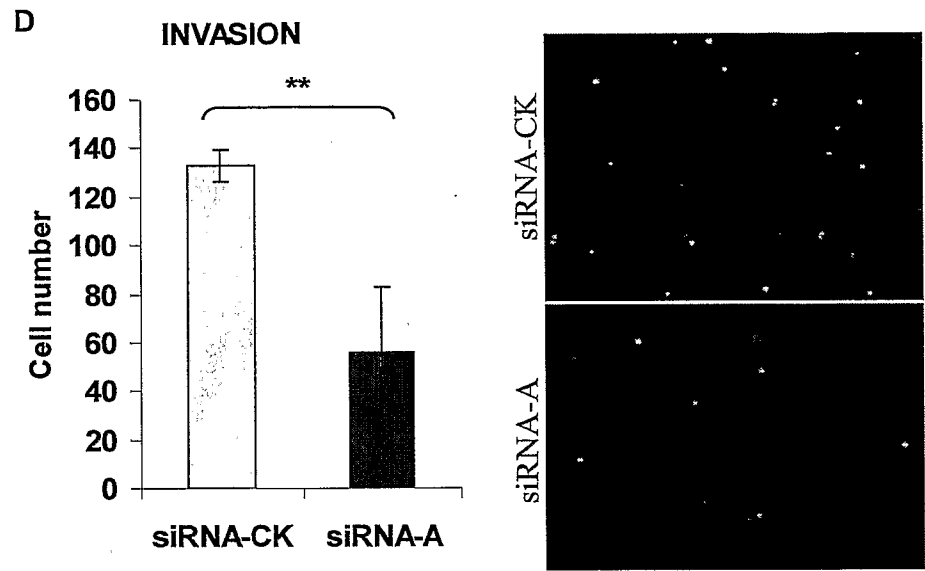
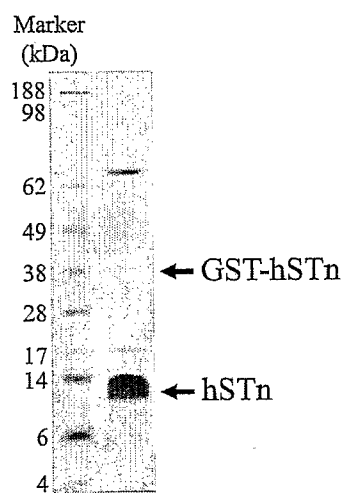
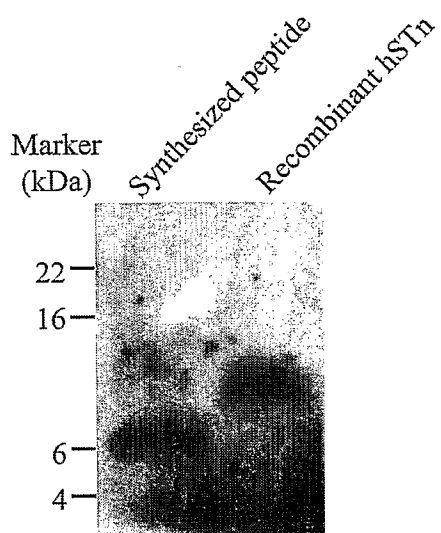


FIGURE 11**FIGURE 12**

FIGURES 13A-H

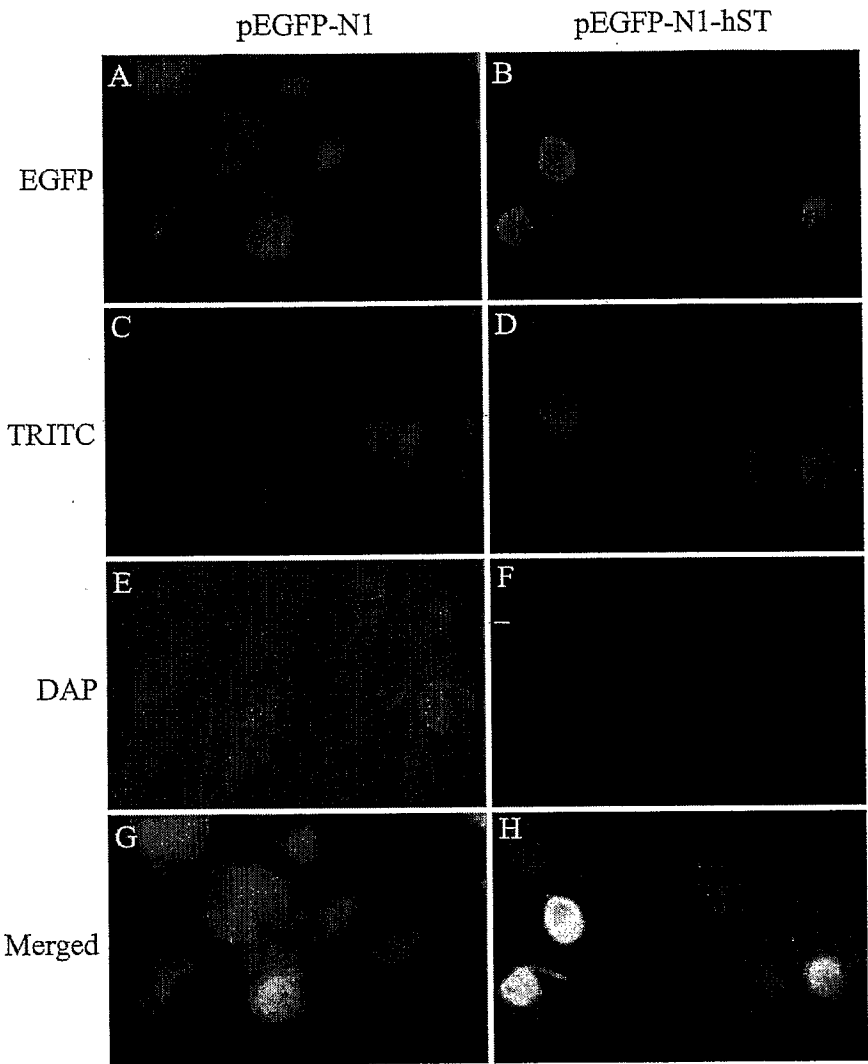


FIGURE 14

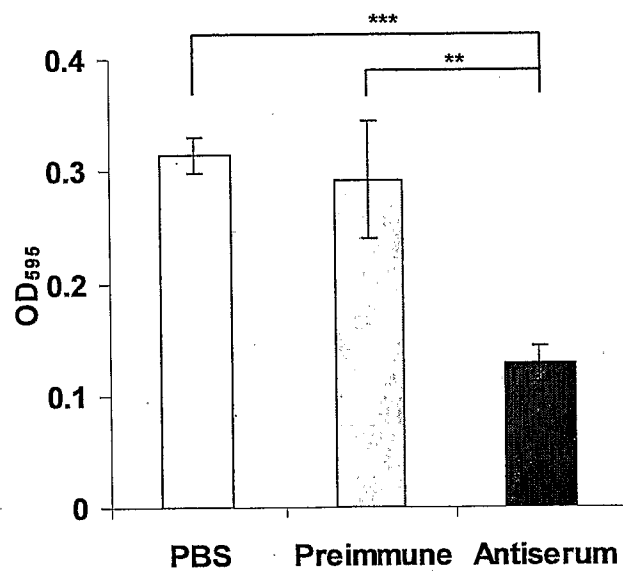


FIGURE 15A

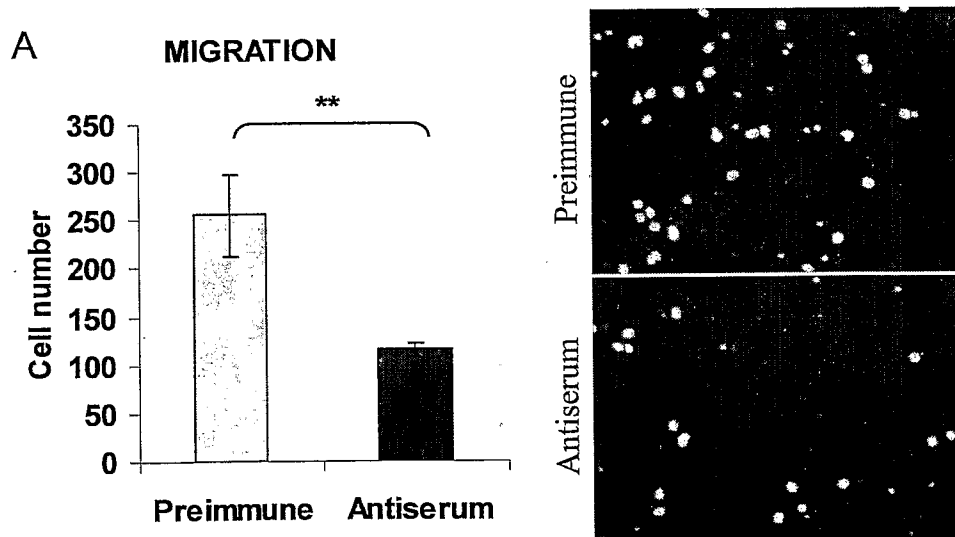


FIGURE 15B

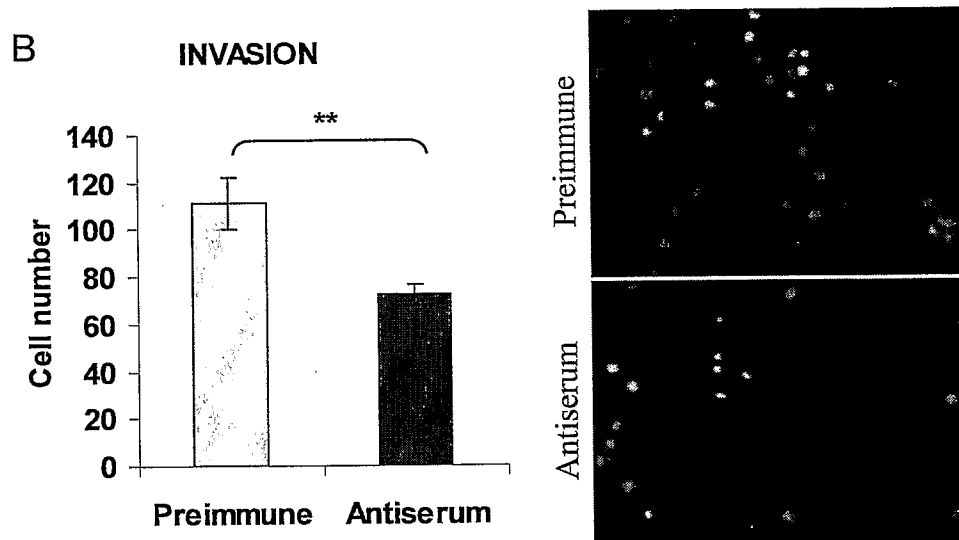
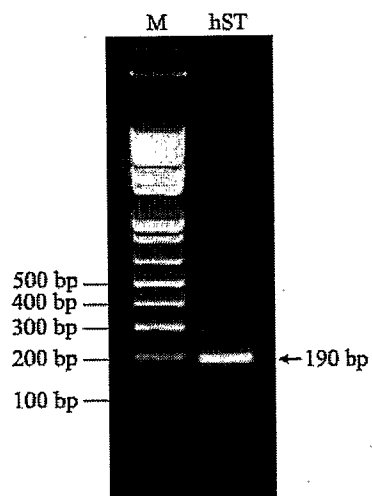
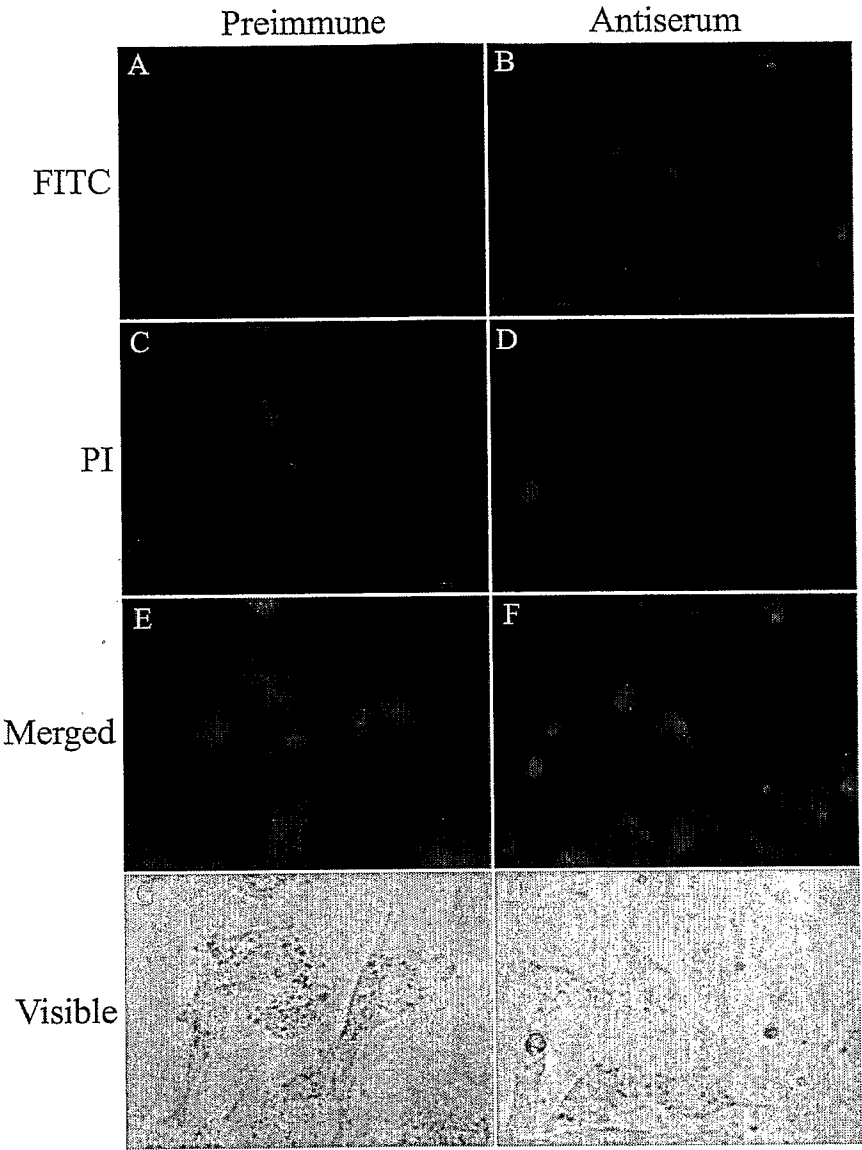


FIGURE 16



FIGURES 17A-H



FIGURES 18A-B

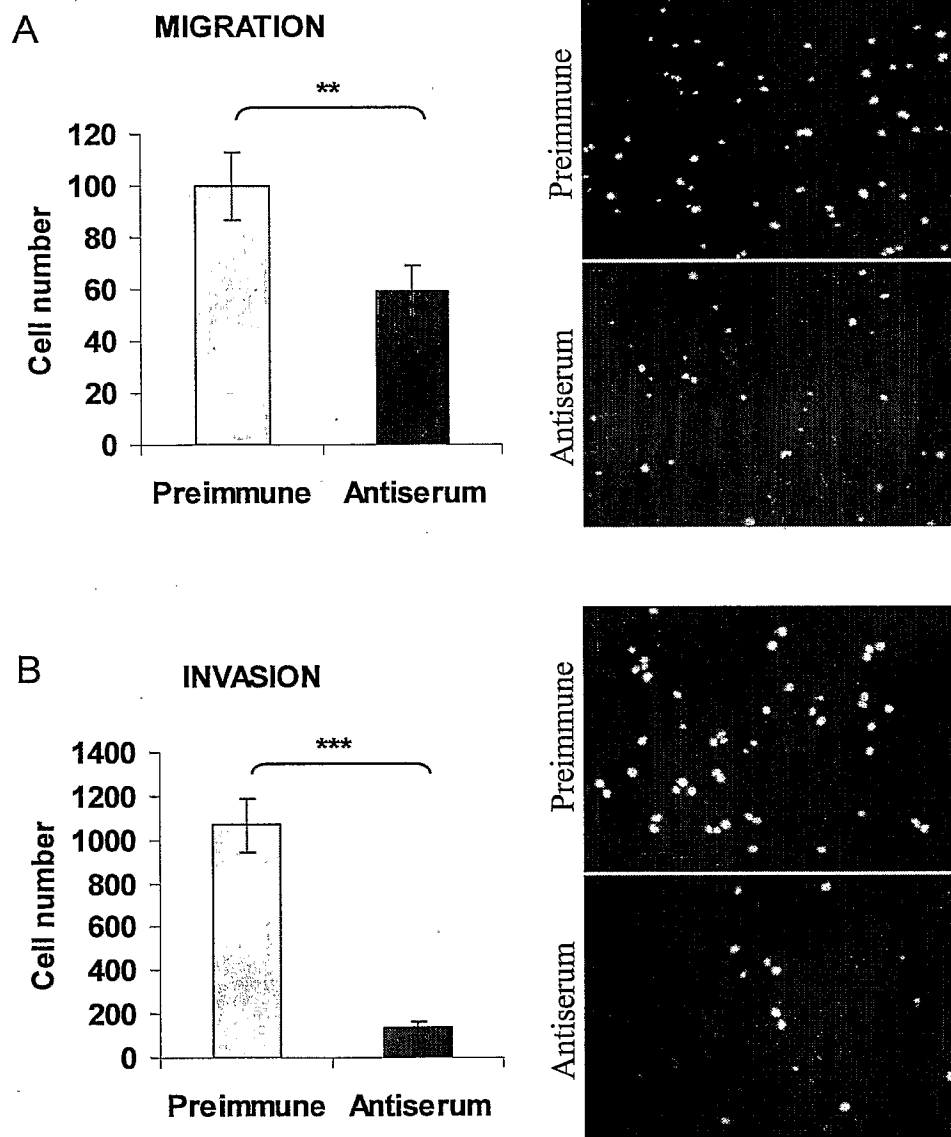


FIGURE 19A

10 20 30 40 50
 MDLSFMAAQL PMMGGAfMDS PNEDFSTEYS LFNSSANVHA AANGQGQPED
 60 70 80
 PPRSSND~~AVL~~ LWIAITATLG NIVVVGVVYA FTF (SEQ ID NO:1)

FIGURE 19B-1

1 gcctagtaga gatctggagc cagaagccca gagacagccg agtgcgccgt gcggtctccg
 61 gacgctcgct gctcagcccg atccccgcca actgtgcagg cggctgaccc gcagcggcag
 121 cggcagcagc gaggactcga gcgctggctg cagcgacacc atggatctct cctttatggc
 181 cgcgcagctg cccatgatgg ggggagcttt catggactcg cccaacgagg acttcagcac
 241 cgagtactcc ctgtttaact cctctgccaa tgtccacgcg gctgccaatg gccagggccca
 301 gccggaagat cctcctcggt cctccaacga cgccgtcttg ctatggattg ccatcatagc
 361 tacgctgggg aacatcgtgg tgggtggcgt ggtgtatgcc ttcaccttct gaggacggca
 421 caccctgcac caccatgggg tgaggcttgg cactgagctc tgacttgctg tcggcctttg
 481 gcttctcctg tgttctagaa ccaggagttt tgaccagggg cggcggccgt ccttctggaa
 541 tttctcccca gcagccctga tttcaaatat cccatgttgt ggtcaagctg agtcagaaga
 601 catggaagta tgggcctcct gcccctagag gcatgacggg gcaaggcctt cagagggcag
 661 attggggatc cttgaaacta cattccagga acatgggacc agatgagaca gctagttaag
 721 tttaaaacat agacatgatt tgatgatcgc ttgctggtgg taaataatca ctcgtgtgac
 781 ttgtttttat gcaaacctat cgaacctagg gcgtggggtg ctggggcaag agcagccctc
 841 agaacttcag tgttctgac ccaattctgg tttcacattc agtcccttgg ccatctagta
 901 gggccattgg atgttcttag tttgactttg aaatggcacc tttgccacca gacacctggg
 961 cccttccaag acccaagtgc attgggagac ccagggatgg ggggttactg gtaatagggt
 1021 ggggttctgg ggggtgtgtt ggtggttttt atctcctggt cggactttct cttcttttta
 1081 agaagaggag aggatgtctt taagagctag atgtgccagg cagtggactc ttcaggccac
 1141 ccacgtgaga atgctgttcc ttctctgagg aatcgtggaa ttttaaagat gaacaagatt
 1201 cacatccact gaattattca acggatgggt cagaaagggg ggtgatttgc ctgtggtcac
 1261 caggcaggtt tgtggaggag tcggaacaga aagatgtgtc ctactctca gctcactgag
 1321 ctctctgccc caactaagca cttccctgag gaggttctg agaagctgcc ctcaggagaa
 1381 tgtccaggca tcttgaaggt ggggtgcgaga ttggcaggct gctcagatac ccgtccttta
 1441 cattcagtgt ggataccgtg ccttctctg aagctgtgaa agtgccttct cccaagccac
 1501 ttcctttaat tctgaaatat cagcatctgg ggtcctggca agcaaggaag cttccaagta
 1561 aaaaccagag agaagggcac acttttcttt cttcattagg aaatcttatt gcacaggaac
 1621 cacccccacc cccaccccc acacctccc aaggcagcat ccagtgagc atagagtggg
 1681 aaaggtccca gaagggggct cactcacctc taggccaga gaggttttct cctcacttta
 1741 tacactgcaa aaacagaaga attgtgtcaa taacaccctc tgtagtgagg aaacttaaaa
 1801 agctggtag gaagctctcg tgtatattta gagacaatta caagaaagct ggacttgcg
 1861 ctgtggtctc aggagaaatg agtgttcttg atgacaggca aaggacatc ttagtgtcc
 1921 agaagcggca ctcttccctg gaagccgcca tgttaatagg attactagcc tggctccaga
 1981 cagtgcctgc tcatggctgc cagttcttac cgatcacatc tgtcactgcc accgtatatc
 2041 atctgccagt gcatcagctt aaggggaggt cactgagtgca aaagaacctg acccttgaca
 2101 atgagggaga agggacatgg accacctgtc tggaattcct ggaatcactg gcaggggtgga
 2161 ggctgggctg gggagttagc cgcggtgtgc gtgaatggct ctgtctccag caagtctctc
 2221 tccatcaaac ccaggtctg cccataagc aagatcttta acagatggat gtctccatga
 2281 gaaaacccaa ggcgagaagc ccagagccat ggcggggtg cttgacgtcc tcatggagtc
 2341 actctgcccc acatgctcaa atcttccctc tggccccaca tccctaggag ggctgaccc
 2401 ctgtaaagat acaggaggca gctccctggc ctccaaatgg cccatggaga tggcagtcgg
 2461 gagacagggt tctgtgtttg ctgcggtgaa gggaggagaa ggcaggagga aaaaggatgg
 2521 cttctagccc tgaagaggac tccagcatcc caggcaccgg gtgcttctg ctgcagtttt
 2581 ccctatggag gcccctcagc ctccagccct aacataaatg tcggttaaat tcagttttca
 2641 agcctctctc ctttttcagt gtcagagcag tagatgggtc agggcattgg aggcctcgac
 2701 cactctgcat tgcagattac agtgacttcc tcgggggttg cccatcttgg tctcctgtgg
 2761 tttcttcatc agcttttttt ttaccagcat ctctcaaata acaatgaaga tagatatgcc
 2821 cattagtgtc tgattaagga gcaaaggctg gatttctggc cacagcgagc tgcactctcc
 2881 ctctgcctc agccgggggtc cgtcttagca gtttgaaaag gggaaaaaga tgccgggtcct
 2941 cactgcttaa gttttgtgtc caggtgccac tagacttgca tgcacactaa ctccttacia
 3001 tcaccacaca gcatcatcgc ccagtgacac agatgaggaa ccagaggctc agaggagtga
 3061 agttgccttc ctgaggtcac acagcatgaa agtgatgagc taggatttga atctgggaag
 3121 ttgggctcta gagccagact gtactgcctt ctgccacact gtactgcctt ctgtgactgg
 3181 gtggcacctc cagggcacat ttacacaagg ccctgaatct gcagaggctg tttctcaaga
 3241 tgcccgctcat ggtgtggcct gggccagctc tggtctccac aggtccctga ctgtcctcag
 3301 agtggaaacat gctcaacctc ccgcccactg ctctctctct gccagattt caggggtgcc

FIGURE 19B-2

```

3361 ggtccccaag gctgcccc ttctttaaga ctgaactcaa gtctccttgg aaggccccgg
3421 tgaagctccc agagactggt tttcttggga tgcaggcaga aggggaccct ccctggccaa
3481 caccaggag cccagcagaa gcacccacac gtagaaagag gctcactaca gccagaagtg
3541 cagagtcaga gtccctgggac catcttgttc tgcaagggtga cccagggtc cccaggacag
3601 gggagaggga tgcctctcat tcagactcta gctggggcct ctgtactggc ttctccttgg
3661 gtggggttgc ctgttacata gctgtgcctc agagaaagggt toctgcattt tctggaatgt
3721 tctctgtgct taccctctgt tgtgcccctc cattgtcctc ctacaagcaa ttaggtgatt
3781 caaaagagca acttaggctg ggtgcagtga ctacaccccg taatcccgcc actttgggag
3841 gccgaggcgg gcagggacag gagttcaaga ccagcctggc caacatgggt aaaccctgtc
3901 tctacaaaaa atacaaaaat taaccagaca ttgtggcatg tgccgtgaat ccagctact
3961 caggaggctg acacaggaga attgcttga caggaggcg gaggctgcag tgagctgaga
4021 ttgtgccact gcactccagc ctgggcaaca caacgagact ctgtctcaaa aaaaaaaaaa
4081 agaaaagaaa aaaaaaagag caacttactg ctttgtgagg ttgtgagtgg ccaccactga
4141 aggtctttga gaagaggcca gacgccgtg tagccaggcc tgtcttaaga ggacttgtgc
4021 ttgtgccact gcactccagc ctgggcaaca caacgagact ctgtctcaaa aaaaaaaaaa
4081 agaaaagaaa aaaaaaagag caacttactg ctttgtgagg ttgtgagtgg ccaccactga
4141 aggtctttga gaagaggcca gacgccgtg tagccaggcc tgtcttaaga ggacttgtgc
4201 ttccagggca ccaggcagga tgatggcgca gctcttctta ctccaagcca acgtgtcctc
4261 tcccttttcc catgaaatca aggtcaagag gcaaataaga ctccctgctc cactctaccc
4321 cccagagaga aatgattctc gctcctttca gatccccag gatctgaggg agaaaggatg
4381 ggaggagggg cagcagcatt tgcctggaaa ggcagcagat gcttttccag ccccggttca
4441 gctggaaggc ttggaggcgg gccagaccac tctggcgtct cctgaagtgg gtccctggag
4501 accgaaggag ctcagtggag tctgtctgtt gtcagcactg ctgcctgatc cctgcaagac
4561 aaatggcact ttcttcttcc agaagcatca tctgccttca ttattagcag taatattatt
4621 cccagttatt attcttaccg gtgccagttt tgcacatctt tttgttgcct tatttgtgtc
4681 tcattttact ctcaaattgc ccctgggggc aggaatgagg atgcagagag atgcacgtta
4741 attactgtcg catttttctg tggagaaaac tgaggccagg gctgaacgct cacagctaag
4801 gagctgtgat tcagaccagc ttctgtcagc tctagaggca cctgcatca tgcccaccag
4861 ggtgatcccc ctgggatgga ccatctcggg atatgaggcc tcggaggctg gggttgagat
4921 ttggtcctga agagcttata gccagattgc cacattcaag tgtaagtcca ggaagggggc
4981 aggcggcagt gcacagggat ttatcagttc cagaacctca cagtataaag aggtcttaga
5041 gagcatctaa tcgagacctt taatttttct gggagagcag ctgaggccgt gtggaaaatt
5101 agtggagagc tgacaagtgt ctgggctcct ggcccagggg tccgtggtcc agcacgttgt
5161 gcgttcagtg ggaagcaaag ggcttgcccg ggattacctg cccagcccc gaggtgggtt
5221 gtgctccctg cagctgccat cggcccgtt tgcttcgtcc tggcagatgc ccagtgattg
5281 tccccgagca agtgccaggg ttgggctgag ctgctatgac agggaggccc agggagttct
5341 gctcagggag ccaaagggaa cagccagatc ctgaatgttc tatgttcacc tgccccagcc
5401 ccaccacccc tggcccactc cacaggcccc tgaccatggt cactcacgga gagggtgga
5461 ggagaagggt gttgaactga gtactgagaa cccagaggac agagcccaca gcttccaagc
5521 aggaaaaggg acctctctga aaaatctgga taaccagaat tatcacagca cctctcatt
5581 cccagcgcgt ccttctgagc tcagaccttg agcattttact gggtttcttt ttgaggaaga
5641 gggaaagtga caaaggacaa aacaatgcaa atcttcatga ctgaagacga tcaaagactc
5701 cctggagcga gaaaacagtt actgccagaa gcagaatgga agagccaaaa agtacacaaa
5761 attggacgca taaattctga aataaaagtg tatgatgtgt tctgagtcac tgtagaagtc
5821 atgcatttat tatcaagata gaaaagagca gagaatgacg tggacattgg toctcggaga
5881 ggctgcatag gtggtgcggt ccccggggga ttctggatgc tggctttttg accgtggcgg
5941 cagcctcgcg cctgcccgga tggctccatc cagacatttg gcaaggctgt catctgtctc
6001 tgggtccttt ttcaagctga tttcctgcct cccaagagg aggtttgagc cccattctgc
6061 cttggaaata aatcctgaca gatgtgcaca gcatatttgc aggggaattg caggcctcca
6121 ttattggggg aatcagagcc cattccacac cgagccccat ccacgtggct gagcaacaga
6181 cctgtgtccc ctgtatctgt ctcactggat gtctttctgc cagtttctta aagaaccaca
6241 ccattactgc atttgccgtt cgaagcgttg tccaacaat gcagatggtt ctgacaaggg
6301 tctatatgct ggcagaaggg agcttccaac cttttaactt gaggaaaatg aatctctgac
6361 aggcttgaga gtgttgccag gtggagcttt tcaggagaca ggggtcttga gccaggaca
6421 cctaactatc gagttttcac taggagactt agtgggtggc ttcattgaggc cagtcgttct
6481 caaggaaggg ctgggggcca tcagcaatgt ctggagacat ttttgattct cacagcttgg
6541 tgggagactg tatttggtat tcattgggtg aggccaggga tgccaccgaa catctacaa
6601 tgcacagggc agccccccac aaataagaat tatccagccc caaatgcca atagtgcga
6661 ggtcagggga ccctgtgctg gccatcctat ctctgattct gaaactatgc atgctttcca
6721 cttttcccca tttgtgagtc attgagtaaa ttaaagctct tctgagcagc agcagtgatc
6781 atggtcactg ccctgcgttc aaataatgag agctgaggac agtgatctgc aactcccagc
6841 atgtcatgtg gtctcttaga aatccatgtg actgtttcca ccatcttggg catttgggg
6901 gacccccaga ctggagggag aaagccctac aaagtggatg ggagtgtggg gctgaacttt
6961 tccctaccct taactttgtg tctctgggac ctccagggac ctggccctc accaatgcat

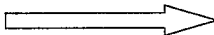
```

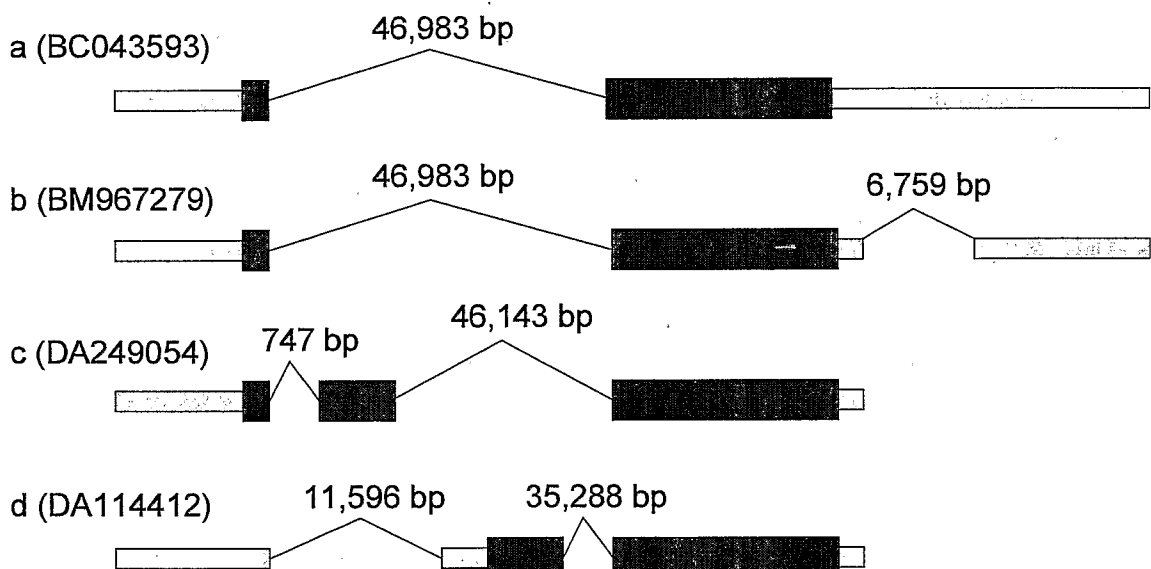
FIGURE 19B-3

7021 atgaagagta tgcttgggga agagccttagg aatgggggtgg gcatgggagt gctgggtagc
 7081 agccttttag caaatctgca tcttctctta tttctgacct ttttccatgt gccagtcct
 7141 atttctgcca gttgaaggca tactaatatt ctttatacta tttaatcttt tgcagaaacc
 7201 ttactattat aacttgctac tctccagata ccaattcttc atgccgagag catcggaat
 7261 gtttttgtgt cttactgatg ttttcatgat caacttgtaa atgtaagcag ttgacttcat
 7321 aaaaggtatt ttaactattc ttggagtcct ttgctacca agcacctggt ttcaccatgc
 7381 gatcactgac ttctctacag tgaagactct ttcttaataa aggatttcgc tgtgctcttt
 7441 tgattaaaaa tatctaacct taaaagacgt aaaaaaaaaa aaaaa (SEQ ID NO:14)

atggatctct cctttatggc cgcgcagctg cccatgatgg ggggagcttt catggactcg cccaacgagg
 acttcagcac cgagtactcc ctgtttaact cctctgcaa tgtccacgcg gctgccaatg gccagggcca
 gccggaagat cctcctcgt cctccaacga cgcgctcttg ctatggattg ccatcatagc tacgctgggg
 aacatcgtgg tgggtggcgt ggtgtatgcc ttcaccttc (SEQ ID NO:15)

FIGURE 19C

Alternatively spliced Saratan mRNA transcripts a-d: 5'  3'

**FIGURE 19D-1**

>Isoform_α.
 MDLSFMAAQLPMMGGAFMDSPNEDFSTEYSLFNSSANVHAAANGQGQPED
 PPRSSNDAVLLWIAIIATLGNIVVGVVYAFTF (SEQ ID NO:102)

>Isoform_β.
 MDLSFMAAQNKDKSRISPSVCDPREDPELAPGKGCWVPGRLPMMGGAFMD
 SPNEDFSTEYSLFNSSANVHAAANGQGQPEDPPRSSNDAVLLWIAIIATL
 GNIVVGVVYAFTF (SEQ ID NO:103)

>Isoform_γ.
 MDQPNSFPVCRCHLEFFQLPMMGGAFMDSPNEDFSTEYSLFNSSANVHAA
 ANGQGQPEDPPRSSNDAVLLWIAIIATLGNIVVGVVYAFTF (SEQ ID NO:104)

CLUSTAL 2.0.5 multiple sequence alignment

Isoform_α.	MDL--SFMAAQ-----LPMMGGAFMDSPNEDFST	27
Isoform_β.	MDL--SFMAAQNKDKSRISPSVCDPREDPELAPGKGCWVPGRLPMMGGAFMDSPNEDFST	58
Isoform_γ.	MDQPNSFPVCRCHLEFFQ-----LPMMGGAFMDSPNEDFST	36
	** ** ..:	*****

FIGURE 19D-2

Isoform α . EYSLFNSSANVHAAANGQGQPEDPPRSSNDAVLLWIAIIATLGNIVVVGVVYAFTF 83 (SEQ NO:102)
 Isoform β . EYSLFNSSANVHAAANGQGQPEDPPRSSNDAVLLWIAIIATLGNIVVVGVVYAFTF 114 (SEQ NO:103)
 Isoform γ . EYSLFNSSANVHAAANGQGQPEDPPRSSNDAVLLWIAIIATLGNIVVVGVVYAFTF 92 (SEQ NO:104)

FIGURE 19E-1

gi|119914061|ref|XP_875290.2| PREDICTED: hypothetical protein [Bos taurus]
 Length=317

GENE ID: 617871 LOC617871 | hypothetical LOC617871 [Bos taurus]

Score = 164 bits (415), Expect = 1e-39, Method: Composition-based stats.
 Identities = 80/83 (96%), Positives = 82/83 (98%), Gaps = 0/83 (0%)

Query 1 MDLSFMAAQLPMMGGAFMDSPNEFSTEYSLFNSSANVHAAANGQGQPEDPPRSSNDAVL 60
 MDLSFMAAQLPMMGGAFMDSPNEFSTEYSLFNSS NVHAA+NGQGQPE+PPRSSNDAVL
 Sbjct 235 MDLSFMAAQLPMMGGAFMDSPNEFSTEYSLFNSSTNVHAASNGQGQPEEPPRSSNDAVL 294
 Query 61 LWIAIIATLGNIVVVGVVYAFTF 83 (SEQ ID NO:2)
 LWIAIIATLGNIVVVGVVYAFTF (SEQ ID NO:3)
 Sbjct 295 LWIAIIATLGNIVVVGVVYAFTF 317 (SEQ ID NO:4)

CLUSTAL W (2.0.5) multiple sequence alignment

query MDLSFMAAQLPMMGGAFMDSPNEFSTEYSLFNSSANVHAAANGQGQPEDPPRSSNDAVL 60
 subject MDLSFMAAQLPMMGGAFMDSPNEFSTEYSLFNSSTNVHAASNGQGQPEEPPRSSNDAVL 60

query LWIAIIATLGNIVVVGVVYAFTF 83 (SEQ ID NO:2)
 subject LWIAIIATLGNIVVVGVVYAFTF 83 (SEQ ID NO:4)

gi|114654788|ref|XP_001147850.1| PREDICTED: hypothetical protein [Pan troglodytes]
 Length=203

GENE ID: 743753 LOC743753 | hypothetical protein LOC743753 [Pan troglodytes]

Score = 155 bits (391), Expect = 7e-37, Method: Composition-based stats.
 Identities = 83/114 (72%), Positives = 83/114 (72%), Gaps = 31/114 (27%)

Query 1 MDLSFMAAQ-----LPMMGGAFMDSPNEFSTEY 29
 MDLSFMAAQ LPMMGGAFMDSPNEFSTEY
 Sbjct 90 MDLSFMAAQNKDKSRISPSVCDPREDPELAPGKGCWVPGRLPMMGGAFMDSPNEFSTEY 149
 Query 30 SLFNSSANVHAAANGQGQPEDPPRSSNDAVLLWIAIIATLGNIVVVGVVYAFTF 83 (SEQ ID NO:5)
 SLFNSSANVHAAANGQGQPEDPPRSSNDAVLLWIAIIATLGNIVVVGVVYAFTF (SEQ ID NO:6)
 Sbjct 150 SLFNSSANVHAAANGQGQPEDPPRSSNDAVLLWIAIIATLGNIVVVGVVYAFTF 203 (SEQ ID NO:7)

CLUSTAL W (2.0.5) multiple sequence alignment

query MDLSFMAAQ-----LPMMGGAFMDSPNEFSTEY 29
 subject MDLSFMAAQNKDKSRISPSVCDPREDPELAPGKGCWVPGRLPMMGGAFMDSPNEFSTEY 60

query SLFNSSANVHAAANGQGQPEDPPRSSNDAVLLWIAIIATLGNIVVVGVVYAFTF 83 (SEQ ID NO:5)
 subject SLFNSSANVHAAANGQGQPEDPPRSSNDAVLLWIAIIATLGNIVVVGVVYAFTF 114 (SEQ ID NO:7)

FIGURE 19E-2

ref|XP_001476208.1| PREDICTED: hypothetical protein LOC268595 [Mus musculus]
 ref|XP_001478491.1| PREDICTED: similar to chromosome 14 open reading frame 132 [Mus musculus]
 gb|EDL18760.1| Mcg1155 [Mus musculus]
 Length=83

GENE ID: 268595 D430019H16Rik | RIKEN Cdna D430019H16 gene [Mus musculus]
 (Over 10 PubMed links)

Score = 153 bits (386), Expect = 4e-36, Method: Compositional matrix adjust.
 Identities = 74/83 (89%), Positives = 77/83 (92%), Gaps = 0/83 (0%)

```
Query 1 MDLSFMAAQLPMMGGAFMDSPNEDFSTEYSLFNSSANVHAAANGQGQPEDPPRSSNDAVL 60
        MDLSFMAAQLPMMGGAFMDS NEDFSTEYSLFNS NVH+ +NG GQPE+ PRSSNDAVL
Sbjct 1 MDLSFMAAQLPMMGGAFMDSSNEDFSTEYSLFNSPGNVHSTSNHGQPEESPRSSNDAVL 60

Query 61 LWIAIIATLGNIVVVGVVYAFTF 83 (SEQ ID NO:105)
        LWIAIIATLGNIVVVGVVYAFTF (SEQ ID NO:106)
Sbjct 61 LWIAIIATLGNIVVVGVVYAFTF 83 (SEQ ID NO:107)
```

CLUSTAL W (2.0.5) multiple sequence alignment

```
query MDLSFMAAQLPMMGGAFMDSPNEDFSTEYSLFNSSANVHAAANGQGQPEDPPRSSNDAVL 60
subject MDLSFMAAQLPMMGGAFMDSSNEDFSTEYSLFNSPGNVHSTSNHGQPEESPRSSNDAVL 60
*****.*****.***:.*:****:*****

query LWIAIIATLGNIVVVGVVYAFTF 83 (SEQ ID NO:105)
subject LWIAIIATLGNIVVVGVVYAFTF 83 (SEQ ID NO:107)
*****
```

gb|EDL81821.1| Rcg28860 [Rattus norvegicus]
 Length=83

Score = 132 bits (331), Expect = 1e-29, Method: Compositional matrix adjust.
 Identities = 74/83 (89%), Positives = 78/83 (93%), Gaps = 0/83 (0%)

```
Query 1 MDLSFMAAQLPMMGGAFMDSPNEDFSTEYSLFNSSANVHAAANGQGQPEDPPRSSNDAVL 60
        MDLSFMAAQLPMMGGAFMDSPNEDFSTEYSLFNS ANVH+ +NG GQ E+ PRSSNDAVL
Sbjct 1 MDLSFMAAQLPMMGGAFMDSPNEDFSTEYSLFNSPANVHSTSNHGQLEESPRSSNDAVL 60

Query 61 LWIAIIATLGNIVVVGVVYAFTF 83 (SEQ ID NO:108)
        LWIAIIAT+GNIVVVGVVYAFTF (SEQ ID NO:109)
Sbjct 61 LWIAIIATVGNIVVVGVVYAFTF 83 (SEQ ID NO:110)
```

CLUSTAL W (2.0.5) multiple sequence alignment

```
query MDLSFMAAQLPMMGGAFMDSPNEDFSTEYSLFNSSANVHAAANGQGQPEDPPRSSNDAVL 60
subject MDLSFMAAQLPMMGGAFMDSPNEDFSTEYSLFNSPANVHSTSNHGQLEESPRSSNDAVL 60
*****.*****.***:.*:*** *:*****

query LWIAIIATLGNIVVVGVVYAFTF 83 (SEQ ID NO:108)
subject LWIAIIATVGNIVVVGVVYAFTF 83 (SEQ ID NO:110)
*****.*****
```

ref|XP_001334647.1| PREDICTED: hypothetical protein [Danio rerio]
 Length=92

GENE ID: 794641 LOC794641 | hypothetical protein LOC794641 [Danio rerio]

Score = 50.8 bits (120), Expect = 3e-05, Method: Compositional matrix adjust.
 Identities = 25/64 (39%), Positives = 42/64 (65%), Gaps = 3/64 (4%)

```
Query 9 QLPMMGGAFMDSPNEDFSTEYSLFNSSANVHAAANGQGQPEDPPRSSNDAVLLWIAIIAT 68
        P + GAFMDSP ++ SL S++N++ + + + E+ + S+DA+ LW+A+IAT
Sbjct 21 HFFAVSGAFMDSP---YNGSASLQTSTSNLNTSISRPAEQEEEGKFSSDAIWLWVAVIAT 77

Query 69 LGNI 72 (SEQ ID NO:111)
        +GNI (SEQ ID NO:112)
Sbjct 78 IGNI 81 (SEQ ID NO:113)
```

FIGURE 19E-3

CLUSTAL W (2.0.5) multiple sequence alignment

```

query      -----MDLS-----FMAAQLEPMGGAFMDSPNEDFSTEYSLFNSSANVHAAANGQGQP 48
mST        MVRRRSGGLEQEEPAVAYDQHFFAVSGAFMDSP---YNGSASLQTSTSNLNTSISRPAEQ 57
           :: .          : * : .***** : . . ** .*: : : : . . :

query      EDPPRSSNDAVLLWIAIIATLGNIVVVGVVYAFTF 83 (SEQ ID NO:111)
subject    EEGKFSSDAIWLWVAVIATIGNIVVVAVVCACAF 92 (SEQ ID NO:114)
           *: : *.** : **:*.*:*****.* * :*
```

FIGURE 19F

CLUSTAL 2.0.5 multiple sequence alignment

```

Homo_sapiens      -----
Bos_Taurus        MDVSKVILSCDADVSRVWKSLSAEQKSGGVHLEMKGLGPREQERLIQEDTLDVGVAYRVS 60
Pan_troglodytes.  -----
Mus_musculus      -----
Rattus_norvegicus. -----
Danio_rerio.      -----

Homo_sapiens      -----
Bos_Taurus        LLFLQSPVLGEVVCVPLSKATPTGQLCRQLDQNOQQOGLGTGGSTDDTQAQKATPLPKR 120
Pan_troglodytes.  -----MPSR 4
Mus_musculus      -----
Rattus_norvegicus. -----
Danio_rerio.      -----

Homo_sapiens      -----
Bos_Taurus        EASGSAPGVPPSPALGGGAGSGHVEGRVPVPPSPSPAPPPGSKPPTRPSPAESSEPEQGT 180
Pan_troglodytes.  EYTG-----LWEERWRGKPGHSDLGPQARDNVVSAPRARCP-----QPR 42
Mus_musculus      -----
Rattus_norvegicus. -----
Danio_rerio.      -----

Homo_sapiens      -----MDLSFM 6
Bos_Taurus        ARARDRPPPPSCRCRAVSRTPARCVAFSGQLCGRLDSSGSDSEDSSAGGSDTMDLSFM 240
Pan_troglodytes.  AEAQRQPRAPC--GLRTLAAQPD-----PRQLCRRLTRSGSGSSEDSSAGCSDTMDLSFM 95
Mus_musculus      -----MDLSFM 6
Rattus_norvegicus. -----MDLSFM 6
Danio_rerio.      -----MVRR-----RSGGLEQEEPAVA 17
                       : :.

Homo_sapiens      AAQ-----LPMMGAFMDSPNEDFSTEYSLFNSS 35
Bos_Taurus        AAQ-----LPMMGAFMDSPNEDFSTEYSLFNSS 269
Pan_troglodytes.  AAQNKDKSRISPSVCDPREDPELAPGKGVWPGRLPMMGGAFMDSNEDFSTEYSLFNSS 155
Mus_musculus      AAQ-----LPMMGAFMDSSNEDFSTEYSLFNSS 35
Rattus_norvegicus. AAQ-----LPMMGAFMDSPNEDFSTEYSLFNSS 35
Danio_rerio.      YDQH-----FPAVSGAFMDSP---YNGSASLQTST 44
                   *          : * : .***** : . . ** .*.

Homo_sapiens      ANVHAAANGQGQPEDPPRSSNDAVLLWIAIIATLGNIVVVGVVYAFTF 83 (SEQ NO:1)
Bos_Taurus        TNVHAAASNGQGQPEEPPRSSNDAVLLWIAIIATLGNIVVVGVVYAFTF 317 (SEQ NO:4)
Pan_troglodytes.  ANVHAAANGQGQPEDPPRSSNDAVLLWIAIIATLGNIVVVGVVYAFTF 203 (SEQ NO:7)
Mus_musculus      GNVHSTSNHGQPEESPRSSNDAVLLWIAIIATLGNIVVVGVVYAFTF 83 (SEQ NO:107)
Rattus_norvegicus. ANVHSTSNHGQLEESPRSSNDAVLLWIAIIATVGNIVVVGVVYAFTF 83 (SEQ NO:110)
Danio_rerio.      SNLNTSISRPAEQEEEGKFSSDAIWLWVAVIATIGNIVVVAVVCACAF 92 (SEQ NO:114)
                   *: : *.** : **:*.*:*****.* * :
```

FIGURE 19G

MYPYDVPDYASLEF MDLSFMAAQL PMMGGAfMDS PNEDFSTEYS LFNSSANVHA AANGQGQPED
PPRSSNDAVL LWIAIIATLG NIVVVGvVYA FTF (SEQ ID NO:8)

MDLSFMAAQL PMMGGAfMDS PNEDFSTEYS LFNSSANVHA AANGQGQPED PPRSSNDAVL LWIAIIATLG
NIVVVGvVYA FTF EFYPYDVPDYA (SEQ ID NO:9)

MDLSFMAAQL PMMGGAfMDS PNEDFSTEYS LFNSSANVHA AANGQGQPED PPRSSNDAVL L (SEQ ID
NO:10)

FIGURE 19H-1

auggaucucu ccuuuauaggc cgcgcagcug cccaugaugg ggggagcuu cauggacucg cccaacgagg
acuucagcac cgaguacucc cuguuuuacu ccucugccaa uguccacgcg gcugccaaug gccagggcca
gccggaagau ccuccucggu ccuccaacga cgccgucuuug ccauggaug ccaucauagc uacgcugggg
aacaucgugg ugguggggcgu gguguaugcc uucaccuuc (SEQ ID NO:74)

5'- AAAUAAUCACUCGUGUGACUU -3' (SEQ ID NO:75)

5'- AAUACAAGAAAGCUGGACUU -3' (SEQ ID NO:76)

5'- AAGAGCAACUUACUGCUUUGU -3' (SEQ ID NO:77)

5'-AUAUACACUCGUGUGACUU-3' (SEQ ID NO:78)

5'-UUACAAGAAAGCUGGACUU-3' (SEQ ID NO:79)

5'-GAGCAACUUACUGCUUUGU-3' (SEQ ID NO:80)

5'-AUAUACACUCGUGUGACUUCUCAAGAGAAAGUCACACGAGUGAUUUAU-3' (SEQ ID NO:81)

5'-UUACAAGAAAGCUGGACUUCUCAAGAGAAAGUCCAGCUUUCUUGUAA-3' (SEQ ID NO:82)

5'-GAGCAACUUACUGCUUUGUUCUCAAGAGAAACAAAGCAGUAAGUUGCUC-3' (SEQ ID NO:83)

5'-XXXXAUAAUCACUCGUGUGACUUCUCAAGAGAAAGUCACACGAGUGAUUAUXXXX-3' (SEQ ID NO:84)

5'-XXXXGUUACAAGAAAGCUGGACUUCUCAAGAGAAAGUCCAGCUUUCUUGUAAXXXX-3' (SEQ ID NO:85)

5'-XXXXGAGCAACUUACUGCUUUGUUCUCAAGAGAAACAAAGCAGUAAGUUGCUCXXXX-3' (SEQ ID NO:86)

5'-ggauccAUAAUCACUCGUGUGACUUCUCAAGAGAAAGUCACACGAGUGAUUAUUUUUUUGGAAaagucc-3'
(SEQ ID NO:87)

5'-ggauccGUUACAAGAAAGCUGGACUUCUCAAGAGAAAGUCCAGCUUUCUUGUAAUUUUUUGGAAaagucc-3'
(SEQ ID NO:88)

5'-ggauccGAGCAACUUACUGCUUUGUUCUCAAGAGAAACAAAGCAGUAAGUUGCUCUUUUUUGGAAaagucc-3'
(SEQ ID NO:89)

AGUGCGCCGUGCGGUCU (SEQ ID NO:90)

UCUCCGGACGCUCGCUG (SEQ ID NO:91)

5'-AUGGAUCUCUCCUUUAUGGCCG-3' (SEQ ID NO:92)

5'-UAGCAAGACGGCGUCGUUG-3' (SEQ ID NO:93).

5'-AUGGCGCAGCUCUCCUACUC-3', (SEQ ID NO:94)

5'-CGUGCAUCUCUCUGCAUCCUC-3', (SEQ ID NO:95)

FIGURE 19H-2

AUAAUCACUCGUGUGACUU (SEQ ID NO:96)

UUACAAGAAAGCUGGACUU (SEQ ID NO:97)

GAGCAACUUACUGCUUUGU (SEQ ID NO:98)

5'-ACACCAUGGAUCUCUCCUUUAUG-3' (SEQ ID NO:99)

5'-UGACUCGAGUCAUAGCAAGACGGCGUCGUUG-3' (SEQ ID NO:100)

FIGURE 19I-1

```

1  GCCUAGUAGA GAUCUGGAGC CAGAAGCCCCA GAGACAGCCG AGUGCGCCGU GCGGUCUCCG
61  GACGCU CGCU GCUCAGCCCCG AUCCCCGCCA ACUGUCAGG CGGCUAGACC GCAGCGGCAG
121 CGGCAGCAGC GAGGACUCGA GCGCUGGCUG CAGCGACACC AUGGAUCUCU CCUUUAUGGC
181 CGCGCAGCUG CCCAUGAUGG GGGGAGCUUU CAUGGACUCG CCCAAGGAGG ACUUCAGCAC
241 CGAGUACUCC CUGUUUAACU CCUCUGCCAA UGUCCACGCG GCUGCCAAUG GCCAGGGCCA
301 GCCGGAAGAU CCUCCUCGGU CCUCCAACGA CGCCGUCUUG CUAUGGAUUG CCAUCAUAGC
361 UACGCU GGGG AACAU CGUGG UGGUGGGCGU GGUGUAUGCC UUCACCUUCU GAGGACGGCA
421 CACCCUGCAC CACCAUGGGG UGAGGCUUGG CACGUAGCUC UGACUUGCUG UCGGCCUUUG
481 GCUUCUCCUG UGUUCUAGAA CCAGGAGUUU UGACCAGGGG CGGCGGCCGU CCUUCUGGAA
541 UUUUCUCCCCA GCAGCCUGA UUUCAAUAU CCCAUGUUGU GGUCAAGCUG AGUCAGAAGA
601 CAUGGAAGUA UGGGCCUCCU GCCCCUAGAG GCAUGACGGG GCAAGCCUU CAGAGGGCAG
661 AUUGGGGAUC CUUGAAACUA CAUUCAGGA ACAUGGGACC AGAUGAGACA GCUAGUUAAG
721 UUUAAAACAU AGACAUGAUU UGAUGAUCGC UUGCUGGUGG UAAAUAAUCA CUCGUGAGC
781 UUGUUUUUAU GCAAACUUAU CGAACCUAGG GCGUGGGGUG CUGGGGCAAG AGCAGCCUUC
841 AGAACUUCAG UGUUCCUGAC CCAAUUCUGG UUUCACAUUC AGUCCUUGG CCAUCUAGUA
901 GGGCCAUUGG AUGUCCUAG UUGACUUUG AAAUGGCACC UUGGCCACCA GACACUUGU
961 CCCUCCAAAG ACCCAAGUGC AUUGGGAGAC CCAGGGAUGG GGGGUUACUG GUAAUAGGUG
1021 GGGUUUCUGG GGGUGUUGU GUGGUUUUU AUUCUUGGU CGGACUUUCU CUUCUUUUUA
1081 AGAAGAGGAG AGGAUGUCUU UAAGAGCUAG AUGUGCCAGG CAGUGGACUC UUCAGGCCAC
1141 CCACGUGAGA AUGCUGUUUC UUCUCUGAGG AAUCGUGGAA UUUUAAGAU GAACAAGAUU
1201 CACAUCCACU GAAUUAUUA ACGGAUGGGU CAGAAAGGGG GGUGAUUUGC CUGUGGUCAC
1261 CAGGCAGGUU UGUGGAGGAG UCGGAACAGA AAGAUGUGUC CUCACUCUCA GCUCACUGAG
1321 CUCUCUGCCC CAACUAAGCA CUUCCUGAG GAGGUUGCUG AGAAGCUGCC CUCAGGAGAA
1381 UGUCCAGGCA UCUUGAAGGU GGGUGCGAGA UUGGCAGGCU GCUCAGAUAC CCGUCCUUUA
1441 CAUUCAGUGU GGAUACCGUG CAUUCUCCUG AAGCUGUGAA AGUGCUUCUG CCCAAGCCAC
1501 UUCUUUAAAU UCUGAAAUUA CAGCAUCUGG GGUCUUGGCA AGCAAGGAAG CUUCCAAGUA
1561 AAAACAGAG AGAAGGGCAC ACUUUUUUUU CUUCAUUAAG AAAUCUUUUA GCACAGGAAC
1621 CACCCCAACC CCCACCCCCC ACACCUUCCC AAGGCAGCAU CCCAGUGCAG AUAGAGUGGG
1681 AAAGGUCCCA GAAGGGGGCU CACUCACCUC UAGGCCCAGA GAGGCUUUCU CCUCACUUUA
1741 UACACUGCAA AAACAGAAGA AUUGUGUCA UAACACCUC UGUAGUGGAG AAACUUAAAA
1801 AGCUGGUUAG GAAGCUCUCG UGUUAUUUA GAGACAAUA CAAGAAAGCU GGACUUGCCG
1861 CUGUGGUCUC AGGAGAAUUG AGUGUUCUUG AUGACAGGCA AAGGGACAUC UUAUGUUGCC
1921 AGAAGCGGCA CUCUCCUGG GAAGCCGCCA UGUUAUAGG AUUACUAGCC UGGCUCAGCA
1981 CAGUGCCUGC UCAUGGCUGC CAGUUCUUA CCAUCACAUU UGUCACUGCC ACCGUUAUUC
2041 AUCUGCCAGU GCAUCAGCUU AAGGGGAGGU CACGAGUGCA AAAGAACCUG ACCCUUGACA
2101 AUGAGGGAGA AGGGACAUGG ACCACCUGUC UGGAAUUCU GGAUACUG GCAGGGUGGA
2161 GGCUGGGCUG GGGAGUUAGC CGCGGUGUGC GUGAAUGGCU CUGUCUCCAG CAAGUCUCUC
2221 UCCAUCAAAC CCCAGGUCUG CCCCAUAAGC AAGAUUUUA ACAGAUGGAU GUCUCCAUGA
2281 GAAAACCCAA GGCAGAAAGC CCAGAGCCAU GGCGGGGUUG CUUGACGUCC UCAUGGAGUC
2341 ACUCUGCCCC ACAUGCUCAA AUCUUCUCCU UGGCCCCACA UCCCUAGGAG GGCCUGACCC
2401 CUGUAAAGAU ACAGGAGGCA GCUCCUGGC CUCCAAUUGG CCCAUGGAGA UGGCAGUCGG
2461 GAGACAGGGU UCUGUGUUUG CUGCGGUGAA GGGAGGAGAA GGCAGGAGGA AAAAGGAUGG
2521 CUUCUAGCCC UGAAGAGGAC UCCAGCAUCC CAGGCACCCG GUGCUUCUGG CUGCAGUUUU
2581 CCCUAGGAG GCCCCUAGC CUCCAGCCCU AACAUAAUUG UCGGUUAAAU UCAGUUUUUA
2641 AGCCUCUCUC CCUUUUCAGU GUCAGAGCAG UAGAUGGUCC AGGGCAUUGG AGGCCUCGAC
2701 CACUCUGCAU UGCAGAUUAC AGUGACUUC UCGGGGUUGC CCCAUUUG UCUCCUGUGG
2761 UUUUUUUAU AGCUUUUUU UUACCAGCAU CUCUCAAUA ACAUAUGAGA UAGAUAUGCC
2821 CAUUAUGUUC UGAUUAAGGA GCAAAGGCGU GAUUUCUGGC CACAGCGAGC UGCACUCUCC
2881 CUCCUGCCUC AGCCGGGGUC CGUCUUAGCA GUUUGGAAAG GGGAAAAAGA UGCCGGUCCU
2941 CACUGCUUAA GUUUUGUGUC CAGGUGCCAC UAGACUUGCA UGCACACUAA CUCCUUACAA
3001 UCACCACACA GCAUCAUCGC CCCAGUGCAC AGAUGAGGAA CCAGAGGCUC AGAGGAGUGA
3061 AGUUGCCUUC CUGAGGUCAC ACAGCAUGAA AGUGAUGAGC UAGGAUUUGA AUCUGGGAAG'

```

FIGURES 19I-2

3121 UUGGGCUCUA GAGCCAGACU GUACUGCCUU CUGCCACACU GUACUGCCUU CUGUGACUGG
 3181 GUGGCACCUC CAGGGCACAU UUACACAAGG CCCUGAAUCU GCAGAGGCUG UUUCUCAAGA
 3241 UGCCCUGCAU GGUGUGGCCU GGGCCAGCUC UGGCUUCCAC AGGUCCCUGA CUGUCCUCAG
 3301 AGUGGAACAU GCUCAACCUC CCGCCCACUG CUCUCUCUCU GCCAGAUUU CAGGGGUGCC
 3361 GGUCCCCAAG GCCUGCCCCC UUCUUUAAGA CUGAACUCAA GUCUCCUUGG AAGGCCCCGG
 3421 UGAAGCUCCT AGAGACUGGU UUUCUUGGGA UGCAGGCAGA AGGGGACCCU CCCUGGCCAA
 3481 CACCCAGGAG CCCAGCAGAA GCACCCACAC GUAGAAAGAG GCUCACUACA GCCAGAAGUG
 3541 CAGAGUCAGA GUCCUGGGAC CAUCUUGUUC UGCAAGGUGA CCCAGGCUC CCCAGGACAG
 3601 GGGAGAGGGA UCGUCCUCAU UCAGACUCUA GCUGGGGCCU CUGUACUGGC UUCUCCUUGG
 3661 GUGGGGUUGC CUGUUACAU GCUGUGCCUC AGAGAAAGGG UCCUGCAUUU UCUGGAAUGU
 3721 UCUCUGUGCU UACCCUCUG UGUGCCCCUC CAUUGCUCCU CUACAAGCAA UAGGGUGAUU
 3781 CAAAAGAGCA ACUUAGGCUG GGUGCAGUGA CUCACACCCG UAAUCCCGCG ACUUUGGGAG
 3841 GCCGAGGCGG GCAGGGACAG GAGUUCAGA CCAGCCUGGC CAACAUGGUG AAACCCUGUC
 3901 UCUACAAAA AUACAAAAU UAACCAGACA UUGUGGCAUG UGCCUGUAU CCCAGCUACU
 3961 CAGGAGGCUG ACACAGGAGA AUUGCUUGAA CCAGGAGGCG GAGGCUGCAG UGAGCUGAGA
 4021 UUGUGCCACU GCACUCCAGC CUGGGCAACA CAACGAGACU CUGUCUCAA AAAAAAAAAA
 4081 AGAAAAGAAA AAAAAAGAG CAACUACUG CUUUGUGAGG UUGUGAGUG CCACCAUGA
 4141 AGGUCUUUGA GAAGAGGCCA GACGCCGUG UAGCCAGGCC UGUCUUAGA GGACUUGUGC
 4201 UCCAGGGAC CCAGGCAGGA UGAUGGCGCA GCUCUCCUA CUCCAAGCCA ACGCUGUCCU
 4261 UCCCCUUUC CAUGAAAUCA AGGUCAAGAG GCAAAUAAGA CUCCUGCUC CACUCUACCC
 4321 CCCAGAGAGA AAUGAUUCUC GCUCUUUCA GAUCCCCAG GAUCUGAGGG AGAAAGGAUG
 4381 GGAGGAGGGG CAGCAGCAUU UCGCUGGAAA GGCAGCAGAU GCUUUUCCAG CCCCGGUUCA
 4441 GCUGGAAGGC UUGGAGGCCG GCCAGACCAC UCUGGCGUCU CCUGAAGUGG GUCCUGGAG
 4501 ACCGAAGAGG CUCAGUGGAG UCUGUCUGUU GUCAGCACUG CUGCCUGAUC CCUGCAAGAC
 4561 AAAUGGCACU UUCUUCUUC AGAAGCAUCA UCUGCCUUA UUAUAGCAG UAAUUAUUAU
 4621 CCCAGUUAU AUUCUUAACG GUGCCAGUUU UGCACAUUU UUUGUUGCUC UAUUUGUGUC
 4681 UCAUUUACUU CUCAAAUUGC CCCUGGGGGC AGGAAUGAGG AUGCAGAGAG AUGCACGUUA
 4741 AUUACUGUCG CAUUUUUCUG UGGAGAAAC UGAGGCCAGG GCUGAACGCU CACAGCUAAG
 4801 GAGCUGUGAU UCAGACCCAG UUCUGUCAGC UCUGAGGCA CCCUGCAUCA UGCCCACCAG
 4861 GGUGAUCCCC CUGGGAUGGA CCAUCUCGG AUUAUGAGCC UCGGAGGCG UGGUUGAGAU
 4921 UUGGUCUGA AGAGCUUAUA GCCAGAUUGC CACAUUCAAG UGUAAUGCCA GGAAGGGGGC
 4981 AGCGGCGAGU GCACAGGGAU UUAUCAGUUC CAGAACCUCA CAGUGAUAG AGGCUUUAUA
 5041 GAGCAUCUAA UCGAGACCUU UAAUUUUUCG GGGAGAGCAG CUGAGGCCGU GUGGAAAAUU
 5101 AGUGGAGAGC UGACAAGUGU CUGGGCUCCU GGCCAGGGG UCCGUGGUCC AGCAGGUUGU
 5161 GCGUUCAGUG GGAAGCAAAG GCGUUGCCCG GGAUUACCUG CCCAGCCCC GAGGUGGGUU
 5221 GUGCUCCUG CAGCUGCCA UCGGCCGCUU UGCUUCGUCC UGGCAGAUCC CAGUGAUUG
 5281 UCCCCGAGCA AGUGCCAGGG UUGGGCUGAG CUGCUAUGAC AGGGAGGCC AGGGAGUUUCU
 5341 GCUCAAGGAG CCAAAGGGAA CAGCCAGAU CUGAAUGUUC UAUGUUCACC UGCCCCAGCC
 5401 CCACCCACCC UGGCCCACUC CACAGGCCCC UGACCAUGGU CACUCACGGA GAGGGAUGGA
 5461 GGAGAAGGUG GUUGAACUGA GUACUGAGAA CCCAGAGGAC AGAGCCACA GCUUCCAAGC
 5521 AGGAAAAGGG ACCUCUCUGA AAAAUUCUGA UAACCAGAAU UAUCACAGCA CCCUCUCAU
 5581 CCCAGCGCGU CCUUCUGAGC UCAGACCUUG AGCAUUUACU GGGUUCUUU UUGAGGAAGA
 5641 GGGAAAGUGA CAAAGGACAA AACAAUGCAA AUCUUAUGA CUGAAGACGA UCAAAGACUC
 5701 CCUGGAGCGA GAAAACAGUU ACUGCCAGAA GCAGAAUGGA AGAGCCAAA AGUACACAAA
 5761 AUGGACGCCA UAAAUUCUGA AAUAAAAGUG UAUGAUGUGU UCUGAGUCAC UGUAGAAGUC
 5821 AUGCAUUUAU UAUCAAGAU GAAAAGAGCA GAGAAUGACG UGGACAUUGG UCCUCGAGA
 5881 GGCUGCAUAG GUGGUGCGGU CCCCAGGGGA UUCUGGAUGC UGGUUCUUUG ACCGUGGCGG
 5941 CAGCCUCGCG CCUGCCCGGA UGGCUCCAUC CAGACAUUUG GCAAGGCUGU CAUCUGCUCU
 6001 UGGGUCCUUU UUCAAGCUGA UUUCUGCCU CCAAGAGG AGGUUUGAGC CCAUUCUGC
 6061 CUUGGAAUA AUCCUGACA GAUGUGCACA GCAUUAUUUG AGGGAUUUG CAGGCCUGCA
 6121 UUAUUGGGG AAUCAGAGCC CAUCCACAC CGAGCCCAU CCACGUGGCU GAGCAACACC
 6181 CCUGUGUCCC CUGUAUCUGU CUCACUGGCU GUCUUUCUGC CAGUUUCUUA AAGAACCACA
 6241 CCAUUACUGC AUUUGCCGU CGAAGCGUUG UCCCAACAAU GCAGAUUGU CUGACAAGGG
 6301 UCUAUUAGCU GGCAGAAAGG AGCUUCCAAC CUUUUAAU GAGGAAAUG AAUCUCUGAC
 6361 AGGCUUGAGA GUGUUGCCAG GUGGAGCUUU UCAGGAGACA GGGGUCUUGA GCCCAGGACA
 6421 CCUAACUAUC GAGUUUUCAC UAGGAGACUU AGUGGUGGCU UUCAUGAGG CAGUCGUUCU
 6481 CAAGGAAGGG CUGGGGGCCA UCAGCAUUGU CUGGAGACAU UUUUGAUUC CACAGCUUGG
 6541 UGGGACUGC UAUUGGUUUC UCAUGGGUGG AGGCCAGGGA UGCCACCGAA CAUCCUACAA
 6601 UGCACAGGGC AGCCCCCAC AAAUAAGAAU UAUCAGCCC CAAAUGCCA AUAGUGCCGA
 6661 GGUCAGGGGA CCCUGUGCUG GCAUCCUUAU CUCUGAUUCU GAAACUAUGC AUGCUUCCA
 6721 CUUUUCCCA UUUGUGAGUC AUUGAGUAAA UUAAGGCUU UCUGAGCAGC AGCAGUGAUC
 6781 AUGGUCACUG CCCUGCGUUC AAAUAAUGCG AGCUGAGGAC AGUGAUCUGC AACUCCAGC
 6841 AUGUCAUGUG GUCUCUAGA AAUCCAUGU ACUGUUCCA CCAUCUUGG CAUUUGUGG
 6901 GACCCCAGA CUGGAGGGAG AAAGCCUAC AAAGUGGAUG GGAGUGUGG GCUGAACUUU

FIGURE 19I-3

6961 UCCCUACCCU UAACUUUGUG UCUCUGGGAC CUCCAGGGAC CUGGCCCCUC ACCAAUGCAU
 7021 AUGAAGAGUA UGCUUGGGGA AGAGCUUAGG AAUGGGGUGG GCAUGGGAGU GCUGGGUAGC
 7081 AGCCUUUGAG CAAAUUCUGCA UCUUCUCUUA UUUUCUGACCU UUUUCCAUGU GCCCAGUCCU
 7141 AUUUCUGCCA GUUGAAGGCA UACUAAUAU CUUUAUACUA UUUAAUCUUU UGCAGAAACC
 7201 UUACUAUUAU AACUUGCUAC UCUCAGAUU CCAAUUCUUC AUGCCGAGAG CAUCGGAUAAU
 7261 GUUUUUGUGU CUUACUGAUG UUUUCAUGAU CAACUUGUAA AUGUAAGCAG UUGACUUCAU
 7321 AAAAGGUAAU UUAACUAUUC UUGGAGUCCU UUGCUACCCA AGCACCUGGU UUCACCAUGC
 7381 GAUCACUGAC UUCUCUACAG UGAAGACUCU UUCUUAAUAA AGGAUUUCGC UGUGUCUUU
 7441 UGAUUAAAAA UAUCUAACCU UAAAAGACGU* (SEQ ID NO:101)

FIGURE 19J

1 GCCUAGUAGA GAUCUGGAGC CAGAAGCCCA GAGACAGCCG AGUGCGCCGU GCGGUCUCCG
 61 GACGCUCGCU GCUCAGCCCG AUCCCCGCCA ACUGUGCAGG CGGCUGACCC GCAGCGGCAG
 121 CGGCAGCAGC GAGGACUCGA GCGCUGGCUG CAGCGACACC AUGGAUCUCU CCUUUAUGGC
 181 CGCGCAGCUG CCCAUGAUGG GGGGAGCUUU CAUGGACUCG CCCAACGAGG ACUUCAGCAC
 241 CGAGUACUCC CUGUUUAACU CCUCUGCCAA UGUCCACGCG GCUGCCAAUG GCCAGGGCCA
 301 GCCGGAAGAU CCUCCUCGGU CCUCCAACGA CGCCGUCUUG CUAUGGAUUG CCAUCAUAGC
 361 UACGCUGGGG AACAUUGUGG UGGUGGGCGU GGUGUAUGCC UUCACCUUCU GAGGACGGCA
 421 CACCCUGCAC CACCAUGGGA AACCUUACUA UUAUAACUUG CUACUCUCCA GAUACCAAUU
 481 CUUCAUGCCG AGAGCAUCGG AAAUGUUUUU GUGUCUUACU GAUGUUUCA UGAUCAACUU
 541 GUAAAUGUAA GCAGUUGACU UCAUAAAAGG UAUUUUAACU AUUCUUGGAG UCCUUUGCUA
 601 CCCAAGCACC UGGUUUCACC AUGCGAUCAC UGACUUCUCU ACAGUGAAGA CUCUUUCUUA
 661 AUAAAGGAUU UCGCUGUGCU CUUUUGAUUA AAAUAUACUA ACCUAAAAG ACGU (SEQ:120)

FIGURE 19K-1

1 GCCUAGUAGA GAUCUGGAGC CAGAAGCCCA GAGACAGCCG AGUGCGCCGU GCGGUCUCCG
 61 GACGCUCGCU GCUCAGCCCG AUCCCCGCCA ACUGUGCAGG CGGCUGACCC GCAGCGGCAG
 121 CGGCAGCAGC GAGGACUCGA GCGCUGGCUG CAGCGACACC AUGGAUCUCU CCUUUAUGGC
 181 CGCGCAGAAC AAAGACAAU CCCGGAUUUC UCCAUCAGUC UGUGACCCUA GAGAAGACCC
 241 AGAGCUGGCU CCAGGGAAGG GCUGCGUUUG GCCUGGGAGA CUGCCCAUGA UGGGGGGAGC
 301 UUUCAUGGAC UCGCCCAACG AGGACUUCAG CACCGAGUAC UCCUGUUUA ACUCCUCUGC
 361 CAUUGUCCAC GCGGCUGCCA AUGGCCAGGG CCAGCCGGAA GAUCCUCCG GGUCCUCCAA
 421 CGACGCCGUC UUGCUAUGGA UUGCCAUCAU AGCUACGCUG GGAACAACUG UGGUGGUGGG
 481 CGUGGUGUUA GCCUUCACCU UCUGAGGACG GCACACCCUG CACCACCAUG GGGUGAGGCU
 541 UGGCACGUAG CUCUGACUUG CUGUCGGCCU UUGGCUUCUC CUGUGUUCUA GAACCAGGAG
 601 UUUUGACCAG GGGCGGCGGC CGUCCUUCUG GAAUUUCUCC CCAGCAGCCC UGAUUUCAA
 661 UAUCCCAUGU UGUGGUCAAG CUGAGUCAGA AGACAUGGAA GUAUGGGCCU CCUGCCCCUA
 721 GAGGCAUGAC GGGGCAAGGC CUUCAGAGGG CAGAUUGGG AUCCUUGAAA CUACAUCUCC
 781 GGAACAUGGG ACCAGAUGAG ACAGCUAGUU AAGUUUAAAA CAUAGACAUG AUUUUGAUAU
 841 CGCUUGCUGG UGGUAAAUAU UACUCUGUGU GACUUGUUUU UAUGCAAACU UAUCGAACCU
 901 AGGGCGUGGG GUGCUGGGGC AAGAGCAGCC CUCAGAACUU CAGUGUUCU GACCCAAUUC
 961 UGGUUUCACA UUCAGUCCCU UGGCCAUCUA GUAGGGCCAU UGGAUGUCC UAGUUUGACU
 1021 UUGAAAUGGC ACCUUUGCCA CCAGACACCU GGUCCCUUCC AAGACCCAAG UGCAUUGGGA
 1081 GACCCAGGGA UGGGGGGUUA CUGGUAAUAG UGGGGGUUUC UGGGGGUGUU GUUGGUGGUU
 1141 UUAUUCUCCU GGUCGGACUU UCUCUUCUUU UUAAGAAGAG GAGAGGAUGU CUUUAAGAGC
 1201 UAGAUGUGCC AGGCAGUGGA CUCUUCAGGC CACCCACGUG AGAAUGCUGU UUCUUCUCUG
 1261 AGGAUUCGUG GAAUUUUAAA GAUGAACAAAG AUUCACAUC ACUGAAUUAU UCAACGGAUG
 1321 GGUCAGAAAG GGGGGUGAUU UGCCUGUGGU CACCAGGCAG GUUUGUGGAG GAGUCGGAAC
 1381 AGAAAGAUGU GUCCUCACUC UCAGCUCACU GAGCUCUCUG CCCAACUAA GCACUUCUCCU
 1441 GAGGAGGUUG CUGAGAAGCU GCCCUCAGGA GAAUGUCCAG GCAUCUUGAA GUGGGGUGCG
 1501 AGAUUGGCAG GCUGCUCAGA UACCCGUCU UUAUUAUCAG UGUGGAUACC GUGCAUUCUC
 1561 CUGAAGCUGU GAAAGUGCUU CUGCCCAAGC CACUUCUUA AAUUCUGAAA UAUCAGCAUC
 1621 UGGGGUCCUG GCAAGCAAGG AAGCUUCCAA GUAAAAACCA GAGAGAAGGG CACACUUUUC
 1681 UUUUUAUUAU AGGAAAUUUU AUUGCACAGG AACCACCCCC ACCCCACCC CCCACACCUU
 1741 CCCAAGGCAG CAUCCAGUG CAGAUAGAGU GGGAAAGGUC CCAGAAGGGG GCUCACUCAC
 1801 CUCUAGGCCC AGAGAGGCUU UCUCUUCACU UUAUACACUG CAAAACAGA AGAAUUGUGU
 1861 CAAUAACACC CUCUGUAGUG GAGAAACUUA AAAAGCUGGU UAGGAAGCUC UCGUGUAUUA
 1921 UUAGAGACAA UUACAAGAA GCUGGACUUG CCGCUGUGGU CUCAGGAGAA AUGAGUGUUC
 1981 UGAUGACAG GCAAAGGGAC AUCUUAGUUG UCCAGAAGCG GCACUCUUC CUGGAAGCCG
 2041 CCAUGUUAU AGGAUUAUAU GCCUGGCCU AGACAGUGCC UGCUCAUGGC UGCCAGUUCU
 2101 UACCGAUCAC AUCUGUCACU GCCACCGUUA AUCAUCUGCC AGUGCAUCAG CUUAAGGGGA
 2161 GGUCACGAGU GCAAAAGAAC CUGACCCUUG ACAUAGAGG AGAAGGGACA UGGACCACCU

FIGURE 19K-2

2221 GUCUGGAAUU CCUGGAAUCA CUGGCAGGGU GGAGGCUGGG CUGGGGAGUU AGCCGCGGUG
 2281 UGCGUGAAUG GCUCUGUCUC CAGCAAGUCU CUCUCCAUCA AACCCCAGGU CUGCCCCAUA
 2341 AGCAAGAUCU UUAACAGAUU GAUGUCUCCA UGAGAAAACC CAAGGCGAGA AGCCCAGAGC
 2401 CAUGGCGGGG UUGCUUGACG UCCUCAUGGA GUCACUCUGC CCCACAUGCU CAAAUUUUCC
 2461 CUCUGGCCCC ACAUCCCUAG GAGGGCCUGA CCCUGUAAA GAUACAGGAG GCAGCUCCCU
 2521 GGCCUCCAAA UGGCCCAUGG AGAUGGCAGU CGGGAGACAG GGUUCUGUGU UUGCUGCGGU
 2581 GAAGGGAGGA GAAGGCAGGA GGA AAAAGGA UGGCUUCUAG CCCUGAAGAG GACUCCAGCA
 2641 UCCCAGGCAC CCGGUGCUUC UGGCUGCAGU UUUCCCUAUG GAGGCCCCUC AGCCUCCAGC
 2701 CCUAACAUAU AUGUCGGUUA AAUUCAGUUU UCAAGCCUCU CUCCCUUUUC AGUGUCAGAG
 2761 CAGUAGAUGG UCCAGGGCAU UGGAGGCCUC GACCACUCUG CAUUGCAGAU UACAGUGACU
 2821 UCCUCGGGGU UGCCCCAUCU UGGUCUCCUG UGGUUUCUUC AUCAGCUUUU UUUUUACAG
 2881 CAUCUCUCAA AUAACAAUGA AGAUAGAUAU GCCCAUUAU GUCUGAUUAA GGAGCAAAGG
 2941 CUGGAUUUCU GGCCACAGCG AGCUGCACUC UCCUCCUGC CUCAGCCGGG GUCCGUCUUA
 3001 GCAGUUUGGA AAGGGGAAAA AGAUGCCGGU CCUCACUGCU UAAGUUUUGU GUCCAGGUGC
 3061 CACUAGACUU GCAUGCACAC UAACUCCUUA CAAUCACCAC ACAGCAUCAU CGCCCCAGUG
 3121 CACAGAUGAG GAACCAGAGG CUCAGAGGAG UGAAGUUGCC UUCUGAGGU CACACAGCAU
 3181 GAAAGUGAUG AGCUAGGAUU UGAAUCUGGG AAGUUGGGCU CUAGAGCCAG ACUGUACUGC
 3241 CUUCUGCCAC ACUGUACUGC CUUCUGUGAC UGGUGGCAC CUCCAGGCA CAUUUACACA
 3301 AGGCCCUGAA UCUGCAGAGG CUGUUUCUCA AGAUGCCCGU CAUGGUGUGG CCUGGGCCAG
 3361 CUCUGGCUUC CACAGGUCCC UGACUGUCCU CAGAGUGGAA CAUGCUCUAC CUCCCGCCCA
 3421 CUGCUCUCUC UCUGCCAGA UUUCAGGGGU GCCGUGCCCC AAGGCCUGCC CCUUCUUUA
 3481 AGACUGAACU CAAGUCUCCU UGGAAGGCC CGGUGAAGCU CCCAGAGACU GGUUUUCUUG
 3541 GGAUGCAGGC AGAAGGGGAC CCUCCUGGC CAACACCCAG GAGCCCAGCA GAAGCACCCA
 3601 CACGUAAGAA GAGGCUCACU ACAGCCAGAA GUGCAGAGUC AGAGUCCUGG GACCAUCUUG
 3661 UUCUGCAAGG UGACCCAGG CUCCCAAGGA CAGGGGAGAG GGAUCGUCU CAUUCAGACU
 3721 CUCAGUGGGG CCUCUGUACU GGCUCUCCU UGGUGGGGU UGCCUGUAC AUAGCUGUGC
 3781 CUCAGAGAAA GGGUCCUGCA UUUCUGGAA UGUUCUCUGU GCUUACCCCU CUGUGUGCCC
 3841 CUCCAUUGCU CCUCUACAAG CAAUAGGUG AUUCAAAGA GCAACUAGG CUGGGUGCAG
 3901 UGACUCACAC CCGUAAUCCC GGCACUUGG GAGGCCGAGG CGGGCAGGGA CAGGAGUUA
 3961 AGACCAGCCU GGCCAACAUG GUGAAACCCU GUCUCUACAA AAAAUACAAA AAUUAACAG
 4021 ACAUUGUGGC AUGUGCCUGU AAUCCAGCU ACUCAGGAGG CUGACACAGG AGAAUUGCUU
 4081 GAACCAGGAG GCGGAGGCUG CAGUGAGCUG AGAUUGUGCC ACUGCACUCC AGCCUGGGCA
 4141 ACACAACGAG ACUCUGUCUC AAAAAAAAAA AAAAGAAAAG AAAAAAAAAA GAGCAACUUA
 4201 CUGCUUUGUG AGGUUGUGAG UGGCCACCAC UGAAGGUCUU UGAGAAGAGG CCAGACGCCG
 4261 CUGUAGCCAG GCCUGUCUUA AGAGGACUUG UGCUUCCAGG GACCCAGGCA GGAUGAUGGC
 4321 GCAGCUCUUC CUACUCCAAG CCAACGCUGU CCUUCUUUU UCCCAUGAAA UCAAGGUCAA
 4381 GAGGCAAAUA AGACUCCUG CUCCACUUA CCCCCAGAG AGAAUAGAU CUCGCUCCUU
 4441 UCAGAUCCCC CAGGAUCUGA GGGAGAAAG AUGGGAGGAG GGGCAGCAGC AUUUCGUGG
 4501 AAAGGCAGCA GAUGCUUUUC CAGCCCGGU UCAGCUGGAA GGCUGGAGG CCGGCCAGAC
 4561 CACUCUGGCG UCUCUGAAG UGGGUCCUG GAGACCGAAG AGGCUCAGUG GAGUCUGUCU
 4621 GUUGUCAGCA CUGCUGCCUG AUCCUGCAA GACAAUUGG ACUUUCUUC UUCAGAGCA
 4681 UCAUCUGCCU UCAUUAUAG CAGUAAUUA AUUCCAGUU AUUAUUCUUA CCGGUGCCAG
 4741 UUUUGCACAU CUUUUUGUG CUCUAUUUGU GUCUCAUUA CUUCUAAA UGCCCCUGG
 4801 GGCAGGAAUG AGGAUGCAGA GAGAUGCAGC UUAUUACUG UCGCAUUUUU CUGUGGAGAA
 4861 AACUGAGGCC AGGGCUGAAC GCUCACAGCU AAGGAGCUGU GAUUCAGACC CAGUUCUGUC
 4921 AGCUCUAGAG GCACCCUGCA UCAUGCCAC CAGGGUGAUC CCCCUGGGAU GGACCAUCUC
 4981 GGAUAUAGAG GCCUCGAGG CUGGGGUUGA GAUUUGGUCC UGAAGAGCUU AUAGCCAGAU
 5041 UGCCACAUUC AAGUGUAAGU CCAGGAAAG GGCAGGCGC AGUGCACAGG GAUUUAUCAG
 5101 UUCAGAACU UACAGUGAU AAGAGGCUU AGAGAGCAU UAAUCGAGAC CUUUAAUUUU
 5161 UCGGGGAGAG CAGCUGAGGC CGUGUGGAAA AUUAGUGGAG AGCUGACAAG UUGUGGGCU
 5221 CCUGGCCAG GGUCCGUGG UCCAGCACGU UGUGCGUUA GUGGGAAGCA AAGGGCUUGC
 5281 CCGGGAUUAC CUGCCCCAGC CCCGAGGUG GUUGUGUCC CUGCAGCUGC CAUCGGCCCCG
 5341 CUUUGCUUCG UCCUGGCAGA UGCCAGUGA UUGUCCCCGA GCAAGUGCCA GGUUGGGCU
 5401 GAGCUGCUAU GACAGGGAG CCCAGGGAGU UCUGCUCAGG GAGCCAAAGG GAACAGCCAG
 5461 AUCCUGAAUG UUCUAUGUUC ACCUGCCCCA GCCCCACCA CCCUGGCCA CUCCACAGGC
 5521 CCCUGACCAU GGUCACUCAC GGAGAGGGAU GGAGGAGAAG GUGGUUGAAC UGAGUACUGA
 5581 GAACCCAGAG GACAGAGCCC ACAGCUUCCA AGCAGGAAAA GGGACCUCUC UGAAAAUUCU
 5641 GGAUAACCAAG AAUUAUCACA GCACCCUCUC AUUCCAGCG CGUCCUUCG AGCUCAGACC
 5701 UUGAGCAUUU ACUGGGUUUC UUUUUGAGGA AGAGGGAAG UGACAAAGGA CAAAACAAUG
 5761 CAAAUUCUUA UGACUGAAGA CGAUCAAAGA CUCCUGGAG CGAGAAAACA GUUACUGCCA
 5821 GAAGCAGAAU GGAAGAGCCA AAAAGUACAC AAAAUGGACG CCAUAAAUUC UGAAAUAAAA
 5881 GUGUAUGAUG UGUUCUGAGU CACUGUAGAA GUCAUGCAU UAUAUUAAG AUAGAAAAGA
 5941 GCAGAGAAUG ACGUGGACAU UGUCCUCGG AGAGGCUGCA UAGGUGGUGG GGUCCCCGGG
 6001 GGAUUCUGGA UGCUGGUCU UUGACCGUG GGCAGCCUC GCGCUGCCC GGAUGGCUC

FIGURE 19K-3

```

6061 AUCCAGACAU UUGGCAAGGC UGUCAUCUGC UCUUGGGUCC UUUUUC AAGC UGAUUUCCUG
6121 CCUCCCCAAG AGGAGGUUUG AGCCCCAUUC UGCCUUGGAA AUAAAUCCUG ACAGAUGUGC
6181 ACAGCAUAUU UGCAGGGAAU UUGCAGGCCU CGAUUAUUGG GGGAAUCAGA GCCCAUUGCA
6241 CACCGAGCCC CAUCCACGUG GCUGAGCAAC ACCCCUGUGU CCCUGUAUC UGUCUCACUG
6301 GCUGUCUUUC UGCCAGUUUC UUAAGAACC ACACCAUAC UGCAUUUGCC GUUCGAAGCG
6361 UUGUCCCAAC AAUGCAGAUG GUUCUGACAA GGGUCUAUUA GCUGGCAGAA GGGAGCUUCC
6421 AACCUIUUUA CUUGAGGAAA AUGAAUCUCU GACAGGCUUG AGAGUGUUGC CAGGUGGAGC
6481 UUUUCAGGAG ACAGGGGUCU UGAGCCAGG ACACCUAACU AUJGAGUUUU CACUAGGAGA
6541 CUUAGUGGUG GCUUUCAUGA GGCCAGUCGU UCUC AAGGAA GGGCUGGGGG CCAUCAGCAA
6601 UGUCUGGAGA CAUUUUUGAU UCUCACAGCU UGGUGGGGAC UGCUAUUGGU AUCUCAUGGG
6661 UGGAGGCCAG GGAUGCCACC GAACAUCUA CAUUGCACAG GGCAGCCCCC CACAAAUAAG
6721 AAUUAUCCAG CCCC AAAUG CCAUAGUGC CGAGGUCAGG GGACCCUGUG CUGGCCAUCC
6781 UAUCUCUGAU UCUGAAACUA UGCAUGCUUU CCACUUUCC CCAUUUGUGA GUCAUUGAGU
6841 AAAUUAAGC UCUUCUGAGC AGCAGCAGUG AUCAUGGUCA CUGCCCUUGC UUCAAAUAAU
6901 GCGAGCUGAG GACAGUGAUC UGCAACUCCC AGCAUGUCAU GUGGUCUCUU AGAAAUCCAU
6961 GUGACUGUUU CCACCAUCUU GGGCAUUUGU GGGGACCCCC AGACUGGAGG GAGAAAGCCC
7021 UACAAAGUGG AUGGGAGUGU GGGGCUGAAC UUUUCCCUAC CCUUAACUUU GUGUCUUGG
7081 GACCUCCAGG GACCUGGCCC CUCACCAUUG CAUAUGAAGA GUAUGCUUGG GGAAGAGCUU
7141 AGGAAUGGGG UGGGCAUGGG AGUGCUGGGU AGCAGCCUUU GAGCAAUCU GCAUCUUCUC
7201 UUAUUUCUGA CCUUUUUCCA UGUGCCAGU CCUAUUUCUG CCAGUUGAAG GCAUACUAAU
7261 AUUCUUUAUA CUAUUUAUC UUUUGCAGAA ACCUUAUUAU UAUAACUUGC UACUCUCCAG
7321 AUACCAAUUC UUCAUGCCGA GAGCAUCGGA AAUGUUUUUG UGUCUUAUC AUGUUUUCAU
7381 GAUCAAUCUG UAAUGUAAG CAGUUGACUU CAUAAAAGGU AUUUUAACUA UUCUUGGAGU
7441 CCUUUCUAC CCAAGCACCU GGUUUCACCA UGCGAUCACU GACUUCUCUA CAGUGAAGAC
7501 UCUUUCUUA UAAAGGAUUU CGCUGUGCUC UUUUGAUUA AAUAUCUAA CCUAAAAGA
7561 CGU* (SEQ ID NO:121)

```

FIGURE 19L-1

```

1 GCCUAGUAGA GAUCUGGAGC CAGAAGCCCC GAGACAGCCG AGUGCGCCGU GCGGUCUCCG
61 GACGCUCGCU GCUCAGCCCG AUCCCCGCCA ACUGUGCAGG CGGCUGACCC GCAGCGGCAG
121 CGGCAGCAGC GAGGACUCGA GCGCUGGCUG CAGCGACACC AUGGAUCUCU CCUUUAUGGC
181 CGCGCAGGUC AAGUGAGCCU GGGGAUAAC AAGUGAGGAU GCCAAUCAU GCAUGGACCA
241 GCCUACUCC UUUCCAGUCU GCGGUUGCCA CUUGGAGUUC UCCAGCUGC CCAUGAUGGG
301 GGGAGCUUUC AUGGACUCGC CCAACGAGGA CUUCAGCACC GAGUACUCCC UGUUUAACUC
361 CUCUGCCAAU GUCCACGCGG CUGCCAAUGG CCAGGGCCAG CCGGAAGAU C UCCUCGGUC
421 CUCCAACGAC GCGGUCUUGC UAUGGAUUC CAUCAUAGCU ACGCUGGGGA ACAUCGUGGU
481 GGUGGGCGUG GUGUAUGCCU UCACCUUG UG AGGACGGCAC ACCCUGCACC ACCAUGGGGU
541 GAGGCUUGGC ACGUAGCUCU GACUUGCUGU CGGCCUUUGG CUUCUCCUGU GUUCUAGAAC
601 CAGGAGUUUU GACCAGGGGC GCGGGCCGUC CUUCUGGAAU UUCUCCCCAG CAGCCUGAU
661 UUCAAAUUC CCAUGUUGUG GUCAAGCUGA GUCAGAAGAC AUGGAAGUAU GGGCCUCCUG
721 CCCUAGAGG CAUGACGGGG CAAGGCCUUC AGAGGGCAGA UUGGGGAUCC UUGAAACUAC
781 AUUCCAGGAA CAUGGGACCA GAUGAGACAG CUAGUUAAGU UUA AAACAU GACAUGAUUU
841 GAUGAUCGCU UGCUGGUGU AAAUAUAC AAUAUAGC UCGUGUGACU UGUUUUUAUG CAUAUCUUAUC
901 GAACCUAGGG CGUGGGGUGC UGGGGCAAGA GCAGCCCUCA GAACUUCAGU GUUCCUGACC
961 CAAUUCUGGU UUCACAUUCA GUCCCUUGGC CAUCUAGUAG GGCCAUUGGA UGUUCCUAGU
1021 UUGACUUUGA AAUGGCACCU UUGCCACCAG ACACCUGGUC CCUUC AAGA CCCAAGUGCA
1081 UUGGGAGACC CAGGGAUGGG GGGUUAUGG UAAUAGGUGG GGUUUCUGGG GGUUGUUGU
1141 GUGGUUUUUA UCUCUGGUC GGACUUUCU UUCUUUUUA GAAGAGGAGA GGAUGUCUUU
1201 AAGAGCUAGA UGUGCCAGGC AGUGGACUCU UCAGGCCACC CACGUGAGAA UGCUGUUUCU
1261 UCUCUGAGGA AUCGUGGAU UUUAAAAGAU AACAAGAUUC ACAUCCACUG AAUUAUUCAA
1321 CGGAUGGGUC AGAAAGGGGG GUGAUUUGCC UGUGGUCACC AGGCAGGUUU GUGGAGGAGU
1381 CGGAACAGAA AGAUGUGUCC UCACUCUCAG CUCACUGAGC UCUCUGCCCC AACUAAGCAC
1441 UUUCCUGAGG AGGUUGCUGA GAAGCUGCCC UCAGGAGAAU GUCCAGGCAU CUUGAAGGUG
1501 GGUGCGAGAU UGGCAGGCUG CUCAGAUACC CGUCCUUUAC AUUCAGUGUG GAUACCGUGC
1561 AUUCUCCUGA AGCUGUGAAA GUGCUUCUGC CCAAGCCACU UCCUUAUUU CUGAAUAUC
1621 AGCAUCUGGG GUCCUGGCAA GCAAGGAAGC UUCAAGUAA AAACCAGAGA GAAGGGCACA
1681 CUUUUCUUUC UUCAUAGGA AAUCUUAUUG CACAGGAACC ACCCCCACCC CCACCCCCCA
1741 CACCUUCCCA AGGCAGCAUC CCAGUGCAGA UAGAGUGGGA AAGGUCCAG AAGGGGGCUC
1801 ACUCACCUCU AGGCCAGAG AGGCUUUCUC CUCACUUUAU ACACUGCAA AACAGAAGAA
1861 UUGUGUCAAU AACACCCUCU GUAGUGGAGA AACUUAAAA GCUGGUUAGG AAGCUCUCGU
1921 GUUAUUUUA AGACAAUUA AAGAAAGCUG GACUUGCCGC UGUGGUCUCA GGAGAAUGA
1981 GUGUUCUUGA UGACAGGCAA AGGGAUUCU UAGUUGUCCA GAAGCGGCAC UCUUCCUGG

```

FIGURE 19L-2

2041 AAGCCGCCAU GUUAAUAGGA UUACUAGCCU GGCUCACAGAC AGUGCCUGCU CAUGGCUGCC
 2101 AGUUCUUAACC GAUCACAUCU GUCACUGCCA CCGUAUAUCA UCUGCCAGUG CAUCAGCUUA
 2161 AGGGGAGGUC ACAGAGUGCAA AAGAACCUGA CCCUUGACAA UGAGGGAGAA GGGACAUGGA
 2221 CCACCUGUCU GGAAUUCUG GAAUCACUGG CAGGGUGGAG GCUGGGCUGG GGAGUUAGCC
 2281 GCGGUGGCG UGAAUGGCUC UGUCUCCAGC AAGUCUCUCU CCAUCAAAACC CCAGGUCUGC
 2341 CCCAUAAGCA AGAUCUUUAA CAGAUGGAUG UCUCUAGAG AAAACCCAAG GCGAGAAGCC
 2401 CAGAGCCAUG GCGGGGUUGC UUGACGUCCU CAUGGAGUCA CUCUGCCCCA CAUGCUCAAA
 2461 UCUUCCUCU GGGCCCAU CCCUAGGAGG GCCUGACCCC UGUAAAGAU CAGGAGGCAG
 2521 CUCCUGGCC UCCAAAUAGG CCAUGGAGAU GGCAGUCGGG AGACAGGGUU CUGUGUUUGC
 2581 UGCGGUGAAG GGAGGAGAAG GCAGGAGGAA AAAGGAUGGC UUCUAGCCCU GAAGAGGACU
 2641 CCAGCAUCCC AGGCACCCGG UGCUUCUGG UGCAGUUUUC CCUAUGGAG CCCUCAGCC
 2701 UCCAGCCCUA ACAUAAUGU CGGUUAAAU CAGUUUCAA GCCUCUCUC CUUUUCAGUG
 2761 UCAGAGCAGU AGAUGGUCCA GGGCAUUGGA GGCCUCGACC ACUCUGCAU GCAGAUUACA
 2821 GUGACUCCU CGGGGUUGCC CCAUCUUGGU CUCCUGUGGU UUCUUAUCA GCUUUUUUU
 2881 UACCAGCAUC UCUCAAAUAA CAAUGAAGAU AGAUUAGCCC AUUAGUGUC GAUUAAGGAG
 2941 CAAAGGCUGG AUUUCUGGCC ACAGCGAGCU GCACUCUCCC UCCUGCCUCA GCCGGGUGCC
 3001 GUCUAGCAG UUUGGAAAGG GGAUAAAGU GCCGGUCCU ACUGCUUAG UUUUGUGUC
 3061 AGGUGCCACU AGACUUGCAU GCACACUAAC UCCUUAACA CACCACACAG CAUCAUCGCC
 3121 CCAGUGCACA GAUGAGGAAC CAGAGGCUCA GAGGAGUGAA GUUGCCUUC UGAGGUCACA
 3181 CAGCAUGAAA GUGAUGAGCU AGGAUUUGAA UCUGGGAAGU UGGGCUCUAG AGCCAGACUG
 3241 UACUGCCUUC UGCCACACUG UACUGCCUUC UGUGACUGGG UGGCACCUC AGGGCACAUC
 3301 UACACAAGGC CCUGAAUCUG CAGAGGCUGU UUCUCAAGAU GCGGUCUAG GUGUGGCCUG
 3361 GGCCAGCUCU GGCUCCACA GGUCCUGAC UGUCCUCAGA GUGGAACAU CUAACCCUC
 3421 CGCCACUGC UCUCUCUCUG CCCAGAUUUC AGGGGUGCCG GUCCCCAAG CCUGCCCCCU
 3481 UCUUUAAGAC UGAACUCAAG UCUCUUGGA AGGCCCCGU GAAGCUCCA GAGACUGGU
 3541 UUCUUGGAU GCAGGCAGAA GGGGACCCUC CCUGGCAAC ACCCAGGAGC CCAGCAGAAG
 3601 CACCCACACG UAGAAAGAGG CUCACUACAG CCAGAAGUGC AGAGUCAGAG UCCUGGGACC
 3661 AUCUUGUUCU GCAAGGUGAC CCCAGGCUC CCAGGACAGG GGAGAGGGAU CGUCCUCAU
 3721 CAGACUCUAG CUGGGGCCUC UGUACUGGC UCUCCUUGGG UGGGGUUGCC UGUUACAUG
 3781 CUGUGCCUCA GAGAAAGGGU CCUGCAUUU CUGGAAUGUU CUCUGUGCU ACCCCUCUG
 3841 GUGCCCCUC AUUGCUCUC UACAAGCAA UAGGUGAUUC AAAAGAGCAA CUUAGGCUUG
 3901 GUGCAGUGAC UCACACCCGU AAUCCCGCA CUUUGGAGG CCGAGGCGG CAGGGACAGG
 3961 AGUUAAGAC CAGCCUGGCC AACAUGGUA AACCUCUGU CUACAAAAA UACAAAAAU
 4021 AACCAGACAU UGUGGCAUGU GCCUGUAUC CCAGCUACUC AGGAGGCUA CACAGGAGAA
 4081 UUGCUUGAAC CAGGAGGCGG AGGCUCAGU GAGCUGAGAU UGUGCCACUG CACUCCAGCC
 4141 UGGGCAACAC AACGAGACUC UGUCUAAAA AAAAAAAAAA GAAAAGAAA AAAAAAGAGC
 4201 AACUUAUCG AACUUGAGGU UGUGAGUGG CACCACUGAA GGUCUUUGAG AAGAGGCGAG
 4261 ACGCCGUGU AGCCAGGCU GUCUUAAGAG GACUUGUGU UCCAGGGACC CAGGCAGGAU
 4321 GAUGGCGCAG CUCUCCUAC UCCAAGCCAA CGCUGUCCU CCCCUCUCC AUGAAUCAA
 4381 GGUCAAGAG CAAUAAGAC UCCUGCUCC ACUCUACCCC CCAGAGAGAA AUGAUUCUG
 4441 CUCCUUUCAG AUCCCCAGG AUCUGAGGGA GAAAGGAUG GAGGAGGGG AGCAGCAUUU
 4501 CGCUGGAAAG GCAGCAGAU CUCUCCAGC CCCGCUUCAG CUGGAAGGCU UGGAGGCCGG
 4561 CCAGACCACU CUGGCUCUC CUGAAGUGG UCCUGGAGA CCGAAGAGC UGAGUUGAGU
 4621 CUGUCUGUUG UCAGCACUG UGCCUGAUCC CUGCAAGACA AAUGGCACU UCCUUCUUA
 4681 GAAGCAUCAU CUGCCUUAU UAUUAGCAG AAUAUUAUUC CCAGUUAUUA UUCUUAACCG
 4741 UGCCAGUUUU GCACAUCUUU UUGUUGUCU AUUUGUGUCU CAUUUACUUC UCAAAUUGC
 4801 CCUGGGGGCA GGAUAGAGGA UGCAGAGAGA UGCACGUUA UUACUGUCG AUUUUUCUGU
 4861 GGAGAAAACU GAGGCCAGG CUGAACGCUC ACAGCUAAG AGCUGUGAU CAGACCCAGU
 4921 UCUGUCAGCU CUAGAGGCAC CCUGCAUCU GGCACAGAU GCGUAGGCU UGGGAUGGAC
 4981 CAUCUCGGGA UAUGAGGCCU CGGAGGCUG GGUAGAGAU UGUUCCUUA GAGCUUAUAG
 5041 CCAGAUUGCC ACAUUAAGU GUAAGUCCAG GAAAGGGGCA GCGGCGAGU CACAGGGAUU
 5101 UAUCAGUUC AGAACCUCAC AGUGAUAGA GGCUUUAGAG AGCAUCUAAU CGAGACCUUU
 5161 AAUUUUUCGG GGAGAGCAGC UGAGGCCGUG UGGAAAUAU GUGGAGAGCU GACAAGUGUC
 5221 UGGGCUCCUG GCCCAGGGGU CCGUGGUCCA GCACGUUGG CGUUCAGUGG GAAGCAAAGG
 5281 GCUUGCCCGG GAUUACCGC CCCAGCCCCG AGGUGGGUUG UGCUCUCCU AGCUGCCAUC
 5341 GGCCCGCUUU GCUUCGUCCU GGCAGAUCC CAGUGAUUGU CCCCAGCAA GUGCCAGGU
 5401 UGGGCUAGAG UGCUAUGACA GGGAGGCCCA GGGAGUUCUG CUCAGGGAG CAAAGGGAAC
 5461 AGCCAGAUCC UGAUUGUUC AUGUUCACCU GCCCAGCCC CACCCACCCU GGCCACUCC
 5521 ACAGGCCCCU GACCAUGGUC ACUCACGGAG AGGGAUGGAG GAGAAGGUGG UUGAACUGAG
 5581 UACUGAGAAC CCAGAGGACA GAGCCACAG CUUCCAAGCA GGAAAAGGGA CCUCUCUGAA
 5641 AAUUCUGGAU AACCAGAAU AUCACAGCAC CCUCUCAUUC CCAGCGGUC CUUCUGAGCU
 5701 CAGACCUUGA GCAUUUACUG GGUUUCUUU UGAGGAAGAG GGAAAGUGAC AAAGGACAAA
 5761 ACAUUGCAAU UCUUCAUGAC UGAAGACGAU CAAAGACUCC CUGGAGCGAG AAAACAGUUA
 5821 CUGCCAGAAG CAGAAUGGAA GAGCCAAAAA GUACACAAAA UGGACGCCAU AAUUCUGAA

FIGURE 19L-3

5881 AUAAAAGUGU AUGAUGUGUU CUGAGUCACU GUAGAAGUCA UGCAUUUAUU AUCAAGAUAG
 5941 AAAAGAGCAG AGAAUGACGU GGACAUUGGU CCUCGGAGAG GCUGCAUAGG UGGUGCGGUC
 6001 CCCGGGGGAU UCUGGAUGCU GGUCUUUUGA CCGUGGCGGC AGCCUCGCGC CUGCCCGGAU
 6061 GGCUCCAUCC AGACAUUUGG CAAGGCUGUC AUCUGCUCUU GGGUCCUUUU UCAAGCUGAU
 6121 UUCUGCCUC CCCAAGAGGA GGUUUGAGCC CCAUUCUGCC UUGGAAAUAA AUCCUGACAG
 6181 AUGUGCACAG CAUAUUUGCA GGGAAUUUGC AGGCCUCGAU UAUUGGGGGA AUCAGAGCCC
 6241 AUUCCACACC GAGCCCCAUC CACGUGGCUG AGCAACACCC CUGUGUCCCC UGUUUCUGUC
 6301 UCACUGGCUG UCUUUCUGCC AGUUUCUUA AGAACCACAC CAUACUGCA UUUGCCGUUC
 6361 GAAGCGUUGU CCCAACAAUG CAGAUGGUUC UGACAAGGGU CUAUAUGCUG GCAGAAGGGA
 6421 GCUUCCAACC UUUUAACUUG AGGAAAUGA AUCUCUGACA GGCUUGAGAG UGUUGCCAGG
 6481 UGGAGCUUUU CAGGAGACAG GGGUCUUGAG CCCAGGACAC CUAACUAUG AGUUUUACAU
 6541 AGGAGACUUA GUGGUGGCUU UCAUGAGGCC AGUCGUUCUC AAGGAAGGGC UGGGGGCCAU
 6601 CAGCAAUGUC UGGAGACAUU UUUGAUUCUC ACAGCUUGGU GGGGACUGCU AUUGGUAUCU
 6661 CAUGGGUGGA GGCCAGGGAU GCCACCGAAC AUCCUACAAU GCACAGGGCA GCCCCCCACA
 6721 AAUAAGAAUU AUCCAGCCCC AAAAUGCCAA UAGUGCCGAG GUCAGGGGAC CCUGUGCUGG
 6781 CCAUCCUAUC UCUGAUUCUG AAACUAUGCA UGCUUCCAC UUUUCCCCAU UUGUGAGUCA
 6841 UUGAGUAAAU UAAAGCUCUU CUGAGCAGCA GCAGUGAUA UGGUCACUGC CCUGCGUUA
 6901 AAUAUUGCGA GCUGAGGACA GUGAUCUGCA ACUCCCAGCA UGUGAUGUGG UCUCUUAGAA
 6961 AUCCAUGUGA CUGUUUCCAC CAUCUUGGGC AUUUGUGGGG ACCCCCAGAC UGGAGGGAGA
 7021 AAGCCCUACA AAGUGGAUGG GAGUGUGGGG CUGAACUUUU CCCUACCCUU AACUUUGUGU
 7081 CUCUGGGACC UCCAGGGACC UGGCCCCUCA CCAUGCAUA UGAAGAGUAU GCUUGGGGAA
 7141 GAGCUUAGGA AUGGGUGGG CAUGGGAGUG CUGGGUAGCA GCCUUUGAGC AAUUCUGCAU
 7201 CUUCUCUUAU UUCUGACCUU UUUCCAUGUG CCCAGUCCUA UUUCUGCCAG UUGAAGGCAU
 7261 ACUAAUAUUC UUUUAUACUA UUAUUCUUU GCAGAAACCU UACUAUUAU ACUUGCUACU
 7321 CUCCAGAUAC CAUUCUUA UGCCGAGAGC AUCGGAUUG UUUUUGUGUC UUACUGAUGU
 7381 UUUCAUGAUC AACUUGUAAA UGUUAGCAGU UGACUUAUA AAAGGUUUU UAACUAUUCU
 7441 UGGAGUCCUU UGCUACCCAA GCACCUGGUU UCACCAUGCG AUCACUGACU UCUCUACAGU
 7501 GAAGACUCUU UCUUAAUAAA GGAUUUCGCU GUGCUCUUUU GAUUAAAAAU AUCUAACCUU
 7561 AAAAGACGU* (SEQ ID NO:122)

FIGURE 20A-1

1 ccacttggag gcgtctggag ccgaagcctg gagccagcca gggcgctgag agggctggag
 61 actcgccacc acgccagggc tgctgaggtc acggtcgccg gggcccgagg gcacctgctc
 121 gcggttcgc cgcctctctg ataattgtgc gggaggctgg actcgacgag cgcacggcga
 181 cgacagcagc agcaacagcg aggactggag caggggcagc agcagcagca gcagcagcag
 241 cagcagcagc cacaccagcc gtgacaccat ggatctgtct tttatggccg ctcatgtgcc
 301 catgatggga ggagcgttca tggactcatc caatgaggac ttcagtaccg agtactccct
 361 gtttaactcc cctggcaatg tccactcgac gtccaacggc cacggccagc ctgaggagtc
 421 tcctcgctcc tccaacgatg ctgttctgct atggattgcc atcatagcaa cactggggaa
 481 catagtgggt gtaggggtgg tgtatgcctt caccttctga ggatgctgaa gctgtgggga
 541 tgccctccc actctggact cacctagtaa ggtttttcat gtggctccat ctttctccca
 601 gatgctggag acacctatgt tcttgaatgg tccctggagg ggcaggcacc atacatctgt
 661 aatttctctc aggcagatcg ggttagcat attccatgtt gaggtgagc tcattcaaga
 721 cacctgaag ttggtgcttt cagctcctaa aagcttgatg gcttgagtc ttcaaacagc
 781 agactgaaaa attcttgaag tttcggtcag tcccagaatg atggttttga acatcctatt
 841 ggggtgggaa aatgcagatc tggtttggtg acattttctt agcagtcaat tatcactgta
 901 tgtcatattt ttttctgca gactttttga acccagtatg tgggtactgg gacaagaaga
 961 actttcacaa gttagaatc ctacctttta attttatgtc ccaccctcgg ggacactgac
 1021 tgcgatgttg taaggtcgca gagtgttgag aaattgactt gtaaatgata ctttgtcacc
 1081 agacacttgg tcccttgca actccaatg cattagaaaa ccagacact ggaatgaact
 1141 ttctgatctt tgtaagaaaa gcagaagacg tcttcacatt ttagatcgtc caggcagtta
 1201 gcattctctc ctccgaggat gatttaatca tgtcccctc atcactggag agatgtcagg
 1261 ggcaggggac tgatgatggt gtggttaagat aatgaaggag tcagcgctga gtctgacctc
 1321 catctagacc aagttctcta gtgagccgt cccacaccca cccagcagc tgaggcgagc
 1381 agggaatcag gagatgttca gtctgggcat ttacaggtca tctggtctct gtgctgtaaa
 1441 ttggttttagc tgctgtgccc tccgtcacag gtttgattcc ccacaccctc cagtgtatcc
 1501 agccatactt gcaaagtctt tgggggaagc tgtagtagca gagaactggc ccaacctttg
 1561 ttgccaagaa gtaggggcat agagtgtctt ctgagcctcc tagaatctc agcagggcag
 1621 gacgtggggc tgacactcag ctgtttgtgt tgggcatgg aagttgtgg atgtttgctc
 1681 tggcttcctg cctcctcgtc acaccctga gcctttccct gtgcccaggg tggggcactc
 1741 tcatgagaga acccaagcca gcattgggaa cgggagtgtg ctgtgaccca gtgctatctg
 1801 agtgctctt gggagctgca gctgggtgtc tgtcccctac cttatctcca tggagatggc
 1861 aggttgggga aggatctgtg tctgctgtac tgagagacag gacaggaaag ctaaggacag
 1921 cttctagcct tgtaagcatg tcagtgtctt tacacacaca cacacacaca cacacacaca

FIGURE 20A-2

```

1981 cacacacaca ggcacacagg cacacacaga gagagagaga gagagagaga gagagagaga
2041 gagagagaga gaggcagatg acatctttcc tgctctccta gatcagacag agtggttacat
2101 tgtagacagt gttacattgt cacggacatg ggaatcccct tgccaacctg gcatcaactg
2161 gtcacattct gccctcctat ttgctccagt ttcttatctg aacttcctca ccacagataa
2221 cagtgaggaa caaatgacag gtcccagcca gtagtagcaa ggtgcagccc actgtccttc
2281 ctgcctctgc ttggctctgt cttggtcatc agctaaggga cgagcaggcc tgagctcagt
2341 gcttgagtgt tgtgagcagg catggtgcta gatttttatg cacagccttc atccttccat
2401 cacctcatag cagcaaggtc cgtcgtgccc agaaaggaac cagaggccca aagaagtggg
2461 gtaacttccg tcaggccaca cagtgtagaa gcgaggatcc tctctctcca ctgaggtaca
2521 tttgaacagg acttgagtca tcataagggtg tttcccaagg tatcccagct tgggatggct
2581 cagaccatca gaatgggaca gcttcggctt tgctacacc ccagggataa tgaaccttcc
2641 cccaccactt aaaatcgacc ttcaatctcc atgtaggctt ccaagaaggg ttcccgaagt
2701 agggtttctt ggtcacatgt ataaggggcc ccagcaagg cttgcaagcc ccaaggaact
2761 gcacattagg aaggtttcag tgaccagcac tgcagtcaga ttgtgtagca agctgaccga
2821 agctggggta cagtctttct gtgagcccca aagctctcca gagcagagga tagcatcttc
2881 ttggggaccc cagctagtac ccattgcct tctccctagg taaggcgggt ccagagccc
2941 ctttaatttg acaggaaggc ccttaccttt tgggtatctg acccttggag aagtgtctgc
3001 cttgctgaaa tgttccctgt ggtccatgaa atccaattag aggaacttat gatttcttt
3061 tctgccttgc ccagagagcc cactgtccct ccttgtagtc gtttctagga tcccaaggat
3121 caggaaatgg ggcaggctcag tgccacacag ccagagaggg agtatgccag cttctggtcc
3181 tggctaggat gcttgtccac tgactaggat ctgtgtggga gctagagcca cacacgcagc
3241 tgaatatgtg atctgggcat tgtggccctc tctgcaggg gtatggccac ctttgtcat
3301 caatgggtag taaaatggcc tatctctcct gctgcagcca gtcttgctta tctctcctgc
3361 tgcagccagt cctacttaca agatcatata tttcccaagt cgggaccaag gaggagcagt
3421 cagagaccaa ctggcacatg ggcacctgtg ggcacctc ctagcctctt tctccaggg
3481 ccaactctgt cagggtctgg aggagagaga ggggtgtaag cttcagctga gagccccgag
3541 gtacagaaca caggctcggg ggggtcctt gaagaaaccg ggggatgaaa ttatgagctc
3601 tggttacccc ttgtgcaacc tctgagatc agagcttgag caagaactgg tatgggggag
3661 ggaggggagg aaccacaaaa ccacagggc agagggcaat gcacatcact accagcgtga
3721 caaaacctg ccatcactca ggacagatgg aggtacacc tggaacacaa gtctgcattg
3781 ttgggcctgg ctactgtgg gtgcgctctg tggtttactg atgggaaata aaacccaagg
3841 gccgaaaagt tgctcagtgg gtagagcatg agaaccctga gttcagatcc tcagagccca
3901 cattaaagta aaaccagctc tgggaggagg ggggaagata ggcaggttca gggggttcgc
3961 tggccagcta ctagaaccaa aaaagatgag ctctggttcc gtgagagacg atggttcttc
4021 ttttgtttgc tttgagtttt gttttgcttt tttaaatggt gagtggataa atgattaagg
4081 aggatccttg atgtcaaccc ctggtctcca catgcaagca agtacaagta cctacacatg
4141 tacatgtgtg cacacatgca cagtagaca agatagagag agagagaggg agaaggaaat
4201 ccaggtgact caggcctttc cgaggacaag cttgagcaca gtagtgtctc catgactgtc
4261 aaggccacat cattggggag acatcgtctc aacctcccca tctttctgcc gatggctgcc
4321 tcaaaagcat ttcgatctg cttttccatg gggctaactc ccagccgatg tgcgttcagc
4381 gtgcattggc cacagggtac ttoggagcga gccctggttc ttcattggagc cctctgtctg
4441 gtggagtagc ccaagtctat gtctgcctcc tgtggtcttt cttgattttt ttttaagact
4501 tgcagcctgt atgagacctc atgtgactga ggatagctgt gctttccacc caaaacaaaa
4561 tcatacactt aaaacattac gaggtttctt tttcttttgt cctttgtaac tcagttgctc
4621 agtttggggg gggaaacttg cagatggtgt cactgtgctg aaatgtcaaa gcataagacc
4681 tgagaatcat tggcatttgc agttcaatgt actgttatta cattaccaat cgttctgtgt
4741 gcaggcttgc caagagggca tggatgccca acctcagggt ggtttgaggg cattgccagg
4801 tgggctgata tgtagaagag cagagtcttg tcccagatt tcccaccaga gaattcccgt
4861 gattcttcat ttcaggcgcc attgttcccc aggggacatt tggcaatgcc tagagatcat
4921 tcttgtctcg ggtttacaga ggagtgcgga ggtcctgcct ctagaggggt gaggccagag
4981 gtgctgtcta tccacacaca caggatggac cctatggctg agtgctaccc agccccaaa
5041 tgccaacggc accgaagatg ggagactcca tcagacctca ggtccaaat cgatggggag
5101 ttttccaaag ctgcacctg catgtgtgag tcagatcaca cgtgggtcag agaaaatcag
5161 agcacagagg ctgctcaagg tgagagaatg tcagccatgt gactgtctcc tgggtggaga
5221 tgacctacg atacagccag actgctgagg ctgagcctgc ctcccacctg tgactctgtc
5281 caggacctct gaggcctgcc tctcaccagg gcacctgaga aacatgatca gtggagagct
5341 cctgtgggg acaggaacag catctgacaa atccaggtct cctttctgac cctctgtgtc
5401 cccagccct gctcgccaaa gctgcagaca ttctaactct tcttctgtc taactctttc
5461 cggaaattctt acccatcacc tgtgtctcac ggattccagc cgttcaccag gagagcatca
5521 gccaaagctt tgtgtttgtg acgattgctt agccagcttg taaatgtgac atggaggtac
5581 ttttaactgtt gttctgggtc cctgctgccc aggcacctgg ttctctgtgt attactggct
5641 tctattttgc agcagaaagg ttctttgtct atgtaagatt aacgtgtgct ctcttgaagg
5701 tatgatggtg acatttccct gtgttagctt tcttgcctat gttgtcattt tctatttcaa
5761 ttaaacaaaa ggaagattta a (SEQ ID NO:117)

```

FIGURE 20B-1

Score = 317 bits (165), Expect = 3e-82
 Identities = 229/261 (87%), Gaps = 0/261 (0%)
 Strand=Plus/Plus

```

Human  155  GACACCATGGATCTCTCCTTTATGGCCGCGCAGCTGCCCATGATGGGGGGAGCTTTCATG  214
          |||||
Mouse  263  GACACCATGGATCTGTCTTTATGGCCGCTCAGTTGCCCATGATGGGAGGAGCGTTCATG  322

Human  215  GACTCGCCCAACGAGGACTTCAGCACCGAGTACTCCCTGTTTAACTCCTCTGCCAATGTC  274
          |||||
Mouse  323  GACTCATCCAATGAGGACTTCAGTACCGAGTACTCCCTGTTTAACTCCCCTGGCAATGTC  382

Human  275  CACGCGGCTGCCAATGGCCAGGGCCAGCCGGAAGATCCTCCTCGGTCTCTCCACGACGCC  334
          |||||
Mouse  383  CACTCGACGTCCAACGGCCACGGCCAGCCTGAGGAGTCTCCTCGCTCCTCCAACGATGCT  442
  
```

FIGURE 20B-2

```

Human  335  GTCTTGCTATGGATTGCCATCATAGCTACGCTGGGGAACATCGTGGTGGTGGGCGTGGTG  394
          ||
Mouse  443  GTTCTGCTATGGATTGCCATCATAGCAACACTGGGGAACATAGTGGTGGTAGGGGTGGTG  502

Human  395  TATGCCTTCACCTTCTGAGGA  415 (SEQ ID NO:118)
          |||||
Mouse  503  TATGCCTTCACCTTCTGAGGA  523 (SEQ ID NO:119)
  
```

FIGURE 20C

```

      10      20      30      40      50      60
MDLSFMAAQL PMMGGAFMDS SNEDFSTEYS LFNSPGNVHS TSNGHGQPEE SPRSSNDAVL
      70      80
LWIAIIATLG NIVVVGVVYA FTF (SEQ ID NO:116)
  
```

FIGURE 20D

CLUSTAL 2.0.5 multiple sequence alignment:

```

human      MDLSFMAAQLPMMGGAFMDSPNEDFSTEYSLFNSSANVHAAANGQGQPEDPPRSSNDAVL 60
Mouse      MDLSFMAAQLPMMGGAFMDSSNEDFSTEYSLFNSPGNVHSTSNHGQPEESPRSSNDAVL 60
          *****
human      LWIAIIATLGNIVVVGVVYAFTF 83 (SEQ ID NO:1)
Mouse      LWIAIIATLGNIVVVGVVYAFTF 83 (SEQ ID NO:116)
          *****
  
```

INTERNATIONAL SEARCH REPORT

International application No.

PCT/NZ2008/000069

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl.

C12N 15/12 (2006.01) *A61K 48/00* (2006.01) *C07K 14/82* (2006.01)
A61K 38/10 (2006.01) *A61P 35/00* (2006.01) *C07K 16/32* (2006.01)
A61K 38/18 (2006.01) *A61P 35/04* (2006.01) *C12N 15/63* (2006.01)
A61K 39/395 (2006.01) *C07H 21/02* (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)

IPC: C12N, C07K

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)

Genomequest: SEQ search based upon the claimed sequences

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US2004/0253606 A1 (AZIZ <i>et al</i>) 16 th December 2004 Paragraphs [0002], [0031], [0032], [0268]-[0307], [0328]-[0333] and SEQ ID NO.'s: 2179, 3356, 6224 and 7340	2-6, 8-9, 12- <u>24, 27</u>
Y		10-11
X	US2005/0196754 A1 (DRMANAC <i>et al</i>) 8 th September 2005 Paragraphs [0229]-[0235], [0321]-[0360] and SEQ ID NO.'s 1233, 1234, 31600 and 31602	5, 7, 12, 15-21 and 25-31
X	GenBank, 09 January 2003, "Homo sapiens chromosome 14 open reading frame 132, mRNA", Database accession number BC043593	1, 3-6 and 12



Further documents are listed in the continuation of Box C



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	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"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
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

22 May 2008

Date of mailing of the international search report

27 JUN 2009

Name and mailing address of the ISA/AU

AUSTRALIAN PATENT OFFICE
 PO BOX 200, WODEN ACT 2606, AUSTRALIA
 E-mail address: pct@ipaustalia.gov.au
 Facsimile No. +61 2 6283 7999

Authorized officer

GAVIN BARTELL

AUSTRALIAN PATENT OFFICE
 (ISO 9001 Quality Certified Service)
 Telephone No : (02) 6222 3647

INTERNATIONAL SEARCH REPORT

International application No.

PCT/NZ2008/000069

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GenBank, 20 December 1999, "DKFZp547H184_r1 547 (synonym: hfbr1) Homo sapiens cDNA clone DKFZp547H184 5', mRNA sequence", Database accession number AL134275	4 and 12
X	GenBank, 03 November 2005, "Homo sapiens cDNA clone BRAMY3003467 5', mRNA sequence", Database accession number DA184725	2
Y	US2007/0032441 A1 (McSWIGGEN <i>et al</i>) 8 th February 2007 The whole document, in particular the abstract	10-11

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/NZ2008/000069

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			
US	2004253606	AU	2003298786	WO	2004048938		
US	2005196754	AU	22924/01	AU	25918/01	AU	25936/01
		AU	25955/01	AU	25965/01	AU	25983/01
		AU	27284/01	AU	27344/01	AU	27348/01
		AU	27385/01	AU	31177/01	AU	31178/01
		AU	31288/01	AU	32657/01	AU	32967/01
		AU	32969/01	AU	32971/01	AU	32989/01
		AU	33003/01	AU	33017/01	AU	33041/01
		AU	33047/01	AU	33293/01	AU	34582/01
		AU	34847/01	AU	34848/01	AU	34865/01
		AU	34944/01	AU	36553/01	AU	36658/01
		AU	36660/01	AU	36663/01	AU	36721/01
		AU	38347/01	AU	39955/01	AU	41511/01
		AU	41873/01	AU	43142/01	AU	45280/01
		AU	45491/01	AU	49251/01	AU	49728/01
		AU	50872/01	AU	51194/01	AU	52926/01
		AU	53620/01	AU	55214/01	AU	55454/01
		AU	59041/01	AU	63005/01	AU	63006/01
		AU	64724/01	AU	74871/01	AU	77840/01
		AU	96235/01	AU	2001234944	AU	2001263005
		AU	2002250419	AU	2002326877	AU	2002332767
		AU	2002360460	AU	2002366951	AU	2002368010
		AU	2003213064	AU	2003240755	AU	2003259696
		AU	2003295649	AU	2003295902	AU	2003298604
		AU	2003302733	AU	2004304767	AU	2007234602
		CA	2395443	CA	2395731	CA	2395736
		CA	2395749	CA	2395763	CA	2395770
		CA	2398215	CA	2398251	CA	2398396
		CA	2398546	CA	2398548	CA	2398554
		CA	2399644	CA	2399649	CA	2399652
		CA	2399672	CA	2399681	CA	2399686
		CA	2399776	CA	2401505	CA	2402293

INTERNATIONAL SEARCH REPORT

International application No.

Information on patent family members

PCT/NZ2008/000069

CA	2402563	CA	2404501	CA	2404555
CA	2405104	CA	2406039	CA	2406121
CA	2409164	CA	2425827	CA	2429343
CA	2440257	CA	2441006	CA	2442077
CA	2453344	CA	2458818	CA	2460425
CA	2469941	CA	2476555	CA	2506356
CA	2507156	CA	2522421	EP	1240178
EP	1242443	EP	1242580	EP	1242596
EP	1248848	EP	1250346	EP	1252333
EP	1254246	EP	1254256	EP	1254265
EP	1254266	EP	1254268	EP	1254269
EP	1254270	EP	1261735	EP	1261736
EP	1261743	EP	1268510	EP	1268543
EP	1268762	EP	1274716	EP	1276751
EP	1276754	EP	1276902	EP	1282635
EP	1285084	EP	1325120	EP	1341804
EP	1345965	EP	1365801	EP	1381621
EP	1409715	EP	1427747	EP	1430146
EP	1432800	EP	1483407	EP	1492557
EP	1504099	EP	1567863	EP	1572235
EP	1572987	EP	1574520	EP	1578989
EP	1591448	EP	1628991	EP	1661986
EP	1683809	MX	PA02006193	US	6436703
US	6465620	US	6569662	US	6586390
US	6610536	US	6635742	US	6667391
US	6673904	US	6783969	US	6806254
US	6818754	US	6824973	US	7029677
US	7067301	US	7109030	US	7317099
US	7319141	US	7378253	US	2002009786
US	2002061567	US	2002127199	US	2002128187
US	2002137044	US	2002142302	US	2002142953
US	2002146692	US	2002146757	US	2002150898
US	2002197679	US	2003022329	US	2003022825
US	2003032034	US	2003044792	US	2003073099
US	2003087370	US	2003092112	US	2003096279
US	2003100746	US	2003104413	US	2003104529
US	2003113847	US	2003124573	US	2003138804

INTERNATIONAL SEARCH REPORT

International application No.

Information on patent family members

PCT/NZ2008/000069

US	2003144491	US	2003152941	US	2003158400
US	2003165881	US	2003165921	US	2003166152
US	2003170818	US	2003180722	US	2003180745
US	2003211987	US	2003215453	US	2003219743
US	2003219744	US	2003224379	US	2003228584
US	2003232054	US	2003235883	US	2004005317
US	2004005592	US	2004013657	US	2004022786
US	2004023870	US	2004034208	US	2004048249
US	2004048817	US	2004053245	US	2004053248
US	2004053250	US	2004068097	US	2004072173
US	2004106116	US	2004110940	US	2004116683
US	2004121948	US	2004122220	US	2004132022
US	2004137434	US	2004152884	US	2004162419
US	2004219521	US	2004220383	US	2004236090
US	2004248156	US	2005027114	US	2005059073
US	2005095237	US	2005153334	US	2005164202
US	2005170374	US	2005175607	US	2005175994
US	2005192215	US	2005202422	US	2005208488
US	2005208498	US	2005221342	US	2005226812
US	2005239060	US	2005266423	US	2005281813
US	2006052591	US	2006073514	US	2006074044
US	2006104977	US	2006149049	US	2006258567
US	2006263803	US	2007042392	US	2007049743
US	2007060743	US	2007264261	US	2008050393
US	2008051565	US	2008076715	WO	0152616
WO	0153312	WO	0153326	WO	0153453
WO	0153454	WO	0153455	WO	0153456
WO	0153466	WO	0153485	WO	0153500
WO	0153515	WO	0154477	WO	0155332
WO	0155333	WO	0155334	WO	0155335
WO	0155336	WO	0155337	WO	0155339
WO	0155435	WO	0155437	WO	0155442
WO	0157175	WO	0157187	WO	0157188
WO	0157190	WO	0157233	WO	0157255
WO	0157260	WO	0157261	WO	0157262
WO	0157265	WO	0157266	WO	0157267
WO	0164834	WO	0164835	WO	0164839

INTERNATIONAL SEARCH REPORT

International application No.

Information on patent family members

PCT/NZ2008/000069

		WO	0164840	WO	0166558	WO	0166689
		WO	0174836	WO	0175064	WO	0175067
		WO	0175093	WO	0177169	WO	0179254
		WO	0179446	WO	0179447	WO	0179449
		WO	0187917	WO	0188088	WO	0188091
		WO	0188092	WO	0216439	WO	0231111
		WO	02051868	WO	02059260	WO	02066600
		WO	02070539	WO	02072138	WO	02077180
		WO	03023013	WO	03029405	WO	03048326
		WO	03054152	WO	03068935	WO	03082915
		WO	2004009834	WO	2004015078	WO	2004047612
		WO	2004047758	WO	2004053061	WO	2004078918
		WO	2005060375	WO	2005069854		
US	2007032441	AU	10904/95	AU	10973/95	AU	15665/95
		AU	18214/95	AU	26422/95	AU	36657/99
		AU	38111/01	AU	38724/99	AU	42229/93
		AU	42489/93	AU	43454/01	AU	44120/96
		AU	47698/93	AU	48760/99	AU	52096/98
		AU	60255/98	AU	67085/98	AU	68041/98
		AU	72905/98	AU	73755/94	AU	74733/00
		AU	76662/96	AU	80886/94	AU	98253/98
		AU	2002365084	AU	2002367785	AU	2003207708
		AU	2003210895	AU	2003211058	AU	2003211082
		AU	2003213005	AU	2003213054	AU	2003213057
		AU	2003213090	AU	2003213119	AU	2003213163
		AU	2003213203	AU	2003215161	AU	2003215203
		AU	2003215345	AU	2003216245	AU	2003216255
		AU	2003216265	AU	2003216311	AU	2003216315
		AU	2003216323	AU	2003216324	AU	2003217550
		AU	2003217594	AU	2003219712	AU	2003219751
		AU	2003219781	AU	2003219817	AU	2003219818
		AU	2003219833	AU	2003220136	AU	2003221258
		AU	2003228667	AU	2003245512	AU	2003247204
		AU	2004266311	AU	2004274951	AU	2004288143
		AU	2005201389	AU	2005212433	AU	2006261653
		AU	2006279454	AU	2006336384	CA	2135499
		CA	2135646	CA	2140343	CA	2168566

INTERNATIONAL SEARCH REPORT

International application No.

Information on patent family members

PCT/NZ2008/000069

CA	2174339	CA	2175824	CA	2176035
CA	2180740	CA	2183992	CA	2190513
CA	2207593	CA	2235368	CA	2279548
CA	2288640	CA	2326695	CA	2330574
CA	2398282	CA	2400415	CA	2403243
CA	2442092	CA	2447161	CA	2448320
CA	2455447	CA	2455506	CA	2456444
CA	2457528	CA	2459532	CA	2463595
CA	2468048	CA	2471421	CA	2476112
CA	2476394	CA	2477014	CA	2526831
CA	2528963	CA	2537085	CA	2541643
CA	2542527	CA	2542835	CA	2543013
CA	2543029	CA	2543030	CA	2554212
CA	2597724	CA	2619876	CN	1312856
EP	0641212	EP	0642589	EP	0654077
EP	0716693	EP	0725788	EP	0728204
EP	0728205	EP	0737247	EP	0746614
EP	0763106	EP	0786522	EP	0799313
EP	0859837	EP	0980424	EP	0986637
EP	1073732	EP	1075508	EP	1212416
EP	1251170	EP	1260586	EP	1263947
EP	1265995	EP	1288296	EP	1321521
EP	1383782	EP	1386004	EP	1390385
EP	1390472	EP	1408111	EP	1423404
EP	1423406	EP	1430157	EP	1432724
EP	1432725	EP	1436314	EP	1442143
EP	1448580	EP	1448590	EP	1458741
EP	1463842	EP	1465910	EP	1470256
EP	1470257	EP	1472265	EP	1472267
EP	1472269	EP	1474433	EP	1476457
EP	1476458	EP	1476459	EP	1476574
EP	1478730	EP	1493818	EP	1495041
EP	1499628	EP	1499631	EP	1501853
EP	1502950	EP	1521768	EP	1522583
EP	1549660	EP	1572067	EP	1572128
EP	1622572	EP	1627061	EP	1644498
EP	1644500	EP	1664299	EP	1670915

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/NZ2008/000069

EP	1675948	EP	1675949	EP	1675950
EP	1675951	EP	1675953	EP	1682660
EP	1682661	EP	1694838	EP	1710307
EP	1713915	EP	1735443	EP	1741781
EP	1767632	EP	1817415	EP	1844147
EP	1891217	EP	1922300	GB	2396155.
GB	2396616	GB	2396864	GB	2397062
GB	2397818	GB	2406568	GB	2406569
GB	2413557	GB	2415961	ID	28053
JP	2000342285	JP	2000342286	JP	2004147666
JP	2005118056	JP	2006271387	JP	2007300926
MX	9302726	MX	9302808	MX	9304329
MX	9406022	MX	9603638	MX	9605716
MX	9704440	US	5496698	US	5525468
US	5610054	US	5612215	US	5616488
US	5616490	US	5622854	US	5627053
US	5631359	US	5631360	US	5639647
US	5646020	US	5646042	US	5658780
US	5672501	US	5672511	US	5686599
US	5693532	US	5693535	US	5714383
US	5716824	US	5731295	US	5767263
US	5783425	US	5795778	US	5804683
US	5811300	US	5817796	US	5831071
US	5837542	US	5837855	US	5840876
US	5869253	US	5877021	US	5879938
US	5891683	US	5891684	US	5902880
US	5972699	US	5972704	US	5977343
US	5985621	US	6017756	US	6022962
US	6054576	US	6057156	US	6093555
US	6103890	US	6117657	US	6127535
US	6132966	US	6132967	US	6146886
US	6159692	US	6162909	US	6183959
US	6194150	US	6239272	US	6258585
US	6280936	US	6303773	US	6316612
US	6346398	US	6353098	US	6362323
US	6365374	US	6379954	US	6410224
US	6432704	US	6436644	US	6437117

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/NZ2008/000069

US	6440719	US	6448009	US	6451540
US	6469158	US	6482932	US	6489465
US	6509460	US	6528639	US	6528640
US	6548657	US	6555668	US	6566127
US	6617438	US	6623962	US	6649751
US	6673918	US	6797815	US	6818447
US	6818759	US	6849726	US	6852535
US	7022828	US	7034009	US	7041817
US	7109165	US	7176304	US	2001006801
US	2001012618	US	2002013458	US	2002028919
US	2002082225	US	2002102568	US	2002103366
US	2002177568	US	2002192685	US	2002193579
US	2002197684	US	2003003469	US	2003004122
US	2003008295	US	2003050259	US	2003060611
US	2003064945	US	2003065155	US	2003068301
US	2003073207	US	2003088087	US	2003092003
US	2003092646	US	2003104985	US	2003105051
US	2003105308	US	2003113891	US	2003119017
US	2003124513	US	2003125270	US	2003125291
US	2003130186	US	2003140362	US	2003143708
US	2003143732	US	2003144489	US	2003148507
US	2003148985	US	2003153521	US	2003165900
US	2003165969	US	2003166896	US	2003166917
US	2003170644	US	2003170891	US	2003171311
US	2003175950	US	2003186909	US	2003190635
US	2003191077	US	2003203870	US	2003206887
US	2003207837	US	2003216335	US	2004006035
US	2004009510	US	2004019001	US	2004023292
US	2004054156	US	2004077565	US	2004102389
US	2004102413	US	2004110296	US	2004127446
US	2004138163	US	2004142895	US	2004147735
US	2004161758	US	2004192626	US	2004198682
US	2004209831	US	2004209832	US	2004219671
US	2004220128	US	2004249178	US	2005014172
US	2005020525	US	2005032733	US	2005048529
US	2005054596	US	2005054598	US	2005070497
US	2005075304	US	2005079610	US	2005080031

INTERNATIONAL SEARCH REPORT

International application No.

Information on patent family members

PCT/NZ2008/000069

US	2005096284	US	2005106726	US	2005119211
US	2005119212	US	2005124566	US	2005124567
US	2005124568	US	2005124569	US	2005130181
US	2005136436	US	2005137153	US	2005137155
US	2005142578	US	2005143333	US	2005148530
US	2005153914	US	2005153915	US	2005153916
US	2005158735	US	2005159376	US	2005159378
US	2005159379	US	2005159380	US	2005159381
US	2005159382	US	2005164224	US	2005164966
US	2005164967	US	2005164968	US	2005170371
US	2005171039	US	2005171040	US	2005176024
US	2005176025	US	2005176038	US	2005176663
US	2005176664	US	2005176665	US	2005176666
US	2005182006	US	2005182007	US	2005182008
US	2005182009	US	2005187174	US	2005191618
US	2005191638	US	2005196765	US	2005196767
US	2005196781	US	2005203040	US	2005209179
US	2005209180	US	2005209182	US	2005222064
US	2005222066	US	2005227935	US	2005227936
US	2005227937	US	2005233329	US	2005233344
US	2005233996	US	2005233997	US	2005233998
US	2005239731	US	2005239739	US	2005256068
US	2005260620	US	2005261212	US	2005261219
US	2005266422	US	2005267058	US	2005277133
US	2005277608	US	2005282188	US	2005287128
US	2005288242	US	2006019913	US	2006019917
US	2006025361	US	2006127891	US	2006142225
US	2006142226	US	2006142557	US	2006148743
US	2006154271	US	2006160757	US	2006211642
US	2006216747	US	2006217331	US	2006217332
US	2006217334	US	2006217335	US	2006217336
US	2006217337	US	2006223990	US	2006240554
US	2006241075	US	2006247194	US	2006247428
US	2006247429	US	2006270623	US	2006275903
US	2006276422	US	2006276635	US	2006281175
US	2006287266	US	2006287267	US	2006292691
US	2006293271	US	2006293272	US	2007004663

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/NZ2008/000069

US	2007004664	US	2007004665	US	2007004667
US	2007026394	US	2007042029	US	2007042983
US	2007049543	US	2007093437	US	2007099858
US	2007160980	US	2007161596	US	2007167393
US	2007173467	US	2007173473	US	2007179104
US	2007185043	US	2007185049	US	2007203333
US	2007270360	US	2007270579	US	2008020058
US	2008032942	US	2008033156	US	2008039412
US	2008039414	WO	0116312	WO	0159103
WO	0166721	WO	9323057	WO	9323569
WO	9402595	WO	9504818	WO	9511910
WO	9513378	WO	9513380	WO	9519429
WO	9523225	WO	9531541	WO	9618736
WO	9715662	WO	9833893	WO	9850530
WO	9955847	WO	9955857	WO	02081494
WO	02081628	WO	02094185	WO	02096927
WO	02097114	WO	03057709	WO	03070193
WO	03070197	WO	03070742	WO	03070743
WO	03070744	WO	03070750	WO	03070881
WO	03070884	WO	03070885	WO	03070886
WO	03070887	WO	03070888	WO	03070895
WO	03070896	WO	03070897	WO	03070903
WO	03070910	WO	03070911	WO	03070912
WO	03070914	WO	03070917	WO	03070918
WO	03070966	WO	03070968	WO	03070969
WO	03070970	WO	03070972	WO	03070983
WO	03072590	WO	03072704	WO	03072705
WO	03074654	WO	03089650	WO	03102131
WO	03106476	WO	2004092383	WO	2004097020
WO	2004111237	WO	2005003350	WO	2005007854
WO	2005007855	WO	2005007859	WO	2005014811
WO	2005019453	WO	2005028649	WO	2005028650
WO	2005035759	WO	2005040379	WO	2005041859
WO	2005044981	WO	2005045032	WO	2005045034
WO	2005045035	WO	2005045036	WO	2005045037
WO	2005045038	WO	2005045039	WO	2005045040
WO	2005045041	WO	2005078097	WO	2005105995

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/NZ2008/000069

WO	2006060598	WO	2006078798	WO	2006128141
WO	2007022369	WO	2007067981	WO	2007076328
WO	2007084865	WO	2007086881	WO	2007086883
WO	2007092059	WO	2007147143	WO	2008011431
WO	2008030239				

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

END OF ANNEX