



- (72) GRAVEL, Roy A., CA
(72) ROZEN, Rima, CA
(72) LECLERC, Daniel, CA
(72) GOYETTE, Philippe, CA
(72) CAMPEAU, Eric, CA
(71) MARTINEX R&D, INC., CA
(51) Int.Cl.⁶ C12N 15/54, A61K 31/195, A01K 67/027, C12N 9/10,
C07K 16/40, A61K 31/205, A61K 38/45, C12Q 1/48, G01N 33/573,
A61K 31/68, C12Q 1/68, A61K 31/70, A61K 38/03
(30) 1996/11/27 (60/031,964) US
(30) 1997/06/20 (60/050,310) US
(54) **METHIONINE SYNTHETASE HUMAINE : CLONAGE ET
METHODES D'EVALUATION DU RISQUE DE
MALFORMATIONS DU TUBE NEURAL, DE MALADIES
CARDIOVASCULAIRES ET DE CANCER**
(54) **HUMAN METHIONINE SYNTHASE: CLONING, AND
METHODS FOR EVALUATING RISK OF NEURAL TUBE
DEFECTS, CARDIOVASCULAR DISEASE, AND CANCER**

(57) L'invention décrit une méthode de détection du risque accru d'hyperhomocystéinémie ainsi que du risque réduit ou accru de malformations du tube neural ou de maladies cardiovasculaires. L'invention décrit également des méthodes thérapeutique visant à réduire le risque de malformations du tube neural, de cancer du côlon et de cancers connexes. On décrit aussi les séquences de la protéine et du gène de la méthionine synthétase humaine ainsi que des composés et des trousseaux pour appliquer les méthodes décrites dans l'invention.

(57) The invention features a method for detecting an increased likelihood of hyperhomocysteinemia and, in turn, an increased or decreased likelihood of neural tube defects or cardiovascular disease. The invention also features therapeutic methods for reducing the risk of neural tube defects, colon cancers and related cancers. Also provided are the sequences of the human methionine synthase gene and protein and compounds and kits for performing the methods of the invention.

**HUMAN METHIONINE SYNTHASE: CLONING, AND METHODS
FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,
CARDIOVASCULAR DISEASE, AND CANCER**

Abstract of Invention

The invention features a method for detecting an increased likelihood of hyperhomocysteinemia and, in turn, an increased or decreased likelihood of neural tube defects or cardiovascular disease. The invention also features therapeutic methods for reducing the risk of neural tube defects, colon cancers and related cancers. Also provided are the sequences of the human methionine synthase gene and protein and compounds and kits for performing the methods of the invention.

-1-

HUMAN METHIONINE SYNTHASE: CLONING, AND METHODS
FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,
5 CARDIOVASCULAR DISEASE, AND CANCER

Field of the Invention

The invention relates to the diagnosis and treatment of patients at risk for methionine synthase deficiency and associated altered risk for diseases such as neural tube defects, cardiovascular disease, and cancer.

10 Background of the Invention

Methionine synthase (EC 2.1.1.13, 5-methyltetrahydrofolate-homocysteine methyltransferase) catalyses the remethylation of homocysteine to methionine in a reaction in which methylcobalamin serves as an intermediate methyl carrier. This occurs by transfer of the methyl group of 5-methyltetrahydrofolate to the enzyme-bound cob(I)alamin to form
15 methylcobalamin with subsequent transfer of the methyl group to homocysteine to form methionine. Over time, cob(I)alamin may become oxidized to cob(II)alamin rendering the enzyme inactive. Regeneration of the functional enzyme occurs through the methionine synthase-mediated methylation of the
20 cob(II)alamin in which S-adenosylmethionine is utilized as methyl donor. In *E. coli*, two flavodoxins have been implicated in the reductive activation of methionine synthase (Fujii, K. and Huennekens, F.M. (1974) *J. Biol. Chem.*,

-2-

249, 6745-6753). A methionine synthase-linked reducing system has yet to be identified in mammalian cells.

Deficiency of methionine synthase activity results in hyperhomocysteinemia, homocystinuria, and megaloblastic anemia without methylmalonic aciduria (Rosenblatt, D.S. (1995) *The Metabolic and Molecular Bases of Inherited Disease*. McGraw-Hill, New York, pp. 3111-3128; Fenton, W.A. and Rosenberg, L.E. (1995) *The Metabolic and Molecular Bases of Inherited Disease*. McGraw-Hill, New York, pp. 3129-3149). Two classes of methionine synthase-associated genetic diseases have been proposed based on complementation experiments between patient fibroblast cell lines (Watkins, D. and Rosenblatt, D.S. (1988) *J. Clin. Invest.*, **81**, 1690-1694). One complementation group, *cblE*, has been postulated to be due to deficiency of the reducing system required for methionine synthesis (Rosenblatt, D.S., Cooper, B.A., Pottier, A., Lue-Shing, H., Matiaszuk, N. and Grauer, K. (1984) *J. Clin. Invest.*, **74**, 2149-2156). Cells from patients in the *cblE* group fail to incorporate ¹⁴C-methyltetrahydrofolate into methionine in whole cells but have significant methionine synthase activity in cell extracts in the presence of a potent reducing agent. The second complementation group, *cblG* group, is thought to result from defects of the methionine synthase apoenzyme. Mutant cells from this group show deficient methionine synthase activity in both whole cells and cell extracts (Watkins, D. and Rosenblatt, D.S. (1988) *J. Clin. Invest.*, **81**, 1690-1694; Watkins, D. and Rosenblatt, D.S. (1989) *Am. J. Med. Genet.*, **34**, 427-434). Moreover, some *cblG* patients show defective binding of cobalamin to methionine synthase in cells incubated with radiolabelled cyanocobalamin (Sillaots, S.L., Hall, C.A., Hurteloup, V., and Rosenblatt, D.S. (1992) *Biochem. Med. Metab. Biol.*, **47**, 242-249).

-3-

The cobalamin-dependent methionine synthase of *E. coli* has been crystallized and the structure of its active site determined (Luschinsky, C.L., Drummond, J.T., Matthews, R.G., and Ludwig, M.L. (1992) *J. Molec. Biol.*, **225**, 557-560; Drennan, C.L., Huang, S., Drummond, J.T., Matthews, R.G., and Ludwig, M.L. (1994) *Science*, **266**, 1669-1674.). The gene encoding methionine synthase has not been cloned from mammals.

Summary of the Invention

We have cloned a gene for mammalian methionine synthase from humans and discovered that mutations in this gene are associated with hyperhomocysteinemia. Hyperhomocysteinemia is a condition that has been implicated in cardiovascular disease and neural tube defects. The presence of such mutations in methionine synthase gene are, thus, associated with increased risk for cardiovascular disease, altered risk for neural tube defects, and decreased risk of colon cancer. The invention features methods for risk detection and treatment of patients with hyperhomocysteinemia, cardiovascular disease, neural tube defects, and cancer. The invention also features compounds and kits which may be used to practice the methods of the invention, methods and compounds for treating or preventing these conditions and methods of identifying therapeutics for the treatment and prevention of these conditions.

In the first aspect, the invention provides purified wild-type mammalian methionine synthase gene, and mutated and polymorphic versions of the mammalian methionine synthase gene, fragments of the wild-type, mutated, and polymorphic gene, and sense and antisense sequences which may be used in the methods of the invention. Preferably, the gene is human. The proteins encoded therefrom are also an aspect of the invention as is a

-4-

methionine synthase polypeptide having conservative substitutions. Preferably, the protein is a recombinant or purified protein having a mutation conferring hyperhomocysteinemia when present in a mammal. In addition, nucleic acids, including genomic DNA, mRNA, and cDNA, and the nucleic acid set forth in

5 SEQ ID NO: 1, or degenerate variants thereof, are provided. The shorter nucleic acid sequences are appropriate for use in cloning, characterizing mutations, the construction of mutations, and creating deletions. In one embodiment, the nucleic acid set forth in SEQ ID NO: 1 is a probe that hybridizes at high stringency to sequences found within the nucleic acid of

10 SEQ ID NO: 1. In further embodiments, the probe has a sequence complementary to at least 50% of at least 60 nucleotides, or the sequence is complementary to at least 90% of at least 18 nucleotides. Protein fragments also are provided. The shorter peptides may be used, for example, in the generation of antibodies to the methionine synthase protein. In some

15 embodiments of this aspect of the invention nucleic acid fragments useful for detection of mutations in the region of the methionine synthase gene which encodes the cobalamin binding domain, and for detecting those mutations which indicate an increased likelihood of hyperhomocysteinemia, are preferred. Most preferred fragments are those useful for detecting the 2756 A → G, Δbp

20 2640-2642, and 2758 C→G mutations/polymorphisms. Given Applicants' discovery, one skilled in the art may readily determine which nucleic acids, detection methods, and mutations are most useful. Mutant proteins encoded by these mutations, including, but not limited to, H920D, ΔIle 881, and D919G are also provided by the invention. Such mutant and polymorphic polypeptides

25 may have decreased or increased biological activity, relative to wild-type methionine synthase.

-5-

In a related aspect, the invention provides antibodies that specifically bind mammalian methionine synthase, and a method for generating such an antibody. The antibody may specifically bind a wild-type methionine synthase, or a mutant or polymorphic methionine synthase. A method for detecting a
5 wild-type, mutant, or polymorphic methionine synthase using the antibody is also provided by the invention.

In a second aspect, the invention provides a method for detecting an increased or decreased risk for hyperhomocysteinemia in a fetus or individual patient. Such a fetus or patient is at increased or decreased risk for neural tube
10 defects and/or cardiovascular disease and at a decreased risk of developing colon cancer. The method includes detection of mutations in the methionine synthase gene present in the fetus, the individual patient, and/or the blood relatives of the fetus and patient. The presence of mutations, particularly in the cobalamin binding domain, indicate an altered (e.g., increased or decreased)
15 risk of hyperhomocysteinemia, neural tube defects, cancer, and cardiovascular disease.

In a related aspect, the invention provides kits for the detection of mutations in the human methionine synthase gene. Such kits may include, for example, nucleic acid sequences, including probes, useful for PCR, SSCP, or
20 RFLP detection of such mutations. Antibodies specific for proteins having mutations, correlated with an increased likelihood of hyperhomocysteinemia, may also be included in the kits of the invention.

In a fourth aspect, the invention features a method for screening for compounds which alter methionine synthase expression or ameliorate or
25 exacerbate conditions of hyperhomocysteinemia. In various embodiments, the invention includes monitoring mutant or wild-type mammalian methionine synthase biological activity by monitoring methionine synthase enzymatic

-6-

activity, or monitoring methionine synthase gene expression levels, by monitoring methionine synthase gene transcription, RNA stability, RNA translation and/or protein stability. In preferred embodiments the methionine synthase gene or protein being monitored is a gene or protein having a mutation associated with hyperhomocysteinemia, and samples are selected from purified or partially purified methionine synthase, cell lysate, a cell, or an animal. Standard assay techniques known to those skilled in the art may be employed in the various embodiments. Compounds detected using this screen can be used to prevent or treat cardiovascular disease and neural tube defects or, in the alternative, to prevent or treat colon cancer. Kits for performing the above screens are also a part of the invention.

In a related aspect, the invention provides nucleic acids encoding wild-type, polymorphic, and mutated methionine synthase, in which the nucleic acid is operably linked to regulatory sequences, comprising a promoter, for the expression of the encoded polypeptides. In one embodiment, the promoter is inducible. The invention also provides cells, including prokaryotic and eukaryotic cells, comprising the nucleic acids. The eukaryotic cells may be yeast cells or mammalian cells.

In another related aspect, the invention features a transgenic mammal having a methionine synthase transgene. The gene may be wild-type, or may contain a mutation or polymorphism. The mammal may have a mutation associated with hyperhomocysteinemia in its methionine synthase gene in an expressible genetic construction or may have a deletion or knockout mutation in one or both alleles sufficient to abolish methionine synthase expression from the locus. In addition, or as a replacement, the mammal may have the methionine synthase gene from another species. For example, in one preferred embodiment the transgenic mammal is a rodent such as a mouse and the

-7-

transgene is from a human. Cells from these transgenic or knockout animals are also provided by the invention. Such transgenic mammals may be used to screen for drugs for the treatment of diseases related to hyperhomocysteinemia.

In a sixth aspect, the invention features a method for treating patients
5 with neural tube defects, colon cancer or related cancers by the delivery of antisense methionine synthase nucleic acid sufficient to lower the levels of methionine synthase polypeptide biological activity.

In a related aspect, the invention provides a method for treating or preventing cardiovascular disease, neural tube defects and cancer. The method
10 comprises detecting an altered risk of such defects by analyzing methionine synthase nucleic acid, potential test subjects being a mammal, a potential parent, either male or female, a pregnant mammal, or a developing embryo or fetus, and then by exposing the subject (e.g., patient or pregnant mammal) to metabolites or cofactors such as, but not limited to, folate, cobalamin, S-
15 adenosyl methionine, betaine, or methionine. In another related aspect, the invention features a method of pretreating or treating colon cancer or neural tube defects by inhibiting or activating methionine synthase biological activity in a mammal, pregnant mammal, embryo, or fetus. In preferred embodiments, this inhibiting or activating may be effected by exposing the subject to nucleic
20 acids, peptides or small molecule-based inhibitors or activators of methionine synthase or substrates. The exposure is to quantities of the compound sufficient to reduce the probability of the subject developing the disease or to confer an increased likelihood of a decrease in the disease symptoms of the subject.

25 By "methionine synthase," "methionine synthase protein," or "methionine synthase polypeptide" is meant a polypeptide, or fragment thereof, which has at least 50% amino acid identity to boxes 1-4 of the human

-8-

methionine synthase polypeptide (SEQ ID NO: 2) (see Fig. 1). It is understood that polypeptide products from splice variants of methionine synthase gene sequences are also included in this definition. Preferably, the methionine synthase protein is encoded by nucleic acid having a sequence which

5 hybridizes to a nucleic acid sequence present in SEQ ID NO: 1 (human methionine synthase cDNA) under stringent conditions. Even more preferably the encoded polypeptide also has methionine synthase biological activity.

By “methionine synthase nucleic acid” or “methionine synthase gene” is meant a nucleic acid, such as genomic DNA, cDNA, or mRNA, that

10 encodes methionine synthase, a methionine synthase protein, methionine synthase polypeptide, or portion thereof, as defined above. A methionine synthase nucleic acid also may be a methionine synthase primer or probe, or antisense nucleic acid that is complementary to a methionine synthase nucleic acid.

15 By “wild-type methionine synthase” is meant a methionine synthase nucleic acid or methionine synthase polypeptide having the nucleic acid and/or amino acid sequence most often observed among members of a given animal species and not statistically associated with a disease phenotype. Wild-type methionine synthase is biologically active methionine synthase. A wild-type

20 methionine synthase is, for example, a human methionine synthase polypeptide having the sequence of SEQ ID NO: 1.

By “mutant methionine synthase,” “methionine synthase mutation(s),” “mutations in methionine synthase,” “polymorphic methionine synthase,” “methionine synthase polymorphism(s),” “polymorphisms in

25 methionine synthase,” is meant a methionine synthase polypeptide or nucleic acid having a sequence that deviates from the wild-type sequence in a manner sufficient to confer an altered risk for a disease phenotype, or enhanced

-9-

protection against a disease, in at least some genetic and/or environmental backgrounds. Such mutations may be naturally occurring or artificially induced. They may be, without limitation, insertion, deletion, frameshift, or missense mutations. A mutant methionine synthase protein may have one or more mutations, and such mutations may affect different aspects of methionine synthase biological activity (protein function), to various degrees.

Alternatively, a methionine synthase mutation may indirectly affect methionine synthase biological activity by influencing, for example, the transcriptional activity of a gene encoding methionine synthase, or the stability of methionine synthase mRNA. For example, a mutant methionine synthase gene may be a gene which expresses a mutant methionine synthase protein or may be a gene which alters the level of methionine synthase protein in a manner sufficient to confer a disease phenotype in at least some genetic and/or environmental backgrounds.

By “biologically active” methionine synthase is meant a methionine synthase protein or methionine synthase gene that provides at least one biological function equivalent to that of the wild-type methionine synthase polypeptide or methionine synthase gene. Biological activities of a methionine synthase polypeptide include, and are not limited to, the ability to catalyze the methylation of homocysteine to generate methionine. Preferably, a biologically active methionine synthase will display activity equivalent to at least 35% of wild-type activity, more preferably, a biologically active methionine synthase will display at least 40-55% of wild-type activity, still more preferably, a biologically active methionine synthase will display at least 60-75% of wild-type activity, and most preferably, a biologically active methionine synthase will display at least 80-90% of wild-type activity. A biologically active methionine synthase also may display more than 100% of wild-type activity.

-10-

Preferably, the biological activity of the wild-type methionine synthase is determined using the methionine synthase nucleic acid of SEQ ID NO: 1 or methionine synthase polypeptide of SEQ ID NO: 2. The degree of methionine synthase biological activity may be intrinsic to the methionine synthase polypeptide itself, or may be modulated by increasing or decreasing the number of methionine synthase polypeptide molecules present intracellularly.

By "high stringency conditions" is meant hybridization in 2X SSC at 40°C with a DNA probe length of at least 40 nucleotides. For other definitions of high stringency conditions, see F. Ausubel et al., Current Protocols in Molecular Biology, pp. 6.3.1-6.3.6, John Wiley & Sons, New York, NY, 1994, hereby incorporated by reference.

By "analyzing" or "analysis" is meant subjecting a methionine synthase nucleic acid or methionine synthase polypeptide to a test procedure that allows the determination of whether a methionine synthase gene is wild-type or mutant. For example, one could analyze the methionine synthase genes of an animal by amplifying genomic DNA using the polymerase chain reaction, and then determining the DNA sequence of the amplified DNA.

By "probe" or "primer" is meant a single- or double-stranded DNA or RNA molecule of defined sequence that can base pair to a second DNA or RNA molecule that contains a complementary sequence (the "target"). The stability of the resulting hybrid depends upon the extent of the base pairing that occurs. The extent of base-pairing is affected by parameters such as the degree of complementarity between the probe and target molecules, and the degree of stringency of the hybridization conditions. The degree of hybridization stringency is affected by parameters such as temperature, salt concentration, and the concentration of organic molecules such as formamide, and is determined by methods known to one skilled in the art. Probes or primers

-11-

specific for methionine synthase nucleic acid preferably will have at least 35% sequence identity, more preferably at least 45-55% sequence identity, still more preferably at least 60-75% sequence identity, still more preferably at least 80-90% sequence identity, and most preferably 100% sequence identity. Probes
5 may be detectably-labelled, either radioactively, or non-radioactively, by methods well-known to those skilled in the art. Probes are used for methods involving nucleic acid hybridization, such as: nucleic acid sequencing, nucleic acid amplification by the polymerase chain reaction, single stranded conformational polymorphism (SSCP) analysis, restriction fragment
10 polymorphism (RFLP) analysis, Southern hybridization, Northern hybridization, in situ hybridization, electrophoretic mobility shift assay (EMSA).

By "pharmaceutically acceptable carrier" means a carrier which is physiologically acceptable to the treated mammal while retaining the
15 therapeutic properties of the compound with which it is administered. One exemplary pharmaceutically acceptable carrier is physiological saline. Other physiologically acceptable carriers and their formulations are known to one skilled in the art and described, for example, in Remington's Pharmaceutical Sciences, (18th edition), ed. A. Gennaro, 1990, Mack Publishing Company,
20 Easton, PA.

By "substantially identical" is meant a polypeptide or nucleic acid exhibiting at least 50%, preferably 85%, more preferably 90%, and most preferably 95% identity to a reference amino acid or nucleic acid sequence. For polypeptides, the length of comparison sequences will generally be at least
25 16 amino acids, preferably at least 20 amino acids, more preferably at least 25 amino acids, and most preferably 35 amino acids. For nucleic acids, the length of comparison sequences will generally be at least 50 nucleotides,

-12-

preferably at least 60 nucleotides, more preferably at least 75 nucleotides, and most preferably 110 nucleotides.

Sequence identity is typically measured using sequence analysis software with the default parameters specified therein (e.g., Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705). This software program matches similar sequences by assigning degrees of homology to various substitutions, deletions, and other modifications. Conservative nucleotide substitutions typically include substitutions which generate changes within the following groups: glycine, alanine, valine, isoleucine, leucine; aspartic acid, glutamic acid, asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine.

By “substantially pure polypeptide” is meant a polypeptide that has been separated from the components that naturally accompany it. Typically, the polypeptide is substantially pure when it is at least 60%, by weight, free from the proteins and naturally-occurring organic molecules with which it is naturally associated. Preferably, the polypeptide is a methionine synthase polypeptide that is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight, pure. A substantially pure methionine synthase polypeptide may be obtained, for example, by extraction from a natural source (e.g., a fibroblast or liver cell) by expression of a recombinant nucleic acid encoding a methionine synthase polypeptide, or by chemically synthesizing the protein. Purity can be measured by any appropriate method, e.g., by column chromatography, polyacrylamide gel electrophoresis, or HPLC analysis.

A protein is substantially free of naturally associated components when it is separated from those contaminants which accompany it in its natural

-13-

state. Thus, a protein which is chemically synthesized or produced in a cellular system different from the cell from which it naturally originates will be substantially free from its naturally associated components. Accordingly, substantially pure polypeptides not only includes those derived from eukaryotic organisms but also those synthesized in *E. coli* or other prokaryotes.

By “substantially pure DNA” is meant DNA that is free of the genes which, in the naturally-occurring genome of the organism from which the DNA of the invention is derived, flank the gene. The term therefore includes, for example, a recombinant DNA which is incorporated into a vector; into an autonomously replicating plasmid or virus; or into the genomic DNA of a prokaryote or eukaryote; or which exists as a separate molecule (e.g., a cDNA or a genomic or cDNA fragment produced by PCR or restriction endonuclease digestion) independent of other sequences. It also includes a recombinant DNA which is part of a hybrid gene encoding additional polypeptide sequence.

By “transgene” is meant any piece of DNA which is inserted by artifice into a cell, and becomes part of the genome of the organism which develops from that cell. Such a transgene may include a gene which is partly or entirely heterologous (i.e., foreign) to the transgenic organism, or may represent a gene homologous to an endogenous gene of the organism.

By “transgenic” is meant any cell which includes a DNA sequence which is inserted by artifice into a cell and becomes part of the genome of the organism which develops from that cell. As used herein, the transgenic organisms are generally transgenic mammals (e.g., rodents such as rats or mice) and the DNA (transgene) is inserted by artifice into the nuclear genome. Preferably the inserted DNA encodes a protein in at least some cells of the organism.

-14-

By “knockout mutation” is meant an alteration in the nucleic acid sequence that reduces the biological activity of the polypeptide normally encoded therefrom by at least 80% relative to the unmutated gene. The mutation may, without limitation, be an insertion, deletion, frameshift mutation, or a missense mutation. Preferably, the mutation is an insertion or deletion, or is a frameshift mutation that creates a stop codon.

By “transformation” is meant any method for introducing foreign molecules into a cell. Lipofection, DEAE-dextran-mediated transfection, microinjection, protoplast fusion, calcium phosphate precipitation, retroviral delivery, electroporation, and biolistic transformation are just a few of the methods known to those skilled in the art which may be used. For example, biolistic transformation is a method for introducing foreign molecules into a cell using velocity driven microprojectiles such as tungsten or gold particles. Such velocity-driven methods originate from pressure bursts which include, but are not limited to, helium-driven, air-driven, and gunpowder-driven techniques. Biolistic transformation may be applied to the transformation or transfection of a wide variety of cell types and intact tissues including, without limitation, intracellular organelles (e.g., and mitochondria and chloroplasts), bacteria, yeast, fungi, algae, animal tissue, and cultured cells.

By “transformed cell” is meant a cell into which (or into an ancestor of which) has been introduced, by means of recombinant DNA techniques, a DNA molecule encoding (as used herein) a methionine synthase polypeptide.

By “positioned for expression” is meant that the DNA molecule is positioned adjacent to a DNA sequence which directs transcription and translation of the sequence (i.e., facilitates the production of, e.g., a methionine synthase polypeptide, a recombinant protein or a RNA molecule).

-15-

By “promoter” is meant a minimal sequence sufficient to direct transcription. Also included in the invention are those promoter elements which are sufficient to render promoter-dependent gene expression controllable for cell type-specific, tissue-specific, temporal-specific, or inducible by
5 external signals or agents; such elements may be located in the 5' or 3' or intron sequence regions of the native gene.

By “operably linked” is meant that a gene and one or more regulatory sequences are connected in such a way as to permit gene expression when the appropriate molecules (e.g., transcriptional activator proteins) are
10 bound to the regulatory sequences.

By “conserved region” is meant any stretch of six or more contiguous amino acids exhibiting at least 30%, preferably 50%, and most preferably 70% amino acid sequence identity between two or more of the methionine synthase family members, (e.g., between human and bacterial
15 methionine synthase). Examples of conserved regions within methionine synthase are Boxes 1-4 (Fig. 1).

By “detectably-labeled” is meant any means for marking and identifying the presence of a molecule, e.g., an oligonucleotide probe or primer, a gene or fragment thereof, or a cDNA molecule. Methods for detectably-
20 labeling a molecule are well known in the art and include, without limitation, radioactive labeling (e.g., with an isotope such as ^{32}P or ^{35}S) and nonradioactive labeling (e.g., chemiluminescent labeling, e.g., fluorescein labeling).

By “antisense” as used herein in reference to nucleic acids, is meant a nucleic acid sequence that is complementary to the coding strand of a gene,
25 preferably, a methionine synthase gene. An antisense nucleic acid is capable of preferentially lowering the activity of a mutant methionine synthase polypeptide encoded by a mutant methionine synthase gene.

-16-

By “purified antibody” is meant antibody which is at least 60%, by weight, free from proteins and naturally occurring organic molecules with which it is naturally associated. Preferably, the preparation is at least 75%, more preferably 90%, and most preferably at least 99%, by weight, antibody, e.g., a methionine synthase amino-terminus-specific antibody. A purified antibody may be obtained, for example, by affinity chromatography using recombinantly-produced protein or conserved motif peptides and standard techniques.

By “specifically binds” is meant an antibody that recognizes and binds a human methionine synthase polypeptide but that does not substantially recognize and bind other non-methionine synthase molecules in a sample, e.g., a biological sample, that naturally includes protein. A preferred antibody binds to the methionine synthase polypeptide sequence of SEQ ID NO: 2 (Fig. 3).

By “neutralizing antibodies” is meant antibodies that interfere with any of the biological activities of a wild-type or mutant methionine synthase polypeptide, for example, the ability of methionine synthase to catalyze the transfer of a methyl group to homocysteine. The neutralizing antibody may reduce the ability of a methionine synthase polypeptide to catalyze the transfer preferably by 10 % or more, more preferably by 25% or more, still more preferably by 50% or more, yet preferably by 70% or more, and most preferably by 90% or more. Any standard assay for the biological activity of methionine synthase, may be used to assess potentially neutralizing antibodies that are specific for methionine synthase.

By “expose” is meant to allow contact between an animal, cell, lysate or extract derived from a cell, or molecule derived from a cell, and a test compound.

-17-

By “treat” is meant to submit or subject an animal (e.g. a human), cell, lysate or extract derived from a cell, or molecule derived from a cell to a test compound.

By “test compound” is meant a chemical, be it naturally-occurring or artificially-derived, that is surveyed for its ability to modulate an alteration in reporter gene activity or protein levels, by employing one of the assay methods described herein. Test compounds may include, for example, peptides, polypeptides, synthesized organic molecules, naturally occurring organic molecules, nucleic acid molecules, and components thereof.

By “assaying” is meant analyzing the effect of a treatment, be it chemical or physical, administered to whole animals or cells derived therefrom. The material being analyzed may be an animal, a cell, a lysate or extract derived from a cell, or a molecule derived from a cell. The analysis may be for the purpose of detecting altered protein biological activity, altered protein stability, altered protein levels, altered gene expression, or altered RNA stability. The means for analyzing may include, for example, the detection of the product of an enzymatic reaction, (e.g., the formation of methionine as a result of methionine synthase activity), antibody labeling, immunoprecipitation, and methods known to those skilled in the art for detecting nucleic acids.

By “modulating” is meant changing, either by decrease or increase, in biological activity.

By “a decrease” is meant a lowering in the level of biological activity, as measured by a lowering/increasing of: a) the formation of methionine as a result of methionine synthase activity; b) protein, as measured by ELISA; c) reporter gene activity, as measured by reporter gene assay, for example, lacZ/ β -galactosidase, green fluorescent protein, luciferase, etc.; or d)

-18-

mRNA, levels of at least 30%, as measured by PCR relative to an internal control, for example, a “housekeeping” gene product such as β -actin or glyceraldehyde 3-phosphate dehydrogenase (GAPDH) or an externally added nucleic acid standard. In all cases, the lowering is preferably by at least 10%
5 more preferably by at least 25%, still more preferably by at least 50%, and even more preferably by at least 70%.

By “an increase” is meant a rise in the level of biological activity, as measured by a lowering/increasing of: a) the formation of methionine as a result of methionine synthase activity; b) protein, as measured by ELISA; c)
10 reporter gene activity, as measured by reporter gene assay, for example, lacZ/ β -galactosidase, green fluorescent protein, luciferase, etc.; or d) mRNA, levels of at least 30%, as measured by PCR relative to an internal control, for example, a “housekeeping” gene product such as β -actin or glyceraldehyde 3-phosphate dehydrogenase (GAPDH) or an externally added nucleic acid standard.
15 Preferably, the increase is by 10% or more, more preferably by 25% or more, even more preferably by 2-fold, and most preferably by at least 3-fold.

By “alteration in the level of gene expression” is meant a change in gene activity such that the amount of a product of the gene, i.e., mRNA or polypeptide, is increased or decreased, or that the stability of the mRNA or the
20 polypeptide is increased or decreased.

By “reporter gene” is meant any gene which encodes a product whose expression is detectable and/or quantifiable by immunological, chemical, biochemical or biological assays. A reporter gene product may, for example, have one of the following attributes, without restriction: fluorescence
25 (e.g., green fluorescent protein), enzymatic activity (e.g., lacZ/ β -galactosidase, luciferase, chloramphenicol acetyltransferase), toxicity (e.g., ricin A), or an ability to be specifically bound by a second molecule (e.g., biotin or a

-19-

detectably labelled antibody). It is understood that any engineered variants of reporter genes, which are readily available to one skilled in the art, are also included, without restriction, in the forgoing definition.

By “protein” or “polypeptide” or “polypeptide fragment” is meant
5 any chain of more than two amino acids, regardless of post-translational modification (e.g., glycosylation or phosphorylation), constituting all or part of a naturally-occurring polypeptide or peptide, or constituting a non-naturally occurring polypeptide or peptide.

By “missense mutation” is meant the substitution of one purine or
10 pyrimidine base (i.e. A, T, G, or C) by another within a nucleic acid sequence, such that the resulting new codon encodes an amino acid distinct from the amino acid originally encoded by the reference (e.g. wild-type) codon.

By “deletion mutation” is meant the deletion of at least one
nucleotide within a polynucleotide coding sequence. A deletion mutation alters
15 the reading frame of a coding region unless the deletion consists of one or more contiguous 3-nucleotide stretches (i.e. “codons”). Deletion of a codon from a nucleotide coding region results in the deletion of an amino acid from the resulting polypeptide.

By “frameshift mutation” is meant the insertion or deletion of at least
20 one nucleotide within a polynucleotide coding sequence. A frameshift mutation alters the codon reading frame at and/or downstream from the mutation site. Such a mutation results either in the substitution of the encoded wild-type amino acid sequence by a novel amino acid sequence, or a premature termination of the encoded polypeptide due to the creation of a stop codon, or
25 both.

Detailed Description of the Drawings

The drawings will first be briefly described.

-20-

Fig. 1 is a diagram showing four homologous regions among methionine synthases. Boxes 1 to 4 were used to design degenerate oligonucleotides for the initial cloning experiments. Ec: *Escherichia coli*, accession number J04975; Ss: *Synechocystis sp.*, accession number D64002; MI1 and MI2: *Mycobacterium leprae*, accession number U000175 (see Drennan *et al.*, 1994); Hi: *Haemophilus influenzae*, accession number U32730; Ce: *Caenorhabditis elegans*, accession number Z46828; Hs: *Homo sapiens*, this work. Identical residues are indicated by a star above the alignment. Amino acid position for each protein is shown at left.

Fig. 2 is a diagram showing overlapping PCR fragments generated to clone human methionine synthase. Oligonucleotides are described in Table 1. Primers in parentheses designate mispriming outcomes that generated valid internal sequence. iPCRc: inverse PCR on cDNA, iPCRg: inverse PCR on genomic DNA.

Fig. 3 is a diagram showing nucleotide sequence (SEQ ID NO: 1) and deduced amino acid sequence (SEQ ID NO: 2) of human methionine synthase. The nucleotide residue numbering is shown in the left margin, and the amino acid residue numbering is shown in the right margin.

Fig. 4 is a photograph showing mapping of the human methionine synthase gene using FISH. Signals are clearly visible at 1q43 (arrows).

Figs. 5A-5C is a series of photographs showing diagnostic tests for mutations in the methionine synthase gene. Numbers above the gel lanes correspond to patients cell lines whereas the letter "c" identifies wild-type controls. Fig. 5A: *Hae*III restriction analysis of genomic DNA PCR products using primers #1796 and #305A. The 2756A→G change creates a *Hae*III site. Expected fragments, 2756A allele: 189 bp, 2756G allele: 159 and 30 bp (the 30 bp fragment was run off the gel). Fig. 5B: Heteroduplex analysis of PCR

-21-

products amplified from RT reactions of patient 1892 and 3 controls. RT-PCR was done with primers #1772 and #1773. Expected PCR product: 338bp, heteroduplexes can be seen above this band in patient 1892 (heterozygous for Δ 2640-2642). C. *Sau*96I restriction analysis of genomic DNA PCR products.

5 PCR was done as in (A). The 2758C→G mutation abolishes a *Sau*96I restriction endonuclease site in patient 2290. Expected fragments, control allele: 159, 30 bp, mutant allele: 189 bp (the 30 bp fragment has been run off the gel).

Fig 6. shows an amino acid sequence comparison among methionine synthases in the Box 2 region. Identical residues are indicated by a star above the alignment. Dots show partially conserved residues, for which at least 6/7 identical or similar residues can be aligned (A,G,S,T; D,E,N,Q; V,L,I,M; K,R; and F,W,Y (Bordo,D. and Argos,P. (1991) *J. Molec. Biol.*, **217**, 721-729)). Mutations identified in this work are shown below the alignment. For

15 abbreviations, see Fig.1; Mm: *Mus musculus*. The seven amino acids conserved in cobalamin-binding proteins (Drennan, C.L., Huang, S., Drummond, J.T., Matthews, R.G., and Ludwig, M.L. (1994) *Science*, **266**, 1669-1674) are underlined.

Detailed Description

20 We used specific regions of homology within the methionine synthase sequences of several lower organisms to clone a human methionine synthase cDNA (SEQ ID NO:1) by a combination of RT-PCR and inverse PCR. The enzyme (SEQ ID NO:2) is 1265 amino acids in length and contains the seven residue structure-based sequence fingerprint identified for cobalamin-

25 containing enzymes. The gene was localized to chromosome 1q43 by the FISH technique. We have identified one missense mutation and a 3 base pair

-22-

deletion in patients of the *cblG* complementation group of inherited homocysteine/folate disorders by SSCP and sequence analysis, as well as an amino acid substitution present in high frequency in the general population.

We conclude that the cDNA that we have identified corresponds to human methionine synthase on the basis of homology to known methionine synthases and by the identification of mutations in patients with a deficiency of enzyme activity. The most striking sequence conservation was found in four boxes of 9-13 amino acids. Box 2 has been proposed to correspond to part of the cobalamin binding domain (Drennan, C.L., Huang, S., Drummond, J.T., Matthews, R.G., and Ludwig, M.L. (1994) *Science*, **266**, 1669-1674). It contains 13 consecutive residues that are identical in all known methionine synthases. Three amino acids within box 2 and four others C-terminal to it correspond to residues proposed by Drennan et al. (Drennan, C.L., Huang, S., Drummond, J.T., Matthews, R.G., and Ludwig, M.L. (1994) *Science*, **266**, 1669-1674) as a structure-based sequence fingerprint for cobalamin binding. The three amino acids appear to make direct contact with the lower face of the corrin ring and dimethylbenzimidazole tail of cobalamin, determined from the crystal structure of the *E. coli* enzyme at 3 Å resolution (Drennan, C.L., Huang, S., Drummond, J.T., Matthews, R.G., and Ludwig, M.L. (1994) *Science*, **266**, 1669-1674). All seven residues are identical in the human sequence (Fig. 6).

A survey of the NCBI databases for homology to the human methionine synthase using BLASTP yielded the various methionine synthases listed in Fig. 1, as well as the glutamate mutase (S41332, Q05488) and methylmalonyl-CoA mutase (P11653)(adenosyl-cobalamin dependent mutases) used to deduce the sequence fingerprint for cobalamin binding (Drennan, C.L., Huang, S., Drummond, J.T., Matthews, R.G., and Ludwig, M.L.

-23-

(1994) *Science*, **266**, 1669-1674). Homology was also found with the cobalamin binding region of the corrinoid: coenzyme M methyltransferase of *Methanosarcina barkeri* (U36337), the 5-methyltetrahydrofolate corrinoid/iron sulfur protein methyltransferase of *Clostridium thermoaceticum* (L34780) and
 5 the B12-dependent 2-methyleneglutarate mutase of *Clostridium barkeri* (S43552, S43237). Further, homology was found with the B12-binding site domain of the recently identified putative methionine synthase of *Agrobacterium tumefaciens* (U48718; partial N-terminal sequence is given, up to region of box 4). Significantly, homology with the B12-binding site domain
 10 was also found in the Hg resistance protein of *Myxococcus xanthus* (Z21955). This protein has not been described as having a cobalamin prosthetic group.

The two mutations we have identified as candidates for causing *cblG* disease are located in the vicinity of the cobalamin binding domain by comparison with *E. coli* methionine synthase (Fig. 6). Ile881 corresponds by
 15 sequence alignment to Val855 in the *E. coli* enzyme. Val855 is within a beta sheet strand that is part of an α/β domain that is a variant of the Rossmann nucleotide binding fold. The H920D substitution is found in a region which, in the *E. coli* enzyme, is in an α helix at the C-terminal end of the α/β domain. It is interesting that the polymorphism we have identified is at the adjacent
 20 residue (D919G). The functional role of the polymorphism and deleterious mutations will have to be examined in expression experiments to confirm their precise effect on the protein.

Through the cloning of a cDNA for human methionine synthase and mutations therein, we can now determine the properties of the human enzyme
 25 and complete the characterization of mutations in patients with severe synthase deficiency. This analysis has allowed us to tie mutations in the gene to disturbances in homocysteine metabolism which are known to result in

hyperhomocysteinemia is a risk factor for cardiovascular disease (Boushey, C.J., Beresford, S.A., Omenn, G.S., and Motulsky, A.G. (1995) *JAMA*, **274**, 1049-1057) and neural tube defects (Steegers- Theunissen, R.P., Boers, G.H., Trijbels, F.J., Finkelstein, J.D., Blom, H.J., Thomas, C.M., Borm, G.F., Wouters, M.G., and Eskes, T.K. (1994) *Metab. Clin. Exp.*, **43**, 1475-1480; and Mills, J.L., McPartlin, J.M., Kirke, P.N., Lee, Y.J., Conley, M.R., Weir, D.G. and Scott, J.M. (1995) *Lancet*, **345**, 149-151).

Our observations indicate the importance of methionine synthase as one of several genes involved in homocysteine metabolism. Results with other pathway genes underscores the significance of our findings. For example, a recently-identified mutation in methylenetetrahydrofolate reductase, the enzyme that synthesizes the 5-methyltetrahydrofolate substrate for the methionine synthase reaction, results in mild hyperhomocysteinemia (Frosst, P., Blom, H.J., Milos, R., Goyette, P., Sheppard, C.A., Matthews, R.G., Boers, G.J., den Heijer, M., Kluijtmans, L.A., van den Heuvel, L.P., et al. (1995) *Nat. Genet.*, **10**, 111-113). Evidence is accumulating that this mutation, present in 35-40% of alleles, is a risk factor in both cardiovascular disease and neural tube defects (Rozen, R. (1996) *Clin. Invest. Med.*, **19**, 171-178). We believe that genetic variants of methionine synthase similarly lead to mild hyperhomocysteinemia with consequent impact on these two multifactorial disorders.

We used specific regions of homology within the methionine synthase sequences, including a portion of the cobalamin binding site determined from the *E. coli* enzyme, to design degenerate oligonucleotides for RT-PCR-dependent cloning of human methionine synthase. We confirmed the identification of the cDNA sequences for human methionine synthase by the high degree of homology to the enzymes in other species and the identification

-25-

of mutations in patients from the *cblG* complementation group. We also mapped the gene to human chromosome 1.

The assays described herein can be used to test for compounds that modulate methionine synthase activity and hence may have therapeutic value in the prevention of neural tube defects, prevention and/or treatment of colon cancer, cardiovascular disease, hyperhomocysteinemia, and megaloblastic anemia without methylmalonic aciduria.

Screens for compounds that modulate methionine synthase enzymatic activity

Screens for potentially useful therapeutic compounds that modulate methionine synthase activity may be readily performed, for example, by assays that measure the incorporation of [14C]5-methyltetrahydrofolate into methionine or protein, or assays that measure the conversion of [14C]-homocysteine into methionine or protein. Examples of such assays, which employ whole cells or cell lysates, are well-known to those skilled in the art (see, e.g., Schuh, S., et al., *N. Engl. J. Med.* 1984, 310:686-69; Rosenblatt, D. S., et al., *J. Clin. Invest.* 1984, 74:2149-2156; Watkins, D., and Rosenblatt, D. S., *J. Clin. Invest.* 1988, 81:1690-1694; and Watkins, D., and Rosenblatt, D. S., *Am. J. Med. Genet.* 1989, 34:427-434), and may be readily adapted for high throughput screening.

ELISA for the detection of compounds that modulate methionine synthase expression

Enzyme-linked immunosorbant assays (ELISAs) are easily incorporated into high-throughput screens designed to test large numbers of compounds for their ability to modulate levels of a given protein. When used in the methods of the invention, changes in a given protein level of a sample,

-26-

relative to a control, reflect changes in the methionine synthase expression status of the cells within the sample. Protocols for ELISA may be found, for example, in Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, NY, 1997. Lysates from cells treated with potential

5 modulators of methionine synthase expression are prepared (see, for example, Ausubel et al., *supra*), and are loaded onto the wells of microtiter plates coated with “capture” antibodies specific for methionine synthase. Unbound antigen is washed out, and a methionine synthase-specific antibody, coupled to an agent to allow for detection, is added. Agents allowing detection include

10 alkaline phosphatase (which can be detected following addition of colorimetric substrates such as *p*-nitrophenolphosphate), horseradish peroxidase (which can be detected by chemiluminescent substrates such as ECL, commercially available from Amersham) or fluorescent compounds, such as FITC (which can be detected by fluorescence polarization or time-resolved fluorescence). The

15 amount of antibody binding, and hence the level of a methionine synthase polypeptide within a lysate sample, is easily quantitated on a microtiter plate reader.

As a baseline control for methionine synthase expression, a sample that is not exposed to test compound is included. Housekeeping proteins are

20 used as internal standards for absolute protein levels. A positive assay result, for example, identification of a compound that decreases methionine synthase expression, is indicated by a decrease in methionine synthase polypeptide within a sample, relative to the methionine synthase level observed in cells which are not treated with a test compound.

-27-

Reporter gene assays for compounds that modulate methionine synthase expression

Assays employing the detection of reporter gene products are extremely sensitive and readily amenable to automation, hence making them ideal for the design of high-throughput screens. Assays for reporter genes may employ, for example, colorimetric, chemiluminescent, or fluorometric detection of reporter gene products. Many varieties of plasmid and viral vectors containing reporter gene cassettes are easily obtained. Such vectors contain cassettes encoding reporter genes such as lacZ/ β -galactosidase, green fluorescent protein, and luciferase, among others. Cloned DNA fragments encoding transcriptional control regions of interest (e.g. that of the mammalian methionine synthase gene) are easily inserted, by DNA subcloning, into such reporter vectors, thereby placing a vector-encoded reporter gene under the transcriptional control of any gene promoter of interest. The transcriptional activity of a promoter operatively linked to a reporter gene can then be directly observed and quantitated as a function of reporter gene activity in a reporter gene assay.

Cells are transiently- or stably-transfected with methionine synthase control region/reporter gene constructs by methods that are well known to those skilled in the art. Transgenic mice containing methionine synthase control region/reporter gene constructs are used for late-stage screens *in vivo*. Cells containing methionine synthase/reporter gene constructs are exposed to compounds to be tested for their potential ability to modulate methionine synthase expression. At appropriate timepoints, cells are lysed and subjected to the appropriate reporter assays, for example, a colorimetric or chemiluminescent enzymatic assay for lacZ/ β -galactosidase activity, or fluorescent detection of GFP. Changes in reporter gene activity of samples

-28-

treated with test compounds, relative to reporter gene activity of appropriate control samples, indicate the presence of a compound that modulates methionine synthase expression.

Quantitative PCR of methionine synthase mRNA as an assay for compounds that modulate methionine synthase expression

The polymerase chain reaction (PCR), when coupled to a preceding reverse transcription step (rtPCR), is a commonly used method for detecting vanishingly small quantities of a target mRNA. When performed within the linear range, with an appropriate internal control target (employing, for example, a housekeeping gene such as actin), such quantitative PCR provides an extremely precise and sensitive means of detecting slight modulations in mRNA levels. Moreover, this assay is easily performed in a 96-well format, and hence is easily incorporated into a high-throughput screening assay. Cells are treated with test compounds for the appropriate time course, lysed, the mRNA is reverse-transcribed, and the PCR is performed according to commonly used methods, (such as those described in Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, NY, 1997), using oligonucleotide primers that specifically hybridize with methionine synthase nucleic acid. Changes in product levels of samples exposed to test compounds, relative to control samples, indicate test compounds that modulate methionine synthase expression.

Secondary screens of test compounds that appear to modulate methionine synthase activity

After test compounds that appear to have methionine synthase-modulating activity are identified, it may be necessary or desirable to subject

-29-

these compounds to further testing. At late stages testing will be performed *in vivo* to confirm that the compounds initially identified to affect methionine synthase activity will have the predicted effect *in vivo*.

Test Compounds

5 In general, novel drugs for prevention of neural tube defects, or prevention and/or treatment of colon cancer or cardiovascular disease are identified from large libraries of both natural product or synthetic (or semi-synthetic) extracts or chemical libraries according to methods known in the art. Those skilled in the field of drug discovery and development will understand
10 that the precise source of test extracts or compounds is not critical to the screening procedure(s) of the invention. Accordingly, virtually any number of chemical extracts or compounds can be screened using the exemplary methods described herein. Examples of such extracts or compounds include, but are not limited to, plant-, fungal-, prokaryotic- or animal-based extracts, fermentation
15 broths, and synthetic compounds, as well as modification of existing compounds. Numerous methods are also available for generating random or directed synthesis (e.g., semi-synthesis or total synthesis) of any number of chemical compounds, including, but not limited to, saccharide-, lipid-, peptide-, and nucleic acid-based compounds. Synthetic compound libraries are
20 commercially available from Brandon Associates (Merrimack, NH) and Aldrich Chemical (Milwaukee, WI). Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant, and animal extracts are commercially available from a number of sources, including Biotics (Sussex, UK), Xenova (Slough, UK), Harbor Branch Oceanographics Institute (Ft. Pierce,
25 FL), and PharmaMar, U.S.A. (Cambridge, MA). In addition, natural and synthetically produced libraries are produced, if desired, according to methods

-30-

known in the art, e.g., by standard extraction and fractionation methods.

Furthermore, if desired, any library or compound is readily modified using standard chemical, physical, or biochemical methods.

In addition, those skilled in the art of drug discovery and
5 development readily understand that methods for dereplication (e.g., taxonomic
dereplication, biological dereplication, and chemical dereplication, or any
combination thereof) or the elimination of replicates or repeats of materials
already known for their therapeutic activities for homocysteinemia,
megaloblastic anemia without methylmalonic aciduria, cardiovascular disease,
10 colon cancer, and neural tube defects should be employed whenever possible.

When a crude extract is found to modulate methionine synthase
biological activity, further fractionation of the positive lead extract is necessary
to isolate chemical constituents responsible for the observed effect. Thus, the
goal of the extraction, fractionation, and purification process is the careful
15 characterization and identification of a chemical entity within the crude extract
that modulates methionine synthase biological activity. The same assays
described herein for the detection of activities in mixtures of compounds can be
used to purify the active component and to test derivatives thereof. Methods of
fractionation and purification of such heterogenous extracts are known in the
20 art. If desired, compounds shown to be useful agents for treatment are
chemically modified according to methods known in the art. Compounds
identified as being of therapeutic value may be subsequently analyzed using
mammalian models of homocysteinemia, megaloblastic anemia without
methylmalonic aciduria, cardiovascular disease, colon cancer, and neural tube
25 defects.

Therapy

Compounds identified using any of the methods disclosed herein, may be administered to patients or experimental animals with a pharmaceutically-acceptable diluent, carrier, or excipient, in unit dosage form.

- 5 Conventional pharmaceutical practice may be employed to provide suitable formulations or compositions to administer such compositions to patients or experimental animals. Although intravenous administration is preferred, any appropriate route of administration may be employed, for example, parenteral, subcutaneous, intramuscular, intracranial, intraorbital, ophthalmic, intraventricular, intracapsular, intraspinal, intracisternal, intraperitoneal, intranasal, aerosol, or oral administration. Therapeutic formulations may be in the form of liquid solutions or suspensions; for oral administration, formulations may be in the form of tablets or capsules; and for intranasal formulations, in the form of powders, nasal drops, or aerosols.

- 15 Methods well known in the art for making formulations are found in, for example, "Remington's Pharmaceutical Sciences." Formulations for parenteral administration may, for example, contain excipients, sterile water, or saline, polyalkylene glycols such as polyethylene glycol, oils of vegetable origin, or hydrogenated naphthalenes. Biocompatible, biodegradable lactide polymer, lactide/glycolide copolymer, or polyoxyethylene-polyoxypropylene copolymers may be used to control the release of the compounds. Other potentially useful parenteral delivery systems for antagonists or agonists of the invention include ethylene-vinyl acetate copolymer particles, osmotic pumps, implantable infusion systems, and liposomes. Formulations for inhalation may contain excipients, for example, lactose, or may be aqueous solutions containing, for example, polyoxyethylene-9-lauryl ether, glycocholate and
- 25

-32-

deoxycholate, or may be oily solutions for administration in the form of nasal drops, or as a gel.

Examples

The following examples are to illustrate, not limit the invention.

5 **Example 1: Cloning human methionine synthase cDNA.**

 An initial survey of the NCBI databases yielded several sequences corresponding to methionine synthase from different organisms. Comparison of these sequences generated four very conserved regions identified as Boxes 1-4 in Fig. 1 (SEQ ID Nos:3-25). Degenerate oligonucleotides (SEQ ID
10 Nos:26-66) were synthesized corresponding to these conserved sequences (Table 1). These were used as primers for RT-PCR with human and mouse mRNA. These experiments yielded PCR products which were subcloned, sequenced and aligned as shown in Fig. 2. In subsequent experiments, oligonucleotide primers were specified from the non-degenerate internal
15 sequences of the subclones and additional PCR products encompassing the conserved boxes were obtained. In later experiments, additional sequences were obtained by inverse PCR ("PCR", Fig. 2) to obtain upstream or downstream sequences from those already determined. At the 3' end, a mouse sequence was obtained from the dbEST database (Accession Number
20 W33307). This sequence was used as the source of primers for additional PCR experiments. Throughout the experiments, the sequences of the PCR products were considered provisionally authentic if they were homologous to the methionine synthase sequences obtained from the databases. The sequences were taken as error free by comparison of the sequences of at least two, and
25 usually three, independent PCR reactions. Sequences were linked into a common sequence if RT-PCRs bridging independently isolated sequences were

-33-

successful. Through this approach the complete coding sequence was determined through exclusive use of PCR reactions.

The coding sequence of human methionine synthase contains 3795 bp (SEQ ID NO:1) encoding a polypeptide of 1265 amino acids in length (SEQ ID NO:2) (Fig. 3), exceeding the length of published methionine synthases by 11-29 residues. The putative initiation codon is in a sequence of good context for the initiation of translation in eukaryotic cells (GACAACATGT, underlined nucleotides matching Kozak consensus (Kozak, M. (1991) *J. Biol. Chem.*, **266**, 19867-19870)). The predicted MW of methionine synthase is 141,000, comparing favorably with the published size of 151,000 based on SDS-polyacrylamide electrophoresis of the pig enzyme (Chen, Z., Crippen, K., Gulati, S., and Banerjee, R. (1994) *J. Biol. Chem.*, **269**, 27193-27197). It shares 58% identity with the *E. coli* and 65% identity with the *C. elegans* enzyme.

15 **Example 2: Chromosomal location.**

Using FISH, the gene encoding methionine synthase was mapped to chromosome band 1q43, close to the telomeric region of the long arm (Fig. 4). A total of 50 cells with at least one signal were observed. A signal was seen on 1 chromatid in 26 cells, on two chromatids in 14 cells, on 3 chromatids in 7 cells, and on 4 chromatids in 3 cells. These results confirm the previous assignment of the gene to chromosome 1 by Mellman et al. (Mellman, I.S., Lin, P.F., Ruddle, F.H., and Rosenberg, L.E. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 405-409), who used cobalamin binding as a marker for the enzyme in human-hamster hybrids.

Example 3: Mutations in the *cbiG* complementation group.

Patients with deficiency of methionine synthase activity have been grouped into the *cbiG* complementation group in cell fusion experiments (Watkins, D. and Rosenblatt, D.S. (1988) *J. Clin. Invest.*, **81**, 1690-1694).

- 5 Fibroblast cultures from patients assigned to *cbiG* were examined by RT-PCR based SSCP analysis. Three mutations were identified by sequencing PCR fragments showing band shifts by SSCP (Fig. 5). In each case, the change was confirmed by an independent diagnostic test on genomic DNA or a separate preparation of cDNA from patient fibroblasts. One of the mutations, 2756A→G
10 (D919G), was confirmed by a diagnostic test that monitored the presence of a HaeIII site created by the mutation (Fig. 5A). Using this test, it was identified as a polymorphism since it was seen in 8 of 52 control alleles (15%). In two other cases, candidate deleterious mutations were identified. One is a 3 bp deletion, bp 2640-2642, that results in the deletion of an isoleucine codon
15 (ΔIle881). It was confirmed by heteroduplex analysis of cDNA generated by RT-PCR (Fig. 5B). The second is a point mutation, 2758C→G. It results in the amino acid substitution H920D. It was confirmed in genomic DNA by the loss of a Sau96I site (Fig. 5C). The latter two mutations were heterozygous in the patient cell lines. Their second mutation has not been identified. The
20 candidate deleterious mutations were not seen in panels of 68 or 52 control alleles, respectively.

Example 4: Additional Roles for Methionine Synthase Polymorphism (Asp919Gly or D919G) in Disease.

- The following data suggest that the D919G polymorphism
25 contributes to altered metabolism of homocysteine, methionine, folates, Vit. B12, and S-adenosylmethionine.

-35-

First in a Montreal study (n=303), in which mother-child pairs (cases and controls) were examined, we observed that infants who were homozygous for the polymorphism (Gly/Gly; Table 2) were at decreased risk for NTD. Measurements of serum folate, RBC folate, plasma homocysteine and serum
5 cobalamin did not give any statistically significant differences, except the trend was toward low folate levels in Gly/Gly individuals (cases and controls).

A second study (n=255) in California also examined the methionine synthase polymorphism as a risk factor for neural tube defects (Table 3). This study shows a similar decreased risk of neural tube defects in children
10 homozygous for Gly/Gly. Since the study encompassed a mix of whites and Hispanics, the data were reexamined stratified according to ethnic origin. Both groups showed a protective effect of Gly/Gly.

In summary, two independent studies suggest a protective effect of Gly/Gly against the risk of neural tube defects. This is likely to be mediated by
15 a mild reduction in methionine synthase activity.

Next, in a study of colon cancer, (212 cases and 345 controls), we observed a decreased risk for colon cancer in the individuals who were homozygous for the polymorphism (relative risk = 0.62); see Table 4. In the same study, we observed significantly decreased levels of plasma folate in
20 individuals who were homozygous for the polymorphism; see Table 5.

The Boston study described in Tables 4 and 5 is presented again in Table 6 with the data stratified according to alcohol intake. As shown in the table, Gly/Gly individuals with a low to medium alcohol intake had a relative risk associated with colon cancer of 0.11. The combined data (low + high
25 alcohol) gave a risk level of 0.62 (Table 6).

In summary, drug therapy targeted to a reduction in methionine synthase activity may be protective in individuals at risk for colon cancer or at

-36-

risk for neural tube defects. Additional polymorphisms or mutations may also exert a protective effect against the risk of neural tube defects or colon cancer. Conversely, it is understood that some polymorphisms and/or mutations may enhance the risk of neural tube defects or colon cancer, for example, by
5 increasing methionine synthase activity.

Example 5: Role of Polymorphism on Homocysteine and Folate Levels.

Third, in a study of individuals participating in the U.S. NHLBI Family Heart Study, we observed both an increase in plasma homocysteine following a methionine load and a decrease in plasma folate in individuals who
10 were homozygous for this polymorphism; see Table 7.

Example 6: Methionine Synthase Assays for the Detection of Compounds that Modulate Methionine Synthase Activity and Expression

Potentially useful therapeutic compounds that modulate (e.g. increase or decrease) methionine synthase activity or expression may be
15 isolated by various screens that are well-known to those skilled in the art. Such compounds may modulate methionine synthase expression at the pre- or post-transcriptional level, or at the pre- or post-translational level.

Example 7: Materials and Methods

Cell lines. The skin fibroblast lines are from patients with
20 methionine synthase deficiency. They were assigned to the *cbIG* complementation group in cell fusion experiments assayed by ¹⁴C-methyltetrahydrofolate incorporation into cellular macromolecules (Watkins, D. and Rosenblatt, D.S. (1988) *J. Clin. Invest.*, **81**, 1690-1694). Control fibroblasts were from other laboratory stocks or the Montreal Children's

-37-

Hospital Cell Repository for Mutant Human Cell Strains. Of the patients for which non-polymorphic mutations were found, WG 1892, a Caucasian male, was diagnosed at the age of 4 years with developmental delay, tremors, gait instability, megaloblastic anemia and homocystinuria; and WG2290, also a
 5 Caucasian male, was diagnosed at age 3 months with failure to thrive, severe eczema, megaloblastic anemia and surprisingly both homocystinuria and methylmalonic aciduria.

Materials. The T/A cloning kit was from Invitrogen. The GeneClean III Kit was obtained from Bio 101 Inc. and the Wizard Mini-Preps were from
 10 Promega. The random-primed DNA labelling kit was from Boehringer-Mannheim. Taq polymerase, Superscript II reverse transcriptase, AMV reverse transcriptase, Trizol reagent, DNazol reagent, T4 DNA ligase, and restriction enzymes were purchased from Gibco BRL. The Sequenase kit for manual sequencing was from United States Biochemicals. The α -[³⁵S]dATP (12.5
 15 Ci/mole) was from Dupont or ICN. The oligonucleotide primers were synthesized by R. Clarizio of the Montreal Children's Hospital Research Institute Oligonucleotide Synthesis Facility or the Sheldon Biotechnology Centre, McGill University.

Homology matches. Comparisons were made between the
 20 published *E. coli* cobalamin-dependent methionine synthase sequence and sequences in the NCBI databases (dbEST and GenBank) using the BLAST programs.

PCR cloning and DNA sequencing. DNA was prepared from fibroblast pellets by the method of Hoar et al. (Hoar,D.I., Haslam,D.B., and
 25 Starozik,D.M. (1984) *Prenat. Diag.*, **4**, 241-247). Total cellular RNA was isolated by the method of Chirgwin et al. (Chirgwin,J.M., Przybyla,A.E., MacDonald,R.J., and Rutter,W.J. (1979) *Biochemistry*, **18**, 5294-5299) and

-38-

reverse-transcribed using oligo-dT₁₅ as primer. PCR was conducted using degenerate oligonucleotides as primers, paired so as to link the sequences of different homology boxes. The PCRs were conducted as described previously (Triggs-Raine,B.L., Akerman,B.R., Clarke,J.T., and Gravel,R.A. (1991) *Am. J. Hum. Genet.*, **49**, 1041-1054) except that the temperature of incubation was modified to accommodate the use of reduced temperatures in the annealing step or by step-down PCR (Hecker,K.H. and Roux,K.H. (1996) *Biotechniques*, **20**, 478-485.(Abstract)). In some experiments, inverse PCR was used to determine sequence upstream or downstream of known sequence (Ochman,H., Medhora,M.M., Garza,D., and Hartl,D.L. (1990) *PCR Protocols: A Guide to Methods and Applications*, Academic Press, San Diego, pp. 219-227). In these instances, genomic DNA or cDNA prepared by reverse transcription of RNA was digested with different four base restriction endonucleases, ligated with T4 DNA ligase, and amplified by PCR using adjacent oligonucleotides priming in opposite directions. Templates for inverse PCR at the cDNA level were generated with 12.5 µg RNA reversed transcribed using AMV-RT. Second strand synthesis was carried out using the random-primed DNA labelling kit adding 1 µl of each dNTP. Samples were incubated 30 min. at 37°C. Template was then treated as genomic DNA for digestion and ligation. Inverse PCR was used to obtain the 5' and 3' ends of the cDNA and to define an intron sequence adjacent to a splice junction for the design of a mutation diagnostic test. The PCR products were purified with GeneClean and were subcloned in the pCR2.1 vector and transformed into *E. coli* as per the supplier's protocol (TA Cloning Kit). The candidate clones were sequenced manually or by the DNA Core Facility of the Canadian Genetic Diseases Network or the McGill University Sheldon Biotechnology Centre.

-39-

Mutation analysis. Genomic DNA and RNA were isolated from control or patient fibroblast pellets using the DNazol or Trizol reagents, respectively, as per the manufacturer. The cDNA template for PCR was prepared by reverse transcription of 3-5 ug total RNA in reactions containing

5 400 U of Superscript II reverse transcriptase and 100ng random hexamers in a total reaction volume of 20 ul. SSCP analysis was performed as described previously (Triggs-Raine,B.L., Akerman,B.R., Clarke,J.T., and Gravel,R.A. (1991) *Am. J. Hum. Genet.*, **49**, 1041-1054) in reactions containing 4 µl of template, 1 µl of each dTTP, dCTP, dGTP (0.625 mM), 0.5 µl of dATP (0.625

10 mM), 1 µl α -[³⁵S]-dATP (12.5 Ci/mole). The radio labelled PCR products mixed with sequencing stop solution were heat denatured and quick chilled on ice prior to loading (Triggs-Raine,B.L., Akerman,B.R., Clarke,J.T., and Gravel,R.A. (1991) *Am. J. Hum. Genet.*, **49**, 1041-1054). As well, an aliquot of each sample was run without prior heating to identify the duplex product.

15 The fragments were subjected to electrophoresis in a 6% acrylamide/10% glycerol gel in 1X TBE for 18 hrs at 8 watts at room temperature. The gel was dried and exposed to Biomax film (Kodak). Fragments that displayed band shifts were sequenced directly.

Two mutations were confirmed directly in PCR amplification

20 products from genomic DNA and one mutation was confirmed in reversed transcribed mRNA. The PCR reactions for mutation confirmation were performed using 4 µl of cDNA template or 500 ng genomic DNA, 500 ng of specific primers, 2.5 U Taq polymerase and 1.5 mM MgCl₂ in a 50 µl volume. Heteroduplex analysis was accomplished by preheating PCR products to 95°C

25 for five minutes and subjecting the samples to electrophoresis in a 9% polyacrylamide gel (Triggs-Raine,B.L., Akerman,B.R., Clarke,J.T., and Gravel,R.A. (1991) *Am. J. Hum. Genet.*, **49**, 1041-1054). Other diagnostic

-40-

assays were accomplished by digesting a 15 µl sample of the PCR products with the indicated restriction endonuclease prior to electrophoresis.

Chromosomal localization. Human metaphase spreads were obtained from short-term cultures of phytohemagglutinin-stimulated peripheral blood lymphocytes. The cells were synchronized with thymidine and treated with BrdU during the late S-phase before harvesting for simultaneous observation of the hybridized sites and chromosome banding. The protocol for FISH was essentially as described previously (Lemieux, N., Malfoy, B., and Forrest, G.L. (1993) *Genomics*, **15**, 169-172; Zhang, X.X., Rozen, R., Hediger, M.A., Goodyer, P., and Eydoux, P. (1994) *Genomics*, **24**, 413-414). Briefly, a 5 kb DNA fragment of the methionine synthase genomic DNA (generated by PCR using oligonucleotides #1782 and #1780) was labelled by nick translation with biotin-16-dUTP (Boehringer-Mannheim), ethanol precipitated and dissolved in hybridization buffer at a final concentration of 8 ng/µl. The slides were denatured in 70% formamide, 2 x SSC at 70°C for 2 min. The biotinylated probe was denatured in the hybridization buffer at 95°C for 10 min, quickly cooled on ice, then applied on slides. Post-washing was done by rinsing in 50% formamide, 2 x SSC at 37°C. The slides were incubated with rabbit anti-biotin antibody (Enzo Biochemicals), biotinylated goat anti-rabbit antibodies (BRL) and streptavidin-FITC. They were stained with propidium iodide and mounted in p-phenylenediamine, pH 11. Cells were observed under the microscope (Zeiss), then captured through a CCD camera and processed using a FISH software (Applied Imaging).

Table 1. Oligonucleotides used for cDNA cloning, chromosome mapping and mutation detection.

5	Oligonucleotides ^a		Sequence	Location
			^b	
10	D1729	(SEQ ID NO:26)	5'-GAYGGNGCNATGGGNACNATGATHCA	100-125
	D1730	(SEQ ID NO:27)	5'-GCNACNGTNAARGGNGAYGTNCAYGAYAT	2332-
	D1731	(SEQ ID NO:28)	5'-RTTYTTNCCDATRTCTGNACRTCNCCTTT	2360
	D1733	(SEQ ID NO:29)	5'-RTGNAGRTAYTCNGCRAANGCYTCNGC	2341
15	D1754	(SEQ ID NO:30)	5'-ATRTGRTCTGNGGTNGTNCRCARCANCCNCC	3400
	D1755	(SEQ ID NO:31)	5'-GGNGGNTGYTGYGGNACNACNCCNGAYCAYAT	992-961
	M1806A	(SEQ ID NO:32)	5'-GTCTGTGTCATAGCCCAGAATGGG	961-992
	M1806B	(SEQ ID NO:33)	5'-TCAGTCTGTGTCATAGCCCAGAAT	3772
20	305A	(SEQ ID NO:34)	5'-GAACTAGAAGACAGAAATTCTCTA	3775
	407A	(SEQ ID NO:35)	5'-TTCCGAGGTCAGGAATTTAAAGATCA	(intronic)
	407B	(SEQ ID NO:36)	5'-GTGTTCTTCGTTTAGCTTCTCCCG	407A
	407D	(SEQ ID NO:37)	5'-CCCCAGCCAGCAAGTATTCCTTAT	407B
25	1107A	(SEQ ID NO:38)	5'-CTAGGTTGTATTTCTTGAGGATC	(SEQ ID NO:37)
	1406D	(SEQ ID NO:39)	5'-GGAGCTGGAAAAATGTTTCTACCTC	(SEQ ID NO:38)
	1406E	(SEQ ID NO:40)	5'-ACAGGAGGGAAGAAAGTCATTTCAG	3833
	1706A	(SEQ ID NO:41)	5'-CCTTCAATTATATTGAGAGGTCGGG	1406D
30	1707A	(SEQ ID NO:42)	5'-CAACCCGAAGGTCTGAAGAAAACC	2194
	1707B	(SEQ ID NO:43)	5'-CCCGCGCTCCAAGACCTGTCG	1406E
	1707C	(SEQ ID NO:44)	5'-CGACAGGTCTTGGAGCGCGGG	1986
	1758	(SEQ ID NO:45)	5'-GGAGTCATGACTCCTAAATCAATAACTC	1706A
35	1760	(SEQ ID NO:46)	5'-GACGACTACAGCAGCATCATGGT	2105
	1766	(SEQ ID NO:47)	5'-AAAAATCATTTTCATCCAGGGAA	1707A
	1772	(SEQ ID NO:48)	5'-ATAGGCAAGAACATAGTTGGAGTAGT	1707B
	2384	(SEQ ID NO:49)	5'-TTTCATCTAACAGCTGGGAACACAC	1707C
40	1774	(SEQ ID NO:50)	5'-TGCCTCTCAGACTTCATCGCTCCC	1758
	1780	(SEQ ID NO:51)	5'-TGCAGCCTGGGGCACAGCAGC	2405
	1782	(SEQ ID NO:52)	5'-ATGGATTGGCTGTCTGAACCTCAC	1760
	1796	(SEQ ID NO:53)	5'-CATGGAAGAATATGAAGATATTAGAC	3377
45	1803	(SEQ ID NO:54)	5'-ACCATCATCCTCATAGGCCTTGCT	1766
	1806C	(SEQ ID NO:55)	5'-CAGACCTGCGAAGGTTGCGGTAC	2505
	1806F	(SEQ ID NO:56)	5'-GAAGTGTTTGCTCCTCCAATCAAC	1772
	2568			2384
50	1803	(SEQ ID NO:54)	5'-ACCATCATCCTCATAGGCCTTGCT	1774
	1806C	(SEQ ID NO:55)	5'-CAGACCTGCGAAGGTTGCGGTAC	3264
	1806F	(SEQ ID NO:56)	5'-GAAGTGTTTGCTCCTCCAATCAAC	1780
	2568			3148
55	1803	(SEQ ID NO:54)	5'-ACCATCATCCTCATAGGCCTTGCT	1782
	1806C	(SEQ ID NO:55)	5'-CAGACCTGCGAAGGTTGCGGTAC	2847
	1806F	(SEQ ID NO:56)	5'-GAAGTGTTTGCTCCTCCAATCAAC	1796
	2568			2752
60	1803	(SEQ ID NO:54)	5'-ACCATCATCCTCATAGGCCTTGCT	1803
	1806C	(SEQ ID NO:55)	5'-CAGACCTGCGAAGGTTGCGGTAC	3331
	1806F	(SEQ ID NO:56)	5'-GAAGTGTTTGCTCCTCCAATCAAC	1806C
	2568			3504

-42-

	1808	(SEQ ID NO:57)	5'-GAGCAGCTTTCAGTATCTTATCACAT	2458-
	2433			
	1827	(SEQ ID NO:58)	5'-ACAAGTTGTGTTCCCTCCATTCCAGT	1657-
	1633			
5	1828	(SEQ ID NO:59)	5'-AGAGCGCTGTAATGTTGCAGGATCA	1125-
	1149			
	1907B	(SEQ ID NO:60)	5'-TGTTTTTCAATGCCCTTCACAAGGG	2057-
	2033			
	1907C	(SEQ ID NO:61)	5'-TAAAAAGTATGGAGCTGCTATGGTG	1464-
10	1488			
	2606A	(SEQ ID NO:62)	5'-GACCAGACAGTAACATATGTCCTTC	1078-
	1054			
	2606B	(SEQ ID NO:63)	5'-ACATTACAGCGCTCTCCAATGTTAAC	1139-
	1114			
15	2706A	(SEQ ID NO:64)	5'-TGAGGTTGAGAAATGGCTTGGACC	3750-
	3773			
	2706B	(SEQ ID NO:65)	5'-GCCACAGATATGTTCTTCCTCAATG	3749-
	3725			
	3107A	(SEQ ID NO:66)	5'-TGTGGAGAGCACGTCTTCTCTGCC	-55 - -32
20				

^a Numbers with the prefix "D" refer to oligonucleotides with degenerate bases shown as N (any base), H (A, C, or T), D (A, G, or T), Y (T or C), or R (A or G); those with the prefix "M" refer to mouse sequences (see Fig. 3).

^b From the first methionine codon, see Fig. 3.

-43-

Table 2. MS Polymorphism in Neural Tube Defects - Montreal Study

Genotype	Cases		Case mothers		Controls		Control mothers		Odds ratio*	95% C.I.
	N	%	N	%	N	%	N	%		
Asp/Asp	38	69	40	66	59	61	55	61	0.07	0.004-1.29
Asp/Gly	17	31	20	33	28	29	34	38		
Gly/Gly	0	0.9	1	2	10	10	1	1		
N	55		61		97		90			

* Odds ratio calculated for genotypes Asp/Asp vs Gly/Gly
(to permit the calculation, the 0 cell was increased to 0.5)

Table 3. MS Polymorphisms in Neural Tube Defects - California Study

Ethnic Group	Genotype 2756A-G	Cases		Controls		Odds ratio*	95% C.I.
		N	%	N	%		
Overall	Asp/Asp	64	67	104	64	1.0	0.56 - 1.72 0.05 - 1.92
	Asp/Gly	30	32	49	30	0.99	
	Gly/Gly	1	1	7	4	0.23	
White only	Asp/Asp	21	66	38	66	2.0	0.44 - 2.9 0.11 - 5.6
	Asp/Gly	10	31	16	28	1.1	
	Gly/Gly	1	3	3	5	0.60	
Hispanic only	Asp/Asp	43	68	66	63	1.0	0.45 - 1.8
	Asp/Gly	20	32	33	31	0.9	
	Gly/Gly	0	0	4	4	0	

* Odds ration calculated for genotypes Asp/Asp vs Gly/Gly

-44-

Table 4. Frequency of MS genotype and relative risk (RR) of colorectal cancer by MS genotype

MS Genotype	Cases		Controls		RR	95% CI
	n	%	n	%		
Asp/Asp	145	(68)	234	(68)	1.0	
Asp/Gly	61	(29)	95	(28)	1.02	0.69 - 1.50
Gly/Gly	6	(3)	16	(5)	0.62	0.24 - 1.64
Total	212		345			

Table 5. Mean of homocysteine and folate (geometric) by case control status and MS Genotype in a colon cancer study

MS genotype	Cases		Controls		Cases & Controls	
	n	mean	n	mean	n	mean
<u>Folate (Bio-Kit) ng/ml</u>						
Asp/Asp	115	3.8	201	3.9 *	316	3.9 **
Asp/Gly	49	4.1 *	80	3.8 *	129	3.9 **
Gly/Gly	6	2.1	12	2.3	18	2.2
<u>Homocysteine (μM)</u>						
Asp/Asp	66	12.5	160	12.1	226	12.3
Asp/Gly	30	10.8	50	11.6	80	11.2
Gly/Gly	4	13.4	9	11.7	13	12.5

* = $p < 0.05$ ** = $p < 0.01$

Table 6. Age Adjusted Relative Risk of Colon Cancer According to MS Polymorphism and Alcohol Intake Status Among US Physicians

-45-

		Genotype			
		2756A->G	Case	Control	
Alcohol intake	Asp919Gly	s	s	Odds	95% C.I.
		N	N	ratio	
Low - Medium 0 - 0.8 drinks/day	Asp/Asp	1013	2e+09	1.0	
	Asp/Gly	7113		0.87	0.54 - 1.4
	Gly/Gly	9		0.11	0.01 - 0.82
	N				
5 High 1 - 2+ drinks/day	Asp/Asp	3721	7e+06	0.74	0.46 - 1.19
	Asp/Gly	563		1.15	0.60 - 2.18
	Gly/Gly			3.83	0.72 -20.47
	N				

-46-

Table 7. Mean Homocysteine and Folate Status by MS Genotype

(Date of Analysis: February 18, 1997)

		Methionine Synthase Genotype			
		Asp/As p	Asp/Gly		Gly/Gly
		Result	Result	P	Result P
5	N	252	111		17
	Fasting Hcy (μ M)	8.5	8.4	0.76	8.7
	post-methionine load Hcy	17.9	19.6		0.74
	(μ M)	7.4	0.05* 6.9		22.3
	Folate (microb test)		0.37		0.03*
					6.3
					0.37

The embodiments of the invention in which an exclusive property or privilege is claimed are defined as follows:

1. Substantially pure nucleic acid encoding a mammalian methionine synthase polypeptide.

2. The nucleic acid of claim 1, wherein said nucleic acid encodes a human polypeptide.

3. The nucleic acid of claim 1, wherein said nucleic acid is DNA.

4. The DNA of claim 3, wherein said DNA is genomic DNA.

5. The DNA of claim 3, wherein said DNA is cDNA.

6. The nucleic acid of claim 1, wherein said nucleic acid has the sequence of SEQ ID NO: 1, or degenerate variants thereof, and wherein said nucleic acid encodes the amino acid sequence of SEQ ID NO: 2.

7. A substantially pure nucleic acid sequence that hybridizes at high stringency to sequences found within the nucleic acid of SEQ ID NO: 1.

8. The sequence of claim 7, wherein said sequence has a sequence complementary to at least 50% of at least 60 nucleotides of the nucleic acid encoding the methionine synthase polypeptide, said sequence sufficient to allow nucleic acid hybridization under high stringency conditions.

9. The sequence of claim 8, wherein said sequence is complementary to at least 90% of at least 18 nucleotides of the nucleic acid encoding the methionine synthase polypeptide.

10. The nucleic acid of claim 7, wherein said nucleic acid contains a mutation or a polymorphism, and wherein said nucleic acid encodes a mutant or polymorphic polypeptide.

11. The mutation of claim 10, wherein said mutation or polymorphism is a point mutation or a deletion mutation.

12. The mutation of claim 10, wherein said mutation is a deletion of base pairs 2640-2642.

13. The mutation of claim 10, wherein said mutation is a 2758 C->G mutation.

14. The polymorphism of claim 10, wherein said polymorphism is a 2756 A-> G mutation.

15. The nucleic acid of claim 3, wherein said nucleic acid is operably linked to regulatory sequences for expression of said polypeptide, and wherein said regulatory sequences comprise a promoter.

16. A cell containing the nucleic acid of claim 15.

17. The cell of claim 16, wherein said cell is a prokaryotic cell.

18. The cell of claim 16, wherein said cell is a eukaryotic cell.

19. The cell of claim 18, wherein said cell is a yeast cell.

20. The cell of claim 16, wherein said cell is a mammalian cell.

21. The promoter of claim 15, wherein said promoter is inducible.

22. A non-human transgenic mammal containing the nucleic acid of claim 15.

23. The mammal of claim 22, wherein said nucleic acid comprises a mutation associated with hyperhomocysteinemia.

24. A non-human mammal wherein one or both genetic alleles encoding a methionine synthase polypeptide are mutated.

25. The mammal of claim 24, wherein one or both genetic alleles encoding a methionine synthase polypeptide are disrupted, deleted, or otherwise rendered nonfunctional.
26. Cells from the mammal of claim 22 or claim 24.
27. The mammal of claim 22 or claim 24, wherein said mammal is a rodent.
28. A substantially pure mammalian methionine synthase polypeptide.
29. The polypeptide of claim 28, wherein said polypeptide is recombinant.
30. The polypeptide of claim 28, wherein said polypeptide is a human polypeptide.
31. The polypeptide of claim 28, wherein said polypeptide is the polypeptide set forth in SEQ ID NO: 2.
32. A polypeptide having conservative amino acid substitutions, wherein said polypeptide has methionine synthase biological activity.
33. A mutant or polymorphic polypeptide, wherein said polypeptide has less methionine synthase biological activity than the polypeptide of SEQ ID NO: 2.
34. The polypeptide of claim 33, wherein said polypeptide has a H920D amino acid substitution.
35. The polypeptide of claim 33, wherein said polypeptide has a D1le881 deletion

36. The polypeptide of claim 33, wherein said polypeptide has a D919G substitution.

37. A mutant or polymorphic polypeptide, wherein said polypeptide has higher methionine synthase biological activity than the polypeptide set forth in SEQ ID NO: 2.

38. An antibody that specifically binds a mammalian methionine synthase polypeptide.

39. The polypeptide of claim 38, wherein said polypeptide is a mutant or polymorphic polypeptide.

40. A method of generating an antibody that specifically binds a methionine synthase polypeptide, said method comprising administering a methionine synthase polypeptide, or fragment thereof, said administering to an animal capable of generating an immune response, and isolating said antibody from said animal.

41. A method of detecting the presence of a methionine synthase polypeptide, said method comprising contacting a sample with the antibody that specifically binds a methionine synthase polypeptide and assaying for binding of said antibody to said polypeptide.

42. A method for detecting sequence variants for methionine synthase in a mammal, said method comprising analyzing the nucleic acid of a test subject to determine whether said test subject contains a mutation in a methionine synthase gene, wherein the presence of said mutation is an indication that said animal has an increased or decreased likelihood of developing hyperhomocysteinemia, cardiovascular disease, neural tube defects, or cancer.

43. The method of claim 42, wherein said method detects an increased risk for hyperhomocysteinemia.

44. The method of claim 42, wherein said likelihood is increased.

45. The method of claim 42, wherein said likelihood is decreased.

46. The method of claim 42, wherein said animal has an increased or decreased likelihood of developing cardiovascular disease.

47. The method of claim 42, wherein said animal has an increased or decreased likelihood of developing neural tube defects.

48. The method of claim 42, wherein said animal has an increased or decreased likelihood of developing cancer.

49. The method of claim 42, wherein primers used for detecting said mutation are selected from:

GAYGGNGCNATGGGNACNATGATHCA (SEQ ID NO:26);

GCNACNGTNAARGGNGAYGTNCAYGAYAT (SEQ ID NO:27);

RTTYTTNCCDATRTCRTGNACRTCNCCTTT (SEQ ID NO:28);

RTGNAGRATAYTCNGCRAANGCYTCNGC (SEQ ID NO:29);

ATRTGRTCNCGGNGTNGTNCRCARCANCCNCC (SEQ ID NO:30);

GGNGGNTGYTGYGGNACNACNCCNGAYCAYAT (SEQ ID NO:31);

GTCTGTGTCATAGCCCAGAATGGG (SEQ ID NO:32);

TCAGTCTGTGTCATAGCCCAGAAT (SEQ ID NO:33);

GAAGTAGAAGACAGAAATTCTCTA (SEQ ID NO:34);

TTCCGAGGTCAGGAATTTAAAGATCA (SEQ ID NO:35);

GTGTTCTTCGTTTAGCTTCTCCCG (SEQ ID NO:36);

CCCCAGCCAGCAAGTATTCCTTAT (SEQ ID NO:37);

CTAGGTTGTATTCCTTGAGGATC (SEQ ID NO:38);

GGAGCTGGAAAAATGTTTCTACCTC (SEQ ID NO:39);

ACAGGAGGGAAGAAAGTCATTGAG (SEQ ID NO:40);

CCTTCAATTATATTGAGAGGTCGGG (SEQ ID NO:41);

CAACCCGAAGGTCTGAAGAAAACC (SEQ ID NO:42);

CCCGCGCTCCAAGACCTGTGCG (SEQ ID NO:43);

CGACAGGTCTTGAGCGCGGG (SEQ ID NO:44);

GGAGTCATGACTCCTAAATCAATAACTC (SEQ ID NO:45);
GACGACTACAGCAGCATCATGGT (SEQ ID NO:46);
AAAAATCATTTCATCCAGGGAA (SEQ ID NO:47);
ATAGGCAAGAACATAGTTGGAGTAGT (SEQ ID NO:48);
TTTCATCTAACAGCTGGGAACACAC (SEQ ID NO:49);
TGCCTCTCAGACTTCATCGCTCCC (SEQ ID NO:50);
TGCAGCCTGGGGCACAGCAGC (SEQ ID NO:51);
ATGGATTGGCTGTCTGAACCTCAC (SEQ ID NO:52);
CATGGAAGAATATGAAGATATTAGAC (SEQ ID NO:53);
ACCATCATCCTCATAGGCCTTGCT (SEQ ID NO:54);
CAGACCTGCGAAGGTTGCGGTAC (SEQ ID NO:55);
GAAGTGGTTGCTCCTCCAATCAAC (SEQ ID NO:56);
GAGCAGCTTTCAGTATCTTATCACAT (SEQ ID NO:57);
ACAAGTTGTGTTCCCTCCATTCCAGT (SEQ ID NO:58);
AGAGCGCTGTAATGTTGCAGGATCA (SEQ ID NO:59);
TGTTTTTCAATGCCCTTCACAAGGG (SEQ ID NO:60);
TAAAAAGTATGGAGCTGCTATGGTG (SEQ ID NO:61);
GACCAGACAGTAACATATGTCCTTC (SEQ ID NO:62);
ACATTACAGCGCTCTCCAATGTTAAC (SEQ ID NO:63);
TGAGGTTGAGAAATGGCTTGGACC (SEQ ID NO:64);
GCCACAGATATGTTCTTCCTCAATG (SEQ ID NO:65); or
TGTGGAGAGCACGTCTTCTCTGCC (SEQ ID NO:66).

50. The method of claim 42, wherein said step further comprises the step of sequencing nucleic acid encoding methionine synthase from said test subject.

51. The method of claim 42, wherein said method further comprises the step of using nucleic acid primers specific for the methionine synthase gene and wherein said primers are used for DNA amplification by the polymerase chain reaction.

52. The method of claim 42, wherein said analyzing includes single strand conformational polymorphism (SSCP) analysis.

53. The method of claim 42, wherein said method is carried out by restriction fragment length (RFLP) polymorphism analysis.

54. The method of claim 42, wherein said method is for the diagnosis of an altered risk for cardiovascular disease.

55. The method of claim 42, wherein said method is for the diagnosis of an altered risk for neural tube defects.

56. The method of claim 42, wherein said method is for the diagnosis of an altered risk for colon cancer.

57. A kit for the analysis of mammalian methionine synthase nucleic acid, said kit comprising nucleic acid probes for analyzing the nucleic acid of a mammal, said analyzing sufficient to determine whether said mammal contains a mutation in said methionine synthase nucleic acid.

58. A kit for the analysis of mammalian methionine synthase nucleic acid, said kit comprising antibodies for analyzing the methionine synthase polypeptide of a mammal, said analyzing sufficient to determine whether said mammal contains a mutation in said methionine synthase nucleic acid.

59. A method of treating or preventing cancer, cardiovascular disease, or neural tube defects, said method comprising inhibiting methionine synthase biological activity.

60. The method of claim 59, wherein said method is for the prevention of neural tube defects.

61. The method of claim 59, wherein said method is for the prevention of cardiovascular disease.

62. The method of claim 59, wherein said method is for the prevention of cancer.

63. The method of claim 59, wherein said method comprises administering a therapeutically effective dose of a methionine synthase inhibitor to a mammal, said inhibitor being a methionine synthase anti-sense nucleic acid.

64. The method of claim 59, wherein said method comprises administering a therapeutically effective dose of a methionine synthase inhibitor to a mammal, said inhibitor being a peptide comprising a portion of a mammalian methionine synthase polypeptide.

65. The method of claim 59, wherein said method comprises administering a therapeutically effective dose of a methionine synthase inhibitor to a mammal, said inhibitor being a small molecule.

66. A method of treating or preventing cardiovascular disease, said method comprising administering to the subject to a therapeutically effective dose of a metabolite or cofactor selected from the group: folate, cobalamin, S-adenosyl methionine, betaine, or methionine.

67. The method of claim 63, claim 65, or claim 66, wherein said mammal is pregnant.

68. A method of preventing neural tube defects, cancer, and cardiovascular disease, said method comprising:

a) detecting an increased risk of neural tube defects, cancer, and cardiovascular disease, wherein said detecting is by analyzing methionine synthase nucleic acid from one or more test subjects selected from: a mammal; a potential parent, either male or female; a pregnant mammal; or a developing embryo or fetus, wherein said analyzing is done by the method of claim D; and

b) exposing said mammal, potential parent, said pregnant mammal, and/or said developing embryo or fetus to a therapeutically effective dose of a metabolite or cofactor selected from the group: cobalamin; S-adenosyl methionine; betaine; or methionine, wherein said exposing is via the administration of said dose to said mammal, potential parent, said pregnant mammal, and/or said developing embryo or fetus.

69. The method of claim 68, wherein said method is for preventing neural tube defects, and said risk is a risk of neural tube defects.

70. The method of claim 68, wherein said method is for preventing cancer, and said risk is a risk of cancer.

71. The method of claim 68, wherein said method is for preventing cardiovascular disease, and said risk is a risk of cardiovascular disease.

72. A method for screening for a compound that modulates methionine synthase biological activity, said method comprising the steps of:

- a) contacting a sample containing wild-type, mutated, or polymorphic methionine synthase with said compound, and
- b) assaying for methionine synthase enzymatic activity, wherein increased enzymatic activity indicates an inducer of methionine synthase biological activity, and decreased enzymatic activity indicates an inhibitor of methionine synthase biological activity.

73. The method of claim 72, wherein said sample is selected from: purified or partially purified methionine synthase, a cell lysate, a cell, or a mammal.

74. A method for screening for a compound that modulates methionine synthase biological activity, said method comprising the steps of:

- a) contacting a sample with said compound, and
- b) assaying for methionine synthase expression, wherein increased expression indicates an inducer of methionine synthase biological activity, and decreased expression indicates an inhibitor of methionine synthase biological activity.

FIG. 1

BOX 1:

** *** *

Ec	(20)	DGGMGTMIQ	(SEQ ID NO: 3)
Ss	(20)	DGAMGTNLQ	(SEQ ID NO: 4)
Ml2	(5)	DGAMGTQLQ	(SEQ ID NO: 5)
Hi	(20)	DGAMGTMIQ	(SEQ ID NO: 6)
Ce	(22)	DGAMGTMIQ	(SEQ ID NO: 7)
Hs	(34)	DGGMGTMIQ	(SEQ ID NO: 8)

BOX 2:

Ec	(752)	ATVKGDVHDIGKN	(SEQ ID NO: 9)
Ss	(729)	ATVKGDVHDIGKN	(SEQ ID NO: 10)
Ml2	(726)	ATVKGDVHDIGKN	(SEQ ID NO: 11)
Hi	(142)	ATVKGDVHDIGKN	(SEQ ID NO: 12)
Ce	(766)	ATVKGDVHDIGKN	(SEQ ID NO: 13)
Hs	(778)	ATVKGDVHDIGKN	(SEQ ID NO: 14)

BOX 3:

** ** *

Ec	(1095)	LAEAFAEYLH	(SEQ ID NO: 15)
Ss	(1085)	MAEALAEWTH	(SEQ ID NO: 16)
Ml1	(56)	LTEALAEYWH	(SEQ ID NO: 17)
Hi	(490)	LAEAMAEYLH	(SEQ ID NO: 18)
Ce	(1084)	LAEAYAAYLH	(SEQ ID NO: 19)
Hs	(1133)	LAEAFAEELH	(SEQ ID NO: 20)

BOX 4:

Ec	(262)	GGCCGTPQHI	(SEQ ID NO: 21)
Ss	(243)	GGCCGTRPDHI	(SEQ ID NO: 22)
Ml2	(226)	GGCCGTPPDHI	(SEQ ID NO: 23)
Ce	(264)	GGCCGTPPDHI	(SEQ ID NO: 24)
Hs	(321)	GGCCGSTPDHI	(SEQ ID NO: 25)

5' box 1 box 4 box 2 box 3 3

01729 01754 01730 01733

01755 01731 1774 M1808A

01729 2506A 1774 M1808B

01729 2506B 1808C M1808B

01729 1827 1808D 2708A

01729 1708A 1808E 2708B

01729 1758 1774 2708B

01729 1766 1774 1107A

01755 1758

11758 1758

1827 1827 1907C 1907B

1407B 1407A 1406E 1806F

1407B 1407B 1406E 1808

1707B 1407D 1406D 1806F

1707C 1707A

3107A 1407D

Fig. 3 page 1 of 2

FIG. 3: Methionine Synthase cDNA (SEQ ID NO: 1) and amino acid (SEQ ID NO: 2) sequences.

317

GETCACCTGTGGAGACGCTCTTCTCTCCCGCCCTCTGCGCAAGGAGAGACTCGACAAC

1 ATGTCAACCGCGCTCCAAGACCTTCTGCAACCCCAAGGCTCTGAAGAAAACCTCTGCGGATGAGATCAATGCCATTCTGCAGAGAGGATTATGGTCTGGATGGAGGATGGGACCATG 40
N S P A L Q D L S Q P E G L K K T L R D E I M A I L Q K R I M V L D G G H G T M

121 ATCCAGCGGGAGAAGCTAAACGAACAACACTTCCAGGCTCAGGAATTTAAGATCATGCCAGGCGCTGAAAGGCAACAATGACATTTTAAGTATAACTCAGGCTGATGTCAATTACCAA 80
I Q R E K L N E E H F R G Q E F K D H A R P L K G N N D I L S I T Q P D V I Y Q

241 ATCCATAAGCAATACTTCTGGCTGGGCGAGATATCATTTGAACAAATACTTTTACGACTAGTATGCCCAAGCTGACTATGGCTTGAACACTTGGCTTACCGGATGAACATGTGC 120
I H K E Y L L A G A D I I E T N T F S S T S I A G A D Y G L E H L A Y R H M H C

361 TCTGCAGGAGTGGCCAGAAAAGCTGCCGAGGAGTAACTCTCCAGACAGGAATTAAGAGTTTGTGGCAGGCGCTCTGGTCCGACTAATAAGACACTCTCTGTGTCCCATCTGTGGAA 160
S A G V A R K A A E E V T L Q T G I K R F V A G A L G P T N K T L S V S P S V E

481 AGGCGGATTAAGGAACATCACTTTGATGAGCTTGTGAAGCATACCAAGAGCAGGCGAAGGACTTCTGGATGGCGGGTGTATCTTACTCATGAACTATTTTCATCTGCTG 200
R P D Y R N I T F D Y E L V E A Y Q E G A L L D G G G V D I L L I E T I F D T A

601 AATCCCAAGGCGAGCTTGTTCAGCTCCAAAATCTTTTGAGGAGAAATATGCTCCCGGCTATCTTTATTTAGGAGCAGTCTGTATGAAAGTGGCGGACTCTTTCCGACAGACA 240
M A K A A L F A L Q N L F E E K Y A P R P I F I S G T I V D K S G R T L S G Q T

721 GGAGGGGATTTGTCACTGCTGTCTCATGAGAACCACTCTGCAATGATTAATTTGCTTTGGGTGAGCTGAGATGAGACCTTTTATTGAAATATGGAAGTGTACAAACAGCC 280
G E G F V I S V S N G E P L C I G L R C A L G A A E M R P P I E I I G K C T T A

841 TATGCTCTCTTATCCCAATGCAGGTCTTCCCAACACTTTGGTGACTATCATGAACCCCTTCTATGATGGCCAGCAGCTAAAGGATTTTGTATGGATGGCTTGGTCAATATAGT 320
Y V L C Y P N A G L P N T F G D Y D E T P S N M A K H L K D F A M O G L V N I Y

961 GGAGCATCTGTGGGTCAACACCAGATCATATCAGGAAATTTGCTGAGCTGTGAAGTGTAAAGTGTAAAGTGTAAAGTGTAAAGTGTAAAGTGTAAAGTGTAAAGTGTAAAGTGT 360
G G C C G S T P D N I R E I A E A V K N C K P R V P P A T A F E G H N L L S G L

1081 GAGCCCTTCAAGATTGGACGTAACCAACTTTGTAACTTGGAGAGCGCTGTAAATGTTGAGGATCAAGGAAGTTTGTAAACTCATCATGGCAGGAAGCTATGAAGAAGCTTGTGT 400
E P F R I G P Y T N F V N I G E R C N V A G S R K F A K L I M A G N Y E E A L C

1201 GTTGCCTCAAGTGCAGGTGAAATGGGAGGCTGCTGAT 440
V A K V Q V E N G A Q V L D V M N D D G H L D G P S A M T R F C N L I A S E P D

1321 ATCCCAAGGTACCTTTTGTGATGAGCTCTCTCAATTTTGTGTGATGAGCTGGTTAAAGTGTCTGCAAGGGAAGTGCATTTGCAATAGCATAGTGTGAAGGAGGAGGAGCAG 480
I A K V L C I D S S N F A V I E A G L G G Q K C I V N S I S L K E G E D D

1441 TTCTTGAGAGGCGCAGGAAGATTAAAGATGAGCTGCTATGCTGCTCATGGCTTTTGTCAAGAGGACAGGCAACAGAAACAGACACAAATCAGAGTGTGACCCCGGCTTAC 520
F L E K A R K I K K Y G A A M V V N A F D E E G G A T E T O T K I R V C T R A Y

1561 CATCTGCTTGTGAAAAAAGTGGGCTTAAATCCAAATGACATATTTTGGCCCTAATATCCTAAGCTTGGGACTGGAATGGAGGAACACAATTTGATGCAATTAATTTTATCCATGCA 560
N L L L G F N P N D I I F D P N I L T I G M E E N N L Y A I N F I H A

1681 ACAAGTCAATTAAGAAACATTACCTGGAGCAGATAAGTGGAGGTCTTTTCAACTTGTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 600
T K V I K E Y L P G A R I S G G L S M L S F S F R G M E A I R E A M H G V F L Y

1801 CATGCAATCAAGTGTGATGAGCTGAGATGAGATGAGTGAATCTGCAACCTTCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 640
H A I K S G M D H E I S G M H L P V I T D D I H K E L L G L C E D L I W N K D P

1921 GAGGCGACTGAGAAGCTCTTACCTTATGCCAGACTCAAGGACAGGAGGAGGAAGTCAATTCAGACTCATGAGTGGAGAAATGGCCCTGTGCAAGAACGCGCTTGAATATGCCCTTGTG 680
E A T E K L L E Y A G T Q G T G G K K V I A T D E W R N G P V E K E R E E T R V

2041 AAGGGCATTGAAAAACATATTATGAGTACTCAGCAAGCAGTTTAAAGTCAAGGCTTAAAGTCAAGGCTTAAAGTCAAGGCTTAAAGTCAAGGCTTAAAGTCAAGGCTTAAAGT 720
K G I E K H I I E D T E E A R L N Q K K Y P R P L N I I E G P L N N G H K I V G

2161 GATCTTTTGGAGCTGGAAGATGTTTCTACCTCAGGTTATAAGTCAAGGCTTAAAGTCAAGGCTTAAAGTCAAGGCTTAAAGTCAAGGCTTAAAGTCAAGGCTTAAAGTCAAG 760
D L F G A G K M F L P Q V I K S A R V N K K A V G H L I P F N E K E R E E T R V

2281 CTTACGCGCAGTGAAGAGAG 800
L N G T V E E E D P Y Q G T I V L A T V K G O V N D I G K N I V G V V L G C N H

2401 TTCCGAGTTATTGATTAGGAGTCACTGCTCATGAT 840
F R V I D L G V M T P C D K I L K A A L D H K A D I I G L S G L I T P S L D E M

2521 ATTTTCTTCCCAAGGAAATGGAGAGATTAGCTATAAGGATTCCATTGTTGATTGGAGGAGCAACCACTTCAAAAACCCACAGCAGTTAAATAGCTCCGAGATACAGTGCACCTGTA 880
I F V A K E M E R L A I R I P L L I G G A T T S K T H T A V K I A P R Y S A P V

2641 ATCCATGCTCTGAGCGCTCAAGAGTGTGGTGGTGTGTTCCGAGCTGTTAGATGAAAATCTAAGGAGTAAATAGCTTGAAGAAATCATGGAAGAAATATGAACATATTAGACAGGAGCAT 920
I H V L D A S K S V V C S G L D E N L K D E Y F E E I N E E Y E D I R Q D H

2761 TATGAGTCTCTCAAGGAGAGGATACTTACCTTAAGTCAAGCCAGAAAAGTGGTTTCAAAATGGATTGGCTGTCTGAACCTCACCCAGTGAAGCCACGTTTATTGGGAGCCAGGTC 960
Y E S L K E R R Y L P L S G A R K S G F R M D U L S E P H P V K P T F I G T Q V

2881 TTTCAAGCACTATCAGCTGCACAAGCTGGTGGACTACATTCAGTGGAGCCCTTTCTTTGATCTGTGGCAGCTCCGGGGCAAGTACCCGAATCGAGGCTTCCCAAGATATTAAACACAAA 1000
F E D Y D L G K L V D Y I O W K P F F D V U Q L R G K Y P M R G F P K I F N D K

3001 ACAGTAGGTGCAGAGGCGAG 1040
T V G G E A R K V Y D D A N N M L N T L I S O K K L R A G V G F P A G S L Y

3121 CAABACACATTCAGCTGTAGCAGAGGCTGTGTGCGCCAGGCTGAGAGCCCATAGCCACTTTCTATGGTTAAGCCCAACAGGCTGACAGGAGCTCTCCAGCAGGAGCCATCTACTAC 1080
Q D D I H L Y A E A A V P Q A A E P I A T F Y G L R Q G A E K D S A S T E P Y V

3241 TGGTCTCAGACTTCATCGCTCCCTTGCATTCTGGCATCGTGACTACCTCGGCTGTTTGGCTTGGCTGCTTGGGTTAGAGAGGCTGAGCAAGGCTATCAGGATGATGCTGAGGAC 1120
C L S D F I A P L N S G I R D Y L G L F A V A C F G V G E E L S K A Y E D D G D D

4/7

Fig. 3 page 2 of 2

3361 TACAGCAGCATCATGGTCAAGGGGCTGGGGACCGGCTGCCAGAGGCTTTGCAGAGAGCTCCATCAAGAGTTCCCGAGAACTGTGGGCTACTGTGGCAGTCAACAGCTGGACGTC
YSSIMVKALGDRLEAF AEELHERVRRLELMAYCGSEQLDV 1160

3481 GCAGACCTGCGAAGTTGCGGTACAAGGCGATCGGCGGGCTCTGGCTACCCAGCCAGCCGACGAGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAGCC
ADLERRLRYKEIRPAPGYPSPDHTTEKLIMURLADIEQSTG 1200

3601 ATTAGGTTAACAGAAATCATTAGCAATGGCACCTGCTTCAGCAGTCTCAGGCTCTACTTCTCCAATTGAAGTCCAAATATTTTGGCTGTGGCCAGATTTCCAAGCATCAGTTGAGGAT
IRLTESLANAPASAVSGLYFSHLKSKYFAVGEKISKDQVED 1240

3721 TATGCATTGAGGAAGACATATCTCTGGCTGAGGTTGAGAAATGGCTTGGACCCATTTTGGGATATGATACAGACTAACTTTTTTTTTTTTTTGGCTTTTTTATCTTGATGATCCTCA
YALRKNISSVAEVEKWLGPILGYDTG 1265

3841 AGGAATACAACCTAG

FIG. 4

5/7

1

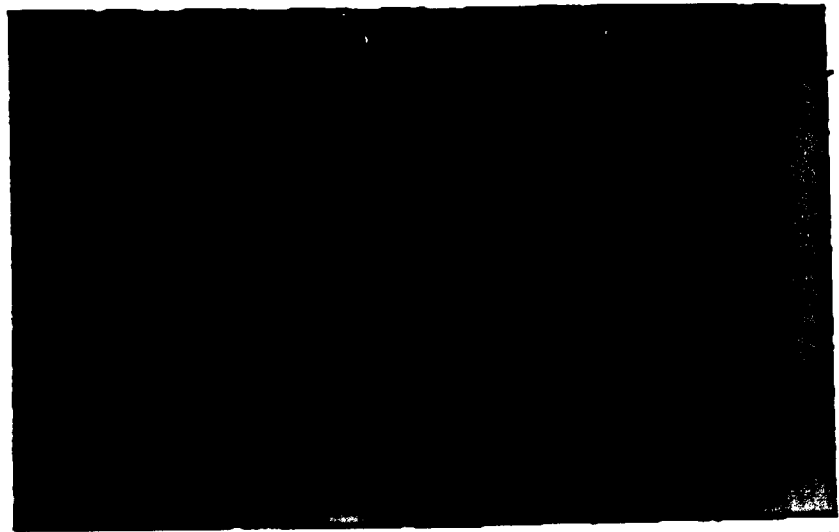
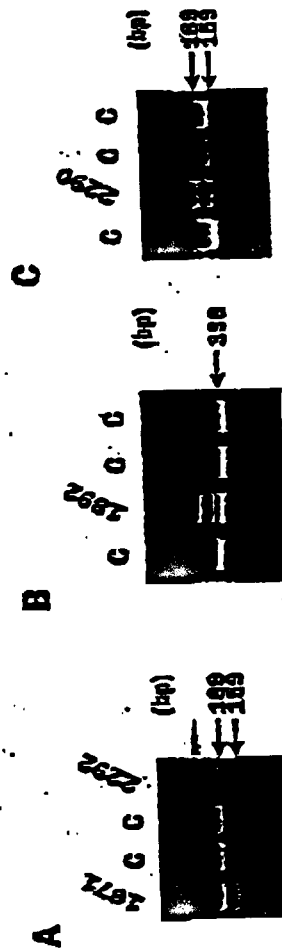


FIG. 5

6/7



box 2

Ec (SEQ ID NO: 73)
 Hi (SEQ ID NO: 72)
 Ce (SEQ ID NO: 71)
 Ml (SEQ ID NO: 70)
 Ss (SEQ ID NO: 69)
 Mm (SEQ ID NO: 68)
 Hs (SEQ ID NO: 67)