

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
10 August 2006 (10.08.2006)

PCT

(10) International Publication Number
WO 2006/084007 A2

(51) International Patent Classification:
C07H 21/04 (2006.01)

(21) International Application Number:
PCT/US2006/003645

(22) International Filing Date: 1 February 2006 (01.02.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/648,837 1 February 2005 (01.02.2005) US

(71) Applicant (for all designated States except US): **THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK** [US/US]; West 116th Street And Broadway, New York, New York 10027 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **TALL, Alan R.** [US/US]; 254 Wilson Drive, Cresskill, New Jersey 07626 (US). **WELCH, Carrie** [US/US]; 260 Kelbourne Avenue, Sleepy Hollow, New York 10591 (US). **SEIDELMANN, Sara Bretschger** [US/US]; 30 WEST 63rd STREET, Apt. 7B, New York, New York 10023 (US).

(74) Agent: **WHITE, John P.**; COOPER & DUNHAM LLP, 1185 Avenue Of The Americas, New York, New York 10036 (US).

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

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

Published:

— without international search report and to be republished upon receipt of that report

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

(54) Title: NOVEL GENE UNDERLYING ATHEROSCLEROSIS LOCUS I

(57) Abstract: This invention provides an isolated nucleic acid encoding Adamtsl-1 protein and related compositions. This invention further provides a method of inhibiting expression of Adamtsl-1 in a cell. This invention further provides a method for inhibiting the onset of atherosclerosis in a subject and a method of treating atherosclerosis in a subject. This invention further provides transgenic animals. Finally, this invention provides for a method of screening for possible candidate agents for treating a human afflicted with atherosclerosis.



WO 2006/084007 A2

NOVEL GENE UNDERLYING ATHEROSCLEROSIS LOCUS I

5 This application claims benefit of U.S. Provisional Application No. 60/648,837, filed February 1, 2005, the contents of which are hereby incorporated by reference.

10 This invention was made with support under United States Government Grant No. HL54591 from the National Institutes of Health. Accordingly, the United States Government has certain rights in the subject invention.

15 Throughout this application, various publications are referenced. Full citations for these publications may be found immediately preceding the claims. The disclosures of these publications are hereby incorporated by reference into this application in order to more fully describe the state of the art as of the date of the invention described and claimed herein.

20

Background of the Invention

25 Atherosclerosis, a progressive disease of the large arteries, is the primary cause of coronary heart disease and stroke. While environmental factors, including smoking, diet, and exercise, contribute to atherosclerosis susceptibility, an estimated 40-60% of variance is due to genetic factors [1]. However, the genes contributing to inter-individual variation in susceptibility are largely unknown [2].

30

One method of studying the genetics of complex diseases is through animal models. Several atherosclerosis loci have been mapped in the murine model [3-7], but none of the underlying genes have been identified by positional
5 cloning.

Summary of the Invention

This invention provides an isolated nucleic acid encoding Adamtsl-1 protein, an isolated Adamtsl-1 protein, an antibody that specifically binds to Adamtsl-1 protein and a vector comprising a nucleic acid encoding Adamtsl-1. This invention also provides a cell comprising a vector, which vector comprises a nucleic acid encoding Adamtsl-1.

This invention further provides a method for inhibiting expression of Adamtsl-1 in a cell wherein Adamtsl-1 expression would otherwise occur comprising contacting the cell with an agent that inhibits expression of Adamtsl-1 in the cell.

This invention further provides a method for inhibiting the onset of atherosclerosis in a subject at risk for becoming afflicted therewith, comprising administering to the subject a prophylactically effective amount of an agent that inhibits the expression of Adamtsl-1 in the subject's cells, thereby inhibiting the onset of atherosclerosis in the subject.

This invention further provides a method of treating a subject afflicted with atherosclerosis comprising administering to the subject a therapeutically effective amount of an agent that inhibits the expression of Adamtsl-1 in the subject's cells, thereby treating the subject.

This invention further provides a transgenic animal whose somatic cells are characterized by a genotype comprising (i) an LDL-receptor knockout and (ii) a polymorphism for

Adamts1-1, wherein the polymorphism for Adamts1-1 predisposes the animal to developing atherosclerotic lesions.

5 This invention further provides a transgenic animal whose somatic cells are characterized by a genotype comprising (i) an Apo-E knockout and (ii) a polymorphism for Adamts1-1, wherein the polymorphism for Adamts1-1 predisposes the animal to developing atherosclerotic lesions.

10 This invention further provides a transgenic animal whose somatic cells are characterized by a genotype comprising a polymorphism for Adamts1-1, wherein the polymorphism for Adamts1-1 predisposes the animal to developing
15 atherosclerotic lesions.

This invention further provides a method for determining whether an agent is a candidate agent for treating a human afflicted with atherosclerosis, which comprises: (a)
20 administering the agent to a transgenic animal whose somatic cells are characterized by a genotype comprising (i) an LDL-receptor knockout and (ii) a polymorphism for Adamts1-1, wherein the polymorphism for Adamts1-1 predisposes the animal to developing atherosclerotic
25 lesions, and wherein the transgenic animal has atherosclerotic lesions; (b) after a sufficient period of time, determining the severity of atherosclerotic lesions present in the transgenic animal; and (c) comparing the severity of atherosclerotic lesions determined in step (b)
30 with the severity of atherosclerotic lesions present in the same type of transgenic animal to which no agent was administered, whereby the agent is determined to be a

candidate agent for treating a human afflicted with atherosclerosis if the severity of atherosclerotic lesions determined in step (b) is less than the severity of lesions present in the transgenic animal to which no agent was administered.

This invention further provides a method for determining whether an agent is a candidate agent for treating a human afflicted with atherosclerosis, which comprises: (a) administering the agent to a transgenic animal whose somatic cells are characterized by the genotype comprising (i) an Apo-E knockout and (ii) a polymorphism for Adamtsl-1, wherein the polymorphism for Adamtsl-1 predisposes the animal to developing atherosclerotic lesions, and wherein the transgenic animal has atherosclerotic lesions; (b) after a sufficient period of time, determining the severity of atherosclerotic lesions present in the transgenic animal; and (c) comparing the severity of atherosclerotic lesions determined in step (b) with the severity of atherosclerotic lesions present in the same type of transgenic animal to which no agent was administered, whereby the agent is determined to be a candidate agent for treating a human afflicted with atherosclerosis if the severity of atherosclerotic lesions determined in step (b) is less than the severity of lesions present in the transgenic animal to which no agent was administered.

Finally, this invention provides a method for determining whether an agent is a candidate agent for treating a human afflicted with atherosclerosis, which comprises: (a) administering the agent to a transgenic animal whose somatic cells are characterized by the genotype comprising

a polymorphism for Adamtsl-1, wherein the polymorphism for Adamtsl-1 predisposes the animal to developing atherosclerotic lesions, and the transgenic animal has atherosclerotic lesions; (b) after a sufficient period of time, determining the severity of atherosclerotic lesions present in the transgenic animal; and (c) comparing the severity of atherosclerotic lesions determined in step (b) with the severity of atherosclerotic lesions present in the same type of transgenic animal to which no agent was administered, whereby the agent is determined to be a candidate agent for treating a human afflicted with atherosclerosis if the severity of atherosclerotic lesions determined in step (b) is less than the severity of lesions determined in the transgenic animal to which no agent was administered.

Brief Description of the Figures

Figure 1

Accelerated atherosclerotic lesion development in *Athsq1* congenic mice. (a) Genetic map of mouse chromosome 4 showing the *Athsq1* region. (b) B6-*Ldlr*^{-/-} mice carrying one (b/m, 43-81 cM) or two (m/m, 43-54 cM) copies of the MOLF-derived *Athsq1* interval exhibit increasing aortic root lesion area with increasing gene copy number compared to non-congenic littermate controls (b/b) after six-week WTD feeding. (c-k) Atherosclerotic lesions from the proximal aorta of B6-*Ldlr*^{-/-} mice carrying zero, one or two copies of the MOLF-derived *Athsq1* susceptibility locus and fed a Western-type diet for six weeks. (c-e) H&E stain; (f-h) Verhoeff's stain for elastin (black); (i-k) Masson's trichrome stain for collagen (blue). (c, f, i) Typical focal fatty streak lesions of *Athsq1* b/b non-congenic littermate controls. (d, g, j) More advanced fatty-fibrous lesions of *Athsq1* b/m heterozygotes. Some of the lesions have elastin- and collagen-containing fibrous caps (arrows) and large necrotic cores (NC). (e, h, k) Dramatically-advanced fibrous lesions of *Athsq1* m/m homozygotes. The lesion area is nearly continuous around the vessel wall with a thick fibrous cap (closed arrow), abundant intraplaque connective tissue (dotted arrow), and large necrotic cores (NC). (c, d, f, g, i, j) 100X magnification; (e, h, k) 40X magnification.

Figure 2

Genetic and physical mapping of *Athsq1*. (a) Localization to a <3.4 Mb interval of chromosome 4 in a set of overlapping congenic strains fed a WTD for twelve weeks.

Filled boxes indicate B6 alleles and open boxes indicate MOLF alleles. Positions of genetic markers are given in megabases (Mb) from the centromere as taken from *Ensembl*.

(b) Further localization to a <1.2 Mb interval using mice

5 fed a WTD for six weeks. Lesion area is in $\mu\text{m}^2/\text{section} \times 10^{-4} \pm \text{SEM}$ and necrotic core area is in $\mu\text{m}^2/\text{section} \times 10^{-3} \pm \text{SEM}$. NS, not significant. ND, not determined. (c)

Reduction of the *Athsq1*-containing interval to a 0.5 Mb interval by SNP analysis. Filled bars represent sub-

10 congenic intervals not conferring lesion susceptibility and the hatched bar represents extension of the sub-congenic interval conferring susceptibility. The recombination

region is indicated with dotted lines showing that only *Adamts11* and the unknown testis-specific EST *Tex30* are

15 within the smallest congenic interval associated with lesion susceptibility.

Figure 3

Expression of *Adamts11* and *Tex30*. (a) Tissue expression

20 pattern of *Adamts11* and (b) *Tex30* in a mouse cDNA panel normalized for housekeeping genes. Lanes (1) heart, (2) brain, (3) spleen, (4) lung, (5) liver, (6) skeletal muscle, (7) kidney, (8) testis, (9) 7-day embryo, (10) 11-day embryo, (11) 15-day embryo, (12) 17-day embryo. (c)

25 Presence of *Adamts11* and absence of *Tex30* in cDNA derived from atherosclerotic aorta following six-week WTD feeding.

The acidic ribosomal phosphoprotein P0 (*Arbp/36B4*) housekeeping gene serves as a positive control. Similar results were observed in multiple mouse aortas. (d)

30 Relative expression of *Adamts11* in thioglycollate-elicited peritoneal macrophages from *Athsq1* b/m and *Athsq1* m/m

congenics versus *Athsq1* b/b controls using quantitative PCR. (e) Relative expression of *Adamts11*, *CD68*, *Mac2*, *CD36*, and *Abca7* in atherosclerotic aorta following six-week WTD feeding in *Athsq1* m/m (n=8) congenics versus *Athsq1* b/b (n=5) controls using quantitative PCR. Data were normalized to the *Arbp* housekeeping gene.

Figure 4

Predicted domain structure of ADAMTSL1 and related murine proteins.

Figure 5

Increased abundance of versican in *Athsq1* congenic atherosclerotic plaques. Atherosclerotic lesions stained for versican (brown) from the proximal aorta of B6-*Ldlr*^{-/-} mice carrying zero or two copies of the MOLF-derived *Athsq1* susceptibility locus and fed a Western-type diet for (a-b) 3.5 weeks; (c-d) 6 weeks; (e-f) 12 weeks. (a, c, e) Absence of staining for versican in lesions of *Athsq1* b/b non-congenic littermate controls. (b, d, f) Accumulation of versican in lesions of *Athsq1* m/m homozygotes, most dramatic at the 6 week time point. Intense staining was observed in the media (M) underlying atherosclerotic plaques and to a lesser extent near the surface and within cellular regions of the plaque (Pl). 400X magnification.

Figure 6

Human *Adamts1-1* nucleotide (SEQ. ID. NO:1) and amino acid sequence (SEQ ID. NO:2).

Figure 7

(a) B6 (SEQ. ID. NO:3) and MOLF (SEQ. ID. NO:5) Adamtsl-1 nucleotide alignment. (b) B6 (SEQ. ID. NO:4) and MOLF (SEQ. ID. NO:6) amino acid alignment.

5

Figure 8

Effect of the *Athsq1* locus on lesion development in a uniform genetic background.

10 Figure 9

ADAMTSL1 protein expression. A. Full-length ADAMTSL1 (250 kDa) is increased in aortic extracts from B6-Ldlr/- congenic mice (m/m) compared to non-congenic controls (b/b). In contrast, no difference was observed in the levels of the known splice variant, punctin (67 kDa). The 75 kDa band corresponds to an unknown species present in aortic extracts but not macrophages. Briefly, 50 µg of protein were loaded on a 6% SDS-PAGE gel. Following overnight transfer, blots were incubated overnight with a polyclonal antibody raised against recombinant punctin at 1:250 dilution. B. Full-length ADAMTSL1 (250 kDa) is secreted by macrophages. Thioglycollate-elicited peritoneal macrophages were collected from B6-Ldlr/- mice and cultured for 24 hours in DMEM plus 0.1% FBS. Excess albumin was removed from the media by immunoprecipitation using an anti-BSA agarose matrix (Sigma). Western blotting was performed as for A. For both A and B, similar results were observed in three independent sets of samples.

30 Figure 10

Quantification of versican staining in atherosclerotic lesions derived from *Athsq1* congenic or non-congenic

littermates following six or twelve weeks of Western diet (WTD) feeding. B, B6-Ldlr^{-/-} mice homozygous for B6 alleles at the congenic interval; M, B6-Ldlr^{-/-} mice homozygous for MOLF alleles. Morphometric analysis was performed using ImagePro Plus (NIH).

Figure 11

Intact versican is increased in aortic homogenates from *Athsql* congenic mice carrying copies of the MOLF-derived susceptibility allele. Proximal aortas were collected from 6-week WTD-fed mice. Two aortas were pooled for each homogenate. Immunoblots were incubated with a polyclonal antibody raised against a fusion protein of the β GAG domain of the core mouse protein. 97,99A is a pool from non-congenic controls and 15,16A from congenics. Similar results were observed in three independent sets of samples.

Detailed Description of the Invention

Terms

5 "Adamtsl-1" is used herein to mean a disintegrin-like and metalloprotease with thrombospondin-like repeats-1 like-1. "Adamtsl-1" and "Adamtsl1" are used synonymously.

10 "Administering" an agent can be effected or performed using any of the various methods and delivery systems known to those skilled in the art. The administering can be performed, for example, intravenously, orally, nasally, via the cerebrospinal fluid, via implant, transmucosally, transdermally, intramuscularly, and subcutaneously. The
15 following delivery systems, which employ a number of routinely used pharmaceutically acceptable carriers, are only representative of the many embodiments envisioned for administering compositions according to the instant methods.

20

Injectable drug delivery systems include solutions (e.g., water or buffer), suspensions, gels, microspheres and polymeric injectables, and can comprise excipients such as solubility-altering agents (e.g., ethanol, propylene glycol
25 and sucrose) and polymers (e.g., polycaprylactones and PLGA's). Implantable systems include rods and discs, and can contain excipients such as PLGA and polycaprylactone.

Oral delivery systems include tablets and capsules. These
30 can contain excipients such as binders (e.g., hydroxypropylmethylcellulose, polyvinyl pyrrolidone, other cellulosic materials and starch), diluents (e.g., lactose

and other sugars, starch, dicalcium phosphate and cellulosic materials), disintegrating agents (e.g., starch polymers and cellulosic materials) and lubricating agents (e.g., stearates and talc).

5

Transmucosal delivery systems include patches, tablets, suppositories, pessaries, gels and creams, and can contain excipients such as solubilizers and enhancers (e.g., propylene glycol, bile salts and amino acids), and other
10 vehicles (e.g., polyethylene glycol, fatty acid esters and derivatives, and hydrophilic polymers such as hydroxypropylmethylcellulose and hyaluronic acid).

Dermal delivery systems include, for example, aqueous and
15 nonaqueous gels, creams, multiple emulsions, microemulsions, liposomes, ointments, aqueous and nonaqueous solutions, lotions, aerosols, hydrocarbon bases and powders, and can contain excipients such as solubilizers, permeation enhancers (e.g., fatty acids,
20 fatty acid esters, fatty alcohols and amino acids), and hydrophilic polymers (e.g., polycarbophil and polyvinylpyrrolidone). In one embodiment, the pharmaceutically acceptable carrier is a liposome or a transdermal enhancer.

25

Solutions, suspensions and powders for reconstitutable delivery systems include vehicles such as suspending agents (e.g., gums, zanthans, cellulosics and sugars), humectants (e.g., sorbitol), solubilizers (e.g., ethanol, water, PEG
30 and propylene glycol), surfactants (e.g., sodium lauryl sulfate, Spans, Tweens, and cetyl pyridine), preservatives and antioxidants (e.g., parabens, vitamins E and C, and

ascorbic acid), anti-caking agents, coating agents, and chelating agents (e.g., EDTA).

5 "Agent" shall mean any chemical entity, including, without limitation, a glycomer, a protein, an antibody, a lectin, a nucleic acid, a small molecule, and any combination thereof.

10 "Antibody" shall include, by way of example, both naturally occurring and non-naturally occurring antibodies. Specifically, this term includes polyclonal and monoclonal antibodies, and antigen-binding fragments (e.g., Fab fragments) thereof. Furthermore, this term includes chimeric antibodies (e.g., humanized antibodies) and wholly synthetic antibodies, and antigen-binding fragments thereof.

15 "Antisense nucleic acid" shall mean any nucleic acid which, when introduced into a cell (directly or via expression of another nucleic acid directly introduced into the cell), specifically hybridizes to at least a portion of an mRNA in
20 the cell encoding a protein (i.e., target protein) whose expression is to be inhibited, and thereby inhibits the target protein's expression.

25 "Bacterial cell" shall mean any bacterial cell. One example of a bacterial cell is *E. coli*.

30 "Candidate agent for treating a human afflicted with atherosclerosis" shall mean an agent which has shown some potential for treating a human afflicted with atherosclerosis (e.g., via the instant assays), and which is suitable for further animal tests and/or human tests wherein its efficacy for treating atherosclerosis is

determined.

"DNAzyme" shall mean a catalytic nucleic acid that is DNA or whose catalytic component is DNA, and which specifically recognizes and cleaves a distinct target nucleic acid sequence, which can be either DNA or RNA. Each DNAzyme has a catalytic component (also referred to as a "catalytic domain") and a target sequence-binding component consisting of two binding domains, one on either side of the catalytic domain.

"High cholesterol diet" includes, but is not limited to, a diet containing 21% (wt/wt) butterfat and 0.15% (wt/wt) cholesterol (Harlan Teklad Adjusted Calories TD 88137, Madison WI).

"Inhibiting" the onset of a disorder shall mean either lessening the likelihood of the disorder's onset, or preventing the onset of the disorder entirely. In the preferred embodiment, inhibiting the onset of a disorder means preventing its onset entirely.

"Isolated protein" (e.g., Adamtsl-1) shall mean a protein separated, in whole or in part, from other proteins. An isolated protein includes, for example, isolated protein obtained from cells, and isolated protein produced by chemical and/or recombinant synthesis. In one embodiment, an isolated protein (e.g., Adamtsl-1) is free of other protein. In another embodiment, an isolated protein constitutes at least half of the protein in a protein sample.

"Isolated nucleic acid" (e.g., Adamtsl-1-encoding nucleic acid) shall mean the nucleic acid free, in whole or in part,

from nucleic acid encoding other proteins. In one embodiment, isolated nucleic acid means the nucleic acid free of any other nucleic acid. A nucleic acid in a vector is considered isolated, where the nucleic acid is not naturally present in the vector. Isolated nucleic acid can be, for example, obtained from cells, or produced by chemical and/or recombinant synthesis.

"Mammalian cell" shall mean any mammalian cell. Mammalian cells include, without limitation, cells which are normal, abnormal and transformed, and are exemplified by neurons, epithelial cells, muscle cells, blood cells, immune cells, stem cells, osteocytes, endothelial cells and blast cells.

"Nucleic acid" shall mean any nucleic acid molecule, including, without limitation, DNA, RNA and hybrids thereof. The nucleic acid bases that form nucleic acid molecules can be the bases A, C, G, T and U, as well as derivatives thereof. Derivatives of these bases are well known in the art, and are exemplified in PCR Systems, Reagents and Consumables (Perkin Elmer Catalogue 1996-1997, Roche Molecular Systems, Inc., Branchburg, New Jersey, USA).

Examples of a "polymorphism for Adamtsl-1" include, but are not limited to the polymorphisms described in Table 1.

"Polypeptide" and "protein" are used interchangeably herein, and each means a polymer of amino acid residues. The amino acid residues can be naturally occurring or chemical analogues thereof. Polypeptides and proteins can also include modifications such as glycosylation, lipid attachment, sulfation, hydroxylation, and ADP-ribosylation.

"Prophylactically effective amount" means an amount sufficient to prevent, or reduce the likelihood of, the onset of a disorder or a complication associated with a disorder in a subject.

5

"Ribozyme" shall mean a catalytic nucleic acid molecule which is RNA or whose catalytic component is RNA, and which specifically recognizes and cleaves a distinct target nucleic acid sequence, which can be either DNA or RNA. Each ribozyme has a catalytic component (also referred to as a "catalytic domain") and a target sequence-binding component consisting of two binding domains, one on either side of the catalytic domain.

10

15 The "severity of atherosclerotic lesions" in a subject correlates with, and can be determined by comparing factors including, without limitation: atherosclerotic lesion area, necrotic core area, coloration, total number of lesions in the subject or an area and fibrous cap thickness.

20

"siRNA" shall mean small interfering ribonucleic acid. Methods of designing and producing siRNA to decrease the expression of a target protein are well known in the art.

25 "Subject" shall mean any animal, such as a human, non-human primate, mouse, rat, guinea pig or rabbit.

"Therapeutically effective amount" means an amount sufficient to treat a subject afflicted with a disorder or a complication associated with a disorder. The therapeutically effective amount will vary with the subject being treated, the condition to be treated, the agent

30

delivered and the route of delivery. A person of ordinary skill in the art can perform routine titration experiments to determine such an amount. Depending upon the agent delivered, the therapeutically effective amount of agent
5 can be delivered continuously, such as by continuous pump, or at periodic intervals (for example, on one or more separate occasions). Desired time intervals of multiple amounts of a particular agent can be determined without undue experimentation by one skilled in the art.

10

"Treating" a disorder shall mean slowing, stopping or reversing the disorder's progression. In the preferred embodiment, treating a disorder means reversing the disorder's progression, ideally to the point of eliminating
15 the disorder itself.

"Vector" shall mean a nucleic acid encoding a nucleic acid of interest and/or a protein of interest, which nucleic acid, when placed in a cell, permits the expression of the
20 nucleic acid or protein of interest. For example, a bacterial expression vector includes a promoter such as the lac promoter and for transcription initiation the Shine-Dalgarno sequence and the start codon AUG. Similarly, a eukaryotic expression vector includes a heterologous or
25 homologous promoter for RNA polymerase II, a downstream polyadenylation signal, the start codon AUG and a termination codon for detachment of the ribosome. Such vectors may be obtained commercially or assembled from the sequences described methods well-known in the art.

30

Embodiments of the Invention

This invention provides an isolated nucleic acid encoding Adamtsl-1 protein. In the preferred embodiment, the
5 Adamtsl-1 protein is human Adamtsl-1 protein. In another embodiment, the Adamtsl-1 protein is mouse Adamtsl-1 protein. In another embodiment, the nucleic acid encodes a protein having a sequence comprising the amino acid sequence set forth in SEQ. ID. NO:2. In another
10 embodiment, the nucleic acid encodes a protein having a sequence comprising the amino acid sequence set forth in SEQ. ID. NO:4 and SEQ. ID. NO:6. In another embodiment, the nucleic acid is DNA or RNA. In another embodiment, the DNA is cDNA.

15

This invention further provides an isolated Adamtsl-1 protein. In the preferred embodiment, the isolated Adamtsl-1 protein is human Adamtsl-1 protein. In another embodiment, the isolated Adamtsl-1 protein is mouse
20 Adamtsl-1 protein. In another embodiment, the isolated Adamtsl-1 protein has a sequence comprising the amino acid sequence set forth in SEQ. ID. NO:2. In another embodiment, the isolated Adamtsl-1 protein has a sequence comprising the amino acid sequence set forth in SEQ. ID.
25 NO:4 and SEQ. ID. NO:6.

This invention further provides an antibody that specifically binds to Adamtsl-1 protein. This invention further provides a vector comprising a nucleic acid
30 encoding Adamtsl-1. In one embodiment, the vector is adapted for expression of the nucleic acid in a cell and comprises regulatory elements necessary for the expression

of the nucleic acid in the cell operatively linked to the nucleic acid so as to permit expression thereof.

5 This invention further provides a cell comprising a vector, which vector comprises a nucleic acid encoding Adamtsl-1 protein. In one embodiment, the cell is bacterial, amphibian, yeast, fungal, insect or mammalian.

10 This invention further provides a method for inhibiting expression of Adamtsl-1 in a cell wherein Adamtsl-1 would otherwise occur comprising contacting the cell with an agent that inhibits expression of Adamtsl-1 in the cell. In one embodiment, the cell is a macrophage cell. In another embodiment, the agent is selected from the group
15 consisting of an antisense nucleic acid, siRNA, a DNase and a ribozyme.

This invention further provides a method for inhibiting the onset of atherosclerosis in a subject at risk for becoming
20 afflicted therewith, comprising administering to the subject a prophylactically effective amount of an agent that inhibits the expression of Adamtsl-1 in the subject's cells, thereby inhibiting the onset of atherosclerosis in the subject. In one embodiment, the cell is a macrophage
25 cell. In another embodiment, the agent is selected from the group consisting of an antisense nucleic acid, siRNA, a DNase, a ribozyme and an antibody that specifically binds to Adamtsl-1.

30 This invention further provides method of treating a subject afflicted with atherosclerosis comprising administering to the subject a therapeutically effective

amount of an agent that inhibits the expression of Adamtsl-1 in the subject's cells, thereby treating the subject. In one embodiment, the cell is a macrophage cell. In another embodiment, the agent is selected from the group consisting of an antisense nucleic acid, siRNA, a DNase, a ribozyme and an antibody that specifically binds to Adamtsl-1.

This invention further provides a transgenic animal whose somatic cells are characterized by a genotype comprising (i) an LDL-receptor knockout and (ii) a polymorphism for Adamtsl-1, wherein the polymorphism for Adamtsl-1 predisposes the animal to developing atherosclerotic lesions. In one embodiment, the animal is a mouse. In another embodiment, the mouse is fed a high-cholesterol diet.

This invention further provides a transgenic animal whose somatic cells are characterized by a genotype comprising (i) an Apo-E knockout and (ii) a polymorphism for Adamtsl-1, wherein the polymorphism for Adamtsl-1 predisposes the animal to developing atherosclerotic lesions. In one embodiment, the animal is a mouse. In another embodiment, the mouse is fed a high-cholesterol diet.

This invention further provides a transgenic animal whose somatic cells are characterized by a genotype comprising a polymorphism for Adamtsl-1, wherein the polymorphism for Adamtsl-1 predisposes the animal to developing atherosclerotic lesions. In one embodiment, the animal is a mouse. In another embodiment, the mouse is fed a high-cholesterol diet.

In the instant transgenic animals, in one embodiment, the Adamtsl-1 polymorphism is MOLF.

This invention further provides a method for determining whether an agent is a candidate agent for treating a human afflicted with atherosclerosis, which comprises: (a) administering the agent to a transgenic animal whose somatic cells are characterized by a genotype comprising (i) an LDL-receptor knockout and (ii) a polymorphism for Adamtsl-1, wherein the polymorphism for Adamtsl-1 predisposes the animal to developing atherosclerotic lesions, and wherein the transgenic animal has atherosclerotic lesions; (b) after a sufficient period of time, determining the severity of atherosclerotic lesions present in the transgenic animal; and (c) comparing the severity of atherosclerotic lesions determined in step (b) with the severity of atherosclerotic lesions present in the same type of transgenic animal to which no agent was administered, whereby the agent is determined to be a candidate agent for treating a human afflicted with atherosclerosis if the severity of atherosclerotic lesions determined in step (b) is less than the severity of lesions present in the transgenic animal to which no agent was administered. In one embodiment, the transgenic animal is a mouse. In another embodiment, the transgenic mouse is fed a high-cholesterol diet.

This invention further provides a method for determining whether an agent is a candidate agent for treating a human afflicted with atherosclerosis, which comprises: (a) administering the agent to a transgenic animal whose somatic cells are characterized by a genotype comprising

(i) an Apo-E knockout and (ii) a polymorphism for Adamtsl-1, wherein the polymorphism for Adamtsl-1 predisposes the animal to developing atherosclerotic lesions, and wherein the transgenic animal has atherosclerotic lesions; (b) after a sufficient period of time, determining the severity of atherosclerotic lesions present in the transgenic animal; and (c) comparing the severity of atherosclerotic lesions determined in step (b) with the severity of atherosclerotic lesions present in the same type of transgenic animal to which no agent was administered, whereby the agent is determined to be a candidate agent for treating a human afflicted with atherosclerosis if the severity of atherosclerotic lesions present in step (b) is less than the severity of lesions present in the transgenic animal to which no agent was administered. In one embodiment, the transgenic animal is a mouse. In another embodiment, the transgenic mouse is fed a high-cholesterol diet.

Finally, this invention provides a method for determining whether an agent is a candidate agent for treating a human afflicted with atherosclerosis, which comprises: (a) administering the agent to a transgenic animal whose somatic cells are characterized by a genotype comprising a polymorphism for Adamtsl-1, wherein the polymorphism for Adamtsl-1 predisposes the animal to developing atherosclerotic lesions, and wherein the transgenic animal has atherosclerotic lesions; (b) after a sufficient period of time, determining the severity of atherosclerotic lesions present in the transgenic animal; and (c) comparing the severity of atherosclerotic lesions determined in step (b) with the severity of atherosclerotic lesions present in

the same type of transgenic animal to which no agent was administered, whereby the agent is determined to be a candidate agent for treating a human afflicted with atherosclerosis if the severity of atherosclerotic lesions
5 determined in step (b) is less than the severity of lesions present in the transgenic animal to which no agent was administered. In one embodiment, the transgenic animal is a mouse. In another embodiment, the transgenic mouse is fed a high-cholesterol diet.

10

In the instant methods, in one embodiment, the Adamts1-1 polymorphism is MOLF.

15

This invention is illustrated in the Experimental Details section which follows. This section is set forth to aid in an understanding of the invention but is not intended to, and should not be construed to limit in any way the invention as set forth in the claims which follow thereafter.

20

25

30

Experimental Details

Introduction

5 Atherosclerosis is genetically complex and the genes contributing to inter-individual variation in susceptibility are largely unknown. Although several atherosclerosis loci have been mapped in the murine model, none of the underlying genes has been identified by
10 positional cloning. A locus affecting atherosclerotic lesion area, independent of plasma lipid levels, was previously mapped in a cross between strains MOLF/Ei (MOLF) and C57BL/6J(B6)-*Ldlr*^{-/-} [1]. Atherosclerosis susceptibility QTL1 (*Athsql*) was identified on chromosome 4
15 with susceptibility derived from the MOLF strain. In the current study, B6-*Ldlr*^{-/-} congenic mice carrying the MOLF-derived *Athsql* interval exhibited greatly accelerated lesion development in a gene dosage-dependent manner. Sub-congenic strain analysis reduced the *Athsql* interval to 0.5
20 Mb containing two candidate genes. Genetic sequence and expression analysis indicate that extracellular matrix protein *Adamts1-1* (a disintegrin-like and metalloprotease with thromospondin-like repeats-1 like-1) likely underlies *Athsql*. The proposed function of ADAMTS like proteins is
25 regulation of matrix remodeling through direct interaction with ADAMTS proteins and modulation of protease activity. This study demonstrates the feasibility of using complex trait analysis to positionally identify an atherosclerosis susceptibility gene and suggests that variants of *Adamts1-1*
30 have a major impact on early lesion development.

Materials and Methods

Mice

MOLF/Ei (MOLF) and B6.129S7-*Ldlr*^{tm1Her} (B6-*Ldlr*^{-/-}) were purchased from The Jackson Laboratory (Bar Harbor, Maine).
5 MOLF females were mated with B6-*Ldlr*^{-/-} males to produce F₁ mice. Female F₁s were backcrossed to B6-*Ldlr*^{-/-} males to produce N2 mice homozygous for the *Ldlr* knockout allele. The *Athsq1* b/m congenic strain, heterozygous for MOLF and B6 alleles at the *Athsq1* susceptibility locus, was created
10 through successive backcrossing to the B6-*Ldlr*^{-/-} strain. At each generation, the mice were typed for four microsatellite markers representing the proximal, medial, and distal portion of the introgressed interval. In order to speed up the breeding process, animals were screened for
15 ~100 markers to select for animals that inherited more of the B6-*Ldlr*^{-/-} background than others. By the N7 generation, all animals were 100% B6-*Ldlr*^{-/-} background with the exception of one sub-congenic line which was ~99.22% B6-*Ldlr*^{-/-} and 0.78% MOLF with no MOLF alleles on chromosome
20 6 (*Athsq2* [4]). *Athsq1* b/m mice carry an approximately 38 centiMorgan (cM) donor interval between *D4Mit185* (at 43 cM distal to the centromere as listed in the Mouse Genome Database, www.informatics.jax.org/) and *D4Mit42* (at 81 cM). The *Athsq1* m/m congenic strain, homozygous for MOLF alleles
25 at the *Athsq1* susceptibility locus, was created through intercrossing *Athsq1* b/m mice and selecting for those carrying 2 MOLF alleles at the locus. Due to difficulty in breeding mice which carried 2 copies of the distal part of the interval, six-week WTD-fed *Athsq1* m/m congenic mice

carry only an interval from *D4Mit185* to *D4Mit120* (at 54 cM).

All congenic mice were weaned onto standard laboratory chow
5 (PicoLab Rodent 20, #5053) at 21 days of age and switched to a Western-type diet at 8-12 weeks of age. The Western diet contained 21% (wt/wt) butterfat and 0.15% (wt/wt) cholesterol (Harlan Teklad Adjusted Calories TD 88137, Madison, WI). Mice were sacrificed after six weeks
10 or three months of Western diet feeding. The breeding colony was produced and maintained in a specific pathogen-free environment. All mice were given *ad libitum* access to food and water and maintained on a standard 12-h light-dark cycle throughout the study. All experimental protocols
15 were approved by the Institutional Animal Care and Research Advisory Committee.

DNA Extraction and Ldlr Genotyping

DNA was extracted from tail tips by a quick alkaline lysis protocol as previously described [4]. Primer sequences and
20 a protocol for *Ldlr* genotyping were obtained from [http://www.jax.org/resources/documents/imr/protocols/Ldlr K O.html](http://www.jax.org/resources/documents/imr/protocols/Ldlr%20.html) (8/5/1998).

Atherosclerotic Lesion Measurements

Anesthetized mice were killed by cervical dislocation. The
25 hearts were perfused with 0.9% NaCl alone (for frozen sections) or 0.9% NaCl followed by 10% formalin (for paraffin sections), and the aortic root was dissected and fixed in 10% formalin. Sequential, 10-micron (frozen) or 6-micron (paraffin) thick cross-sections were stained with

oil red O (frozen) or hematoxylin and eosin (H&E) (paraffin). Every tenth section, for a total of six sections, was quantified by video microscopy with Image Pro 3.0 Plus software (Media Cybernetics) and an average lesion size was determined for each mouse. Necrotic core area was determined from H&E stained sections as the lesion acellular area. Cap thickness was determined from Verhoeff's stained sections using a scoring system based on numbers of elastic layers as described previously [8]. Frozen sections were used for all twelve-week WTD data and paraffin sections for six-week data.

Sequence analysis

RNA was isolated from mouse tissues. First-strand cDNA was synthesized with oligo(dT)¹²⁻¹⁸ primer using SuperScript™ II RNase H⁻ reverse transcriptase (Invitrogen) according to the manufacturer's protocol. Direct bi-directional sequencing was performed on PCR products using an ABI Genetic Analyzer 3100 (Applied Biosystems). MOLF gene sequences were analyzed using Sequencher version 3.1 and compared to B6 and published sequence (Celera and Ensembl). Gel electrophoresis was used to confirm the size difference caused by 5769 del45nt in the MOLF strain. A fragment from the 3'UTR of *Adamts11* was amplified by PCR, using the following primers: 5'-primer (5'-GGACAGCAAGGAGTTTATAACACATA-3') (SEQ. ID. NO:7); 3'-primer (5'-TTTCTTGGTCTTCCCACACC-3') (SEQ. ID. NO:8) and DNA fragments were analyzed on a 3% agarose gel (124bp and 169bp for MOLF and B6 respectively).

Tissue Expression Analysis

RT-PCR was performed using Mouse Multiple Tissue cDNA Panel I according to the manufacturer's instructions (Clontech) or cDNA prepared from proximal aorta homogenates. Amplification of *Tex30* was carried out with the following primers: 5'-primer (5'- AAATGTTACCGGTGACAT-3') (SEQ. ID. NO:9); 3'-primer (5'- TATTTCTCTGCCTCCTTA-3') (SEQ. ID. NO:10) and *Adamts11* with the following primers: 5'-primer (5'- CTCCCACCACCTGCATGA-3') (SEQ. ID. NO:11); 3'-primer (5'-CCTTGAGGTCCTCCACTGAG-3') (SEQ. ID. NO:12). The expected fragment sizes were 441bp for *Tex30* and 408bp for *Adamts11*.

15 *Quantitative PCR*

Peritoneal macrophages were harvested 4 days after thioglycollate injection, centrifuged at 1000 rpm, and plated with DMEM+10%FBS. Cells were washed 3x with PBS and total RNA was extracted with the use of RNeasy mini kit (QIAGEN) according to the manufacturer's instructions. First strand cDNA synthesis was carried out as described above. *Adamts11* expression was determined with Assays-on-Demand Gene Expression product Mm.00553186 (Applied Biosystems) according to manufacturer's instruction using the Stratagene Mx4000 Multiplex Quantitative PCR System. Assays were performed in duplicate and normalized to the acidic ribosomal phosphoprotein PO (*Arbp/36B4*). Primer and probe sequences for *Arbp* were as follows: forward, 5'- AGATGCAGCAGATCCGCAT-3'; reverse, 5'-GTTCTTGCCCATCAGCACC-3'; and probe, 5'-JOE-CGCTCCGAGGGAAGGCCG-3'.

Statistical Analysis

ANOVA or Mann Whitney (if unequal variance) was performed using STATVIEW 5.0 (Abacus Concepts, Inc.).

5

Table 1A. *Athsql* candidate sequence variants identified in MOLF relative to the B6 mouse strain.

Gene	Nucleotide	Amino Acid
<i>Adamts1-1</i>	396 C>T	105 L>F
	820 T>C	246 I>T
	2859 A>G	926 T>A
	5769 del45nt	3' UTR
	6042 del15nt	3' UTR
	6624 ins8nt	3' UTR
	6647 ins8nt	3' UTR
<i>mCG1697</i>	798 C>A	180 T>N
	1497 A>G	413 S>N

10 Results

A locus affecting atherosclerotic lesion area, independent of traditional risk factors, was previously mapped in a cross between strains MOLF/Ei (MOLF) and C57BL/6J(B6)-*Ldlr*^{-/-} [4]. Atherosclerosis susceptibility QTL1 (*Athsql*) was identified on chromosome 4 with susceptibility derived from the MOLF strain. To confirm the *Athsql* susceptibility locus and test its amenability to fine mapping, congenic mice carrying one or two copies of MOLF chromosome 4 between *D4Mit185* (43 cM) and *D4Mit120* (54 cM) or *D4Mit42* (81 cM) on a B6-*Ldlr*^{-/-} background were created (Fig. 1a). After six weeks of Western-type diet (WTD) feeding, mean atherosclerotic lesion area was significantly greater in

Athsq1 b/m heterozygotes ($93,000 \pm 9,000 \mu\text{m}^2/\text{section}$, mean $\pm$ SEM) and *Athsq1* m/m homozygotes ($220,000 \pm 35,000 \mu\text{m}^2/\text{section}$) compared to non-congenic *Athsq1* b/b controls ($53,000 \pm 5,000 \mu\text{m}^2/\text{section}$; $p < 0.0003$ and $p < 0.0001$, respectively) (Fig. 1b). The difference in lesion areas was independent of plasma cholesterol levels (data not shown). Although the effect of *Athsq1* was female-specific in the original cross [4], the locus affected lesion development in both sexes in the uniform genetic background (Fig. 8).

While non-congenic littermate controls exhibited small, focal, fatty streak lesions (Fig. 1c, f, and i), *Athsq1* b/m heterozygotes developed both fatty streak lesions and more advanced fatty-fibrous lesions characterized by fibrous cap and necrotic core formation (Fig. 1d, g, and j). In stark contrast to *Athsq1* b/b controls, *Athsq1* m/m homozygotes exhibited advanced fibrous lesions, often covering the entire circumference of the vessel wall (Fig. 1e, h, and k). Mean necrotic core area was significantly greater in congenic mice ($10,000 \pm 4,000 \mu\text{m}^2/\text{section}$ for *Athsq1* b/m and $32,000 \pm 11,000 \mu\text{m}^2/\text{section}$ for *Athsq1* m/m) compared to non-congenic controls ($300 \pm 1200 \mu\text{m}^2/\text{section}$; $p < 0.04$ and $p < 0.0006$, respectively). The most advanced lesions from control mice exhibited thin (single elastic layer) fibrous caps while advanced lesions from *Athsq1* b/m and *Athsq1* m/m mice exhibited intermediate (two to four elastic layers) or thick (greater than four elastic layers) fibrous caps (trend for heterozygotes versus controls and $p < 0.0002$ for homozygotes versus controls) [8]. Thus, the congenic mice

exhibited greatly accelerated atherosclerotic lesion development in a gene dosage-dependent manner.

For fine mapping, nine congenic/sub-congenic strains
5 carrying varying lengths of the MOLF-derived chromosome 4
interval were generated (Fig 2). Atherosclerotic lesion
development was compared with non-congenic littermate
controls to determine whether or not a particular strain
carried the *Athsq1* variant. Using the twelve-week WTD
10 feeding protocol employed in the original mapping study
[4], no difference in lesion area was observed between
homozygous congenics carrying a region extending from
D4Mit185 to *D4Mit42* and those carrying a shorter region
from *D4Mit185* to *D4Mit120* (Fig. 2a). Furthermore, analysis
15 of four sub-congenic strains, heterozygous for MOLF
alleles, confirmed this observation and narrowed the *Athsq1*
containing interval to a <3.4 Mb region between *D4Mit152*
and *Adfp-rs* (Fig. 2a). While this region lies proximal to
the original linkage peak, sub-congenic strain analysis
20 strongly supports the presence of a single gene affecting
atherosclerosis susceptibility localizing to this sub-
region. However, the possibility that the effect of a
second, distally located gene was lost in the uniform
genetic background cannot be ruled out. Following a pilot
25 study, a six-week feeding regimen was used for further
mapping studies. Comparison of two additional heterozygous
sub-congenic strains to non-congenic littermates reduced
the interval to a <1.2 Mb region between *D4Mit302* and *Adfp-rs*
(Fig. 2b). To narrow the region, single nucleotide
30 polymorphisms (SNPs) were identified between the B6 and
MOLF strains by comparative sequence analysis. Sequencing
of SNPs in critical sub-congenics narrowed *Athsq1* to an

approximate 0.5 Mb region between SNPs *Adamts1*-479 and *Rraga-22del1nt* (Fig. 2c). This interval included only two expressed sequences: *Adamts11* and an unknown EST (Celera *Tex30*; RIKEN 4930500009) which was designated as *Tex30*.

5

To identify candidate sequence variant(s) underlying *Athsql*, the 5'UTR, coding sequence, and 3'UTR of *Tex30* and *Adamts11* derived from the MOLF strain were sequenced and compared those sequences to B6. Two non-conservative amino acid variants for *Tex30* and three variants for *Adamts11* were detected (Table 1B).

10

Table 1B. *Athsql* candidate sequence variants identified in MOLF relative to the B6 mouse strain.

Gene	Nucleotide	Amino Acid
<i>Adamts11</i>	396 C>T	105 L>F
	820 T>C	246 I>T
	3042 A>G	987 T>A
	5769 del45nt	3'UTR
	6042 del15nt	3'UTR
	6624 ins8nt	3'UTR
	6647 ins7nt	3'UTR
<i>Tex30</i>	798 C>A	180 T>N
	1497 A>G	413 S>N

15

In addition, multiple deletion/insertion polymorphisms were detected in the 3'UTR of *Adamts11* in MOLF versus B6 (Table 1B). Fragments spanning each of the two deletions were amplified and sequenced from MOLF genomic DNA, ruling out the possibility that these polymorphisms represented splice variants as opposed to deletions.

20

Next, the tissue expression patterns of *Tex30* and *Adamts11* were compared by RT-PCR. *Adamts11* was widely expressed in adult as well as embryonic tissues (Fig. 3a). In contrast, *Tex30* was expressed in a highly testis-specific manner (Fig. 3b). *Adamts11*, but not *Tex30*, mRNA was also detected in atherosclerotic lesions (Fig. 3c). The testis specific expression of *Tex30* is supported by data in the corresponding GenBank UniGene Cluster (Mm.56451) where cDNA sources included only testis and round spermatids. Because *Tex30* is not expressed in tissues relevant to atherogenesis, it is unlikely to be the gene underlying *Athsq1*. Therefore, further investigation focused on *Adamts11*.

Using a probe-based quantitative PCR assay, a significant induction of *Adamts11* mRNA was observed in lesion-containing aorta compared to unaffected vessels (32-fold, $p=0.005$). *Adamts11* levels were approximately 4-fold (*Athsq1* b/m) and 5.5-fold (*Athsq1* m/m) higher in isolated peritoneal macrophages derived from congenics versus controls (Fig. 4d, $p<0.006$ and $p<0.001$, respectively). In addition, *Athsq1* m/m mice showed greater expression of *Adamts11* mRNA in lesion-containing aorta compared to *Athsq1* b/b mice (Fig. 3e, $p=0.001$) while the levels of other genes expressed by macrophages were not different (Fig. 3e). Although functional promoter variants between MOLF and B6 have been excluded by sub-congenic SNP analysis (Fig. 2c), the altered expression level of *Adamts11* may be due to an intronic regulatory variant or an effect of one or more of the 3' UTR polymorphisms on transcript stability. These data suggest that increased expression of *Adamts11* in

lesional macrophages may contribute to accelerated atherogenesis.

Discussion

5 *Adamts11* is an extracellular matrix (ECM) glycoprotein that is a member of a non-enzymatic sub-family related to the *Adamts* metalloprotease family. The complete coding sequence of *Adamts11* has not been previously published.

10 However, an N-terminal fragment termed punctin has been described [9]. A 7.2 kb transcript was sequenced, encoding 30 exons, which predicts a protein with structural homology to other family members, including ADAMTSL3 [10] and papilin [11] (Fig. 4). The cDNA sequence has been
15 deposited in Genbank. A conserved cassette of domains (papilin cassette) at the 5' end of the ADAMTS-like molecules is homologous to a non-catalytic cassette of domains at the 3' end of ADAMTS molecules (Fig. 4). Papilin, which has been described in some detail in
20 *Drosophila*, is an ECM glycoprotein secreted by macrophage-like phagocytes that binds ADAMTS protease and possibly modulates protease activity and local matrix composition [11-13]. Homology between ADAMTSL1 and papilin likely reflects conservation of function between the two proteins
25 and suggests that ADAMTSL1 may also modulate the activity of ADAMTS metalloproteases [9].

Currently, the role of ADAMTS proteases in atherogenesis is unknown but they have been implicated in the processing of procollagen [14], cleavage of proteoglycans [15-19],
30 inflammation [20], and cell migration [21, 22]. Specifically, cleavage of versican, a proteoglycan implicated in atherogenesis [23, 24], can occur via

multiple ADAMTS protease family members [16]. Thus, lesions derived from congenic and non-congenic mice were stained using a specific versican antibody. *Athsq1* congenic mice exhibited a dramatic accumulation of versican in lesions while non-congenic littermates exhibited virtually no accumulation (Fig. 5). These data are in agreement with a previous report that versican is nearly absent from early through advanced lesions of B6-*Ldlr*^{-/-} and B6-*apoE*^{-/-} mice [25]. In contrast, versican is prominent in human lesions and has been proposed to increase atherosclerosis by a variety of mechanisms [23, 24]. Versican accumulation in *Athsq1* congenic mice was most prominent at the six-week diet time point when differences in lesion development were most pronounced between congenics and controls. Staining for collagen, hyaluronan, and total proteoglycans showed no differences in comparable lesions derived from the two strains (data not shown), suggesting a specific role for versican. Hence, increased ADAMTSL1 activity may inhibit one or more ADAMTS family members responsible for proteolysis of aortic versican, providing a local niche for intimal expansion and lesion progression.

While a number of atherosclerosis susceptibility loci have been mapped in the mouse [3-7], none of the underlying genes have been previously identified by positional cloning. In this study, genetic exclusion narrowed the *Athsq1* interval to two genes and gene expression patterns were used to identify the underlying gene. Although the precise molecular actions of *Adamts11* in atherogenesis remains to be elucidated, this study provides a more

rigorous method of complex gene identification than
candidate gene selection [26].

5

10

15

20

25

30

References

1. Motulsky AG, B.J., *Genetics of coronary artery disease*, in *The Genetic Basis of Common Disease.*, R.J. King RA, Motulsky AG, Editor. 2002, Oxford University Press: New York, NY. p. 105-126.
5
2. Lusis, A.J., A.M. Fogelman, and G.C. Fonarow. Genetic Basis of Atherosclerosis: Part I: New Genes and Pathways. *Circulation*. 2004;110:1868-1873.
10
3. Dansky, H.M., P. Shu, M. Donavan, J. Montagno, D.L. Nagle, J.S. Smutko, N. Roy, S. Whiteing, J. Barrios, T.J. McBride, J.D. Smith, G. Duyk, J.L. Breslow, and K.J. Moore. A Phenotype-Sensitizing Apoe-Deficient Genetic Background Reveals Novel Atherosclerosis Predisposition Loci in the Mouse. *Genetics*. 2002;160:1599-1608.
15
4. Welch, C.L., S. Bretschger, N. Latib, M. Bezouevski, Y. Guo, N. Pleskac, C.-P. Liang, C. Barlow, H. Dansky, J.L. Breslow, and A.R. Tall. Localization of atherosclerosis susceptibility loci to chromosomes 4 and 6 using the Ldlr knockout mouse model. *PNAS*. 2001;98:7946-7951.
20
25
5. Mehrabian, M., J. Wong, X. Wang, Z. Jiang, W. Shi, A.M. Fogelman, and A.J. Lusis. Genetic Locus in Mice That Blocks Development of Atherosclerosis Despite Extreme Hyperlipidemia. *Circ Res*. 2001;89:125-130.
30

6. Ishimori, N., R.H. Li, P.M. Kelmenson, R. Korstanje, K.A. Walsh, G.A. Churchill, K. Forsman-Semb, and B. Paigen. Quantitative trait loci analysis for plasma HDL-cholesterol concentrations and atherosclerosis susceptibility between inbred mouse strains C57BL/6J and 129S1/SvImJ. *Arteriosclerosis Thrombosis and Vascular Biology*. 2004;24:161-166.
7. Paigen, B., D. Mitchell, K. Reue, A. Morrow, A.J. Lusis, and R.C. Leboeuf. Ath-1, a Gene Determining Atherosclerosis Susceptibility and High-Density-Lipoprotein Levels in Mice. *Proceedings of the National Academy of Sciences of the United States of America*. 1987;84:3763-3767.
8. Kozaki, K., W.E. Kaminski, J. Tang, S. Hollenbach, P. Lindahl, C. Sullivan, J.-C. Yu, K. Abe, P.J. Martin, R. Ross, C. Betsholtz, N.A. Giese, and E.W. Raines. Blockade of Platelet-Derived Growth Factor or Its Receptors Transiently Delays but Does Not Prevent Fibrous Cap Formation in ApoE Null Mice. *Am J Pathol*. 2002;161:1395-1407.
9. Hirohata, S., L.W. Wang, M. Miyagi, L. Yan, M.F. Seldin, D.R. Keene, J.W. Crabb, and S.S. Apte. Punctin, a Novel ADAMTS-like Molecule, ADAMTSL-1, in Extracellular Matrix. *J. Biol. Chem*. 2002;277:12182-12189.
10. Hall, N.G., P. Klenotic, B. Anand-Apte, and S.S. Apte. ADAMTSL-3/punctin-2, a novel glycoprotein in

extracellular matrix related to the ADAMTS family of metalloproteases*1. *Matrix Biology*. 2003;22:501-510.

11. Kramerova, I., N. Kawaguchi, L. Fessler, R. Nelson, Y. Chen, A. Kramerov, M. Kusche-Gullberg, J. Kramer, B. Ackley, A. Sieron, D. Prockop, and J. Fessler. Papilin in development; a pericellular protein with a homology to the ADAMTS metalloproteinases. *Development*. 2000;127:5475-5485.

12. Hortsch M, O.A., Fishman S, Soneral SN, Marikar Y, Dong R, Jacobs JR. The expression of MDP-1, a component of Drosophila embryonic basement membranes, is modulated by apoptotic cell death. *Int J Dev Biol*. 1998;42:33-42.

13. Fessler, J.H., I. Kramerova, A. Kramerov, Y. Chen, and L.I. Fessler. Papilin, a novel component of basement membranes, in relation to ADAMTS metalloproteases and ECM development. *The International Journal of Biochemistry & Cell Biology*. 2004;36:1079-1084.

14. Colige, A., A. Beschin, B. Samyn, Y. Goebels, J. Van Beeumen, B.V. Nusgens, and C.M. Lapiere. Characterization and partial amino acid sequencing of a 107-kDa procollagen I N-proteinase purified by affinity chromatography on immobilized type XIV collagen. *The Journal Of Biological Chemistry*. 1995;270:16724-16730.

15. Tortorella, M.D., T.C. Burn, M.A. Pratta, I. Abbaszade, J.M. Hollis, R. Liu, S.A. Rosenfeld, R.A.

Copeland, C.P. Decicco, and a. Wynn et. Purification and cloning of aggrecanase-1: a member of the ADAMTS family of proteins. *Science*. 1999;284:1664-1666.

- 5 16. Sandy, J.D., J. Westling, R.D. Kenagy, M.L. Iruela-Arispe, C. Verscharen, J.C. Rodriguez-Mazaneque, D.R. Zimmermann, J.M. Lemire, J.W. Fischer, T.N. Wight, and A.W. Clowes. Versican V1 Proteolysis in Human Aorta in Vivo Occurs at the Glu441-Ala442 Bond, a Site That Is
10 Cleaved by Recombinant ADAMTS-1 and ADAMTS-4. *J. Biol. Chem.* 2001;276:13372-13378.
17. Russell, D.L., K.M.H. Doyle, S.A. Ochsner, J.D. Sandy, and J.S. Richards. Processing and Localization of
15 ADAMTS-1 and Proteolytic Cleavage of Versican during Cumulus Matrix Expansion and Ovulation. *J. Biol. Chem.* 2003;278:42330-42339.
18. Matthews, R.T., S.C. Gary, C. Zerillo, M. Pratta, K.
20 Solomon, E.C. Arner, and S. Hockfield. Brain-enriched Hyaluronan Binding (BEHAB)/Brevican Cleavage in a Glioma Cell Line Is Mediated by a Disintegrin and Metalloproteinase with Thrombospondin Motifs (ADAMTS) Family Member. *J. Biol. Chem.* 2000;275:22695-22703.
- 25 19. Abbaszade, I., R.Q. Liu, F. Yang, S.A. Rosenfeld, O.H. Ross, J.R. Link, D.M. Ellis, M.D. Tortorella, M.A. Pratta, and a. Hollis et. Cloning and characterization of ADAMTS11, an aggrecanase from the ADAMTS family.
30 *The Journal Of Biological Chemistry*. 1999;274:23443-23450.

20. Kuno, K., N. Kanada, E. Nakashima, F. Fujiki, F. Ichimura, and K. Matsushima. Molecular cloning of a gene encoding a new type of metalloproteinase-disintegrin family protein with thrombospondin motifs as an inflammation associated gene. *Journal of Biological Chemistry*. 1997;272:556-562.
21. Blelloch, R., S.S. Anna-Arriola, D. Gao, Y. Li, J. Hodgkin, and J. Kimble. The gon-1 Gene Is Required for Gonadal Morphogenesis in *Caenorhabditis elegans**1. *Developmental Biology*. 1999;216:382-393.
22. Rao, C., D. Foernzler, S.K. Loftus, S. Liu, J.D. McPherson, K.A. Jungers, S.S. Apte, W.J. Pavan, and D.R. Beier. A defect in a novel ADAMTS family member is the cause of the belted white-spotting mutation. *Development (Cambridge, England)*. 2003;130:4665-4672.
23. Wight, T.N. and M.J. Merrilees. Proteoglycans in Atherosclerosis and Restenosis: Key Roles for Versican. *Circ Res*. 2004;94:1158-1167.
24. Williams, K.J. Arterial wall chondroitin sulfate proteoglycans: diverse molecules with distinct roles in lipoprotein retention and atherogenesis. *Current Opinion in Lipidology*. 2001;12:477-487.
25. Kunjathoor, V.V., D.S. Chiu, K.D. O'Brien, and R.C. LeBoeuf. Accumulation of Biglycan and Perlecan, but Not Versican, in Lesions of Murine Models of Atherosclerosis. *Arterioscler Thromb Vasc Biol*. 2002;22:462-468.

26. Glazier, A.M., J.H. Nadeau, and T.J. Aitman. Finding Genes That Underlie Complex Traits. *Science*. 2002;298:2345-2349.

5

10

15

20

25

30

What is claimed is:

1. An isolated nucleic acid encoding Adamtsl-1 protein.

5 2. The isolated nucleic acid of claim 1, wherein the Adamtsl-1 protein is human Adamtsl-1 protein.

3. The isolated nucleic acid of claim 1, wherein the Adamtsl-1 protein is mouse Adamtsl-1 protein.

10

4. The isolated nucleic acid of claim 2, wherein the nucleic acid encodes a protein having a sequence comprising the amino acid sequence set forth in SEQ. ID. NO:2.

15

5. The isolated nucleic acid of claim 3, wherein the nucleic acid encodes a protein having a sequence comprising the amino acid sequence set forth in SEQ. ID. NO:4.

20

6. The isolated nucleic acid of claim 1, wherein the nucleic acid is DNA or RNA.

25

7. The isolated nucleic acid of claim 6, wherein the DNA is cDNA.

8. An isolated Adamtsl-1 protein.

30

9. The isolated Adamtsl-1 protein of claim 8, wherein the Adamtsl-1 protein is human Adamtsl-1 protein.

10. The isolated Adamtsl-1 protein of claim 8, wherein the

Adamtsl-1 protein is mouse Adamtsl-1 protein.

11. The isolated Adamtsl-1 protein of claim 9, wherein the
protein has a sequence comprising the amino acid
5 sequence set forth in SEQ. ID. NO:2.

12. The isolated Adamtsl-1 protein of claim 10, wherein
the protein has a sequence comprising the amino acid
sequence set forth in SEQ. ID. NO:4.

10 13. An antibody which specifically binds to Adamtsl-1
protein.

14. A vector comprising a nucleic acid encoding Adamtsl-1
15 protein.

15. The vector of claim 14, wherein the vector is adapted
for expression of the nucleic acid in a cell and
comprises regulatory elements necessary for the
20 expression of the nucleic acid in the cell operatively
linked to the nucleic acid so as to permit expression
thereof.

16. A cell comprising a vector, which vector comprises a
25 nucleic acid encoding Adamtsl-1 protein.

17. The cell of claim 16, wherein the cell is a bacterial,
amphibian, yeast, fungal, insect, or mammalian cell.

30 18. A method for inhibiting expression of Adamtsl-1 in a
cell wherein Adamtsl-1 expression would otherwise
occur comprising contacting the cell with an agent

that inhibits expression of Adamtsl-1 in the cell.

19. The method of claim 18, wherein the cell is a macrophage cell.

5

20. The method of claim 18, wherein the agent is selected from the group consisting of an antisense nucleic acid, siRNA, a DNAzyme and a ribozyme.

10 21. A method for inhibiting the onset of atherosclerosis in a subject at risk for becoming afflicted therewith, comprising administering to the subject a prophylactically effective amount of an agent that inhibits the expression of Adamtsl-1 in the subject's
15 cells, thereby inhibiting the onset of atherosclerosis in the subject.

22. The method of claim 21, wherein the cells are macrophage cells.

20

23. The method of claim 21, wherein the agent is selected from the group consisting of an antisense nucleic acid, siRNA, a DNAzyme, a ribozyme and an antibody that specifically binds to Adamtsl-1.

25

24. A method of treating a subject afflicted with atherosclerosis comprising administering to the subject a therapeutically effective amount of an agent that inhibits the expression of Adamtsl-1 in the
30 subject's cells, thereby treating the subject.

25. The method of claim 24, wherein the cells are

macrophage cells.

26. The method of claim 24, wherein the agent is selected from the group consisting of an antisense nucleic acid, siRNA, a DNase, a ribozyme and an antibody that specifically binds to Adamtsl-1.

27. A transgenic animal whose somatic cells are characterized by a genotype comprising (i) an LDL-receptor knockout and (ii) a polymorphism for Adamtsl-1, wherein the polymorphism for Adamtsl-1 predisposes the animal to developing atherosclerotic lesions.

28. The transgenic animal of claim 27, wherein the animal is a mouse.

29. The transgenic animal of claim 28, wherein the transgenic mouse is fed a high-cholesterol diet.

30. A transgenic animal whose somatic cells are characterized by a genotype comprising (i) an Apo-E knockout and (ii) a polymorphism for Adamtsl-1, wherein the polymorphism for Adamtsl-1 predisposes the animal to developing atherosclerotic lesions.

31. The transgenic animal of claim 30, wherein the animal is a mouse.

32. The transgenic animal of claim 31, wherein the transgenic mouse is fed a high-cholesterol diet.

33. A transgenic animal whose somatic cells are

characterized by a genotype comprising a polymorphism for Adamtsl-1, wherein the polymorphism for Adamtsl-1 predisposes the animal to developing atherosclerotic lesions.

5

34. The transgenic animal of claim 33, wherein the animal is a mouse.

35. The transgenic animal of claim 34, wherein the transgenic mouse is fed a high-cholesterol diet.

10

36. A method for determining whether an agent is a candidate agent for treating a human afflicted with atherosclerosis, which comprises:

15

(a) administering the agent to a transgenic animal whose somatic cells are characterized by a genotype comprising (i) an LDL-receptor knockout and (ii) a polymorphism for Adamtsl-1, wherein the polymorphism for Adamtsl-1 predisposes the animal to developing atherosclerotic lesions, and wherein the transgenic animal has atherosclerotic lesions;

20

(b) after a sufficient period of time, determining the severity of atherosclerotic lesions present in the transgenic animal; and

25

(c) comparing the severity of atherosclerotic lesions determined in step (b) with the

30

severity of atherosclerotic lesions present in the same type of transgenic animal to which no agent was administered, whereby the agent is determined to be a candidate agent for treating a human afflicted with atherosclerosis if the severity of atherosclerotic lesions determined in step (b) is less than the severity of lesions present in the transgenic animal to which no agent was administered.

37. The method of claim 36, wherein the transgenic animal is a mouse.

38. The method of claim 37, wherein the transgenic mouse is fed a high-cholesterol diet.

39. A method for determining whether an agent is a candidate agent for treating a human afflicted with atherosclerosis, which comprises:

(a) administering the agent to a transgenic animal whose somatic cells are characterized by a genotype comprising (i) an Apo-E knockout and (ii) a polymorphism for Adamtsl-1, wherein the polymorphism for Adamtsl-1 predisposes the animal to developing atherosclerotic lesions, and wherein the transgenic animal has atherosclerotic lesions;

(b) after a sufficient period of time, determining the severity of atherosclerotic lesions present

in the transgenic animal; and

- (c) comparing the severity of atherosclerotic lesions determined in step (b) with the severity of atherosclerotic lesions present in the same type of transgenic animal to which no agent was administered, whereby the agent is determined to be a candidate agent for treating a human afflicted with atherosclerosis if the severity of atherosclerotic lesions present in step (b) is less than the severity of lesions present in the transgenic animal to which no agent was administered.

40. The method of claim 39, wherein the transgenic animal is a mouse.

41. The method of claim 40, wherein the transgenic mouse is fed a high-cholesterol diet.

42. A method for determining whether an agent is a candidate agent for treating a human afflicted with atherosclerosis, which comprises:

- (a) administering the agent to a transgenic animal whose somatic cells are characterized by a genotype comprising a polymorphism for Adamtsl-1, wherein the polymorphism for Adamtsl-1 predisposes the animal to developing atherosclerotic lesions, and wherein the transgenic animal has atherosclerotic lesions;

(b) after a sufficient period of time, determining the severity of atherosclerotic lesions present in the transgenic animal; and

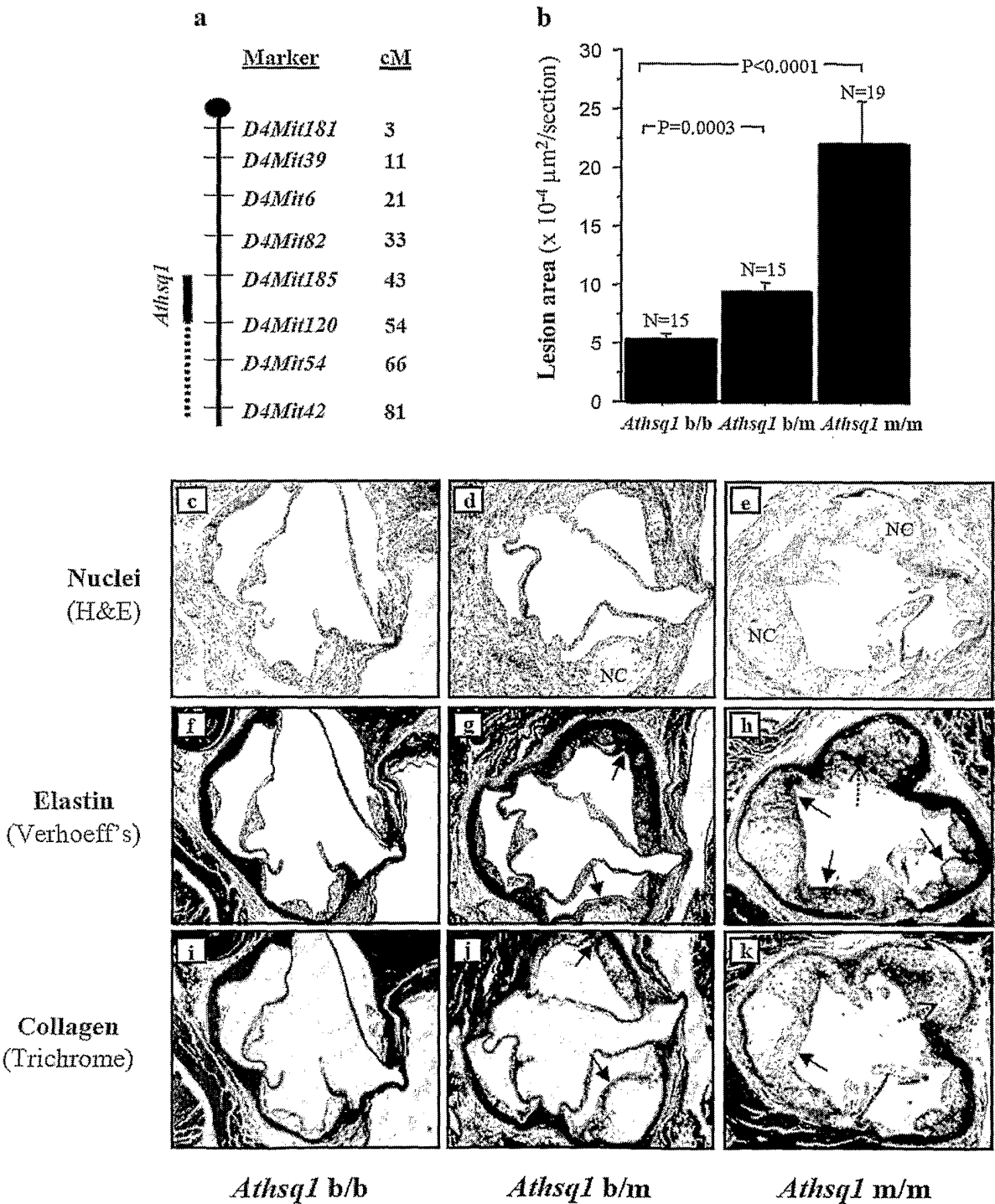
5 (c) comparing the severity of atherosclerotic lesions determined in step (b) with the severity of atherosclerotic lesions present in the same type of transgenic animal to which no agent was administered, whereby the agent is
10 determined to be a candidate agent for treating a human afflicted with atherosclerosis if the severity of atherosclerotic lesions determined in step (b) is less than the severity of lesions present in the transgenic animal to
15 which no agent was administered.

43. The method of claim 42, wherein the transgenic animal is a mouse.

20 44. The method of claim 43, wherein the transgenic mouse is fed a high-cholesterol diet.

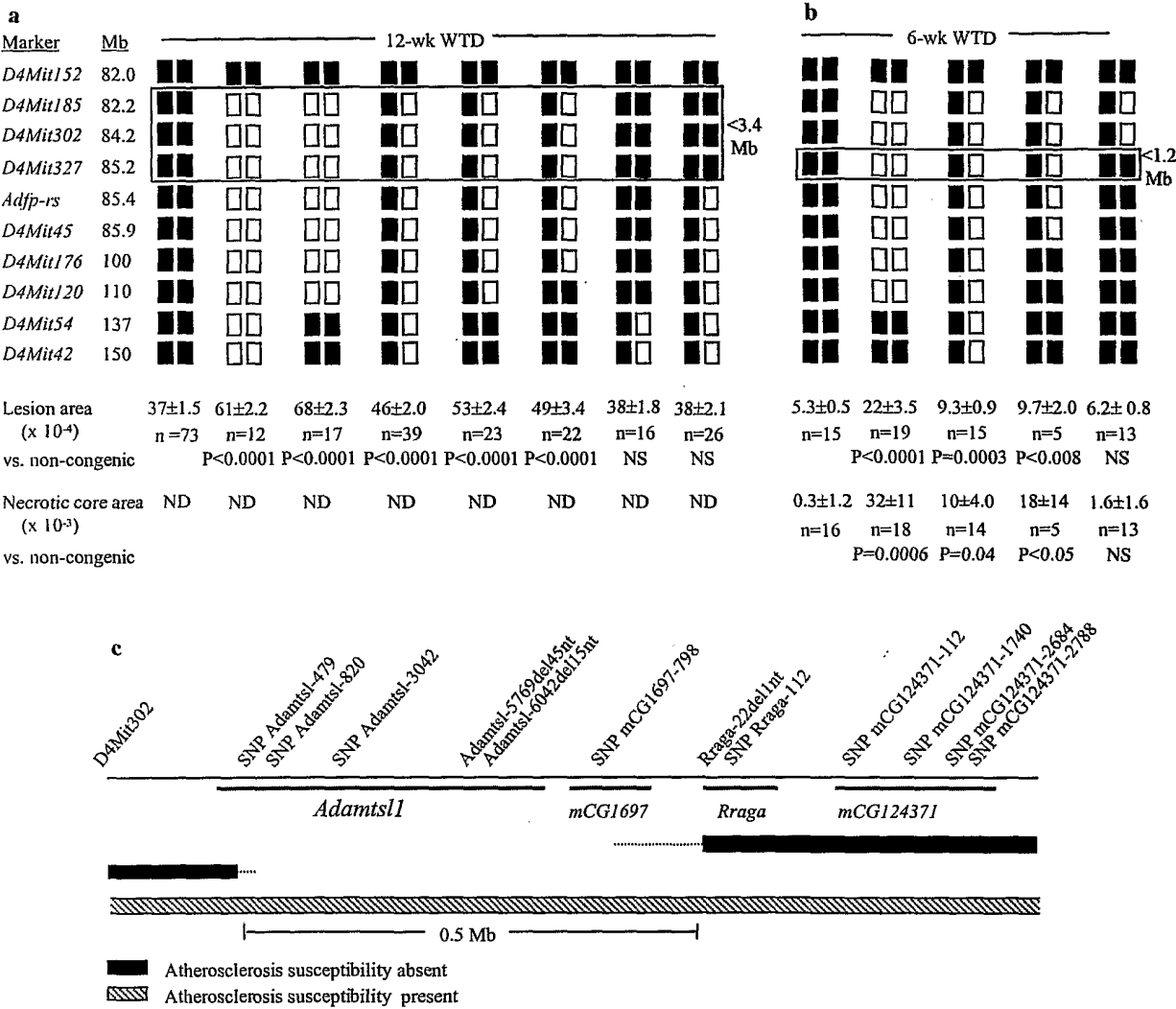
1/24

Figure 1



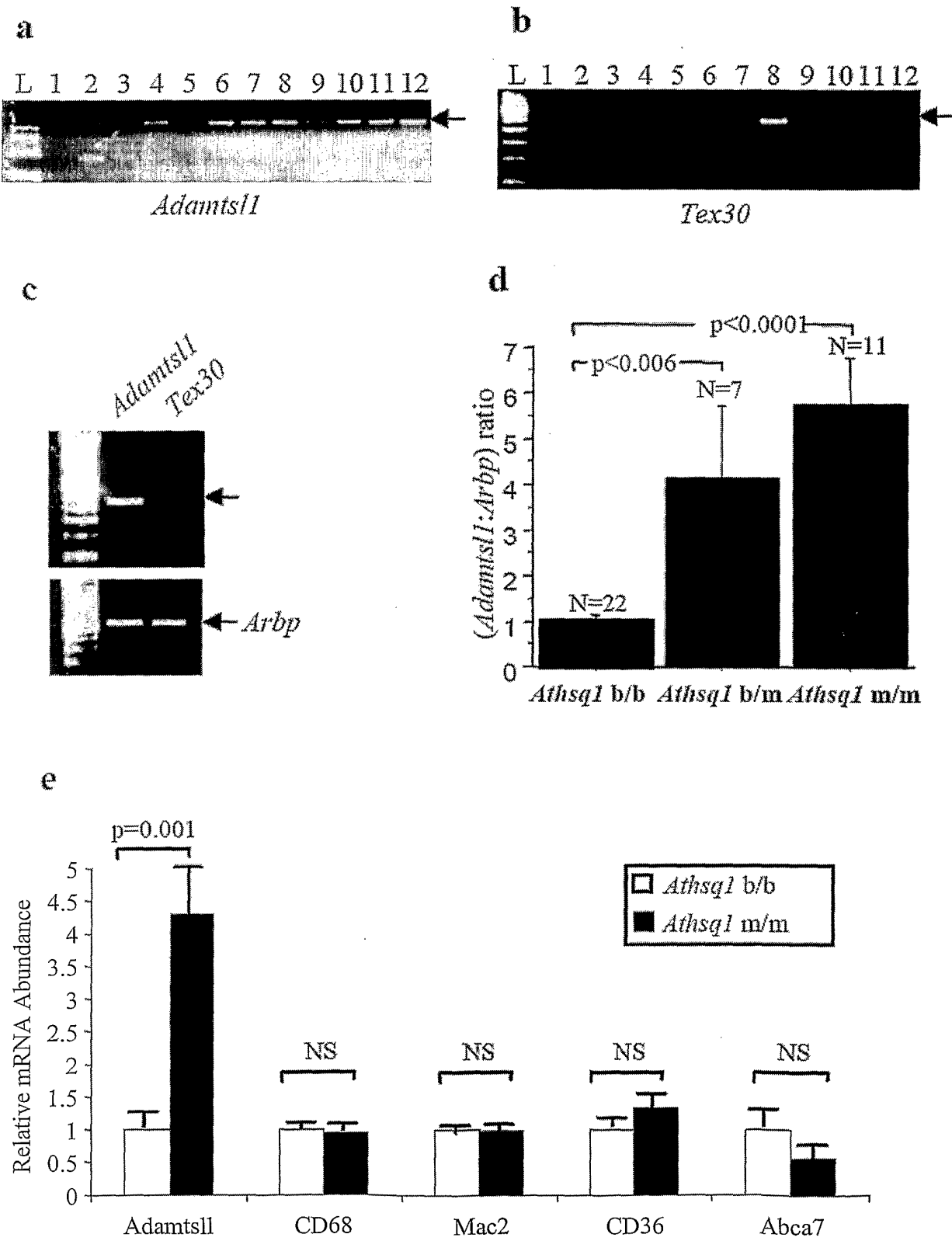
2/24

Figure 2



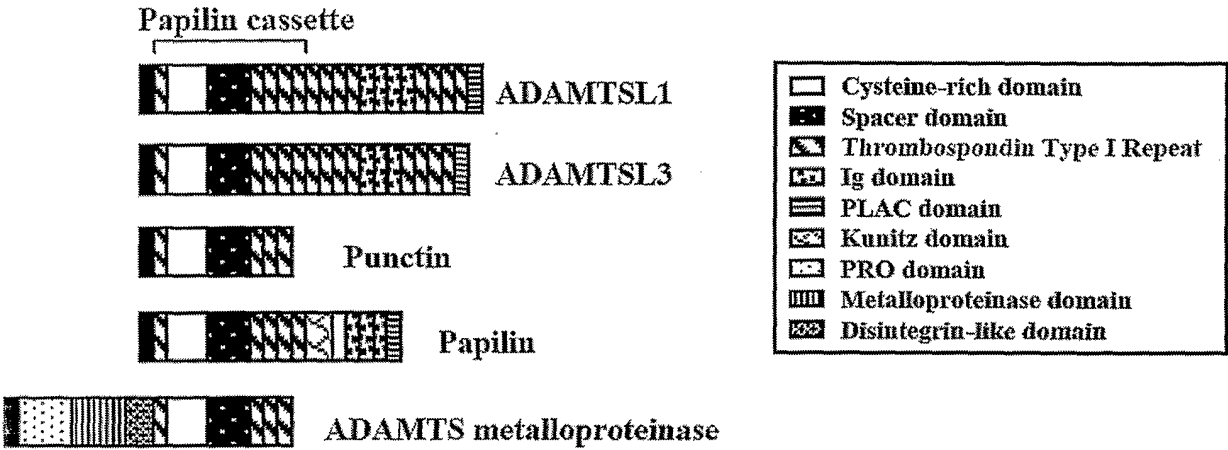
3/24

Figure 3



4/24

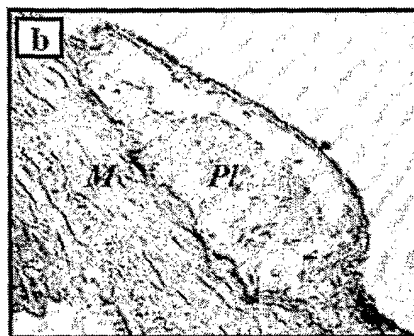
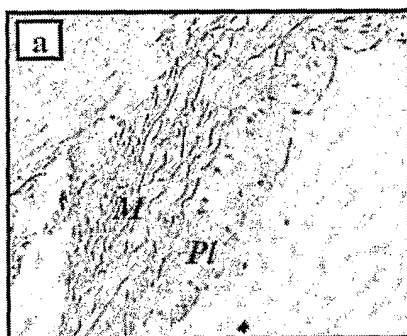
Figure 4



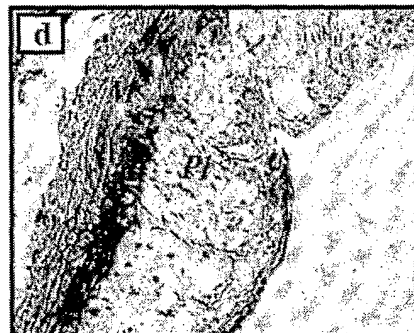
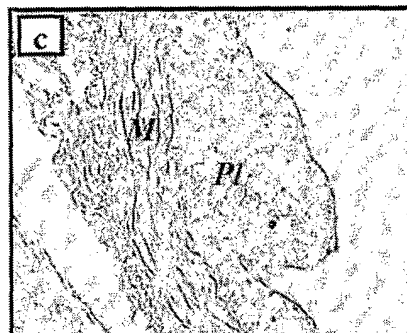
5/24

Figure 5

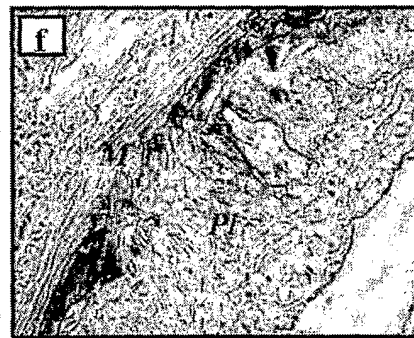
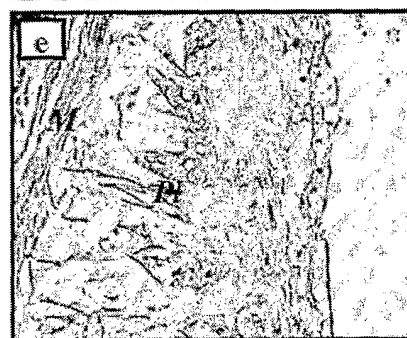
Versican
(3.5-wk WTD)



Versican
(6-wk WTD)



Versican
(12-wk WTD)



Athsq1 b/b

Athsq1 m/m

6/24

Figure 6

>hCT2302210

TGCCTGCTCCGAGAGAGATTCCATTTCAGCTTACCCCCACCCATCCACCCACCCACCCCT
CGGTCAGGAAATGTGAGAGGGGCTGATGGAAGCTGATAGGCAGGACTGGAGTGTAGCAC
CAGTACTGGATGTGACAGCAGGCAGAGGAGCACTTAGCAGCTTATTCAGTGTCCGATTCT
GATTCCGGCAAGGATCCAAGCATGGAATGCTGCCGTCGGGCAACTCCTGGCACACTGCTC
CTCTTTCTGGCTTTCCTGCTCCTGAGTTCAGGACCGCACGCTCCGAGGAGGACCGGGAC
GGCCTATGGGATGCCTGGGGCCCATGGAGTGAATGCTCACGCACCTGCGGGGGTGGGGCC
TCCTACTCTCTGAGGCGCTGCCTGAGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATAC
AGAACATGCAGTAATGTGGACTGCCACCAGAAGCAGGTGATTTCCGAGCTCAGCAATGC
TCAGCTCATAATGATGTCAAGCACCATTGGCCAGTTTTATGAATGGCTTCCTGTGTCTAAT
GACCCTGCAACCCCATGTTCACTCAAGTGCCAAGCCAAAGGAACAACCCCTGGTTGTTGAA
CTAGCACCTAAGGTCTTAGATGGTACGCGTTGCTATACAGAATCTTTGGATATGTGCATC
AGTGGTTTTATGCCAAATTGTTGGCTGCGATCACCAGCTGGGAAGCACCGTCAAGGAAGAT
AACTGTGGGGTCTGCAACGGAGATGGGTCCACCTGCCGGCTGGTCCGAGGGCAGTATAAA
TCCCAGCTCTCCGCAACCAAATCGGATGATACTGTGGTTGCAATTCCCTATGGAAGTAGA
CATATTCGCCTTGTCTTAAAAGGTCCTGATCACTTATATCTGGAAACCAAACCCCTCCAG
GGGACTAAAGGTGAAAACAGTCTCAGCTCCACAGGAACTTTCCTTGTGGACAATTCTAGT
GTGGACTTCCAGAAATTTCCAGACAAAGAGATACTGAGAATGGCTGGACCACTCACAGCA
GATTTTCATTGTCAAGATTTCGTAACCTCGGGCTCCGCTGACAGTACAGTCCAGTTCATCTTC
TATCAACCCATCATCCACCGATGGAGGGAGACGGATTTCTTTCCTTGCTCAGCAACCTGT
GGAGGAGGTTATCAGCTGACATCGGCTGAGTGCTACGATCTGAGGAGCAACCGTGTGGTT
GCTGACCAATACTGTCACTATTACCCAGAGAACATCAAACCCAAACCCAGCTTCAGGAG
TGCAACTTGGATCCTTGTCCAGCCAGTGACGGATACAAGCAGATCATGCCTTATGACCTC
TACCATCCCTTTCCTCGGTGGGAGGCCACCCCATGGACCGCGTGCTCCTCCTCGTGTGGG
GGGGGCTCCAGAGCCGGGCAGTTTCTCTGTGGAGGAGGACATCCAGGGGCATGTCACT
TCACTGGAAGAGTGGAAATGCATGTACACCTCTAAGATGCCCATCGCGCAGCCCTGCAAC
ATTTTTGACTGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGCACAGTGACATGTGGC
CAGGGCCTCAGATACCGTGTGGTCTCTGCATCGACCATCGAGGAATGCACACAGGAGGC
TGTAGCCCAAAAACAAAGCCCCACATAAAAGAGGAATGCATCGTACCCACTCCCTGCTAT
AAACCCAAAGAGAACTTCCAGTCGAGGCCAAGTTGCCATGGTTCAAACAAGCTCAAGAG
CTAGAAGAAGGAGCTGCTGTGTGTCAGAGGAGCCCTCGTTCATCCAGAGGCCTGGTCGGCC
TGCACAGTCACCTGTGGTGTGGGGACCCAGGTGCGAATAGTCAGGTGCCAGGTGCTCCTG
TCTTTCTCTCAGTCCGTGGCTGACCTGCCTATTGACGAGTGTGAAGGGCCCAAGCCAGCA
TCCCAGCGTGCTGTTATGCAGGCCCATGCAGCGGGGAAATTCCTGAGTTCAACCCAGAC
GAGACAGATGGGCTCTTTGGTGGCCTGCAGGATTTTCGACGAGCTGTATGACTGGGAGTAT
GAGGGGTTACCAAGTGCTCCGAGTCCTGTGGAGGAGGTGTCCAGGAGGCTGTGGTGAGC
TGCTTGAACAAACAGACTCGGGAGCCTGCTGAGGAGAACCTGTGCGTGACCAGCCGCCGG
CCCCCAGCTCCTGAAGTCTGCAATTTGGATCCCTGCCAGCAAGGTGGGAAATTGGC
AAGTGGAGTCCATGTAGTCTCACATGTGGGGTCGGCCTACAGACCAGAGACGTCTTCTGC
AGCCACCTGCTTTCAGAGAGATGAATGAAACAGTCATCCTGGCTGATGAGCTGTGTGCG
CAGCCCAAGCCCAGCACGGTGCAAGCTTGTAACCGCTTTAATTGCCCCCCAGCCTGTAC
CCTGCACAGTGGCAGCCGTGTTCCAGAACGTGTGGCGGGGGTGTTCAGAAACGTGAGGTT
CTTTGCAAGCAGCGCATGGCTGATGGCAGCTTCTTGAGCTTCTTGAGACCTTCTGTTCA
GCTTCAAACCTGCCTGCCAGCAAGCATGCAAGAAAGATGACTGTCCCAGCGAGTGCTT
CTCTCAGACTGGACAGAGTGTTCACAAGCTGCGGGGAAGGCACCCAGACTCGAAGCGCC
ATTTGCCGAAAGATGCTGAAAACCGGCATCTCAACGGTTGTCAATTCCACCTGTGCCCC
CCCCTGCCTTTCTTCTCCTCCATCAGGCCCTGTATGCTGGCAACCTGTGCAAGGCCCGGG
CGGCCATCCACGAAGCACAGCCCGCACATCGCGGCCGCCAGGAAGGTCTACATCCAGACT
CGCAGGCAGAGGAAGCTGCACTTCGTGGTGGGGGGCTTCGCCTACCTGCTCCCCAAGACG
GCGGTGGTGCTGCGCTGCCCGGCGCGCAGGGTCCGCAAGCCCCCTCATCACCTGGGAGAAG
GACGGCCAGCACCTCATCAGCTCGACGCACGTACCGTGGCCCCCTTCGGCTATCTCAAG
ATCCACCGCCTCAAGCCCTCGGATGCAGGCGTCTACACCTGCTCAGCGGGCCCGGCCGG
GAGCACTTTGTGATTAAGCTCATCGGAGGCAACCGCAAGCTCGTGGCCCGGCCCTTGAGC
CCGAGAAGTGAGGAAGAGGTGCTTGCGGGGAGGAAGGGCGGCCCGAAGGAGGCCCTGCAG
ACCCACAAACACCAGAACGGGATCTTCTCCAACGGCAGCAAGGCGGAGAAGCGGGGGCCTG
GCCGCCAACCCGGGGAGCCGCTACGACGACCTCGTCTCCCGGCTGCTGGAGCAGGGCGGC

7/24

Figure 6

TGGCCCGGAGAGCTGCTGGCCTCGTGGGAGGCGCAGGACTCCGCGGAAAGGAACACGACC
TCGGAGGAGGACCCGGGTGCAGAGCAAGTGCTCCTGCACCTGCCCTTCACCATGGTGACC
GAGCAGCGGCGCCTGGACGACATCCTGGGGAACCTCTCCCAGCAGCCCCGAGGAGCTGCGC
GACCTCTACAGCAAGCACCTGGTGGCCAGCTGGCCAGGAGATCTTCCGCAGCCACCTG
GAGCACCAAGGACACGCTCCTGAAGCCCTCGGAGCGCAGGACTTCCCCAGTGACTCTCTCG
CCTCATAAACACGTGTCTGGCTTCAGCAGCTCCCTGCGGACCTCCTCCACCGGGGACGCC
GGGGGAGGCTCTCGAAGGCCACACCGCAAGCCCACCATCCTGCGCAAGATCTCAGCGGCC
CAGCAGCTCTCAGCCTCGGAGGTGGTCACCCACCTGGGGCAGACGGTGGCCCTGGCCAGC
GGGACACTGAGTGTTCTTCTGCACTGTGAGGCCATCGGCCACCCAAGGCCTACCATCAGC
TGGGCCAGGAATGGAGAAGAAGTTCAGTTCAGTGACAGGATTCTTCTACAGCCAGATGAT
TCCTTACAGATCTTGGCACCATGTGAAGCAGATGTGGGTTTCTACACTTGCAATGCCACC
AATGCCTTGGGATACGACTCTGTCTCCATTGCCGTACATTAGCAGGAAAGGCCACTAGTG
AAAACGTACGAATGACAGTGATCAACACGGAGAAGCCTGCAGTCACAGTCGATATAGGA
AGCACCATCAAAACAGTGCAGGGAGTGAATGTGACAATCAACTGCCAGGTTGCAGGAGTG
CCTGAAGCTGAAGTCACTTGGTTCAGGAATAAAAGCAAACCTGGGCTCCCCGCACCATCTG
CACGAAGGCTCCTTGCTGCTCACAACGTGTCTCCTCGGATCAGGGCCTGTACTCCTGC
AGGGCGGCCAATCTTCATGGAGAGCTGACTGAGAGCACCCAGCTGCTGATCCTAGATCCC
CCCCAAGTCCCCACACAGTTGGAAGACATCAGGGCCTTGCTCGCTGCCACTGGACCGAAC
CTTCCTTCAGTGCTGACGTCTCCTCTGGGAACACAGCTGGTCCTGGATCCTGGGAATTCT
GCTCTCCTTGGCTGCCCCATCAAAGGTCAACCCTGTCCCTAATATCACCTGGTTTCATGGT
GGTCAGCCAATTGTCACTGCCACAGGACTGACGCATCACATCTTGGCAGCTGGACAGATC
CTTCAAGTTGCAAACCTTAGCGGTGGGTCTCAAGGGGAATTCAGCTGCCTTGCTCAGAA
GAGGCAGGGGTGCTCATGCAGAAAGGCATCTTTAGTGATCCAAGATTACTGGTGGTCTGTG
GACAGACTGGCAACCTGCTCAGCCTCCTGTGGTAACCGGGGGGTTTTCAGCAGCCCCGCTTG
AGGTGCCTCTGAACAGCACGGAGGTCAACCCCTGCCACTGCGCAGGGAAGGTTTCGCCCT
GCGGTGCAGCCCATCGCGTGCAACCGGAGAGACTGCCCTTCTCGGTGGATGGTGACCTCC
TGGTCTGCCTGTACCCGGAGCTGTGGGGGAGGTGTCCAGACCCGCAGGGTGACCTGTCAA
AAGCTGAAAGCCTCTGGGATCTCCACCCCTGTGTCCAATGACATGTGCACCCAGGTGCGC
AAGCGGCCTGTGGACACCCAGGCCTGTAACCAGCAGCTGTGTGTGGAGTGGGCCTTCTCC
AGCTGGGGCCAGTGCAATGGGCCTTGCAATCGGGCCTCACCTAGCTGTGCAACACAGACAA
GTCTTCTGCCAGACACGGGATGGCATCACCTTACCATCAGAGCAGTGCACTGCTCTTCCG
AGGCCTGTGAGCACCCAGAACTGCTGGTCAGAGGCCTGCAGTGTAACCTGGAGAGTCAGC
CTGTGGACCCTGTGCACAGCTACCTGTGGCAACTACGGCTTCCAGTCCCGGCGTGTGGAG
TGTGTGCATGCCCGCACCAACAAGGCAGTGCTGAGCACCTGTGCTCCTGGGGGCCCCGG
CCTGCCAACTGGCAGCGCTGCAACATCACCCCATGTGAAAACATGGAGTGACAGAGACACC
ACCAGTACTGCGAGAAGGTGAAACAGCTGAAACTCTGCCAACTCAGCCAGTTAAATCT
CGCTGCTGTGGAACCTTGTGGCAAAGCGTGAAGATAGGGTGTGGGGAAAAACTCTACCCTG
GCCACACGAAGGACTCACGCAACCACCTCGGACAGAACCTAAGCTTTCTTCATTTTATTT
ATTTATTTCCCCCTCCCCACTCCACACACACCTTCCAACCTCCTCCACCTCCACCTTCA
AGCATAAGGACGTCCGCGTGTCTTTCTTTTCTTTAGCTGGAGGACAGGATGTTGGGAAA
GGAAAGGACAGATGTCTAAAGGAGGTTGCAGAGCAGGCCAGGCAGACAGTGGGGGCTCCC
TTGAAGAGCTTCTCTCCCTCCCAAACCTGGGTCTCAAAGACCTAGAAAGAGGCAGGCACAG
CCCCTGCGGACAGCAGGGAGCCAGAAGGTTTGTAGCCTATTGGTGCAAACATTGGACAAA
TTCCTGTGTCTTTCTAGAAAGCGCACTATCACAACACAGGAGTGTTTTGCTCCTTTGTC
TCCTCTTCCCCATCTATGTCCCTTTAGTCACAGTTAGGACAAATGGGGAGGGGACACCAT
GCTGAGGCAGAACTAGCCCAGAACTCACTCAGTTCTTCTAGTGGGTGAGTGACAGAGAGA
GAAGAACTCAGATCACCAAGTAGGGAGAGGTAAAAAAGCAAACAAAGCAGGCTCTAAGGCA
CACAACATTGCAGAAAAATGAGGAAGGGAGGGGAGGGAAGGGACAGAAAGCAAAAAGGAGCC
TGTGGTGTTCCTCAGTGGGGCAGGGTGAGCAGGGGCTTCCAGGCTGCATGAGGCTCATGG
ACCAGCTCTGATCCCATGCATGTGCGCATGCTCAGAGCCCTGCTGCCACAACAGAGCAC
TGCGCTGCGTGGGAGTCCCCACTTCCCAAGCTATCAGAGTCAACGTCCTGCCTGTGCAGC
TGCAGCAAAGCCAGTGAGAGGTGGGTCTCGCCATGCAGTAAGGCCACCCTGGCACCTCTT
TATCTAAATCCGAAGTCCCCTAGCCCCGCACTAACTAACTGCTGCTGTGGGCCAGGGCCA
TTTTGAGCATGAATGGCCAGGTTTTTTCCTTCTAGGACCTTGCTGCTCCACCGAAGG
GCCAGGGATATGGTTAACTTATCAACATCAACCATTAAGTCACTGTGCCAGAGAG
TATCTGTCAGGCTGTGACAGCTGTGACAGTTGTAGCAACCTCTTCATTCCAGAGTACCCCA

8/24

Figure 6

GGGACCGGGGTGGGACAATGGGTTTATGCGTGTCCACAGTACACCCTCCCTCTCCCAGCC
 TCCACCCCAGGGTCTGCAGGTCTCCGGCATGTAGTATTTATCTAGCAAGGCGGGGTGGT
 GGAGGCAGCACCTTGGCAAAGCAGCTCACACACTGCAGCCCACTCATCAGCTGTGGTGA
 GCGGCTGGAGCAAAGTCAAAGTCATGCAGCAAAATGAAAACCTCTGGGACTCTTCGGCAA
 AATCCTCATTAAGCCGAGCAGCTTTGGCCAAGTAATTTTTGCCTCCTTCCCTCGCGTGGC
 CTGAGTTTAGGAGCAAGGGTGGCCAGAGTCCCTTACCCACAGATAAGCCTCCCCTCATGA
 AATGCCACTCACCCCGGGCTACCATTGACATCAGGGCTGCATTTCCAGCCAGCCTGGAAG
 TAAAATTTGAGAGGAAGACAATATTAATCTGTGTCCCCACCTAGTGAGCTGTGGACAGGT
 TTAAGTTGGGTCTCCTTCTTCTTACCACAAAAACAGGCTCTAAGAAATCATGTTACTAA
 AAAATCAGTGTAAGTCTGTTTAAATAAAAAAGAATGTTTTCT

>hCP1889215

MECCRATPGTLLFLAFLLLSSRTARSEEDRDGLWDAWGPWSECSRTC GGGASYS LRRC
 LSSKSCEGRNIRYRTCSNVDPCPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNPDNPCS
 LKCAKGTTLVVELAPKVLDTGRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNG
 DGSTCRLVRGQYKSQLSATKSDDTVVAIPYGSRHRLVLKGPDLHLYETKTLQGTKGENS
 LSSTGTFLVDNSSVDFQKFPDKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHR
 WRETDFPCSATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCP
 ASDGYKQIMPYDLYHPLPRWEATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKC
 MYTPKMPIAQPCNIFDCPKWLAQEWSPCTVTCGQGLRYRVLCIDHRGMHTGGCSPKTKP
 HIKEECIVPTPCYKPKKEKLPVEAKLPWFKQAQELLEGAAVSEEPSFIPEAWSACTVTCGV
 GTQVRIVRCQVLLSFSQSVDLPIDECEGPKPASQRACYAGPCSGEIEFNPDETDLGLFG
 GLQDFDELYDWEYEGFTKCESCGGGVQEAVVSCLNKQTREPAEENLCVTSRRPPQLLKS
 CNLDPCPARWEIGKWSPCSLTGCVGLQTRDVFCSHLLSREMNETVILADELCRQPKPSTV
 QACNRFNCPA WYPAQWQPCSRTCGGGVQKREVLCKQRMADGSFLELPETFCSASKPACQ
 QACKKDDCPSEWLLSDWTECSTSCGEGTQTRSAICRKMLKTGISTVVNSTLCPLPFSSS
 IRPCLMLATCARPGRPSTKHSPIHAAARKVYIQTRRQRKLHFVVGGFAYLLPKTAVVLRCP
 ARVRKPLITWEKDGOHLISSTHVTVPFGLYKIHRLKPSDAGVYTCSAGPAREHFVIKL
 IGGNRKLVARPLSPRSEEEVLGRKGKPKALQTHKHQNGIFSNGSKAEKRGLAANPGSR
 YDDLVSRLLEQGGWPGELLASWEAQDSAERNTTSEEDPGAEQVLLHLPFTMVTQRRLLDD
 ILGNLSQQPEELRDLYSKHLVAQLAQEIFRSHLEHQDTLLKPSERRTSPVTLSPHKHVS
 FSSSLRTSSTGDAGGGSRRPHRKPTILRKISAAQQLSASEVVTHLGQTVALASGTLVLL
 HCEAIGHPRPTISWARNGEEVQFSDRILLQPDDSLQILAPVEADVGFYTCNATNALGYDS
 VSIAVTLAGKPLVKTSRMTVINTEKPAVTVDIGSTIKTVQGVNVTINCQVAGVPEAEVTW
 FRNKSGLGSPHHLHEGSLLLTNVSSSDQGLYSCRAANLHGELTESTQLLILDPPQVPTQL
 EDIRALLAATGPNLPSVLTSPPLGTQLVLDPGNSALLGCPIKGHPVPNITWFHGGQPIVTA
 TGLTHHILAAGQILQVANLSGGSQGEFSCLAQNEAGVLMQKASLVIQDYWWSVDRLATCS
 ASCGNRGVQQPRLRCLLNSTEVNPAHCAGKVRPAVQPIACNRRDCPSRWMVTSWSACTRS
 CGGGVQTRRVTCQKCLKASGISTPVSNDMCTQVAKRPVDTQACNQQLCWEWAFSSWGQCNG
 PCIGPHLAVQHRQVFCQTRDGITLPSEQSALPRPVSTQNCWSEACSVHWRVSLWTLCTA
 TCGNYGFQSRREVCVHARTNKAVPEHLCSWGPRPANWQRCNITPCENMECRDTTRYCEKV
 KQLKLCQLSQFKSRCCGTCGKA

9/24

Figure 7a

B6 AGCACTGGATGTGACAGCAGCGGGCAGAGGAGCACTTAGCAGCTTATTCAGTGTCCGATT 60
MOLF AGCACTGGATGTGACAGCAGCGGGCAGAGGAGCACTTAGCAGCTTATTCAGTGTCCGATT 60

B6 CGGATTCCGGCAAAGATCCAAGTATGGAATGCTGCCGTCGGGCAGCTCCTGGCACACCGC 120
MOLF CGGATTCCGGCAAAGATCCAAGTATGGAATGCTGCCGTCGGGCAGCTCCTGGCACACCGC 120

B6 TGCTCGTTCTGGCTTTCTACTCCTGAGTTCAGGACTGCACGCTCAGAGGAAGACCGTG 180
MOLF TGCTCGTTCTGGCTTTCTACTCCTGAGTTCAGGACTGCACGCTCAGAGGAAGACCGTG 180

B6 AAGGCCTCTGGGATGCCCTGGGGCCCTTGAGCGAGTGTTCAGAACCTGTGGTGGGGGTG 240
MOLF AAGGCCTCTGGGATGCCCTGGGGCCCTTGAGCGAGTGTTCAGAACCTGTGGTGGGGGTG 240

B6 CCTCCTATTCCCTGAGACGCTGCCTGAGCAGCAAGAGCTGTGAAGGGAGAAATATTCGAT 300
MOLF CCTCCTATTCCCTGAGACGCTGCCTGAGCAGCAAGAGCTGTGAAGGGAGAAATATTCGAT 300

B6 ATAGAACATGCAGTAATGTGGACTGCCACCAGAAGCAGGTGATTTCAGGCTCAACAGT 360
MOLF ATAGAACATGCAGTAATGTGGACTGTCCACCAGAAGCAGGTGATTTCAGGCTCAACAGT 360

B6 GTTCTGCTCATAACGACGTCAGTACCATGGACAGCTTTATGAATGGCTTCCTGTATCTA 420
MOLF GTTCTGCTCATAACGACGTCAGTACCATGGACAGCTTTATGAATGGCTTCCTGTATCTA 420

B6 ATGACCCCGACAATCCATGCTCAAGTGCCAAGCAAAGGAACAGCCTGGTTGTGG 480
MOLF ATGACCCCTGACAATCCATGTTCAAGTGCCAAGCAAAGGAACAGCCTGGTTGTGG 480

B6 AACTAGCACCGAAGGTCTTAGATGGTACGCGATGCTATACAGAATCACTGGATATGTGCA 540
MOLF AACTAGCACCGAAGGTCTTAGATGGTACGCGATGCTATACAGAATCACTGGATATGTGCA 540

B6 TCAGCGGCCTATGCCAAATTGTTGGCTGTGATCACCAGCTGGGGAGCACTGTCAAAGAGG 600
MOLF TCAGCGGCCTATGCCAAATTGTTGGCTGTGATCACCAGCTGGGGAGCACTGTCAAAGAGG 600

B6 ATAACCTGTGGAGTGTGCAACGGAGATGGCTCAACGTGCCGGCTGGTCCGAGGGCAGTATA 660
MOLF ATAACCTGTGGAGTGTGCAACGGAGATGGCTCAACGTGCCGGCTGGTCCGAGGGCAGTATA 660

B6 AATCTCAGCTCTCTGCAAGTAAATCGGATGACACCGTGGTCGCCATTCTTACGGCAGCA 720
MOLF AATCTCAGCTCTCTGCAAGTAAATCGGATGACACCGTGGTCGCCATTCTTACGGCAGCA 720

B6 GACACATCCGCCTTGTCTTAAAGGGCCAGATCATTTGTATCTGGAGACCAAGACCTCC 780
MOLF GACACATCCGCCTTGTCTTAAAGGGCCAGATCATTTGTATCTGGAGACCAAGACCTCC 780

B6 AGGGAACATAAGGTGAAAATAGTCTGAGCTCTACAGGCATTTTCCTTGTGGACAATTCTA 840
MOLF AGGGAACATAAGGTGAAAATAGTCTGAGCTCTACAGGCATTTTCCTTGTGGACAATTCTA 840

B6 CAGTGGACTTCCAGAACTTCCAGACAAAGAGATACTGAGGATGACTGGGCCACTGACAG 900
MOLF CAGTGGACTTCCAGAACTTCCAGACAAAGAGATACTGAGGATGACTGGGCCACTGACAG 900

10/24

Figure 7a

B6	CAGATTTCATCATCAAGATCCATGACTTGGGACCTGCAGACAGCACCGTCCAGTTCATCT	960
MOLF	CAGATTTCATCATCAAGATCCATGACTTGGGACCTGCAGACAGCACCGTCCAGTTCATCT	960

B6	TCTACCAGCCCATCATCCACCGATGGAGGGAAACAGACTTCTTTCCTTGCTCAGCAACGT	1020
MOLF	TCTACCAGCCCATCATCCACCGATGGAGGGAAACAGACTTCTTTCCTTGCTCAGCAACGT	1020

B6	GTGGTGGAGGCTACCAGCTGACGTCTGCCGAGTGCTATGATCTACGGAGTAACCGTGTGG	1080
MOLF	GTGGTGGAGGTTACCAGCTGACGTCTGCCGAGTGCTATGATCTACGGAGTAACCGTGTGG	1080

B6	TGGCCGACCAGTACTGTCACTACTACCCCGAGAACGTCAAGCCCAAACCCAAGCTTCAGG	1140
MOLF	TGGCCGACCAGTACTGTCACTACTACCCCGAGAACGTCAAGCCCAAACCCAAGCTTCAGG	1140

B6	AGTGCAACTTGGATCCTTGTCCAGCCAGTGACGGATACAAGCAGATCATGCCTTATGACC	1200
MOLF	AGTGCAACTTGGATCCTTGTCCAGCCAGTGACGGATACAAGCAGATCATGCCTTATGACC	1200

B6	TCTACCATCCCCCTTCCTCGGTGGGAGGCCACCCCGTGGACCGCGTGCTCCTCCTCATGTG	1260
MOLF	TCTACCATCCCCCTTCCTCGGTGGGAGGCCACCCCGTGGACCGCGTGCTCCTCCTCATGTG	1260

B6	GGGGGGGCATCCAGAGCCGGGCAGTTTCTCTGTGTGGAGGAGGACATCCAGGGGCATGTCA	1320
MOLF	GGGGGGGCATCCAGAGCCGGGCAGTTTCTCTGTGTGGAGGAGGACATCCAGGGGCATGTCA	1320

B6	CTTCAGTGGAAGAGTGGAAATGCATGTACACCCCAAAGATGCCAGTCGTGCAGCCCTGCA	1380
MOLF	CTTCAGTGGAAGAGTGGAAATGCATGTACACCCCAAAGATGCCAGTCGTGCAGCCCTGCA	1380

B6	ACATTTTGTGACTGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGCACAGTGACCTGTG	1440
MOLF	ACATTTTGTGACTGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGCACAGTGACCTGTG	1440

B6	GCCAGGGCCTCAGATACCGCGTGGTACTCTGCATTGACCACCGAGGAATGCACACAGGAG	1500
MOLF	GCCAGGGCCTCAGATACCGCGTGGTACTCTGCATTGACCACCGAGGAATGCACACAGGAG	1500

B6	GCTGCAGTGCTAAAACAAAGCCCCACATAAAGAAGAATGCATCATTTCCACACCCCTGCT	1560
MOLF	GCTGCAGTGCGAAAACAAAGCCCCACATAAAGAAGAATGCATCATTTCCACACCCCTGCT	1560

B6	ACAAGCCCAGAGAGAAACTTCCAATAGAGGCCAAGCTGCCATGGTTCAAGCAAGCCCAGG	1620
MOLF	ACAAGCCCAGAGAGAAACTTCCAATAGAGGCCAAGCTGCCATGGTTCAAGCAAGCCCAGG	1620

B6	AGCTAGAGGAAGGAGCTGCAGTGTCTGAAGAGCCTTCGTTTCATCCAGAAGCTTGGTCGG	1680
MOLF	AGCTAGAGGAAGGAGCTGCAGTGTCTGAAGAGCCTTCGTTTCATCCAGAAGCTTGGTCGG	1680

B6	CCTGTACTGTACCTGTGGCGTGGGACACAGGTGCGAATAATAAGATGCCAGGTGCTCC	1740
MOLF	CCTGTACTGTACCTGTGGCGTGGGACACAGGTGCGAATAATAAGATGCCAGGTGCTCC	1740

B6	TGCCTTTCTCTCAGTCTGTGGCTGACCTGCCTGCGGATGAATGCGAAGGGCCTAAGCCAG	1800
MOLF	TGCCTTTCTCTCAGTCTGTGGCTGACCTGCCTGCGGATGAATGCGAAGGGCCTAAGCCAG	1800

11/24

Figure 7a

```
*****
B6      CATCGCAACGTGCCTGTTATGCGGGACCGTGCAACGGGGAAACTCCGGAATTTAACCCGG 1860
MOLF    CATCGCAACGTGCCTGTTATGCGGGACCGTGCAACGGGGAAACTCCGGAATTTAACCCGG 1860
*****

B6      ATAACGAGACGGACTCTTGGGTGGCCTGCAGGACCTGGACGAGCTGTATGACTGGGAAT 1920
MOLF    ATAACGAGACGGACTCTTGGGTGGCCTGCAGGACCTGGACGAGCTGTATGACTGGGAAT 1920
*****

B6      ATGAGGGCTTCACCAAGTGCTCTGAGTCATGTGGAGGAGGTGTCCAGGAGGCTGTGGTAA 1980
MOLF    ATGAGGGCTTCACCAAGTGCTCTGAGTCATGTGGAGGAGGTGTCCAGGAGGCTGTGGTAA 1980
*****

B6      GCTGCCTGAACAAACAGACCCGGGAACTCGCTGATGAGAACCCTGTGTGTGACCAGCCGAC 2040
MOLF    GCTGCCTGAACAAACAGACCCGGGAGCTCGCTGATGAGAACCCTGTGTGTGACCAGCCGAC 2040
*****

B6      GGCCCCCGCAGCTCCTGAAATCCTGTAACTTGGATCCCTGCCCAGCAAGGTGGGAAATTG 2100
MOLF    GGCCCCCGCAGCTCCTGAAATCCTGTAACTTGGATCCCTGCCCAGCAAGGTGGGAAATTG 2100
*****

B6      GCAAGTGGAGCCCATGTAGCCTCACCTGTGGGGTTGGCCTGCAGACCAGAGATGTTGTCT 2160
MOLF    GCAAGTGGAGCCCATGTAGCCTCACCTGTGGGGTTGGCCTGCAGACCAGAGATGTTGTCT 2160
*****

B6      GTTCCACCTACTCTCCAGAGAGATGAACGAGGTGGTTGTCTGGCCGATGAGCTGTGCC 2220
MOLF    GTTCCACCTACTCTCCAGAGAGATGAACGAGGTGGTTGTCTGGCCGATGAGCTGTGCC 2220
*****

B6      ATCACCCCAAACCCAGCACAGTGCAAGCCTGCAACCGTTTCAATTGCCCCCAGCCTGGT 2280
MOLF    ATCACCCCAAACCCAGCACAGTGCAAGCCTGCAACCGTTTCAATTGCCCCCAGCCTGGT 2280
*****

B6      ACCCTGCACAGTGGCAGCTGTGTTCCAGGACCTGTGGAGGAGGGATTGAGAAACGTGATG 2340
MOLF    ACCCTGCACAGTGGCAGCTGTGTTCCAGGACCTGTGGAGGAGGGATTGAGAAACGTGATG 2340
*****

B6      TTCTCTGCAAGCAGCGTATGGCTGACGGCAGCTTCTTGAGCTTCCAGAGACCTTCTGTT 2400
MOLF    TTCTCTGCAAGCAGCGTATGGCTGACGGCAGCTTCTTGAGCTTCCAGAGACCTTCTGTT 2400
*****

B6      CTGCCTCCAAACCAACCTCCCACCAAGGGTGCAAGAAAGATGACTGCCCCAGTGAGTGGC 2460
MOLF    CAGCCTCCAAACCAACCTCCCACCAAGGGTGCAAGAAAGATGACTGCCCCAGTGAGTGGC 2460
* *****

B6      TTCTCTCAGAGTGGTCAGAGTGCTCTGTGAGCTGTGGGGAGGGCACCCAGACTCGAAGTG 2520
MOLF    TTCTCTCAGAGTGGTCAGAGTGCTCTGTGAGCTGTGGGGAGGGCACCCAGACTCGAAGTG 2520
*****

B6      CCATTTGCCAGAGGGTGCTGAAAACCTGGGGTCTCCACTGTTGTCAATTCTACTTTGTGCC 2580
MOLF    CCATTTGCCAGAGGGTGCTGAAAACCTGGGGTCTCCACTGTTGTCAATTCTACTTTGTGCC 2580
*****

B6      CTCCCCTGCCCTTCTCTTCTCCATCAGACCCTGTATGCTGGCAACCTGTGCAAGGCCTG 2640
MOLF    CTCCCCTGCCCTTCTCTTCTCCATCAGACCCTGTATGCTGGCAACCTGTGCAAGGCCTG 2640
*****

B6      GAAGGCCTTCCACCAAGCACAGTCCTCACATCGCAGCTGCACGAAACATATACATCCAGA 2700
```

12/24

Figure 7a

MOLF	GAAGGCCTTCCACCAAGCACAGTCCTCACATCGCAGCTGCACGAAACATATACATCCAGA	2700

B6	CCCGCAGGCAGAGGAAGCTGCACTTCGTTGTGGGAGGATTTGCTTACCTGCTTCCCAAGA	2760
MOLF	CCCGCAGGCAGAGGAAGCTGCACTTCGTTGTGGGAGGATTTGCTTACCTGCTTCCCAAGA	2760

B6	CTACCGTGGTGCTCCGGTGTCCACACGCAGATTCCGCAAGCCCCTCATCACGTGGGAAA	2820
MOLF	CTACCGTGGTGCTCCGGTGTCCACACGCAGATTCCGCAAGCCCCTCATCACGTGGGAAA	2820

B6	AGGACGGCCAGCACAGCATCAGCTCGGCTCATGTACCGTAGCCCCCTTTGGCTACCTGA	2880
MOLF	AGGACGGCCAGCACAGCATCAGCTCGGCTCATGTACCGTAGCCCCCTTTGGCTACCTGA	2880

B6	AGATCCATCGCCTCAAGCCCTCAGATGCGGGCATCTACACCTGTTCTGCGGGGCCCCGCC	2940
MOLF	AGATCCATCGCCTCAAGCCCTCAGATGCGGGCATCTACACCTGTTCTGCGGGGCCCCGCC	2940

B6	GGGAGCAGTTTGTGATTAAGCTCATTGGAGGCAACCGCAAGCTCGTGGCCAGGCCCTGA	3000
MOLF	GGGAGCAGTTTGTGATTAAGCTCATTGGAGGCAACCGCAAGCTCGTGGCCAGGCCCTGA	3000

B6	GCCTTTGGAGTGAGGAAGAGGAGGCCCTGCAGGTGAGGAAGACCAATCCAAAAGAAGCCT	3060
MOLF	GCCTTTGGAGTGAGGAAGAGGAGGCCCTGCAGGTGAGGAAGGCAATCCAAAAGAAGCCT	3060

B6	TGCAAACCCACAAACACCAAAATGGCATCTTCTCCAACGGCAGCAAAGCTGAGAAGCGCG	3120
MOLF	TGCAAACCCACAAACACCAAAATGGCATCTTCTCCAACGGCAGCAAAGCTGAGAAGCGCG	3120

B6	GCCTCACTGCAGATCCTGGGAACCGCTATGATGACATTGTGTCTAGGCTGTTGGAGCAGG	3180
MOLF	GCCTCACTGCAGATCCTGGGAACCGCTATGATGACATTGTGTCTAGGCTGTTGGAGCAGG	3180

B6	GCGGCTGGCCAGGAGAGCTCCTGGCCTCGTGGGAGGTGCAGGACTCTGCAGAAAGGAACG	3240
MOLF	GCGGTTGGCCAGGAGAGCTCCTGGCCTCGTGGGAGGTGCAGGACTCTGCAGAAAGGAACG	3240

B6	CATCCTCAGAAGAGGATCCCAATGCTGAACAGGCCCTCCTGCATCTCCCCTTACCATGG	3300
MOLF	CATCCTCAGAAGAGGATCCCAATGCTGAACAGGCCCTCCTGCATCTCCCCTTACCATGG	3300

B6	TGGCCGAGCAGAAGCGTCTGGATGACATCCTCAGGAACCTGTCCCAGCAGCCAGAGGAGC	3360
MOLF	TGGCCGAGCAGAAGCGTCTGGATGACATCCTCAGGAACCTGTCCCAGCAGCCAGAGGAGC	3360

B6	TACGGGACCTCTACAGCAAGCACCTGGTGGCCAGCTGGCCAGGACATCTTCCGAAGCC	3420
MOLF	TACGGGACCTCTACAGCAAGCACCTGGTGGCCAGCTGGCCAGGACATCTTCCGAAGCC	3420

B6	ACCTGGAGAACCAGGATCTGCTTCCTAAGCCCTCTGAGCAGAGGTTTCTCCCATGGCCG	3480
MOLF	ACCTGGAGAACCAGGATCTGCTTCCTAAGCCCTCTGAGCAGAGGTTTCTCCCATGGCCG	3480

B6	TCCCAGCTCACAAACATGTCTCTGGCTTCAGCAGCTCCCTAAGGAGCTCATCTGGAGAGG	3540
MOLF	TCCCAGCTCACAAACATGTCTCTGGCTTCAGCAGCTCCCTAAGGAGCTCATCTGGAGAGG	3540

13/24

Figure 7a

B6	CTGGAGGTGGCTCCCGTAGGCCATCGCAAGCCCGCCATCTTGCGGAAGATCTCAGCGG	3600
MOLF	CTGGGGGTGGCTCCCGTAGGCCATCGCAAGCCCGCCATCTTGCGGAAGATCTCAGCGG	3600

B6	CACAGCAACTCTCAGCCTCGGAGGTAGTCACCCACCTGGGGCAGACTGTGGCTCTAGCCA	3660
MOLF	CACAGCAACTCTCAGCCTCGGAGGTAGTCACCCACCTGGGGCAGACTGTGGCTCTAGCCA	3660

B6	GTGGGACCCCTGAGTGTGCTTCTGCACTGTGAGGCGGTAGGCAACCCAAGGCCACCATCC	3720
MOLF	GTGGGACCCCTGAGTGTGCTTCTGCACTGTGAGGCGGTAGGCAACCCAAGGCCACCATCC	3720

B6	ACTGGACCAGGAACGGAGAAGCAGTTCAATTCAGTGACAGGATTCTTCTACAGCCAGATG	3780
MOLF	ACTGGACCAGGAACGGAGAAGCAGTTCAATTCAGTGACAGGATTCTTCTACAGCCAGATG	3780

B6	ATTCCTTACAGATCTTGGCACCAGTGAAGCTGATGTGGGTTTCTATACTTGCAATGCCA	3840
MOLF	ATTCCTTACAGATCTTGGCACCAGTGAAGCTGATGTGGGTTTCTATACTTGCAATGCCA	3840

B6	CAAATGCCTTGGGATATGACTCTGTCTCCATTGCTGTACATTAGCAGGAAAGCCACTAG	3900
MOLF	CAAATGCCTTGGGATATGACTCTGTCTCCATTGCTGTACATTAGCAGGAAAGCCACTAG	3900

B6	TGAAGACATCCCGAATGACTGTGTCTCAACACAGAGAAGCCTACAGTCACAGTGGATATCG	3960
MOLF	TGAAGACATCCCGAATGACTGTGTCTCAACACAGAGAAGCCTACAGTCACAGTGGATATCG	3960

B6	GTGGCACCGTCAGAACCGTCCGAGGAGTGAATGTGACAATTAAGTCCAAGTGGCAGGTG	4020
MOLF	GTGGCACCGTCAGAACCGTCCGAGGAGTGAATGTGACAATTAAGTCCAAGTGGCAGGTG	4020

B6	TACCTGAAGCTGAAGTCACCTGGTTTCAAGAAATAAAGCAAATTGGGCTCCTCCCACCACC	4080
MOLF	TACCTGAAGCTGAAGTCACCTGGTTTCAAGAAATAAAGCAAATTGGGCTCCTCCCACCACC	4080

B6	TGCATGAGGGCTCCTCCCACCACCTGCATGAGGGCTCCTTGCTGCTCACAGATGTGTCCT	4140
MOLF	TGCATGAGGGCTCCTCCCACCACCTGCATGAGGGCTCCTTGCTGCTCACAGATGTGTCCT	4140

B6	TCTCTGATCAAGGCCTGTATTCTGCAGGGCAGCCAATCTACACGGAGAACAGACAGAGA	4200
MOLF	TCTCTGATCAAGGCCTGTATTCTGCAGGGCAGCCAATCTACACGGAGAACAGACAGAGA	4200

B6	GCACGCAGCTCCTGATTCTAGATCCCCGCAAGTCCCCACACAGTTGGAAGACATCAGGG	4260
MOLF	GCACGCAGCTCCTGATTCTAGATCCCCGCAAGTCCCCACACAGTTGGAAGACATCAGGG	4260

B6	CCTTGCTCTTAGCCACTGGACCGAACCTTCCTTCGGTGCTTATGTCCCCTCTGGGAACAC	4320
MOLF	CCTTGCTCTTAGCCACTGGACCGAACCTTCCTTCGGTGCTTATGTCCCCTCTGGGAACAC	4320

B6	AGCTGGTCCTGGATCCTGGGAATTCGCTCTCCTTGGCTGCCCCATCAAAGGTCACCCCA	4380
MOLF	AGCTGGTCCTGGATCCTGGGAATTCGCTCTCCTTGGCTGCCCCATCAAAGGTCACCCCA	4380

B6	CCCCCAATATCACCTGGTTCCAAAATGGCCAGCCAATTGCCACTGCCCCGGGGCTGACCC	4440
MOLF	CCCCCAATATCACCTGGTTCCAAAATGGCCAGCCAATTGCCACTGCCCCGGGGCTGACCC	4440

14/24

Figure 7a

B6	ACCACATCTGGGGAGCTGGACAGATCCTACGAGTTGCAAATCTCAGTGGAGGACCTCAAG	4500
MOLF	ACCACATCTGGGGAGCTGGACAGATCCTACGAGTTGCAAATCTCAGTGGAGGACCTCAAG	4500

B6	GGGAGTTCAGCTGCCTGGCTCAGAATGAAGCAGGGACCCTCCTGCAGAAAGCCTCTTTGG	4560
MOLF	GGGAGTTCAGCTGCCTGGCTCAGAATGAAGCAGGGACCCTCCTGCAGAAAGCCTCTTTGG	4560

B6	TGATCCAAGATTACTGGTGGTCTGTGGATAGACTGGCAACCTGCTCCGCATCTTGTGGTA	4620
MOLF	TGATCCAAGATTACTGGTGGTCTGTGGATAGACTGGCAACCTGCTCCGCATCTTGTGGTA	4620

B6	ATAGAGGCATTTCATCAGCCGCGACTCAGGTGTTTACTGAACACCACGGAAGTTGACCCCTG	4680
MOLF	ATAGAGGCATTTCATCAGCCGCGACTCAGGTGTTTACTGAACACCACGGAAGTTGACCCCTG	4680

B6	AGCACTGTACAGGAAAGCCTCGGCCCGCCGTGCAGCCTGTTGCCTGCAACCGGAGAGACT	4740
MOLF	AGCACTGTACAGGAAAGCCTCGGCCCGCCGTGCAGCCTGTTGCCTGCAACCGGAGAGACT	4740

B6	GTCCTTCTCGATGGATGGTGACGTCCTGGTCCGCCTGTACCCGGAGTTGCGGTGGAGGTG	4800
MOLF	GTCCTTCTCGATGGATGGTGACGTCCTGGTCCGCCTGTACCCGGAGTTGCGGTGGAGGTG	4800

B6	TACAAACCCGAAGGGTGACCTGTCAGAAGTTGAAAGCTTCTGGTATTTCTACCCCGGTGT	4860
MOLF	TACAAACCCGAAGGGTGACCTGTCAGAAGTTGAAAGCTTCTGGTATTTCTACCCCGGTGT	4860

B6	CCAATGACATGTGCAGTCAACTGGCCAAACGGCCTGTGGACACCCAGGCCTGCAACCAGC	4920
MOLF	CCAATGACATGTGCAGTCAACTGGCCAAACGGCCTGTGGACACCCAGGCCTGCAACCAGC	4920

B6	AGCTCTGTGTGGAATGGGCCTTCTCCAGCTGGGGCCAGTGCAATGGGCCTTGTCATAGGCC	4980
MOLF	AGCTCTGTGTGGAATGGGCCTTCTCCAGCTGGGGCCAGTGCAATGGGCCTTGTCATAGGCC	4980

B6	CTCGCCTTGCTGTGCAGCACAGGCAGGTCTTCTGCCAGACCCGGGATGGCATCACCTTAC	5040
MOLF	CTCGTCTTGCTGTGCAGCACAGACAGGTCTTCTGCCAGACCCGGGATGGCATCACCTTAC	5040
	**** *****	
B6	CATCAGAGCAGTGCAGTGCCCTCCCCAGGCCCGTGAGCACCCAAAAGCTGCTGGTCAGAGG	5100
MOLF	CATCAGAGCAGTGCAGTGCCCTCCCCAGGCCCGTGAGCACCCAAAAGCTGCTGGTCAGAGG	5100

B6	CCTGCAGTGTAACACTGGAGAGTCAGCCTCTGGACCCTATGCACGGCAACCTGTGGCAACT	5160
MOLF	CCTGCAGTGTAACACTGGAGGGTCAGCCTCTGGACCCTATGCACGGCAACCTGTGGCAACT	5160

B6	ATGGCTTCCAGTCCCGCGGGGTGGAGTGTGTTACAGTGCGCACCAACAAGGCTGTGCCCG	5220
MOLF	ATGGCTTCCAGTCCCGCGGGGTGGAGTGTGTTACAGTGCGCACCAACAAGGCTGTGCCCG	5220

B6	AGCACCTGTGTTCTTGGGGGCCCCGGCCCGCCAACCTGGCAACGCTGCAACGTCACCTCCAT	5280
MOLF	AGCACCTGTGTTCTTGGGGGCCCCGGCCCGCCAACCTGGCAACGCTGCAACGTCACCTCCAT	5280

B6	GTGAAAACACTGAGTGCAGAGACACCACCAGGTACTGCGAGAAGGTGAGACAGTTGAAAC	5340
MOLF	GTGAAAACACTGAGTGCAGAGACACCACCAGGTACTGCGAGAAGGTGAGACAGTTGAAAC	5340

15/24

Figure 7a

```
*****
B6      TCTGCCAGCTCGGCCAGTTCCGATCTCGCTGCTGTGGAACGTGTGGCAAAGCATGAAGAC 5400
MOLF    TCTGCCAGCTCGGCCAGTTCCGATCTCGCTGCTGTGGAACGTGTGGCAAAGCATGAAGAC 5400
*****

B6      AGGGCACGGGGATGAACTCTACCCTGGCCACACAAAGGACTCACACAACCACCTTGGACA 5460
MOLF    AGGGCACGGGGATGAACTCTACCCTGGCCACACAAAGGACTCACATAACCACCTTGGACA 5460
*****

B6      TGACTTACGCTTTCTTCCTTTTATTTATTTATTTGCCCCCTCCCCACACACCCATGTCCAC 5520
MOLF    TGACTTACGCTTTCTTCCTTTTATTTATTTATTTGCCCCCTCCCCACACACCCATGTCCAC 5520
*****

B6      CTGCACCCATGCTCCCGAGGACTGTGGCATGTTTTCTTATTTCAGTTAGCTGGAAGCTAGT 5580
MOLF    CTGCACCCATGCTCCCGAGGACTGTGGCATGTTTTCTTATTTCAGTTAGCTGGAAGCTAGT 5580
*****

B6      GTGGAGAACAGGAAAAAAAAAAAAAAAAAGATGTCTAAAGGAGGGTTGCAGAGCAGACAGCA 5640
MOLF    GTGGAGAACAGGAAAAAAAAAAAAAAAAAGATGTCTAAAGGAGGGTTGCAGAGCAGACAGCA 5640
*****

B6      GCTCCCTTGCAGAGCTGGCTCCTTTCTGATCAGGGTCTCAGAGACACAGGAAGACCAAGC 5700
MOLF    GCTCCCTTGCAGAGCTGGCTCCTTTCTGATCAGGGTCTCAGAGACACAGGAAGAGCAAGC 5700
*****

B6      CATAGTCTATAGGGACAGCAAGGAGTTTATAACACATAGGTACAGGCACAGGCCAATCAG 5760
MOLF    CATAGCCTATAGGGACAGCAAGGAGTTTATAACACATAGGTACAGGCACAGGCCAATCAG 5760
*****

B6      GAGTGCCGTTCCCTTTGTTTCCTGGTGGTAGGAACAGGGGGGGGGGGGGAAGCAGCATA 5820
MOLF    GAGTGCCG-----CAGCATA 5775
*****

B6      CTAGAGCAGAAACCAGGCCAGACTCCAGTTCTTCTGGCAGTGGTGTGGGAAGACCAAGAA 5880
MOLF    CTAGAGCAGAAACCAGGCCAGACTCCAGTTCTTCTGGCAGTGGTGTGGGAAGACCAAGAA 5835
*****

B6      AGAAGAGAACTGAGCTGTGGGCAAGAGGGTTATGCCAAAAGAGACAAGCAAAATAGGCAG 5940
MOLF    AGAAGAGAACTGAGCTGTGGGCAAGAGGGTTATGCCAAAAGAGACAAGCAAAATAGGCAG 5895
*****

B6      AAGGCATGCCGCTATGCAGAAGACAGAGGCCGGGAAGAGAGAAGGAAAGCGTGCGCACGC 6000
MOLF    AAGGCATGCCGCTATGCAGAAGACAGAGGCCGGGAAGAGAGAAGGAAAGCGTGCGCACGC 5955
*****

B6      ACCTCAGTGGGAGGAGCTTGCTGGCAGGACAGATTTTCAAGCAGTCACGTGAGACCCATA 6060
MOLF    ACCTCAGTGGGAGGAGCTTGCTGGCAGGACAGATTTTCAAGCA-----TA 6000
*****

B6      TGACAAGCCTGTGCCCCGGAACATGTGCACAGTCCTCGAGCTTCAACAGAGTTGGGGAAG 6120
MOLF    TGACAAGTCTGTGCCCCGGAACATGTGCACAGTCCTCGAGGTTCAACAGAGTTGGGGAAG 6060
*****

B6      GGTATTTCAGAGTCCTCTCAGGGTACCCCGCTGCGTGACACGGGTGACACTGCATGACAT 6180
MOLF    GGTATTTCAGAGTCCTCTCAGGGTACCCCGCTGCGTGACACGGGTGACACTGCATGACAT 6120
*****

B6      GCCCTTTGTGTGGTTATACAAGTGTGTCTACCCCTCTTCTGTCAAGTGACTTTTGTGTCT 6240
```

16/24

Figure 7a

MOLF	GCCCTTTGTGTGGTTATACAAGTGTGTCTACCCCTCTTCTGTCAAGTGACTTTTGTGTCT	6180

B6	GCAGCTGTGGACAACCCAACAAAGGGTCAAGGTCAAGTGTGCAGTAAGGCTGCCCCTAGC	6300
MOLF	GCAGCTGTGGACAACCCAACAAAGGGTCAAGGTCAAGTGTGCAGTAAGGCTGCCCCTAGC	6240

B6	ATCTCCCTACTTAAGTCAAGGGTCACTAGCCATGAAC'TGACTGCTGCCGAGGGCCAAGGG	6360
MOLF	ATCTCCCTACTTAAGTCAAGGGTCACTAGCCATGAAC'TGACTGCTGCCGAGGGCCAAGGG	6300

B6	CATTTGATCGTGAATGCCCTGGTTTTCATCCGAATGACTTTGCTGCTCCTTGTTGGGC	6420
MOLF	CATTTGACCGTGAATGCCCTGGTTTTCATCCGAATGACTTTGCTGCTCCTTGTTGGGC	6360

B6	CAGGGACTAGAATCATAGCTCACAAAGGATTCATTAAGCAACCTGTTTGC GGCAAGCAT	6480
MOLF	CAGGGACTAGAATCATAGCTCACAAAGGATTCATTAAGCAACCTGTTTGCAGGCAAGCAT	6420

B6	CTACAGGATGTAGAAAGCCTTACTACCAGATCTCACCAGAGAGGAGGAGGACGGAGGG	6540
MOLF	CTACAGGATGTGGGAAGCCTTACTACCAGATCTCACCAGAGAGGAGGAGGACGGAGGG	6480
	***** *	
B6	GCCCTGCACACCCACAGTGCCTTTCTCTCCCAACCCTCACACATTGGGCCTGGGAGCCC	6600
MOLF	GCCCTGCACACCCACAGTGCCTTTCTCTCCCAACCCTCACACATTGGGCCTGGGAGCCC	6540

B6	TTCTGAACCACAGAAGAGCTGCCA-----ATTTACGCAGCAGGCTGGGAGTAGGGA	6652
MOLF	TTCTGAACCACAGAAGAGCTGCCGTACACCCGATTTACGCAGCAGGCTGGGAGTAGGGA	6600

B6	GT-----GTGATCAGGCGCTATGGCAAAGCTAGAAGTCAGCTAAAGACTCACAAAGACA	6705
MOLF	GTAGGGAGTGTGATCAGGCGCTATGGCAAAGCTAGAAGTCAGCTAAAGACTCACAAAGACA	6660
	**	
B6	GCCGGAGAAAGCTTTGTAACCATCCTCTCGACTACACAGATACTATGGTAATTAGCACAC	6765
MOLF	GCCGGAGAAAGCTTTGTAACCATCCTCTCGACTACACAGATACTATGGTAATTAGCACAC	6720

B6	CAGACTGAGTGAATGGACCACTCAGCGGGCTAAGCAGCACACACAATTAGCAACACACA	6825
MOLF	CAGACTGAGTGAATGGACCACTCAGCGGGCTAAGCAGCACACACAATTAGCAACACACA	6780

B6	TCAGCTATGGCAAAAACAGAAGCCCTTGGCAAATCCTCAGTCTGCAGAGAAGCTTCAACA	6885
MOLF	TCAGCTATGGCAAAAACAGAAGCCCTTGGCAAATCCTCAGTCTGCAGAGAAGCTTCAACA	6840

B6	GTAATAATCGCTTCCTTCCTGAGCAGCCTCAGTTTAGCAGATGAGGCAGGCAGAGCCCAG	6945
MOLF	GTAATAATCGCTTCCTTCCTGAGCAGCCTCAGTTTAGCAGATGAGGCAGGCAGAGCCCAG	6900

B6	TAGCCATCAGATAAGCCGCCCCCTCGCAAACACACTCACCAAATGCTATCACCCCAGGGCC	7005
MOLF	TAGCCATCAGATAAGCCGCCCCCTCACAAACACACTCACCAAATGCTATCACCCCAGGGCC	6960

B6	ACCACTGACATCAGGGCTGCATTTCCCAACCAGCCTGGAAGTTAAGTTTGAAAGAGGATG	7065
MOLF	ACCACTGACATCAGGGCTGCATTTCCCAACCAGCCTGGAAGTTAAGTTTGAAAGAGGATG	7020

17/24

Figure 7a

```
B6      TTAATCTGTGTCTTCTTAGATGCGCTATGGACAGTTAGGGTAGGCCTCTTCCTCATCACA 7125
MOLF    TTAATCTGTGTCTTCTTAGATGCGCTATGGACAGTTAGGGTAGGCCTCTTCCTCATCACA 7080
        *****

B6      AAAATAGACTCTAAGAAATTGTGTTATTAAAAGTCAGTGTAAATGTTTATTAAAATTAAA 7185
MOLF    AAAATAGACTCTAAGAAATTGTGTTATTAAAAGTCAGTGTAAATGTTTATTAAAATTAAA 7140
        *****

B6      GAGAGTGTTTCTATGT 7201
MOLF    GAGAGTGTTTCTATGT 7156
        *****
```

18/24

Figure 7a

B6	MECCRAAPGTPLLVLAFLLSSRTARSEEDREGLWDAWGPWSECSRTC GGGASYS LRRC	60
MOLF	MECCRAAPGTPLLVLAFLLSSRTARSEEDREGLWDAWGPWSECSRTC GGGASYS LRRC	60

B6	LSSKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKYHGQLYEWLPVSNPDNP	120
MOLF	LSSKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKYHGQLYEWLPVSNPDNP	120
***** ; *****		
B6	LKCQAKGTSLVVELAPKVLDTGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVC	180
MOLF	LKCQAKGTSLVVELAPKVLDTGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVC	180

B6	DGSTCRLVRGQYKSQLSASKSDDTVVAIPYGSRHIRLVKGPDLHLYLETKTLQGT	240
MOLF	DGSTCRLVRGQYKSQLSASKSDDTVVAIPYGSRHIRLVKGPDLHLYLETKTLQGT	240

B6	LSSTGIFLVDNSTVDFQKLPDKEILRMTGPLTADFI IKI HDLGPADSTVQFIF	300
MOLF	LSSTGIFLVDNSTVDFQKLPDKEILRMTGPLTADFI IKI HDLGPADSTVQFIF	300

B6	WRETDFFPSCSATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENVKPKPKLQEC	360
MOLF	WRETDFFPSCSATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENVKPKPKLQEC	360

B6	ASDGYKQIMPYDLYHPLPRWEATPWTACSSSCGGGIQSRVSCVEEDIQGHVTSVEE	420
MOLF	ASDGYKQIMPYDLYHPLPRWEATPWTACSSSCGGGIQSRVSCVEEDIQGHVTSVEE	420

B6	MYTPKMPVVQPCNIFDCPKWLAQEWSPCTVTTCGQGLRYRVVLCIDHRGMHTGG	480
MOLF	MYTPKMPVVQPCNIFDCPKWLAQEWSPCTVTTCGQGLRYRVVLCIDHRGMHTGG	480

B6	HIKEECIIPTPCYKPREKLPIEAKLPWFKQAQEELEGA AVSEEPSFIPEAWSACT	540
MOLF	HIKEECIIPTPCYKPREKLPIEAKLPWFKQAQEELEGA AVSEEPSFIPEAWSACT	540

B6	GTQVRIIRCQVLLPFSQSVDLPADECEGPKPASQRACYAGPCNGETPEFNPNDG	600
MOLF	GTQVRIIRCQVLLPFSQSVDLPADECEGPKPASQRACYAGPCNGETPEFNPNDG	600

B6	GLQDLDELYDWEYEGFTKCSSESCGGGVQEAVVSCLNKQTRELADENLCVTSRR	660
MOLF	GLQDLDELYDWEYEGFTKCSSESCGGGVQEAVVSCLNKQTRELADENLCVTSRR	660

B6	CNLDPCPARWEIGKWSPCSLTGCVGLQTRDVVCSHLLSREMNEVVVLADELCH	720
MOLF	CNLDPCPARWEIGKWSPCSLTGCVGLQTRDVVCSHLLSREMNEVVVLADELCH	720

B6	QACNRFNCPPAWYPAQWQLCSRTC GGGIQKRDVLCQRMADGSFLELPETFC	780
MOLF	QACNRFNCPPAWYPAQWQLCSRTC GGGIQKRDVLCQRMADGSFLELPETFC	780

B6	QGCKKDDCPSEWLLSEWSECSVSCGEGTQTRSAICQRVLKTGVSTVVNSTLCP	840
MOLF	QGCKKDDCPSEWLLSEWSECSVSCGEGTQTRSAICQRVLKTGVSTVVNSTLCP	840

B6	IRPCMLATCARPGRPSTKHSPHIAAARNIYIQTRRQKRLHFVVGGFAYLLPKT	900
MOLF	IRPCMLATCARPGRPSTKHSPHIAAARNIYIQTRRQKRLHFVVGGFAYLLPKT	900

19/24

Figure 7a

```
*****
B6      TRRFRKPLITWEKDGQHSISSAHVTVPFGYLIKIHRLKPSDAGIYTCSAGPAREQFVIKL 960
MOLF    TRRFRKPLITWEKDGQHSISSAHVTVPFGYLIKIHRLKPSDAGIYTCSAGPAREQFVIKL 960
*****

B6      IGGNRKLVARPLSLWSEEEEEALQVRKTNPKALQTHKHQNGIFSNGSKAEKRGLTADPGN 1020
MOLF    IGGNRKLVARPLSLWSEEEEEALQVRKANPKALQTHKHQNGIFSNGSKAEKRGLTADPGN 1020
*****

B6      RYDDIVSRLLLEQGGWPGELLASWEVQDSAERNASSEEDPNAEQALLHLPFTMVAEQKRLD 1080
MOLF    RYDDIVSRLLLEQGGWPGELLASWEVQDSAERNASSEEDPNAEQALLHLPFTMVAEQKRLD 1080
*****

B6      DILRNLSSQQPEELRDLYSKHLVAQLAQDIFRSHLENQDLLPKPSEQRFFPMVPAHKHVS 1140
MOLF    DILRNLSSQQPEELRDLYSKHLVAQLAQDIFRSHLENQDLLPKPSEQRFFPMVPAHKHVS 1140
*****

B6      GFSSSLRSSSGEAGGGSRRPHRKPAILRKISAAQQLSASEVVTHLGQTVALASGTLVLL 1200
MOLF    GFSSSLRSSSGEAGGGSRRPHRKPAILRKISAAQQLSASEVVTHLGQTVALASGTLVLL 1200
*****

B6      HCEAVGNPRPTIHWTRNGEAVQFSDRILLQPDDSLQILAPVEADVGFYTCNATNALGYDS 1260
MOLF    HCEAVGNPRPTIHWTRNGEAVQFSDRILLQPDDSLQILAPVEADVGFYTCNATNALGYDS 1260
*****

B6      VSIATVLAGKPLVKTSRMTVLNTEKPTVTVDIGGTVRTVRGVNVTINCQVAGVPEAEVTW 1320
MOLF    VSIATVLAGKPLVKTSRMTVLNTEKPTVTVDIGGTVRTVRGVNVTINCQVAGVPEAEVTW 1320
*****

B6      FRNKSCLGSSSHHLHEGSSSHHLHEGSLLLTDVSFSDQGLYSCRAANLHGEQTESTQLLLILD 1380
MOLF    FRNKSCLGSSSHHLHEGSSSHHLHEGSLLLTDVSFSDQGLYSCRAANLHGEQTESTQLLLILD 1380
*****

B6      PPQVPTQLEDIRALLLATGPNLPSVLMSPGLTQLVLDPGNSALLGCPIKGHPPTNITWFQ 1440
MOLF    PPQVPTQLEDIRALLLATGPNLPSVLMSPGLTQLVLDPGNSALLGCPIKGHPPTNITWFQ 1440
*****

B6      NGQPIATAPGLTHHIWGAGQILRVANLSGGPQGEFSCLAQNEAGTLLQKASLVIQDYWWS 1500
MOLF    NGQPIATAPGLTHHIWGAGQILRVANLSGGPQGEFSCLAQNEAGTLLQKASLVIQDYWWS 1500
*****

B6      VDRLATCSASCGNRGIHQPLRLCLLNTTEVDPEHCTGKPRPAVQPVACNRRDCPSRWMVT 1560
MOLF    VDRLATCSASCGNRGIHQPLRLCLLNTTEVDPEHCTGKPRPAVQPVACNRRDCPSRWMVT 1560
*****

B6      SWSACTRSCGGGVQTRRVTCQKLKASGISTPVSNDMCSQLAKRPVDTQACNQQLCWEWAF 1620
MOLF    SWSACTRSCGGGVQTRRVTCQKLKASGISTPVSNDMCSQLAKRPVDTQACNQQLCWEWAF 1620
*****

B6      SSWGQCNGPCIGPRLAVQHRQVFCQTRDGITLPSEQCSALPRPVSTQNCWSEACSVHWRV 1680
MOLF    SSWGQCNGPCIGPRLAVQHRQVFCQTRDGITLPSEQCSALPRPVSTQNCWSEACSVHWRV 1680
*****

B6      SLWTLCTATCGNYGFQSRRECVHVRTNKAVPEHLCSWGPRPANWQRCNVTPCENTECCR 1740
MOLF    SLWTLCTATCGNYGFQSRRECVHVRTNKAVPEHLCSWGPRPANWQRCNVTPCENTECCR 1740
*****

B6      TTRYCEKVRQLKLCQLGQFRSRCCGTCGKA 1770
```

20/24

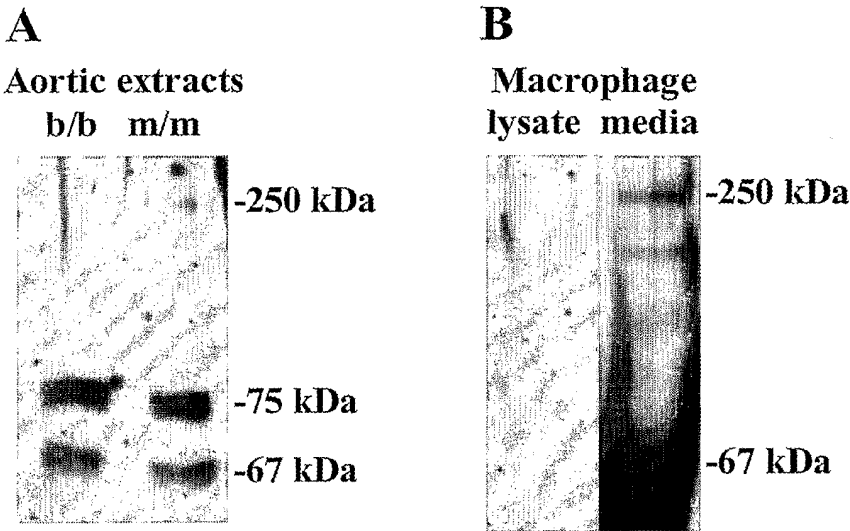
Figure 7a

MOLF

TTRYCEKVRQLKLCQLGQFRSRCCGTCGKA 1770

22/24

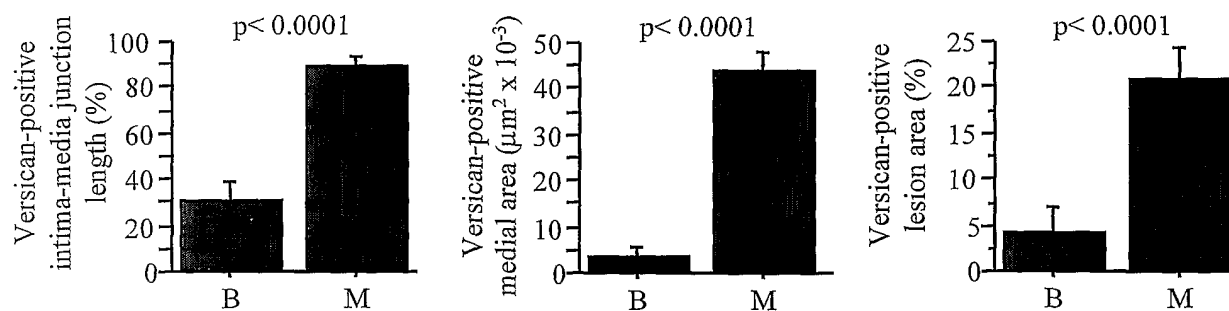
Figure 9



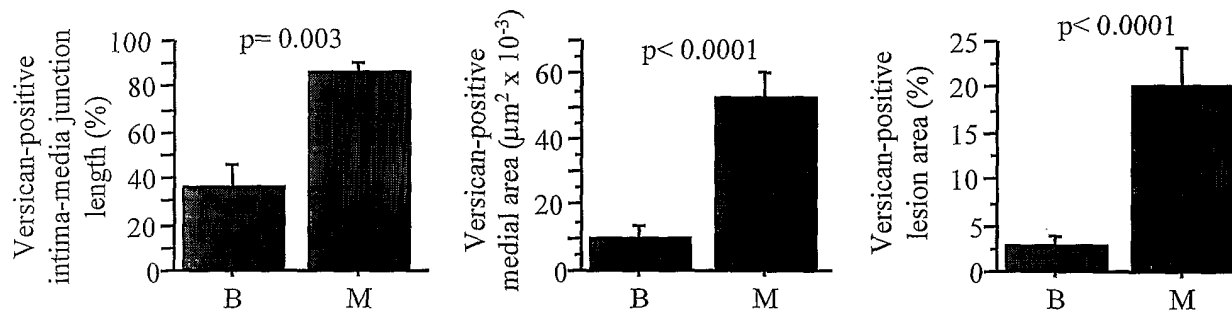
23/24

Figure 10

6-wk WTD

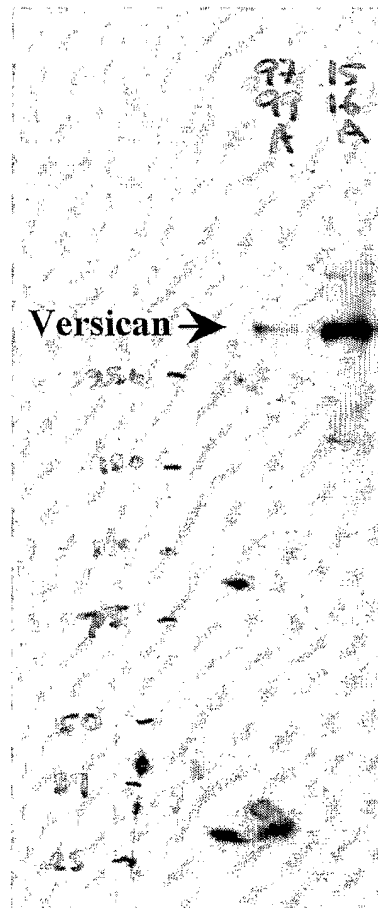


12-wk WTD



24/24

Figure 11



SEQUENCE LISTING

<110> THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK
TALL, Alan R.
WELCH, Carrie
SEIDELMANN, Sara Bretschger

<120> NOVEL GENE UNDERLYING ATHEROSCLEROSIS LOCUS I

<130> 72427-A-PCT/JPW/JW

<150> US 60/648,837
<151> 2005-02-01

<160> 12

<170> PatentIn version 3.3

<210> 1
<211> 7244
<212> DNA
<213> Homo sapiens

<400> 1
 tgcctgctcc gagagagatt ccattcagct tccccccac ccatccacc acccaccct 60
 cggtcaggaa atgtgagagg ggctgatgga agctgatagg caggactgga gtgttagcac 120
 cagtactgga tgtgacagca ggcagaggag cacttagcag cttattcagt gtccgattct 180
 gattccggca aggatccaag catggaatgc tgcgctcggg caactcctgg cacactgctc 240
 ctctttcttg ctttcctgct cctgagttcc aggaccgcac gctccgagga ggaccgggac 300
 ggccctatggg atgcctgggg cccatggagt gaatgctcac gcacctgogg gggtagggcc 360
 tcctactctc tgaggcgctg cctgagcagc aagagctgtg aaggaagaaa tatccgatac 420
 agaacatgca gtaatgtgga ctgcccacca gaagcagggtg atttccgagc tcagcaatgc 480
 tcagctcata atgatgtcaa gcaccatggc cagttttatg aatggcttcc tgtgtctaat 540
 gacctgaca acccatgttc actcaagtgc caagccaaag gaacaaccct ggttgttgaa 600
 ctagcaccta aggtcttaga tggtagcgtg tgctatacag aatctttgga tatgtgcac 660
 agtggtttat gccaaattgt tggctgcgat caccagctgg gaagcaccgt caaggaagat 720
 aactgtgggg tctgcaacgg agatgggtcc acctgcgggc tggtcogagg gcagtataaa 780
 tccagctct ccgcaaccaa atcggatgat actgtggttg caattcccta tggaagtaga 840
 catattogcc ttgtcttaaa aggtcctgat cacttatatc tggaaaccaa aacctccag 900
 gggactaaag gtgaaaacag totcagctcc acaggaactt tccttgtgga caattctagt 960
 gtggacttcc agaaatttcc agacaaagag atactgagaa tggctggacc actcacagca 1020

gatttcattg tcaagattcg taactcgggc tccgctgaca gtacagtcca gttcatcttc	1080
tatcaaccca tcatccaccg atggagggag acggattttct ttctttgctc agcaacctgt	1140
ggaggagggt atcagctgac atcggctgag tgctacgata tgaggagcaa ccgtgtgggt	1200
gctgaccaat actgtcacta ttaccagag aacatcaaac ccaaacccaa gcttcaggag	1260
tgcaacttgg atccttgtcc agccagtgac ggatacaagc agatcatgcc ttatgaacctc	1320
taccatcccc ttctcgggtg ggaggccacc ccatggaccg cgtgctctc ctctgtggg	1380
gggggcatcc agagccgggc agtttctgt gtggaggagg acatccaggg gcatgtcact	1440
tcagtggaag agtggaagt catgtacacc cctaagatgc ccatogcgca gccctgcaac	1500
atcttttgact gccctaaatg gctggcacag gagtggctctc cgtgcacagt gacatgtggc	1560
cagggcctca gataccgtgt ggtcctctgc atcgaccatc gaggaatgca cacaggaggc	1620
tgtagcccaa aaacaaagcc ccacataaaa gaggaatgca tcgtaccac tccctgctat	1680
aaacccaaag agaaacttcc agtcgaggcc aagttgccat ggttcaaaca agctcaagag	1740
ctagaagaag gagctgctgt gtcagaggag ccctcgttca tcccaggagg ctggtcggcc	1800
tgcacagtca cctgtggtgt ggggaccag gtgcgaatag tcagggtgca ggtgctcctg	1860
tctttctctc agtccgtggc tgacctgct attgaagagt gtgaagggcc caagccagca	1920
tcccagcgtg cctgttatgc aggccatgc agcggggaaa ttcttgagtt caaccagac	1980
gagacagatg ggctcttttg tggcctgcag gatttogacg agctgtatga ctgggagtat	2040
gaggggttca ccaagtgtc cgagtcctgt ggaggagggtg tccaggaggc tgtggtgagc	2100
tgcttgaaca aacagactcg ggagcctgct gaggagaacc tgtgcgtgac cagccgccgg	2160
ccccacagc tctgaagtc ctgcaatttg gatccctgcc cagcaagggtg ggaaattggc	2220
aagtggagtc catgtagtct cacatgtggg gtgggctac agaccagaga cgtcttctgc	2280
agccacctgc ttccagaga gatgaatgaa acagtcatcc tggctgatga gctgtgtcgc	2340
cagcccaagc ccagcacggt gcaagcttgt aaccgcttta attgcccccc agcctgggtac	2400
cctgcacagt ggcagccgtg ttocagaacg tgtggcgggg gtgttcagaa acgtgagggt	2460
ctttgcaagc agcgcattgg tgatggcagc ttctggagc ttctgagac cttctgttca	2520
gcttcaaaac ctgcctgcca gcaagcatgc aagaaagatg actgtcccag cgagtggctt	2580
ctctcagact ggacagagtg ttocacaagc tgoggggaag gcaccagac tcgaagcgcc	2640
atttgccgaa agatgctgaa aaccggcatc tcaacggttg tcaattccac cctgtgccc	2700

cccctgcctt	tctcttcctc	catcaggccc	tgtatgctgg	caacctgtgc	aaggcccggg	2760
eggccatcca	cgaagcacag	cccgcacatc	gcgcccgcca	ggaaggtcta	catccagact	2820
cgcaggcaga	ggaagctgca	cttcgtgggtg	gggggcttcg	cctacctgct	ccccaaagacg	2880
gcggtggtgc	tgcgctgccc	ggcgcgcagg	gtccgcaagc	ccctcatcac	ctggggagaag	2940
gacggccagc	acctcatcag	ctcgacgcac	gtcacggtgg	cccccttcgg	ctatctcaag	3000
atccaccgcc	tcaagccctc	ggatgcaggc	gtctacacct	gctcagcggg	ccgggcccgg	3060
gagcactttg	tgattaagct	catcggaggc	aaccgcaagc	tcgtggcccg	gcccttgagc	3120
ccgagaagtg	aggaagaggt	gcttgcgggg	aggaagggcg	gcccgaaggga	ggccctgcag	3180
accacaaaac	accagaacgg	gatcttctcc	aacggcagca	aggcggagaa	gcggggcctg	3240
gccgccaacc	cggggagccg	ctacgacgac	ctcgtctccc	ggctgctgga	gcagggcggc	3300
tggcccggag	agctgctggc	ctcgtgggag	gcgcaggact	ccgcggaaaag	gaacacgacc	3360
tcggaggagg	acccgggtgc	agagcaagtg	ctcctgcacc	tgcccttcac	catggtgacc	3420
gagcagcggc	gcctggacga	catcctgggg	aacctctccc	agcagcccga	ggagctgcgc	3480
gacctctaca	gcaagcacct	ggtggcccag	ctggcccagg	agatcttcg	cagccacctg	3540
gagcaccagg	acacgctcct	gaagccctcg	gagcgcagga	cttccccagt	gactctctcg	3600
cctcataaac	acgtgtctgg	cttcagcagc	tccttcggga	cctcctccac	cggggacgcc	3660
gggggaggct	ctcgaaggcc	acaccgcaag	cccaccatcc	tgcgcaagat	ctcagcggcc	3720
cagcagctct	cagcctcgga	ggtggtcacc	cacctggggc	agacggtggc	cctggccagc	3780
gggacactga	gtgttcttct	gcactgtgag	gccatcggcc	acccaaggcc	taccatcagc	3840
tggggccagga	atggagaaga	agttcagttc	agtgacagga	ttcttctaca	gccagatgat	3900
tccttacaga	tcttggcacc	agtggaaagca	gatgtgggtt	tctacacttg	caatgccacc	3960
aatgccttgg	gatacgactc	tgtctccatt	gccgtcacat	tagcaggaaa	gccactagtg	4020
aaaacgtcac	gaatgacagt	gatcaacacg	gagaagcctg	cagtcacagt	cgatatagga	4080
agcaccatca	aaacagtgca	gggagtgaat	gtgacaatca	actgccagggt	tgcaggagtg	4140
cctgaagctg	aagtcaacttg	gttcaggaat	aaaagcaaac	tgggctcccc	gcaccatctg	4200
cacgaaggct	ccttgctgct	cacaaaogtg	tcctcctcgg	atcagggcct	gtactcctgc	4260
agggcgccca	atcttcatgg	agagctgact	gagagcacc	agctgctgat	cctagatccc	4320
ccccaaagtcc	ccacacagtt	ggaagacatc	agggccttgc	tcgctgccac	tggaccgaac	4380
cttccttcag	tgctgacgtc	tcctctggga	acacagctgg	tcctggatcc	tgggaattct	4440

gctctccttg gctgccccat caaaggtcac cctgtcccta atatcacctg gtttcatggt	4500
ggtcagccaa ttgtcactgc cacaggactg acgcatacaca tcttggcagc tggacagatc	4560
cttcaagttg caaaccttag cgggtgggtct caaggggaat tcagctgoot tgctcagaat	4620
gaggcagggg tgctcatgca gaaggcatct ttagtgatcc aagattactg gtggtctgtg	4680
gacagactgg caacctgctc agcctcctgt ggtaacoggg gggttcagca gccccgcttg	4740
aggtgcctgc tgaacagcac ggagggtcaac cctgcccact ggcagggaa ggttcgcct	4800
gcggtgcagc ccatcgctg caaccggaga gactgccctt ctgggtggat ggtgacctcc	4860
tggctctgcct gtaccocggag ctgtggggga ggtgtccaga cccgcagggg gacctgtcaa	4920
aagctgaaag cctctgggat ctccaccctt gtgtccaatg acatgtgcac ccaggtcgcc	4980
aagcggcctg tggacacca ggcctgtaac cagcagctgt gtgtggagtg ggccttctcc	5040
agctggggcc agtgcaatgg gccttgcatc gggcctcacc tagctgtgca acacagacaa	5100
gtcttctgcc agacacggga tggcatcacc ttaccatcag agcagtgcag tgctcttccg	5160
aggcctgtga gcaccagaa ctgctgggtca gaggcctgca gtgtacactg gagagtcagc	5220
ctgtggaccc tgtgcacagc tacctgtggc aactacggct tccagtcccg gcgtgtggag	5280
tgtgtgcatg cccgcaccaa caaggcagt cctgagcacc tgtgtctctg ggggccccgg	5340
cctgccaaact ggcagcgtg caacatcacc ccatgtgaaa acatggagtg cagagacacc	5400
accaggtact gcgagaaggt gaaacagctg aaactctgcc aactcagcca gtttaaactt	5460
cgctgctgtg gaacttgtgg caaagcgtga agatagggtg tggggaaaaa ctctaccctg	5520
gccacacgaa ggactcacgc aaccacctcg gacagaacct aagctttott cattttattt	5580
atttatattcc ccctccccac tccacacaca cccttccaac ctctccacc tccaccttca	5640
agcataagga cgtccgctg ttttctcttt cagttagctg gaggacagga tgttgggaaa	5700
ggaaaggaca gatgtctaaa ggaggttgca gagcaggcca ggcagacagt gggggctccc	5760
ttgaagagct tcctccctcc caaacctggg tctcaaagac ctagaaagag gcaggcacag	5820
cccctgcgga cagcagggag ccagaagggt tgtagcctat tggtgcaaac attggacaaa	5880
ttcctgtgtc tttcctagaa gcgcactatc acaaacacag gagtgttttg ctcccttgtc	5940
tcctcttccc catctatgtc cctttagtca cagttaggac aaatggggag gggacaccat	6000
gctgaggcag aaactagccc agaactcact cagttcttct agtgggtgag tgcagagaga	6060
gaagaactca gatcaccagt agggagaggt aaaaaagcaa acaaagcagg ctctaaggca	6120

cacaacattg cagaaaaatga ggaagggagg ggaggggaagg gacagaagca aaaaggagcc 6180
 tgtggtgttc ccagtgggg cagggtgagc aggggcttcc aggctgcatg aggctcatgg 6240
 accagctctg atcccatgca tgtgcgcatg ctccagagccc tgctgcccac aacagagcac 6300
 tgcgctgcgt gggagtcccc acttcccaag ctatcagagt caacgtcctg cctgtgcagc 6360
 tgcagcaaag ccagtgagag gtgggtctcg ccatgcagta aggccaccct ggcacctctt 6420
 tatctaaatc cgaagtcccc tagccccgca ctaactaact gctgctgtgg gccagggcca 6480
 ttttgagcat gaatggcca ggttttttgc cttctaggac ctttgcctgt ccaccgaagg 6540
 gccagggact atggttaact tatcaacatc aaccattaa ctagtactg tgccagagag 6600
 tatctgtcag gctgtcaggc tgtcagggtg tagcaacctc ttcattocag agctagccca 6660
 gggaccgggg tgggacaatg ggtttatgcg tgtccacagt acaccctccc tctcccagcc 6720
 tccaccccag ggtctgcagg tcctccggca tgtagtattt atctagcaag gcgggggtgt 6780
 ggaggcagca ccctggcaaa gcagctcaca cactgcagcc aactcatca gctgtggtga 6840
 ggcggtgga gcaaagtcaa agtcatgcag caaaatgaaa actctgggac tottoggcaa 6900
 aatcctcatt aagccgagca gctttggcca agtaattttt gcctccttcc ctogcgtggc 6960
 ctgagtttag gagcaagggg gccagagtc ccttaccac agataagcct cccctcatga 7020
 aatgccactc accccgggct accattgaca tcagggtgct atttccagcc agcctggaag 7080
 taaaatttga gaggaagaca atattaatct gtgtccccac ctagtgagct gtggacaggt 7140
 ttaagttggg tctccttctt cttcaccaca aaaacaggct ctaagaaatc atgttactaa 7200
 aaaatcagtg taaagtctgt ttaaaataaa aaagaatgtt ttct 7244

<210> 2
 <211> 1762
 <212> PRT
 <213> Homo sapiens

<400> 2

Met Glu Cys Cys Arg Arg Ala Thr Pro Gly Thr Leu Leu Leu Phe Leu
 1 5 10 15

Ala Phe Leu Leu Leu Ser Ser Arg Thr Ala Arg Ser Glu Glu Asp Arg
 20 25 30

Asp Gly Leu Trp Asp Ala Trp Gly Pro Trp Ser Glu Cys Ser Arg Thr
 35 40 45

Cys Gly Gly Gly Ala Ser Tyr Ser Leu Arg Arg Cys Leu Ser Ser Lys
50 55 60

Ser Cys Glu Gly Arg Asn Ile Arg Tyr Arg Thr Cys Ser Asn Val Asp
65 70 75 80

Cys Pro Pro Glu Ala Gly Asp Phe Arg Ala Gln Gln Cys Ser Ala His
85 90 95

Asn Asp Val Lys His His Gly Gln Phe Tyr Glu Trp Leu Pro Val Ser
100 105 110

Asn Asp Pro Asp Asn Pro Cys Ser Leu Lys Cys Gln Ala Lys Gly Thr
115 120 125

Thr Leu Val Val Glu Leu Ala Pro Lys Val Leu Asp Gly Thr Arg Cys
130 135 140

Tyr Thr Glu Ser Leu Asp Met Cys Ile Ser Gly Leu Cys Gln Ile Val
145 150 155 160

Gly Cys Asp His Gln Leu Gly Ser Thr Val Lys Glu Asp Asn Cys Gly
165 170 175

Val Cys Asn Gly Asp Gly Ser Thr Cys Arg Leu Val Arg Gly Gln Tyr
180 185 190

Lys Ser Gln Leu Ser Ala Thr Lys Ser Asp Asp Thr Val Val Ala Ile
195 200 205

Pro Tyr Gly Ser Arg His Ile Arg Leu Val Leu Lys Gly Pro Asp His
210 215 220

Leu Tyr Leu Glu Thr Lys Thr Leu Gln Gly Thr Lys Gly Glu Asn Ser
225 230 235 240

Leu Ser Ser Thr Gly Thr Phe Leu Val Asp Asn Ser Ser Val Asp Phe
245 250 255

Gln Lys Phe Pro Asp Lys Glu Ile Leu Arg Met Ala Gly Pro Leu Thr
260 265 270

Ala Asp Phe Ile Val Lys Ile Arg Asn Ser Gly Ser Ala Asp Ser Thr
275 280 285

Val Gln Phe Ile Phe Tyr Gln Pro Ile Ile His Arg Trp Arg Glu Thr
290 295 300

Asp Phe Phe Pro Cys Ser Ala Thr Cys Gly Gly Gly Tyr Gln Leu Thr
305 310 315 320

Ser Ala Glu Cys Tyr Asp Leu Arg Ser Asn Arg Val Val Ala Asp Gln
325 330 335

Tyr Cys His Tyr Tyr Pro Glu Asn Ile Lys Pro Lys Pro Lys Leu Gln
340 345 350

Glu Cys Asn Leu Asp Pro Cys Pro Ala Ser Asp Gly Tyr Lys Gln Ile
355 360 365

Met Pro Tyr Asp Leu Tyr His Pro Leu Pro Arg Trp Glu Ala Thr Pro
370 375 380

Trp Thr Ala Cys Ser Ser Ser Cys Gly Gly Gly Ile Gln Ser Arg Ala
385 390 395 400

Val Ser Cys Val Glu Glu Asp Ile Gln Gly His Val Thr Ser Val Glu
405 410 415

Glu Trp Lys Cys Met Tyr Thr Pro Lys Met Pro Ile Ala Gln Pro Cys
420 425 430

Asn Ile Phe Asp Cys Pro Lys Trp Leu Ala Gln Glu Trp Ser Pro Cys
435 440 445

Thr Val Thr Cys Gly Gln Gly Leu Arg Tyr Arg Val Val Leu Cys Ile
450 455 460

Asp His Arg Gly Met His Thr Gly Gly Cys Ser Pro Lys Thr Lys Pro
465 470 475 480

His Ile Lys Glu Glu Cys Ile Val Pro Thr Pro Cys Tyr Lys Pro Lys
485 490 495

Glu Lys Leu Pro Val Glu Ala Lys Leu Pro Trp Phe Lys Gln Ala Gln

500					505					510						
Glu	Leu	Glu	Glu	Gly	Ala	Ala	Val	Ser	Glu	Glu	Pro	Ser	Phe	Ile	Pro	
515					520					525						
Glu	Ala	Trp	Ser	Ala	Cys	Thr	Val	Thr	Cys	Gly	Val	Gly	Thr	Gln	Val	
530					535					540						
Arg	Ile	Val	Arg	Cys	Gln	Val	Leu	Leu	Ser	Phe	Ser	Gln	Ser	Val	Ala	
545					550					555					560	
Asp	Leu	Pro	Ile	Asp	Glu	Cys	Glu	Gly	Pro	Lys	Pro	Ala	Ser	Gln	Arg	
565					570					575						
Ala	Cys	Tyr	Ala	Gly	Pro	Cys	Ser	Gly	Glu	Ile	Pro	Glu	Phe	Asn	Pro	
580					585					590						
Asp	Glu	Thr	Asp	Gly	Leu	Phe	Gly	Gly	Leu	Gln	Asp	Phe	Asp	Glu	Leu	
595					600					605						
Tyr	Asp	Trp	Glu	Tyr	Glu	Gly	Phe	Thr	Lys	Cys	Ser	Glu	Ser	Cys	Gly	
610					615					620						
Gly	Gly	Val	Gln	Glu	Ala	Val	Val	Ser	Cys	Leu	Asn	Lys	Gln	Thr	Arg	
625					630					635					640	
Glu	Pro	Ala	Glu	Glu	Asn	Leu	Cys	Val	Thr	Ser	Arg	Arg	Pro	Pro	Gln	
645					650					655						
Leu	Leu	Lys	Ser	Cys	Asn	Leu	Asp	Pro	Cys	Pro	Ala	Arg	Trp	Glu	Ile	
660					665					670						
Gly	Lys	Trp	Ser	Pro	Cys	Ser	Leu	Thr	Cys	Gly	Val	Gly	Leu	Gln	Thr	
675					680					685						
Arg	Asp	Val	Phe	Cys	Ser	His	Leu	Leu	Ser	Arg	Glu	Met	Asn	Glu	Thr	
690					695					700						
Val	Ile	Leu	Ala	Asp	Glu	Leu	Cys	Arg	Gln	Pro	Lys	Pro	Ser	Thr	Val	
705					710					715					720	
Gln	Ala	Cys	Asn	Arg	Phe	Asn	Cys	Pro	Pro	Ala	Trp	Tyr	Pro	Ala	Gln	
725					730					735						

Trp Gln Pro Cys Ser Arg Thr Cys Gly Gly Gly Val Gln Lys Arg Glu
740 745 750

Val Leu Cys Lys Gln Arg Met Ala Asp Gly Ser Phe Leu Glu Leu Pro
755 760 765

Glu Thr Phe Cys Ser Ala Ser Lys Pro Ala Cys Gln Gln Ala Cys Lys
770 775 780

Lys Asp Asp Cys Pro Ser Glu Trp Leu Leu Ser Asp Trp Thr Glu Cys
785 790 795 800

Ser Thr Ser Cys Gly Glu Gly Thr Gln Thr Arg Ser Ala Ile Cys Arg
805 810 815

Lys Met Leu Lys Thr Gly Ile Ser Thr Val Val Asn Ser Thr Leu Cys
820 825 830

Pro Pro Leu Pro Phe Ser Ser Ser Ile Arg Pro Cys Met Leu Ala Thr
835 840 845

Cys Ala Arg Pro Gly Arg Pro Ser Thr Lys His Ser Pro His Ile Ala
850 855 860

Ala Ala Arg Lys Val Tyr Ile Gln Thr Arg Arg Gln Arg Lys Leu His
865 870 875 880

Phe Val Val Gly Gly Phe Ala Tyr Leu Leu Pro Lys Thr Ala Val Val
885 890 895

Leu Arg Cys Pro Ala Arg Arg Val Arg Lys Pro Leu Ile Thr Trp Glu
900 905 910

Lys Asp Gly Gln His Leu Ile Ser Ser Thr His Val Thr Val Ala Pro
915 920 925

Phe Gly Tyr Leu Lys Ile His Arg Leu Lys Pro Ser Asp Ala Gly Val
930 935 940

Tyr Thr Cys Ser Ala Gly Pro Ala Arg Glu His Phe Val Ile Lys Leu
945 950 955 960

Ile Gly Gly Asn Arg Lys Leu Val Ala Arg Pro Leu Ser Pro Arg Ser
 965 970 975

Glu Glu Glu Val Leu Ala Gly Arg Lys Gly Gly Pro Lys Glu Ala Leu
 980 985 990

Gln Thr His Lys His Gln Asn Gly Ile Phe Ser Asn Gly Ser Lys Ala
 995 1000 1005

Glu Lys Arg Gly Leu Ala Ala Asn Pro Gly Ser Arg Tyr Asp Asp
 1010 1015 1020

Leu Val Ser Arg Leu Leu Glu Gln Gly Gly Trp Pro Gly Glu Leu
 1025 1030 1035

Leu Ala Ser Trp Glu Ala Gln Asp Ser Ala Glu Arg Asn Thr Thr
 1040 1045 1050

Ser Glu Glu Asp Pro Gly Ala Glu Gln Val Leu Leu His Leu Pro
 1055 1060 1065

Phe Thr Met Val Thr Glu Gln Arg Arg Leu Asp Asp Ile Leu Gly
 1070 1075 1080

Asn Leu Ser Gln Gln Pro Glu Glu Leu Arg Asp Leu Tyr Ser Lys
 1085 1090 1095

His Leu Val Ala Gln Leu Ala Gln Glu Ile Phe Arg Ser His Leu
 1100 1105 1110

Glu His Gln Asp Thr Leu Leu Lys Pro Ser Glu Arg Arg Thr Ser
 1115 1120 1125

Pro Val Thr Leu Ser Pro His Lys His Val Ser Gly Phe Ser Ser
 1130 1135 1140

Ser Leu Arg Thr Ser Ser Thr Gly Asp Ala Gly Gly Gly Ser Arg
 1145 1150 1155

Arg Pro His Arg Lys Pro Thr Ile Leu Arg Lys Ile Ser Ala Ala
 1160 1165 1170

Gln	Gln	Leu	Ser	Ala	Ser	Glu	Val	Val	Thr	His	Leu	Gly	Gln	Thr
1175						1180					1185			
Val	Ala	Leu	Ala	Ser	Gly	Thr	Leu	Ser	Val	Leu	Leu	His	Cys	Glu
1190						1195					1200			
Ala	Ile	Gly	His	Pro	Arg	Pro	Thr	Ile	Ser	Trp	Ala	Arg	Asn	Gly
1205						1210					1215			
Glu	Glu	Val	Gln	Phe	Ser	Asp	Arg	Ile	Leu	Leu	Gln	Pro	Asp	Asp
1220						1225					1230			
Ser	Leu	Gln	Ile	Leu	Ala	Pro	Val	Glu	Ala	Asp	Val	Gly	Phe	Tyr
1235						1240					1245			
Thr	Cys	Asn	Ala	Thr	Asn	Ala	Leu	Gly	Tyr	Asp	Ser	Val	Ser	Ile
1250						1255					1260			
Ala	Val	Thr	Leu	Ala	Gly	Lys	Pro	Leu	Val	Lys	Thr	Ser	Arg	Met
1265						1270					1275			
Thr	Val	Ile	Asn	Thr	Glu	Lys	Pro	Ala	Val	Thr	Val	Asp	Ile	Gly
1280						1285					1290			
Ser	Thr	Ile	Lys	Thr	Val	Gln	Gly	Val	Asn	Val	Thr	Ile	Asn	Cys
1295						1300					1305			
Gln	Val	Ala	Gly	Val	Pro	Glu	Ala	Glu	Val	Thr	Trp	Phe	Arg	Asn
1310						1315					1320			
Lys	Ser	Lys	Leu	Gly	Ser	Pro	His	His	Leu	His	Glu	Gly	Ser	Leu
1325						1330					1335			
Leu	Leu	Thr	Asn	Val	Ser	Ser	Ser	Asp	Gln	Gly	Leu	Tyr	Ser	Cys
1340						1345					1350			
Arg	Ala	Ala	Asn	Leu	His	Gly	Glu	Leu	Thr	Glu	Ser	Thr	Gln	Leu
1355						1360					1365			
Leu	Ile	Leu	Asp	Pro	Pro	Gln	Val	Pro	Thr	Gln	Leu	Glu	Asp	Ile
1370						1375					1380			
Arg	Ala	Leu	Leu	Ala	Ala	Thr	Gly	Pro	Asn	Leu	Pro	Ser	Val	Leu

1385		1390		1395
Thr Ser Pro Leu Gly Thr Gln Leu Val Leu Asp Pro Gly Asn Ser				
1400		1405		1410
Ala Leu Leu Gly Cys Pro Ile Lys Gly His Pro Val Pro Asn Ile				
1415		1420		1425
Thr Trp Phe His Gly Gly Gln Pro Ile Val Thr Ala Thr Gly Leu				
1430		1435		1440
Thr His His Ile Leu Ala Ala Gly Gln Ile Leu Gln Val Ala Asn				
1445		1450		1455
Leu Ser Gly Gly Ser Gln Gly Glu Phe Ser Cys Leu Ala Gln Asn				
1460		1465		1470
Glu Ala Gly Val Leu Met Gln Lys Ala Ser Leu Val Ile Gln Asp				
1475		1480		1485
Tyr Trp Trp Ser Val Asp Arg Leu Ala Thr Cys Ser Ala Ser Cys				
1490		1495		1500
Gly Asn Arg Gly Val Gln Gln Pro Arg Leu Arg Cys Leu Leu Asn				
1505		1510		1515
Ser Thr Glu Val Asn Pro Ala His Cys Ala Gly Lys Val Arg Pro				
1520		1525		1530
Ala Val Gln Pro Ile Ala Cys Asn Arg Arg Asp Cys Pro Ser Arg				
1535		1540		1545
Trp Met Val Thr Ser Trp Ser Ala Cys Thr Arg Ser Cys Gly Gly				
1550		1555		1560
Gly Val Gln Thr Arg Arg Val Thr Cys Gln Lys Leu Lys Ala Ser				
1565		1570		1575
Gly Ile Ser Thr Pro Val Ser Asn Asp Met Cys Thr Gln Val Ala				
1580		1585		1590
Lys Arg Pro Val Asp Thr Gln Ala Cys Asn Gln Gln Leu Cys Val				
1595		1600		1605

Glu Trp Ala Phe Ser Ser Trp Gly Gln Cys Asn Gly Pro Cys Ile
1610 1615 1620

Gly Pro His Leu Ala Val Gln His Arg Gln Val Phe Cys Gln Thr
1625 1630 1635

Arg Asp Gly Ile Thr Leu Pro Ser Glu Gln Cys Ser Ala Leu Pro
1640 1645 1650

Arg Pro Val Ser Thr Gln Asn Cys Trp Ser Glu Ala Cys Ser Val
1655 1660 1665

His Trp Arg Val Ser Leu Trp Thr Leu Cys Thr Ala Thr Cys Gly
1670 1675 1680

Asn Tyr Gly Phe Gln Ser Arg Arg Val Glu Cys Val His Ala Arg
1685 1690 1695

Thr Asn Lys Ala Val Pro Glu His Leu Cys Ser Trp Gly Pro Arg
1700 1705 1710

Pro Ala Asn Trp Gln Arg Cys Asn Ile Thr Pro Cys Glu Asn Met
1715 1720 1725

Glu Cys Arg Asp Thr Thr Arg Tyr Cys Glu Lys Val Lys Gln Leu
1730 1735 1740

Lys Leu Cys Gln Leu Ser Gln Phe Lys Ser Arg Cys Cys Gly Thr
1745 1750 1755

Cys Gly Lys Ala
1760

<210> 3

<211> 7201

<212> DNA

<213> mouse

<400> 3

agcactggat gtgacagcag cgggcagagg agcacttagc agcttattca gtgtccgatt 60

cggattccgg caaagatcca agtatggaat gctgcgctcg ggcagctcct ggcacaccgc 120

tgctogttct ggcttttcta ctcttgagtt ccaggactgc acgctcagag gaagaccgtg 180

aaggcctctg	ggatgcctgg	ggcccttggga	gcgagtgttc	cagaacctgt	ggtgggggtg	240
cctcctattc	cctgagacgc	tgccctgagca	gcaagagctg	tgaagggaga	aatattcgat	300
atagaacatg	cagtaatgtg	gactgcccac	cagaagcagg	tgattttccga	gctcaacagt	360
gttctgctca	taacgacgtc	aagtaccatg	gacagcttta	tgaatggctt	cctgtatcta	420
atgaccccg	caatccatgc	tcactcaagt	gccaagcaaa	aggaaccagc	ctggttgtgg	480
aactagcacc	gaaggtctta	gatggtacgc	gatgctatac	agaatcactg	gatatgtgca	540
tcagcggcct	atgccaaatt	gttggctgtg	atcaccagct	ggggagcact	gtcaaagagg	600
ataactgtgg	agtgtgcaac	ggagatggct	caacgtgccg	gctggtccga	gggcagtata	660
aatctcagct	ctctgcaagt	aaatcggatg	acaccgtggt	cgccattcct	tacggcagca	720
gacacatccg	ccttgtctta	aaagggccag	atcatttgta	tctggagacc	aagaccctcc	780
agggaaactaa	aggtgaaaat	agtctgagct	ctacaggcat	tttccttggtg	gacaattcta	840
cagtggactt	ccagaaaactt	ccagacaaag	agatactgag	gatgactggg	ccactgacag	900
cagatttcat	catcaagatc	catgacttgg	gacctgcaga	cagcacccgtc	cagttcatct	960
tctaccagcc	catcatccac	cgatggaggg	aaacagactt	ctttccttgc	tcagcaacgt	1020
gtggtggagg	ctaccagctg	acgtctgccg	agtgtatga	tctacggagt	aaccgtgtgg	1080
tggccgacca	gtactgtcac	tactaccccg	agaacgtcaa	gccccaaacc	aagcttcagg	1140
agtgcaactt	ggatccttgt	ccagccagtg	acggatacaa	gcagatcatg	cottatgacc	1200
tctaccatcc	ccttcctcgg	tgggaggcca	ccccgtggac	cgctgtctcc	tctcatgtg	1260
gggggggcat	ccagagccgg	gcagtttcct	gtgtggagga	ggacatccag	gggcatgtca	1320
cttcagtga	agagtggaaa	tgcatgtaca	ccccaaagat	gccagtogtg	cagccctgca	1380
acatttttga	ctgccctaaa	tggctggcac	aggagtggtc	tccgtgcaca	gtgacctgtg	1440
gccagggcct	cagataccgc	gtggtactct	gcattgacca	ccgaggaatg	cacacaggag	1500
gctgcagtgc	taaaacaaaag	ccccacataa	aagaagaatg	catcattccc	acaccgtgct	1560
acaagcccag	agagaaaactt	ccaatagagg	ccaagctgcc	atggttcaag	caagcccagg	1620
agctagagga	aggagctgca	gtgtctgaag	agccttcgtt	catcccagaa	gcttggtogg	1680
cctgtactgt	cacctgtggc	gtggggacac	agggtgcgaat	aataagatgc	caggtgctcc	1740
tgcccttctc	tcagtctgtg	gctgacctgc	ctgcggatga	atgogaaggg	cctaagccag	1800
catcgcaacg	tgccctgttat	gcggggaccgt	gcaacgggga	aaactccgga	tttaaccogg	1860

ataacggaga	cggactcttg	ggtggcctgc	aggacctgga	cgagctgtat	gactgggaat	1920
atgagggctt	caccaagtgc	tctgagtcac	gtggaggagg	tgtccaggag	gctgtggtaa	1980
gctgcctgaa	caaacagacc	cggaactcg	ctgatgagaa	cctgtgtgtg	accagccgac	2040
ggcccccgca	gctcctgaaa	tcctgtaact	tggatccctg	cccagcaagg	tgggaaattg	2100
gcaagtggag	cccatgtagc	ctcacctgtg	gggttggcct	gcagaccaga	gatgttgtct	2160
gttccccact	actctccaga	gagatgaacg	aggtggttgt	cctggccgat	gagctgtgcc	2220
atcaccccaa	accagcaca	gtgcaagcct	gcaaccgttt	caattgcccc	ccagcctggt	2280
accctgcaca	gtggcagctg	tgttccagga	cctgtggagg	agggattcag	aaacgtgatg	2340
ttctctgcaa	gcagcgtatg	gctgacggca	gcttcctgga	gcttccagag	accttctgtt	2400
ctgcctccaa	accaacctcc	caccaagggg	gcaagaaaga	tgactgcccc	agtgagtggc	2460
ttctctcaga	gtggtcagag	tgctctgtga	gctgtgggga	gggcacccag	actcgaagtg	2520
ccatttgcca	gaggggtgctg	aaaactgggg	tctccactgt	tgtcaattct	actttgtgcc	2580
ctccccctgc	cttctcttcc	tccatcagac	cctgtatgct	ggcaacctgt	gcaaggcctg	2640
gaaggccttc	caccaagcac	agtccacaca	tcgcagctgc	acgaaacata	tacatccaga	2700
cccgcaggca	gaggaagctg	cacttcgttg	tgggaggatt	tgttacctg	cttcccaaga	2760
ctaccgtggt	gctccggtgt	cccacacgca	gattccgcaa	gccccctc	acgtgggaaa	2820
aggacggcca	gcacagcatc	agctcggctc	atgtcacctg	agcccccttt	ggctacctga	2880
agatccatcg	cctcaagccc	tcagatgcgg	gcattctacac	ctgttctgcg	gggcccgc	2940
gggagcagtt	tgtgattaag	ctcattggag	gcaaccgcaa	gctcgtggcc	aggccccctga	3000
gccttttgag	tgaggaagag	gaggccctgc	aggtgaggaa	gaccaatcca	aaagaagcct	3060
tgcaaaccga	caaacaccaa	aatggcatct	tctccaacgg	cagcaaagct	gagaagcgcg	3120
gcctcactgc	agatcctggg	aaccgctatg	atgacattgt	gtctaggctg	ttggagcagg	3180
gcggctggcc	aggagagctc	ctggcctcgt	gggaggtgca	ggactctgca	gaaaggaacg	3240
catcctcaga	agaggatccc	aatgctgaac	aggccctcct	gcattctccc	ttaccatgg	3300
tggccgagca	gaagcgtctg	gatgacatcc	tcaggaacct	gtcccagcag	ccagaggagc	3360
tacgggacct	ctacagcaag	cacctgggtg	cccagctggc	ccaggacatc	ttccgaagcc	3420
acctggagaa	ccaggatctg	cttccaaagc	cctctgagca	gaggtttctc	cccatggccg	3480
tcccagctca	caaacatgtc	tctggcttca	gcagctccct	aaggagctca	tctggagagg	3540
ctggaggtgg	ctcccgtagg	cctcatcgca	agcccgccat	cttgcggaag	atctcagcgg	3600

cacagcaact	ctcagcctcg	gaggtagtca	cccacctggg	gcagactgtg	gctctagcca	3660
gtgggaccct	gagtgtgctt	ctgcactgtg	aggcggtagg	caacccaagg	cccaccatcc	3720
actggaccag	gaacggagaa	gcagttcaat	tcagtgcacg	gattcttcta	cagccagatg	3780
attccttaca	gatcttggca	ccagtgggaag	ctgatgtggg	tttctatact	tgcaatgcc	3840
caaatgcctt	gggatatgac	tctgtctcca	ttgctgtcac	attagcagga	aagccactag	3900
tgaagacatc	ccgaatgact	gtgctcaaca	cagagaagcc	tacagtcaca	gtggatatcg	3960
gtggcaccgt	cagaaccgtc	cgaggagtga	atgtgacaat	taactgcaa	gtggcagggtg	4020
tacctgaagc	tgaagtcacc	tggttcagaa	ataaaagcaa	attgggctcc	tcccaccacc	4080
tgcattgagg	ctcctccac	cacctgcatg	agggctcctt	gctgctcaca	gatgtgtcct	4140
tctctgatca	aggcctgtat	tcctgcaggg	cagccaatct	acacggagaa	cagacagaga	4200
gcacgcagct	cctgattcta	gatccccgc	aagtccccac	acagttggaa	gacatcaggg	4260
ccttgctctt	agccactgga	ccgaaccttc	cttoggtgct	tatgtccct	ctgggaacac	4320
agctggctct	ggatcctggg	aattccgctc	tccttgggtg	ccccatcaa	ggtcacccca	4380
cccccaatat	cacctgggtc	caaaatggcc	agccaattgc	cactgccccg	gggctgaccc	4440
accacatctg	gggagctgga	cagatcctac	gagttgcaaa	tctcagtgga	ggacctcaag	4500
gggagttcag	ctgcctggct	cagaatgaag	cagggacctc	cctgcagaaa	gcctcttttg	4560
tgatccaaga	ttactggtgg	tctgtggata	gactggcaac	ctgctccgca	tcttgtggta	4620
atagaggcat	tcattcagcc	cgactcaggt	gtttactgaa	caaccaaggaa	gttgaccctg	4680
agcactgtac	aggaaagcct	cggccccgcg	tgacgcctgt	tgcttgcaac	cggagagact	4740
gtccttctcg	atggatgggt	acgtcctggg	cgcctgttac	cgggagttgc	ggtggagggtg	4800
tacaaaaccg	aagggtgacc	tgtcagaagt	tgaaagcttc	tggtatttct	accccggtgt	4860
ccaatgacat	gtgcagtcaa	ctggccaaac	ggcctgtgga	caccagggcc	tgcaaccagc	4920
agctctgtgt	ggaatgggcc	ttctccagct	ggggccagtg	caatgggcct	tgcataggcc	4980
ctgccttgc	tgtgcagcac	aggcaggtct	tctgccagac	cgggatggc	atcaccttac	5040
catcagagca	gtgcagtgcc	ctccccaggc	ccgtgagcac	ccaaaactgc	tggtcagagg	5100
cctgcagtgt	acactggaga	gtcagcctct	ggaccctatg	cacggcaacc	tgtggcaact	5160
atggcttcca	gtcccggcgg	gtggagtgtg	ttcacgtgcg	caccaacaag	gctgtgcccc	5220
agcacctgtg	ttcctggggg	cccgggcccg	ccaactggca	acgctgcaac	gtcactccat	5280

gtgaaaacac	tgagtgcaga	gacaccacca	ggtactgcga	gaaggtgaga	cagttgaaac	5340
tctgccagct	cggccagttc	cgatctcgct	gctgtggaac	gtgtggcaaa	gcatgaagac	5400
agggcacggg	gatgaactct	accctggcca	cacaaaggac	tcacacaacc	accttggaca	5460
tgacttacgc	tttcttcctt	ttattttattt	atttgcccct	ccccacacac	ccatgtccac	5520
ctgcacccat	gctcccagag	actgtggcat	gtttttottat	tcagttagct	ggaagctagt	5580
gtggagaaca	ggaaaaaaaa	aaaaaaagat	gtctaaagga	gggttgacga	gcagacagca	5640
gctcccttgc	agagctggct	cctttctgat	cagggctctca	gagacacagg	aagaccaagc	5700
catagtctat	agggacagca	aggagtttat	aacacatagg	tacaggcaca	ggccaatcag	5760
gagtgccgtt	cctttgtttc	ctggtggtag	gaacaggggg	gggggggggg	aagcagcata	5820
ctagagcaga	aaccaggcca	gactccagtt	cttctggcag	tggtgtggga	agaccaagaa	5880
agaagagaac	tgagctgtgg	gcaagaggg	tatgccaaaa	gagacaagca	aaataggcag	5940
aaggcatgcc	gctatgcaga	agacagaggc	cgggaagaga	gaaggaaagc	gtgcgcacgc	6000
acctcagtgg	gaggagcttg	ctggcaggac	agattttcaa	gcagtcacgt	gagaccata	6060
tgacaagcct	gtgcccggga	acatgtgcac	agtctcgag	cttcaacaga	gttggggaag	6120
ggtattcaga	gtcctctcag	ggtaccgcg	ctgcgtgaca	cgggtgacac	tgcattgacat	6180
gccctttgtg	tggttatata	agtgtgtcta	cccctcttct	gtcaagtgac	ttttgtgtct	6240
gcagctgtgg	acaacccaac	aaagggtaaa	ggtcagctgt	gcagtaaggc	tgcccctagc	6300
atctccctac	ttaagtcaag	ggtcactagc	catgaactga	ctgctgccga	gggccaaggg	6360
catttgatcg	tgaatgcccc	tggttttcca	tcogaatgac	tttgctgctc	cttggtgggc	6420
cagggactag	aatcatagct	cacaaaggat	tcattaagca	acctgtttgc	gggcaagcat	6480
ctacaggatg	tagaaagcct	tactaccaga	tctcaccaga	gaggaggaga	ggacggaggg	6540
gcctgcaca	cccacagtgc	gctttctctc	ccaaccctca	cacattgggc	ctgggagccc	6600
ttcctgaacc	acagaagagc	tgccaattta	cgcagcaggc	tgggagtagg	gagtgtgac	6660
aggcgtatg	gcaaagctag	aagtcagcta	aagactcaca	agacagccgg	agaaagcttt	6720
gtaaccatcc	tctcgactac	acagatacta	tggttaattag	cacaccagac	tgagtgaatg	6780
gaccactcag	cgggctaagc	agcacacaca	attcagcaac	acacatcagc	tatggcaaaa	6840
acagaagccc	ttggcaaata	ctcagtctgc	agagaagctt	caacagtaat	aatcgcttcc	6900
ttcctgagca	gcctcagttt	agcagatgag	gcaggcagag	cccagtagcc	atcagataag	6960
cgcgccctcg	caaacacact	caccaaatac	tatcacccca	gggccaccac	tgacatcagg	7020

gctgcatttc ccaaccagcc tggaagttaa gtttgaaaga ggatgttaat ctgtgtottc 7080
 ttagatgogc tatggacagt tagggtaggc ctcttcctca tcacaaaaat agactctaag 7140
 aaattgtggt attaaaaagtc agtgtaatgt ttatttaaaa ttaaagagag tgtttctatg 7200
 t 7201

<210> 4
 <211> 1770
 <212> PRT
 <213> mouse

<400> 4

Met Glu Cys Cys Arg Arg Ala Ala Pro Gly Thr Pro Leu Leu Val Leu
 1 5 10 15

Ala Phe Leu Leu Leu Ser Ser Arg Thr Ala Arg Ser Glu Glu Asp Arg
 20 25 30

Glu Gly Leu Trp Asp Ala Trp Gly Pro Trp Ser Glu Cys Ser Arg Thr
 35 40 45

Cys Gly Gly Gly Ala Ser Tyr Ser Leu Arg Arg Cys Leu Ser Ser Lys
 50 55 60

Ser Cys Glu Gly Arg Asn Ile Arg Tyr Arg Thr Cys Ser Asn Val Asp
 65 70 75 80

Cys Pro Pro Glu Ala Gly Asp Phe Arg Ala Gln Gln Cys Ser Ala His
 85 90 95

Asn Asp Val Lys Tyr His Gly Gln Leu Tyr Glu Trp Leu Pro Val Ser
 100 105 110

Asn Asp Pro Asp Asn Pro Cys Ser Leu Lys Cys Gln Ala Lys Gly Thr
 115 120 125

Ser Leu Val Val Glu Leu Ala Pro Lys Val Leu Asp Gly Thr Arg Cys
 130 135 140

Tyr Thr Glu Ser Leu Asp Met Cys Ile Ser Gly Leu Cys Gln Ile Val
 145 150 155 160

Gly Cys Asp His Gln Leu Gly Ser Thr Val Lys Glu Asp Asn Cys Gly
 165 170 175

Val Cys Asn Gly Asp Gly Ser Thr Cys Arg Leu Val Arg Gly Gln Tyr
 180 185 190

Lys Ser Gln Leu Ser Ala Ser Lys Ser Asp Asp Thr Val Val Ala Ile
 195 200 205

Pro Tyr Gly Ser Arg His Ile Arg Leu Val Leu Lys Gly Pro Asp His
 210 215 220

Leu Tyr Leu Glu Thr Lys Thr Leu Gln Gly Thr Lys Gly Glu Asn Ser
 225 230 235 240

Leu Ser Ser Thr Gly Ile Phe Leu Val Asp Asn Ser Thr Val Asp Phe
 245 250 255

Gln Lys Leu Pro Asp Lys Glu Ile Leu Arg Met Thr Gly Pro Leu Thr
 260 265 270

Ala Asp Phe Ile Ile Lys Ile His Asp Leu Gly Pro Ala Asp Ser Thr
 275 280 285

Val Gln Phe Ile Phe Tyr Gln Pro Ile Ile His Arg Trp Arg Glu Thr
 290 295 300

Asp Phe Phe Pro Cys Ser Ala Thr Cys Gly Gly Gly Tyr Gln Leu Thr
 305 310 315 320

Ser Ala Glu Cys Tyr Asp Leu Arg Ser Asn Arg Val Val Ala Asp Gln
 325 330 335

Tyr Cys His Tyr Tyr Pro Glu Asn Val Lys Pro Lys Pro Lys Leu Gln
 340 345 350

Glu Cys Asn Leu Asp Pro Cys Pro Ala Ser Asp Gly Tyr Lys Gln Ile
 355 360 365

Met Pro Tyr Asp Leu Tyr His Pro Leu Pro Arg Trp Glu Ala Thr Pro
 370 375 380

Trp Thr Ala Cys Ser Ser Ser Cys Gly Gly Gly Ile Gln Ser Arg Ala

385		390		395		400
Val Ser Cys Val	Glu Glu Asp Ile Gln Gly His Val Thr Ser Val Glu					
	405		410			415
Glu Trp Lys Cys Met Tyr Thr Pro Lys Met Pro Val Val Gln Pro Cys						
	420		425			430
Asn Ile Phe Asp Cys Pro Lys Trp Leu Ala Gln Glu Trp Ser Pro Cys						
	435		440			445
Thr Val Thr Cys Gly Gln Gly Leu Arg Tyr Arg Val Val Leu Cys Ile						
	450		455			460
Asp His Arg Gly Met His Thr Gly Gly Cys Ser Ala Lys Thr Lys Pro						
	465		470			475
His Ile Lys Glu Glu Cys Ile Ile Pro Thr Pro Cys Tyr Lys Pro Arg						
	485		490			495
Glu Lys Leu Pro Ile Glu Ala Lys Leu Pro Trp Phe Lys Gln Ala Gln						
	500		505			510
Glu Leu Glu Glu Gly Ala Ala Val Ser Glu Glu Pro Ser Phe Ile Pro						
	515		520			525
Glu Ala Trp Ser Ala Cys Thr Val Thr Cys Gly Val Gly Thr Gln Val						
	530		535			540
Arg Ile Ile Arg Cys Gln Val Leu Leu Pro Phe Ser Gln Ser Val Ala						
	545		550			555
Asp Leu Pro Ala Asp Glu Cys Glu Gly Pro Lys Pro Ala Ser Gln Arg						
	565		570			575
Ala Cys Tyr Ala Gly Pro Cys Asn Gly Glu Thr Pro Glu Phe Asn Pro						
	580		585			590
Asp Asn Gly Asp Gly Leu Leu Gly Gly Leu Gln Asp Leu Asp Glu Leu						
	595		600			605
Tyr Asp Trp Glu Tyr Glu Gly Phe Thr Lys Cys Ser Glu Ser Cys Gly						
	610		615			620

Gly Gly Val Gln Glu Ala Val Val Ser Cys Leu Asn Lys Gln Thr Arg
625 630 635 640

Glu Leu Ala Asp Glu Asn Leu Cys Val Thr Ser Arg Arg Pro Pro Gln
645 650 655

Leu Leu Lys Ser Cys Asn Leu Asp Pro Cys Pro Ala Arg Trp Glu Ile
660 665 670

Gly Lys Trp Ser Pro Cys Ser Leu Thr Cys Gly Val Gly Leu Gln Thr
675 680 685

Arg Asp Val Val Cys Ser His Leu Leu Ser Arg Glu Met Asn Glu Val
690 695 700

Val Val Leu Ala Asp Glu Leu Cys His His Pro Lys Pro Ser Thr Val
705 710 715 720

Gln Ala Cys Asn Arg Phe Asn Cys Pro Pro Ala Trp Tyr Pro Ala Gln
725 730 735

Trp Gln Leu Cys Ser Arg Thr Cys Gly Gly Gly Ile Gln Lys Arg Asp
740 745 750

Val Leu Cys Lys Gln Arg Met Ala Asp Gly Ser Phe Leu Glu Leu Pro
755 760 765

Glu Thr Phe Cys Ser Ala Ser Lys Pro Thr Ser His Gln Gly Cys Lys
770 775 780

Lys Asp Asp Cys Pro Ser Glu Trp Leu Leu Ser Glu Trp Ser Glu Cys
785 790 795 800

Ser Val Ser Cys Gly Glu Gly Thr Gln Thr Arg Ser Ala Ile Cys Gln
805 810 815

Arg Val Leu Lys Thr Gly Val Ser Thr Val Val Asn Ser Thr Leu Cys
820 825 830

Pro Pro Leu Pro Phe Ser Ser Ser Ile Arg Pro Cys Met Leu Ala Thr
835 840 845

Cys Ala Arg Pro Gly Arg Pro Ser Thr Lys His Ser Pro His Ile Ala
850 855 860

Ala Ala Arg Asn Ile Tyr Ile Gln Thr Arg Arg Gln Arg Lys Leu His
865 870 875 880

Phe Val Val Gly Gly Phe Ala Tyr Leu Leu Pro Lys Thr Thr Val Val
885 890 895

Leu Arg Cys Pro Thr Arg Arg Phe Arg Lys Pro Leu Ile Thr Trp Glu
900 905 910

Lys Asp Gly Gln His Ser Ile Ser Ser Ala His Val Thr Val Ala Pro
915 920 925

Phe Gly Tyr Leu Lys Ile His Arg Leu Lys Pro Ser Asp Ala Gly Ile
930 935 940

Tyr Thr Cys Ser Ala Gly Pro Ala Arg Glu Gln Phe Val Ile Lys Leu
945 950 955 960

Ile Gly Gly Asn Arg Lys Leu Val Ala Arg Pro Leu Ser Leu Trp Ser
965 970 975

Glu Glu Glu Glu Ala Leu Gln Val Arg Lys Thr Asn Pro Lys Glu Ala
980 985 990

Leu Gln Thr His Lys His Gln Asn Gly Ile Phe Ser Asn Gly Ser Lys
995 1000 1005

Ala Glu Lys Arg Gly Leu Thr Ala Asp Pro Gly Asn Arg Tyr Asp
1010 1015 1020

Asp Ile Val Ser Arg Leu Leu Glu Gln Gly Gly Trp Pro Gly Glu
1025 1030 1035

Leu Leu Ala Ser Trp Glu Val Gln Asp Ser Ala Glu Arg Asn Ala
1040 1045 1050

Ser Ser Glu Glu Asp Pro Asn Ala Glu Gln Ala Leu Leu His Leu
1055 1060 1065

Pro	Phe	Thr	Met	Val	Ala	Glu	Gln	Lys	Arg	Leu	Asp	Asp	Ile	Leu
1070						1075					1080			
Arg	Asn	Leu	Ser	Gln	Gln	Pro	Glu	Glu	Leu	Arg	Asp	Leu	Tyr	Ser
1085						1090					1095			
Lys	His	Leu	Val	Ala	Gln	Leu	Ala	Gln	Asp	Ile	Phe	Arg	Ser	His
1100						1105					1110			
Leu	Glu	Asn	Gln	Asp	Leu	Leu	Pro	Lys	Pro	Ser	Glu	Gln	Arg	Phe
1115						1120					1125			
Pro	Pro	Met	Ala	Val	Pro	Ala	His	Lys	His	Val	Ser	Gly	Phe	Ser
1130						1135					1140			
Ser	Ser	Leu	Arg	Ser	Ser	Ser	Gly	Glu	Ala	Gly	Gly	Gly	Ser	Arg
1145						1150					1155			
Arg	Pro	His	Arg	Lys	Pro	Ala	Ile	Leu	Arg	Lys	Ile	Ser	Ala	Ala
1160						1165					1170			
Gln	Gln	Leu	Ser	Ala	Ser	Glu	Val	Val	Thr	His	Leu	Gly	Gln	Thr
1175						1180					1185			
Val	Ala	Leu	Ala	Ser	Gly	Thr	Leu	Ser	Val	Leu	Leu	His	Cys	Glu
1190						1195					1200			
Ala	Val	Gly	Asn	Pro	Arg	Pro	Thr	Ile	His	Trp	Thr	Arg	Asn	Gly
1205						1210					1215			
Glu	Ala	Val	Gln	Phe	Ser	Asp	Arg	Ile	Leu	Leu	Gln	Pro	Asp	Asp
1220						1225					1230			
Ser	Leu	Gln	Ile	Leu	Ala	Pro	Val	Glu	Ala	Asp	Val	Gly	Phe	Tyr
1235						1240					1245			
Thr	Cys	Asn	Ala	Thr	Asn	Ala	Leu	Gly	Tyr	Asp	Ser	Val	Ser	Ile
1250						1255					1260			
Ala	Val	Thr	Leu	Ala	Gly	Lys	Pro	Leu	Val	Lys	Thr	Ser	Arg	Met
1265						1270					1275			
Thr	Val	Leu	Asn	Thr	Glu	Lys	Pro	Thr	Val	Thr	Val	Asp	Ile	Gly

1280		1285		1290
Gly Thr Val Arg Thr Val Arg	Gly Val Asn Val Thr	Ile Asn Cys		
1295	1300	1305		
Gln Val Ala Gly Val Pro Glu	Ala Glu Val Thr Trp	Phe Arg Asn		
1310	1315	1320		
Lys Ser Lys Leu Gly Ser Ser	His His Leu His Glu	Gly Ser Ser		
1325	1330	1335		
His His Leu His Glu Gly Ser	Leu Leu Leu Thr Asp	Val Ser Phe		
1340	1345	1350		
Ser Asp Gln Gly Leu Tyr Ser	Cys Arg Ala Ala Asn	Leu His Gly		
1355	1360	1365		
Glu Gln Thr Glu Ser Thr Gln	Leu Leu Ile Leu Asp	Pro Pro Gln		
1370	1375	1380		
Val Pro Thr Gln Leu Glu Asp	Ile Arg Ala Leu Leu	Leu Ala Thr		
1385	1390	1395		
Gly Pro Asn Leu Pro Ser Val	Leu Met Ser Pro Leu	Gly Thr Gln		
1400	1405	1410		
Leu Val Leu Asp Pro Gly Asn	Ser Ala Leu Leu Gly	Cys Pro Ile		
1415	1420	1425		
Lys Gly His Pro Thr Pro Asn	Ile Thr Trp Phe Gln	Asn Gly Gln		
1430	1435	1440		
Pro Ile Ala Thr Ala Pro Gly	Leu Thr His His Ile	Trp Gly Ala		
1445	1450	1455		
Gly Gln Ile Leu Arg Val Ala	Asn Leu Ser Gly Gly	Pro Gln Gly		
1460	1465	1470		
Glu Phe Ser Cys Leu Ala Gln	Asn Glu Ala Gly Thr	Leu Leu Gln		
1475	1480	1485		
Lys Ala Ser Leu Val Ile Gln	Asp Tyr Trp Trp Ser	Val Asp Arg		
1490	1495	1500		

Leu	Ala	Thr	Cys	Ser	Ala	Ser	Cys	Gly	Asn	Arg	Gly	Ile	His	Gln
1505						1510					1515			
Pro	Arg	Leu	Arg	Cys	Leu	Leu	Asn	Thr	Thr	Glu	Val	Asp	Pro	Glu
1520						1525					1530			
His	Cys	Thr	Gly	Lys	Pro	Arg	Pro	Ala	Val	Gln	Pro	Val	Ala	Cys
1535						1540					1545			
Asn	Arg	Arg	Asp	Cys	Pro	Ser	Arg	Trp	Met	Val	Thr	Ser	Trp	Ser
1550						1555					1560			
Ala	Cys	Thr	Arg	Ser	Cys	Gly	Gly	Gly	Val	Gln	Thr	Arg	Arg	Val
1565						1570					1575			
Thr	Cys	Gln	Lys	Leu	Lys	Ala	Ser	Gly	Ile	Ser	Thr	Pro	Val	Ser
1580						1585					1590			
Asn	Asp	Met	Cys	Ser	Gln	Leu	Ala	Lys	Arg	Pro	Val	Asp	Thr	Gln
1595						1600					1605			
Ala	Cys	Asn	Gln	Gln	Leu	Cys	Val	Glu	Trp	Ala	Phe	Ser	Ser	Trp
1610						1615					1620			
Gly	Gln	Cys	Asn	Gly	Pro	Cys	Ile	Gly	Pro	Arg	Leu	Ala	Val	Gln
1625						1630					1635			
His	Arg	Gln	Val	Phe	Cys	Gln	Thr	Arg	Asp	Gly	Ile	Thr	Leu	Pro
1640						1645					1650			
Ser	Glu	Gln	Cys	Ser	Ala	Leu	Pro	Arg	Pro	Val	Ser	Thr	Gln	Asn
1655						1660					1665			
Cys	Trp	Ser	Glu	Ala	Cys	Ser	Val	His	Trp	Arg	Val	Ser	Leu	Trp
1670						1675					1680			
Thr	Leu	Cys	Thr	Ala	Thr	Cys	Gly	Asn	Tyr	Gly	Phe	Gln	Ser	Arg
1685						1690					1695			
Arg	Val	Glu	Cys	Val	His	Val	Arg	Thr	Asn	Lys	Ala	Val	Pro	Glu
1700						1705					1710			

His Leu Cys Ser Trp Gly Pro Arg Pro Ala Asn Trp Gln Arg Cys
 1715 1720 1725

Asn Val Thr Pro Cys Glu Asn Thr Glu Cys Arg Asp Thr Thr Arg
 1730 1735 1740

Tyr Cys Glu Lys Val Arg Gln Leu Lys Leu Cys Gln Leu Gly Gln
 1745 1750 1755

Phe Arg Ser Arg Cys Cys Gly Thr Cys Gly Lys Ala
 1760 1765 1770

<210> 5
 <211> 7156
 <212> DNA
 <213> mouse

<400> 5
 agcactggat gtgacagcag cgggcagagg agcacttagc agcttattca gtgtccgatt 60
 cggattccgg caaagatcca agtatggaat gctgccgtcg ggcagctcct ggcacaccgc 120
 tgctcgttct ggctttccta ctctgagtt ccaggactgc acgctcagag gaagaccgtg 180
 aaggcctctg ggatgcctgg ggcccttggg gcgagtgttc cagaacctgt ggtgggggtg 240
 cctcctattc cctgagacgc tgcctgagca gcaagagctg tgaaggaaga aatattcgat 300
 atagaacatg cagtaatgtg gactgtccac cagaagcagg tgatttccga gctcaacagt 360
 gttctgctca taacgacgtc aagtaccatg gacagtttta tgaatggctt cctgtatcta 420
 atgaccctga caatccatgt tcaactcaagt gccaaagaaa aggaaccagc ctggttggtg 480
 aactagcacc gaaggtctta gatggtacgc gatgctatac agaactcactg gatatgtgca 540
 tcagcggcct atgccaaatt gttggctgtg atcaccagct ggggagcact gtcaaagagg 600
 ataactgtgg agtgtgcaac ggagatggct caacgtgccg gctgggtccga gggcagtata 660
 aatctcagct ctctgcaagt aaatcggatg acaccgtggg cgccattcct tacggcagca 720
 gacacatccg ccttgtctta aaagggccag atcatttgta tctggagacc aagaccctcc 780
 agggaaactaa aggtgaaaat agtctgagct ctacaggcac tttccttggtg gacaattcta 840
 cagtggactt ccagaaaactt ccagacaaaag agatactgag gatgactggg ccaactgacag 900
 cagatttcat catcaagatc catgacttgg gacctgcaga cagcaccgtc cagttcatct 960
 tctaccagcc catcatccac cgatggaggg aaacagactt ctttccttgc tcagcaacgt 1020

gtggtggagg ttaccagctg acgtctgccg agtgctatga tctacggagt aaccgtgtgg	1080
tggccgacca gtactgtcac tactaccccg agaacgtcaa gcccaaacc aagcttcagg	1140
agtgcaactt ggatccttgt ccggccagtg acggatacaa gcagatcatg ccttatgacc	1200
tctaccatcc ccttcctcgg tgggaggcca ccccgtaggac cgcgtgctcc tcctcgtgtg	1260
gggggggcat ccagagcccg gcagtttcct gtgtggagga ggacatccag gggcatgtca	1320
cttcagtgga agagtggaaa tgcattgtaca ccccaaagat gccagtcgtg cagccctgca	1380
acatTTTTga ctgccctaaa tggctggcac aggagtggtc tccgtgcaca gtgacctgtg	1440
gccagggcct cagataccgc gtggtactct gcattgacca ccgaggaatg cacacaggag	1500
gctgcagtgc gaaaacaaaag ccccacataa aagaagaatg catcattccc acaccctgct	1560
acaagcccag agagaaactt ccaatagagg ccaagctgcc atggttcaag caagcccagg	1620
agctagagga aggagctgca gtgtctgaag agccttcgtt catcccagaa gcttggtcgg	1680
cctgtactgt cacctgtggc gtggggacac aggtgcgaat aataagatgc caggtgctcc	1740
tgcctttctc tcagtctgtg gctgacctgc ctgcggatga atgcgaaggg cctaagccag	1800
catcgcaacg tgcctgttat gcgggaccgt gcaacgggga aactccggaa tttaaccogg	1860
ataacggaga cggactcttg ggtggcctgc aggacctgga cgagctgtat gactgggaat	1920
atgagggcct caccaagtgc tctgagtcac gtggaggagg tgtccaggag gctgtggtaa	1980
gctgcctgaa caaacagacc cgggagctcg ctgatgagaa cctgtgtgtg accagccgac	2040
ggccccgcga gctcctgaaa tcctgtaact tggatccctg ccagcaagg tgggaaattg	2100
gcaagtggag cccatgtagc ctcacctgtg gggttggcct gcagaccaga gatgttgtct	2160
gttccacact actctccaga gagatgaacg aggtggttgt cctggccgat gagctgtgcc	2220
atcccccaa acccagcaca gtgcaagcct gcaaccgttt caattgcccc ccagcctggt	2280
accctgcaca gtggcagctg tgttccagga cctgtggagg agggattcag aaacgtgatg	2340
ttctctgcaa gcagcgtatg gctgacggca gcttcctgga gcttccagag acctctgttt	2400
cagcctccaa accaacctcc caccaagggt gcaagaaaga tgactgcccc agtgagtggc	2460
ttctctcaga gtggtcagag tgctctgtga gctgtgggga gggcaccag actogaagtg	2520
ccatttgcca gaggggtgctg aaaaactggg tctccactgt tgtcaattct actttgtgcc	2580
ctccccctgc cttctcttcc tccatcagac cctgtatgct ggcaacctgt gcaaggcctg	2640
gaaggccttc caccaagcac agtcctcaca tcgcagctgc acgaaacata tacatccaga	2700
ccgcaggca gaggaagctg cacttcgttg tgggaggatt tgcttacctg cttcccaaga	2760

ctaccgtggt gctccggtgt cccacacgca gattccgcaa gcccctcatt acgtgggaaa	2820
aggacggcca gcacagcatt agctcggctc atgtcacctgt agccccctttt ggctacctga	2880
agatccattc cctcaagccc tcagatgcgg gcattctacac ctgtttctgct gggcccgccc	2940
gggagcagtt tgtgattaag ctcatctggag gcaaccgcaa gctcgtggcc agggccctga	3000
gccttttgag tgaggaagag gaggcctgc aggtgaggaa ggccaatcca aaagaagcct	3060
tgcaaaccga caaacaccaa aatggcattt totocaaagg cagcaaagct gagaagcgct	3120
gcctcactgc agatcctggg aaccgctatg atgacattgt gtctaggctg ttggagcagg	3180
gcggttgccc aggagagctc ctggcctcgt gggagggtgca ggactctgca gaaaggaacg	3240
cattctcaga agaggatccc aatgctgaac agggccctcct gcattctccc ttccattgg	3300
tggccgagca gaagcgtctg gatgacattc tcaggaaact gtccagcag ccagaggagc	3360
tacgggacct ctacagcaag cacttggtgg ccagctggc ccaggacatt ttccgaagcc	3420
acctggagaa ccaggatctg ctctcctaac cctctgagca gaggtttcct cccatggcgg	3480
tccagctca caaacatgtc tctggcttca gcagctcct aaggagctca tctggagagg	3540
ctgggggtgg ctcccgtagg cctcattgca agcccgccat ctgctggaag atctcagcgg	3600
cacagcaact ctacgctcgt gaggtagtca cccacctggg gcagactgtg gctctagcca	3660
gtgggacctt gactgtgctt ctgcactgtg aggcggtagg caacccaagg cccaccattc	3720
actggaccag gaacggagaa gcagttcaat tcagtgacag gattcttcta cagccagatg	3780
attccttaca gatcttgcca ccagtggaag ctgatgtggg tttctatact tgcaatgcca	3840
caaatgcctt gggatatgac tctgtctoca ttgctgtcac attagcagga aagccactag	3900
tgaagacatt ccgaatgact gtgctcaaca cagagaagcc tacagtcaca gtggatatcg	3960
gtggcacctt cagaaccgtc cgaggagtga atgtgacaat taactgcca gtggcagggtg	4020
tacctgaagc tgaagtcacc tggttcagaa ataaaagcaa attgggctcc tcccaccacc	4080
tgcatgaggg ctccctccac caactgcatt agggctcctt gctgctcaca gatgtgtcct	4140
tctctgatca aggcctgtat tctgcaggg cagccaattt acacggagaa cagacagaga	4200
gcacgcagct cctgattcta gatccccgc aagtcaccac acagttggaa gacatcaggg	4260
ccttgctctt agccactgga ccgaacctt ctctgggtgt tatgtccct ctgggaacac	4320
agctggctct ggatcctggg aattccgctc tcttggtgtg ccccatcaaa ggtcacccca	4380
cccccaatat cacttggttc caaaatggcc agccaattgc cactgccccg gggctgaccc	4440

accacatctg	gggagctgga	cagatcctac	gagttgcaaa	tctcagtgga	ggacctcaag	4500	
gggagttcag	ctgcctggct	cagaatgaag	cagggaccct	cctgcagaaa	gcctcttttg	4560	
tgatccaaga	ttactggtgg	tctgtggata	gactggcaac	ctgctccgca	tcttgtggta	4620	
atagaggcat	tcatcagccg	cgactcaggt	gtttactgaa	caccacggaa	gttgaccctg	4680	
agcactgtac	aggaaagcct	cggcccgccg	tgacagcctgt	tgccctgcaac	cggagagact	4740	
gtccttctcg	atggatggtg	acgtcctggt	ccgcctgtac	cgggagttgc	ggtggagggtg	4800	
tacaaacccg	aaggggtgacc	tgtcagaagt	tgaaagcttc	tggtattttct	accccggtgt	4860	
ccaatgacat	gtgcagtcaa	ctggccaaac	ggcctgtgga	caccagggcc	tgcaaccagc	4920	
agctctgtgt	ggaatgggcc	ttctccagct	ggggccagtg	caatgggcct	tgcataggcc	4980	
ctcgtcttgc	tgtgcagcac	agacaggtct	tctgccagac	ccgggatggc	atcaccttac	5040	
catcagagca	gtgcagtgcc	ctccccaggc	ccgtgagcac	ccaaaactgc	tggtcagagg	5100	
cctgcagtgt	acactggagg	gtcagcctct	ggaccctatg	cacggcaacc	tgtggcaact	5160	
atggcttcca	gtcccggcgg	gtggagtgtg	ttcacgtgcg	caccaacaag	gctgtgcccg	5220	
agcacctgtg	ttcctggggg	ccccggcccg	ccaactggca	acgctgcaac	gtcactccat	5280	
gtgaaaacac	tgagtgcaga	gacaccacca	ggtactgcga	gaagggtgaga	cagttgaaac	5340	
tctgccagct	cggccagttc	cgatctcgct	gctgtggaac	gtgtggcaaa	gcatgaagac	5400	
agggcacggg	gatgaactct	accctggcca	cacaaaggac	tcacataacc	accttgagca	5460	
tgacttacgc	tttcttcctt	ttatttat	atttgccct	ccccacacac	ccatgtccac	5520	
ctgcacccat	gtcccgagg	actgtggcat	gttttcttat	tcagttagct	ggaagctagt	5580	
gtggagaaca	gaaaaaaaaa	aaaaaaagat	gtctaaagga	gggttgacaga	gcagacagca	5640	
gtcccttgc	agagctggct	cctttctgat	cagggctctca	gagacacagg	aagagcaagc	5700	
catagcctat	agggacagca	aggagt	ttat	aacacatagg	tacaggcaca	ggccaatcag	5760
gagtgcgcga	gcatactaga	gcagaaacca	ggccagactc	cagttcttct	ggcagtgggtg	5820	
tggaagacc	aagaaagaag	agaactgagc	tgtgggcaag	agggttatgc	caaaagagac	5880	
aagcaaaata	ggcagaaggc	atgcgcctat	gcagaagaca	gaggccggga	agagagaagg	5940	
aaagcgtgcg	cacgcacctc	agtgggagga	gcttgctggc	aggacagatt	ttcaagcata	6000	
tgacaagtct	gtgcccgga	acatgtgcac	agtoctcgag	gttcaacaga	ggtggggaag	6060	
ggtattcaga	gtcctctcag	ggtaccccg	ctgcgtgaca	oggggtgacac	tgcatgacat	6120	
gccctttgtg	tggttatata	agtgtgtcta	ccctcttct	gtcaagtgc	tttgtgtct	6180	

```

gcagctgtgg acaacccaac aaaggggtcaa ggtcagctgt gcagtaaggc tgcccctagc 6240
atctccctac ttaagtcaag ggtcactagc catgaaactga ctgctgccga gggccaaggg 6300
catttgaccg tgaatgcccc tggttttcca tccgaatgac tttgctgctc cttggtgggc 6360
cagggactag aatcatagct cacaaaggat tcattaagca acctgtttgc aggcaagcat 6420
ctacaggatg tgggaagcct tactaccaga tctcaccaga gaggaggaga ggacagaggg 6480
gccctgcaca cccacagtgc gctttctctc ccaaccctca cacattgggc ctgggagccc 6540
ttctgaacc acagaagagc tgccgtacac ccgatttacg cagcaggctg ggagtaggga 6600
gtagggagtg tgatcaggcg ctatggcaaa gctagaagtc agctaaagac tcacaagaca 6660
gccggagaaa gctttgtaac catcctctcg actacacaga tactatggta attagcacac 6720
cagactgagt gaatggacca ctcagcgggc taagcagcac acacaattca gcaacacaca 6780
tcagctatgg caaaaacaga agcccttggc aaatcctcag tctgcagaga agcttcaaca 6840
gtaataatcg cttccttcct gagcagcctc agtttagcag atgaggcagg cagagcccag 6900
tagccatcag ataagccgcc cctcacaac acactcacca aatgctatca cccagggcc 6960
accactgaca tcagggtgc atttcccaac cagcctggaa gttaagtttg aaagaggatg 7020
ttaatctgtg tcttcttaga tgcgctatgg acagttaggg taggcctctt cctcatcaca 7080
aaaatagact ctaagaaatt gtgttattaa aagtcagtgt aatgtttatt taaaattaaa 7140
gagagtgttt ctatgt 7156

```

```

<210> 6
<211> 1770
<212> PRT
<213> mouse

```

```

<400> 6

```

```

Met Glu Cys Cys Arg Arg Ala Ala Pro Gly Thr Pro Leu Leu Val Leu
1          5          10          15

```

```

Ala Phe Leu Leu Leu Ser Ser Arg Thr Ala Arg Ser Glu Glu Asp Arg
20          25          30

```

```

Glu Gly Leu Trp Asp Ala Trp Gly Pro Trp Ser Glu Cys Ser Arg Thr
35          40          45

```

```

Cys Gly Gly Gly Ala Ser Tyr Ser Leu Arg Arg Cys Leu Ser Ser Lys
50          55          60

```

Ser Cys Glu Gly Arg Asn Ile Arg Tyr Arg Thr Cys Ser Asn Val Asp
65 70 75 80

Cys Pro Pro Glu Ala Gly Asp Phe Arg Ala Gln Gln Cys Ser Ala His
85 90 95

Asn Asp Val Lys Tyr His Gly Gln Phe Tyr Glu Trp Leu Pro Val Ser
100 105 110

Asn Asp Pro Asp Asn Pro Cys Ser Leu Lys Cys Gln Ala Lys Gly Thr
115 120 125

Ser Leu Val Val Glu Leu Ala Pro Lys Val Leu Asp Gly Thr Arg Cys
130 135 140

Tyr Thr Glu Ser Leu Asp Met Cys Ile Ser Gly Leu Cys Gln Ile Val
145 150 155 160

Gly Cys Asp His Gln Leu Gly Ser Thr Val Lys Glu Asp Asn Cys Gly
165 170 175

Val Cys Asn Gly Asp Gly Ser Thr Cys Arg Leu Val Arg Gly Gln Tyr
180 185 190

Lys Ser Gln Leu Ser Ala Ser Lys Ser Asp Asp Thr Val Val Ala Ile
195 200 205

Pro Tyr Gly Ser Arg His Ile Arg Leu Val Leu Lys Gly Pro Asp His
210 215 220

Leu Tyr Leu Glu Thr Lys Thr Leu Gln Gly Thr Lys Gly Glu Asn Ser
225 230 235 240

Leu Ser Ser Thr Gly Thr Phe Leu Val Asp Asn Ser Thr Val Asp Phe
245 250 255

Gln Lys Leu Pro Asp Lys Glu Ile Leu Arg Met Thr Gly Pro Leu Thr
260 265 270

Ala Asp Phe Ile Ile Lys Ile His Asp Leu Gly Pro Ala Asp Ser Thr
275 280 285

Val Gln Phe Ile Phe Tyr Gln Pro Ile Ile His Arg Trp Arg Glu Thr
 290 295 300

Asp Phe Phe Pro Cys Ser Ala Thr Cys Gly Gly Gly Tyr Gln Leu Thr
 305 310 315 320

Ser Ala Glu Cys Tyr Asp Leu Arg Ser Asn Arg Val Val Ala Asp Gln
 325 330 335

Tyr Cys His Tyr Tyr Pro Glu Asn Val Lys Pro Lys Pro Lys Leu Gln
 340 345 350

Glu Cys Asn Leu Asp Pro Cys Pro Ala Ser Asp Gly Tyr Lys Gln Ile
 355 360 365

Met Pro Tyr Asp Leu Tyr His Pro Leu Pro Arg Trp Glu Ala Thr Pro
 370 375 380

Trp Thr Ala Cys Ser Ser Ser Cys Gly Gly Gly Ile Gln Ser Arg Ala
 385 390 395 400

Val Ser Cys Val Glu Glu Asp Ile Gln Gly His Val Thr Ser Val Glu
 405 410 415

Glu Trp Lys Cys Met Tyr Thr Pro Lys Met Pro Val Val Gln Pro Cys
 420 425 430

Asn Ile Phe Asp Cys Pro Lys Trp Leu Ala Gln Glu Trp Ser Pro Cys
 435 440 445

Thr Val Thr Cys Gly Gln Gly Leu Arg Tyr Arg Val Val Leu Cys Ile
 450 455 460

Asp His Arg Gly Met His Thr Gly Gly Cys Ser Ala Lys Thr Lys Pro
 465 470 475 480

His Ile Lys Glu Glu Cys Ile Ile Pro Thr Pro Cys Tyr Lys Pro Arg
 485 490 495

Glu Lys Leu Pro Ile Glu Ala Lys Leu Pro Trp Phe Lys Gln Ala Gln
 500 505 510

Glu Leu Glu Glu Gly Ala Ala Val Ser Glu Glu Pro Ser Phe Ile Pro
 515 520 525

Glu Ala Trp Ser Ala Cys Thr Val Thr Cys Gly Val Gly Thr Gln Val
 530 535 540

Arg Ile Ile Arg Cys Gln Val Leu Leu Pro Phe Ser Gln Ser Val Ala
 545 550 555 560

Asp Leu Pro Ala Asp Glu Cys Glu Gly Pro Lys Pro Ala Ser Gln Arg
 565 570 575

Ala Cys Tyr Ala Gly Pro Cys Asn Gly Glu Thr Pro Glu Phe Asn Pro
 580 585 590

Asp Asn Gly Asp Gly Leu Gly Leu Gln Asp Leu Asp Glu Leu Tyr Asp
 595 600 605

Trp Glu Tyr Glu Gly Phe Thr Lys Cys Ser Glu Ser Cys Gly Gly Gly
 610 615 620

Val Gln Glu Ala Val Val Ser Cys Leu Asn Lys Gln Thr Arg Glu Leu
 625 630 635 640

Ala Asp Glu Asn Leu Cys Val Thr Ser Arg Arg Pro Pro Gln Leu Leu
 645 650 655

Lys Ser Cys Asn Leu Asp Pro Cys Pro Ala Arg Trp Glu Ile Gly Lys
 660 665 670

Trp Ser Pro Cys Ser Leu Thr Cys Gly Val Gly Leu Gln Thr Arg Asp
 675 680 685

Val Val Cys Ser His Leu Leu Ser Arg Glu Met Asn Glu Val Val Val
 690 695 700

Leu Ala Asp Glu Leu Cys His His Pro Lys Pro Ser Thr Val Gln Ala
 705 710 715 720

Cys Asn Arg Phe Asn Cys Pro Pro Ala Trp Tyr Pro Ala Gln Trp Gln
 725 730 735

Leu Cys Ser Arg Thr Cys Gly Gly Gly Ile Gln Lys Arg Asp Val Leu

740	745	750
Cys Lys Gln Arg Met Ala Asp Gly Ser Phe Leu Glu Leu Pro Glu Thr		
755	760	765
Phe Cys Ser Ala Ser Lys Pro Thr Ser His Gln Gly Cys Lys Lys Asp		
770	775	780
Asp Cys Pro Ser Glu Trp Leu Leu Ser Glu Trp Ser Glu Cys Ser Val		
785	790	795
Ser Cys Gly Glu Gly Thr Gln Thr Arg Ser Ala Ile Cys Gln Arg Val		
805	810	815
Leu Lys Thr Gly Val Ser Thr Val Val Asn Ser Thr Leu Cys Pro Pro		
820	825	830
Leu Pro Phe Ser Ser Ser Ile Arg Pro Cys Met Leu Ala Thr Cys Ala		
835	840	845
Arg Pro Gly Arg Pro Ser Thr Lys His Ser Pro His Ile Ala Ala Ala		
850	855	860
Arg Asn Ile Tyr Ile Gln Thr Arg Arg Gln Arg Lys Leu His Phe Val		
865	870	875
Val Gly Gly Phe Ala Tyr Leu Leu Pro Lys Thr Thr Val Val Leu Arg		
885	890	895
Cys Pro Thr Arg Arg Phe Arg Lys Pro Leu Ile Thr Trp Glu Lys Asp		
900	905	910
Gly Gln His Ser Ile Ser Ser Ala His Val Thr Val Ala Pro Phe Gly		
915	920	925
Tyr Leu Lys Ile His Arg Leu Lys Pro Ser Asp Ala Gly Ile Tyr Thr		
930	935	940
Cys Ser Ala Gly Pro Ala Arg Glu Gln Phe Val Ile Lys Leu Ile Gly		
945	950	955
Gly Asn Arg Lys Leu Val Ala Arg Pro Leu Ser Leu Trp Ser Glu Glu		
965	970	975

Glu Glu Ala Leu Gln Val Arg Lys Ala Asn Pro Lys Glu Ala Leu Gln
 980 985 990

Thr His Lys His Gln Asn Gly Ile Phe Ser Asn Gly Ser Lys Ala Glu
 995 1000 1005

Lys Arg Gly Leu Thr Ala Asp Pro Gly Asn Arg Tyr Asp Asp Ile
 1010 1015 1020

Val Ser Arg Leu Leu Glu Gln Gly Gly Trp Pro Gly Glu Leu Leu
 1025 1030 1035

Ala Ser Trp Glu Val Gln Asp Ser Ala Glu Arg Asn Ala Ser Ser
 1040 1045 1050

Glu Glu Asp Pro Asn Ala Glu Gln Ala Leu Leu His Leu Pro Phe
 1055 1060 1065

Thr Met Val Ala Glu Gln Lys Arg Leu Asp Asp Ile Leu Arg Asn
 1070 1075 1080

Leu Ser Gln Gln Pro Glu Glu Leu Arg Asp Leu Tyr Ser Lys His
 1085 1090 1095

Leu Val Ala Gln Leu Ala Gln Asp Ile Phe Arg Ser His Leu Glu
 1100 1105 1110

Asn Gln Asp Leu Leu Pro Lys Pro Ser Glu Gln Arg Phe Pro Pro
 1115 1120 1125

Met Ala Val Pro Ala His Lys His Val Ser Gly Phe Ser Ser Ser
 1130 1135 1140

Leu Arg Ser Ser Ser Gly Glu Ala Gly Gly Gly Ser Arg Arg Pro
 1145 1150 1155

His Arg Lys Pro Ala Ile Leu Arg Lys Ile Ser Ala Ala Gln Gln
 1160 1165 1170

Leu Ser Ala Ser Glu Val Val Thr His Leu Gly Gln Thr Val Ala
 1175 1180 1185

Leu	Ala	Ser	Gly	Thr	Leu	Ser	Val	Leu	Leu	His	Cys	Glu	Ala	Val
1190						1195					1200			
Gly	Asn	Pro	Arg	Pro	Thr	Ile	His	Trp	Thr	Arg	Asn	Gly	Glu	Ala
1205						1210					1215			
Val	Gln	Phe	Ser	Asp	Arg	Ile	Leu	Leu	Gln	Pro	Asp	Asp	Ser	Leu
1220						1225					1230			
Gln	Ile	Leu	Ala	Pro	Val	Glu	Ala	Asp	Val	Gly	Phe	Tyr	Thr	Cys
1235						1240					1245			
Asn	Ala	Thr	Asn	Ala	Leu	Gly	Tyr	Asp	Ser	Val	Ser	Ile	Ala	Val
1250						1255					1260			
Thr	Leu	Ala	Gly	Lys	Pro	Leu	Val	Lys	Thr	Ser	Arg	Met	Thr	Val
1265						1270					1275			
Leu	Asn	Thr	Glu	Lys	Pro	Thr	Val	Thr	Val	Asp	Ile	Gly	Gly	Thr
1280						1285					1290			
Val	Arg	Thr	Val	Arg	Gly	Val	Asn	Val	Thr	Ile	Asn	Cys	Gln	Val
1295						1300					1305			
Ala	Gly	Val	Pro	Glu	Ala	Glu	Val	Thr	Trp	Phe	Arg	Asn	Lys	Ser
1310						1315					1320			
Lys	Leu	Gly	Ser	Ser	His	His	Leu	His	Glu	Gly	Ser	Ser	His	His
1325						1330					1335			
Leu	His	Glu	Gly	Ser	Leu	Leu	Leu	Thr	Asp	Val	Ser	Phe	Ser	Asp
1340						1345					1350			
Gln	Gly	Leu	Tyr	Ser	Cys	Arg	Ala	Ala	Asn	Leu	His	Gly	Glu	Gln
1355						1360					1365			
Thr	Glu	Ser	Thr	Gln	Leu	Leu	Ile	Leu	Asp	Pro	Pro	Gln	Val	Pro
1370						1375					1380			
Thr	Gln	Leu	Glu	Asp	Ile	Arg	Ala	Leu	Leu	Leu	Ala	Thr	Gly	Pro
1385						1390					1395			

Asn Leu Pro Ser Val Leu Met Ser Pro Leu Gly Thr Gln Leu Val	
1400	1405 1410
Leu Asp Pro Gly Asn Ser Ala Leu Leu Gly Cys Pro Ile Lys Gly	
1415	1420 1425
His Pro Thr Pro Asn Ile Thr Trp Phe Gln Asn Gly Gln Pro Ile	
1430	1435 1440
Ala Thr Ala Pro Gly Leu Thr His His Ile Trp Gly Ala Gly Gln	
1445	1450 1455
Ile Leu Arg Val Ala Asn Leu Ser Gly Gly Pro Gln Gly Glu Phe	
1460	1465 1470
Ser Cys Leu Ala Gln Asn Glu Ala Gly Thr Leu Leu Gln Lys Ala	
1475	1480 1485
Ser Leu Val Ile Gln Asp Tyr Trp Trp Ser Val Asp Arg Leu Ala	
1490	1495 1500
Thr Cys Ser Ala Ser Cys Gly Asn Arg Gly Ile His Gln Pro Arg	
1505	1510 1515
Leu Arg Cys Leu Leu Asn Thr Thr Glu Val Asp Pro Glu His Cys	
1520	1525 1530
Thr Gly Lys Pro Arg Pro Ala Val Gln Pro Val Ala Cys Asn Arg	
1535	1540 1545
Arg Asp Cys Pro Ser Arg Trp Met Val Thr Ser Trp Ser Ala Cys	
1550	1555 1560
Thr Arg Ser Cys Gly Gly Gly Val Gln Thr Arg Arg Val Thr Cys	
1565	1570 1575
Gln Lys Leu Lys Ala Ser Gly Ile Ser Thr Pro Val Ser Asn Asp	
1580	1585 1590
Met Cys Ser Gln Leu Ala Lys Arg Pro Val Asp Thr Gln Ala Cys	
1595	1600 1605
Asn Gln Gln Leu Cys Val Glu Trp Ala Phe Ser Ser Trp Gly Gln	

1610 1615 1620
 Cys Asn Gly Pro Cys Ile Gly Pro Arg Leu Ala Val Gln His Arg
 1625 1630 1635
 Gln Val Phe Cys Gln Thr Arg Asp Gly Ile Thr Leu Pro Ser Glu
 1640 1645 1650
 Gln Cys Ser Ala Leu Pro Arg Pro Val Ser Thr Gln Asn Cys Trp
 1655 1660 1665
 Ser Glu Ala Cys Ser Val His Trp Arg Val Ser Leu Trp Thr Leu
 1670 1675 1680
 Cys Thr Ala Thr Cys Gly Asn Tyr Gly Phe Gln Ser Arg Arg Val
 1685 1690 1695
 Glu Cys Val His Val Arg Thr Asn Lys Ala Val Pro Glu His Leu
 1700 1705 1710
 Cys Ser Trp Gly Pro Arg Pro Ala Asn Trp Gln Arg Cys Asn Val
 1715 1720 1725
 Thr Pro Cys Glu Asn Thr Glu Cys Arg Asp Thr Thr Arg Tyr Cys
 1730 1735 1740
 Glu Lys Val Arg Gln Leu Lys Leu Cys Gln Leu Gly Gln Phe Arg
 1745 1750 1755
 Ser Arg Cys Cys Gly Thr Cys Gly Lys Ala Leu Gly
 1760 1765 1770

<210> 7
 <211> 26
 <212> DNA
 <213> artificial

<220>
 <223> 5' primer used for amplification of 3' UTR of Adamts1-1

<400> 7
 ggacagcaag gagtttataa cacata

26

<210> 8
 <211> 20

<212> DNA
<213> unknown

<220>
<223> 3' primer used for amplification of 3' UTR of Adamts1-1

<400> 8
tttcttggtc ttcccacacc 20

<210> 9
<211> 18
<212> DNA
<213> unknown

<220>
<223> 5' primer used for amplification of Tex30

<400> 9
aaatggtacc ggtgacat 18

<210> 10
<211> 18
<212> DNA
<213> unknown

<220>
<223> 3' primer used for amplification of Tex30

<400> 10
tattttcttg cctcctta 18

<210> 11
<211> 18
<212> DNA
<213> unknown

<220>
<223> 5' primer used for amplification of Adamts1-1

<400> 11
ctcccaccac ctgcatga 18

<210> 12
<211> 20
<212> DNA
<213> unknown

<220>
<223> 3' primer used for amplification of Adamts1-1

<400> 12
ccttgagggtc ctccactgag 20