



(86) Date de dépôt PCT/PCT Filing Date: 2002/01/22

(87) Date publication PCT/PCT Publication Date: 2002/08/01

(85) Entrée phase nationale/National Entry: 2003/07/17

(86) N° demande PCT/PCT Application No.: US 2002/001818

(87) N° publication PCT/PCT Publication No.: 2002/059287

(30) Priorité/Priority: 2001/01/23 (60/263,378) US

(51) Cl.Int.<sup>7</sup>/Int.Cl.<sup>7</sup> C12N 9/00

(71) Demandeur/Applicant:  
LEXICON GENETICS INCORPORATED, US

(72) Inventeurs/Inventors:  
TURNER, C. ALEXANDER JR., US;  
MATHUR, BRIAN, US

(74) Agent: OGILVY RENAULT

(54) Titre : NOUVELLES KINASES HUMAINES ET POLYNUCLEOTIDES CODANT POUR CELLES-CI

(54) Title: NOVEL HUMAN KINASES AND POLYNUCLEOTIDES ENCODING THE SAME

(57) **Abrégé/Abstract:**

Novel human polynucleotide and polypeptide sequences are disclosed that can be used in therapeutic, diagnostic, and pharmacogenomic applications.



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

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
1 August 2002 (01.08.2002)

(10) International Publication Number  
**PCT**  
**WO 02/059287 A2**

(51) International Patent Classification<sup>7</sup>: **C12N 9/00**

(21) International Application Number: **PCT/US02/01818**

(22) International Filing Date: 22 January 2002 (22.01.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/263,378 23 January 2001 (23.01.2001) US

(71) Applicant: **LEXICON GENETICS INCORPORATED**  
[US/US]; 4000 Research Forest Drive, The Woodlands, TX  
77381 (US).

(72) Inventors: **TURNER, C., Alexander, Jr.**; 67 Win-  
ter Wheat Place, The Woodlands, TX 77381 (US).  
**MATHUR, Brian**; 12000 Sawmill Road #2014, The  
Woodlands, TX 77380 (US).

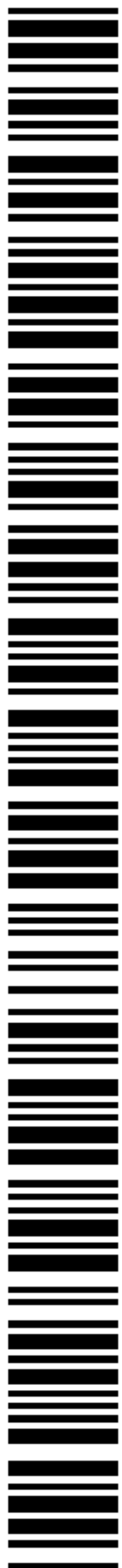
(74) Agents: **ISHIMOTO, Lance, K.** et al.; Lexicon Genetics  
Incorporated, 4000 Research Forest Drive, The Woodlands,  
TX 77381 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,  
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,  
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,  
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,  
MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG,  
SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN,  
YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),  
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),  
European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR,  
GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent  
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR,  
NE, SN, TD, TG).

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

*For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.*



**WO 02/059287 A2**

(54) Title: NOVEL HUMAN KINASES AND POLYNUCLEOTIDES ENCODING THE SAME

(57) Abstract: Novel human polynucleotide and polypeptide sequences are disclosed that can be used in therapeutic, diagnostic, and pharmacogenomic applications.

NOVEL HUMAN KINASES AND  
POLYNUCLEOTIDES ENCODING THE SAME

The present application claims the benefit of U.S. Provisional Application Number 60/263,378, which was filed on  
5 January 23, 2001, and is herein incorporated by reference in its entirety.

1. INTRODUCTION

The present invention relates to the discovery, identification, and characterization of novel human  
10 polynucleotides encoding proteins sharing sequence similarity with animal kinases. The invention encompasses the described polynucleotides, host cell expression systems, the encoded proteins, fusion proteins, polypeptides and peptides, antibodies to the encoded proteins and peptides, and genetically engineered  
15 animals that either lack or overexpress the disclosed polynucleotide sequences, antagonists and agonists of the proteins, and other compounds that modulate the expression or activity of the proteins encoded by the disclosed polynucleotide sequences, which can be used for diagnosis, drug screening,  
20 clinical trial monitoring, the treatment of diseases and disorders, and cosmetic or nutraceutical applications.

2. BACKGROUND OF THE INVENTION

Kinases mediate the phosphorylation of a wide variety of proteins and compounds in the cell. Along with phosphatases,  
25 kinases are involved in a range of regulatory pathways. Given the physiological importance of kinases, they have been subject to intense scrutiny and are proven drug targets.

### 3. SUMMARY OF THE INVENTION

The present invention relates to the discovery, identification, and characterization of nucleotides that encode novel human proteins and the corresponding amino acid sequences of these proteins. The novel human proteins (NHPs) described for the first time herein share structural similarity with animal kinases, including, but not limited to, serine-threonine kinases, carbon catabolite repressing kinases, calcium/calmodulin-dependent protein kinases, and cAMP-dependent protein kinases. Accordingly, the described NHPs encode novel kinases having homologues and orthologs across a range of phyla and species.

The novel human polynucleotides described herein, encode open reading frames (ORFs) encoding proteins of 668 and 608 amino acids in length (see respectively SEQ ID NOS:2 and 4).

The invention also encompasses agonists and antagonists of the described NHPs, including small molecules, large molecules, mutant NHPs, or portions thereof, that compete with native NHPs, peptides, and antibodies, as well as nucleotide sequences that can be used to inhibit the expression of the described NHPs (e.g., antisense and ribozyme molecules, and open reading frame or regulatory sequence replacement constructs) or to enhance the expression of the described NHPs (e.g., expression constructs that place the described polynucleotide under the control of a strong promoter system), and transgenic animals that express a NHP sequence, or "knock-outs" (which can be conditional) that do not express a functional NHP. Knock-out mice can be produced in several ways, one of which involves the use of mouse embryonic stem cell ("ES cell") lines that contain gene trap mutations in a murine homolog of at least one of the described NHPs. When the unique NHP sequences described in SEQ ID NOS:1-4 are "knocked-out" they provide a method of identifying phenotypic

expression of the particular gene, as well as a method of assigning function to previously unknown genes. In addition, animals in which the unique NHP sequences described in SEQ ID NOS:1-4 are "knocked-out" provide a unique source in which to  
5 elicit antibodies to homologous and orthologous proteins, which would have been previously viewed by the immune system as "self" and therefore would have failed to elicit significant antibody responses. To these ends, gene trapped knockout ES cells have been generated in murine homologs of the described NHPs.

10 Additionally, the unique NHP sequences described in SEQ ID NOS:1-4 are useful for the identification of protein coding sequences, and mapping a unique gene to a particular chromosome. These sequences identify biologically verified exon splice junctions, as opposed to splice junctions that may have been  
15 bioinformatically predicted from genomic sequence alone. The sequences of the present invention are also useful as additional DNA markers for restriction fragment length polymorphism (RFLP) analysis, and in forensic biology.

Further, the present invention also relates to processes  
20 for identifying compounds that modulate, *i.e.*, act as agonists or antagonists of, NHP expression and/or NHP activity that utilize purified preparations of the described NHPs and/or NHP products, or cells expressing the same. Such compounds can be used as therapeutic agents for the treatment of any of a wide  
25 variety of symptoms associated with biological disorders or imbalances.

#### 4. DESCRIPTION OF THE SEQUENCE LISTING AND FIGURES

The Sequence Listing provides the sequences of the novel  
30 human ORFs encoding the described novel human kinase proteins.

## 5. DETAILED DESCRIPTION OF THE INVENTION

The NHPs described for the first time herein are novel proteins that are expressed in, *inter alia*, human cell lines and human brain, pituitary, hypothalamus, adipose, cerebellum, adrenal gland, fetal lung, and 6- and 9-week embryo cells. The described sequences were compiled from sequences available in GENBANK, and cDNAs generated from human brain, pituitary, hypothalamus, adipose, cerebellum, adrenal gland, fetal lung, and fetal brain mRNAs (Edge Biosystems, Gaithersburg, MD) that were identified using primers generated from human genomic DNA.

The present invention encompasses the nucleotides presented in the Sequence Listing, host cells expressing such nucleotides, the expression products of such nucleotides, and: (a) nucleotides that encode mammalian homologs of the described genes, including the specifically described NHPs, and the NHP products; (b) nucleotides that encode one or more portions of a NHP that correspond to functional domains, and the polypeptide products specified by such nucleotide sequences, including, but not limited to, the novel regions of any active domain(s); (c) isolated nucleotides that encode mutant versions, engineered or naturally occurring, of the described NHPs, in which all or a part of at least one domain is deleted or altered, and the polypeptide products specified by such nucleotide sequences, including, but not limited to, soluble proteins and peptides; (d) nucleotides that encode chimeric fusion proteins containing all or a portion of a coding region of a NHP, or one of its domains (e.g., a receptor/ligand binding domain, accessory protein/self-association domain, etc.) fused to another peptide or polypeptide; or (e) therapeutic or diagnostic derivatives of the described polynucleotides, such as oligonucleotides, antisense polynucleotides, ribozymes, dsRNA, or gene therapy

constructs, comprising a sequence first disclosed in the Sequence Listing.

As discussed above, the present invention includes the human DNA sequences presented in the Sequence Listing (and  
5 vectors comprising the same), and additionally contemplates any nucleotide sequence encoding a contiguous NHP open reading frame (ORF) that hybridizes to a complement of a DNA sequence presented in the Sequence Listing under highly stringent conditions, e.g., hybridization to filter-bound DNA in 0.5 M  
10 NaHPO<sub>4</sub>, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1x SSC/0.1% SDS at 68°C (Ausubel et al., eds., 1989, Current Protocols in Molecular Biology, Vol. I, Green Publishing Associates, Inc., and John Wiley & Sons, Inc., N.Y., at p. 2.10.3) and encodes a functionally equivalent expression  
15 product. Additionally contemplated are any nucleotide sequences that hybridize to the complement of a DNA sequence that encodes and expresses an amino acid sequence presented in the Sequence Listing under moderately stringent conditions, e.g., washing in 0.2x SSC/0.1% SDS at 42°C (Ausubel et al., 1989, supra), yet  
20 still encode a functionally equivalent NHP product. Functional equivalents of a NHP include naturally occurring NHPs present in other species, and mutant NHPs, whether naturally occurring or engineered (by site directed mutagenesis, gene shuffling, directed evolution as described in, for example, U.S. Patent  
25 Nos. 5,837,458 or 5,723,323 both of which are herein incorporated by reference). The invention also includes degenerate nucleic acid variants of the disclosed NHP polynucleotide sequences.

Additionally contemplated are polynucleotides encoding NHP  
30 ORFs, or their functional equivalents, encoded by polynucleotide sequences that are about 99, 95, 90, or about 85 percent similar to corresponding regions of SEQ ID NOs:1 or 3 (as measured by

BLAST sequence comparison analysis using, for example, the GCG sequence analysis package, as described herein, using default parameters).

The invention also includes nucleic acid molecules,  
5 preferably DNA molecules, that hybridize to, and are therefore the complements of, the described NHP-encoding polynucleotides. Such hybridization conditions can be highly stringent or less highly stringent, as described herein. In instances where the nucleic acid molecules are deoxyoligonucleotides ("DNA oligos"),  
10 such molecules are generally about 16 to about 100 bases long, or about 20 to about 80 bases long, or about 34 to about 45 bases long, or any variation or combination of sizes represented therein that incorporate a contiguous region of sequence first disclosed in the Sequence Listing. Such oligonucleotides can be  
15 used in conjunction with the polymerase chain reaction (PCR) to screen libraries, isolate clones, and prepare cloning and sequencing templates, etc.

Alternatively, such NHP oligonucleotides can be used as hybridization probes for screening libraries, and assessing gene  
20 expression patterns (particularly using a microarray or high-throughput "chip" format). Additionally, a series of the described NHP oligonucleotide sequences, or the complements thereof, can be used to represent all or a portion of the described NHP sequences. An oligonucleotide or polynucleotide  
25 sequence first disclosed in at least a portion of one or more of the sequences of SEQ ID NOS:1-4 can be used as a hybridization probe in conjunction with a solid support matrix/substrate (resins, beads, membranes, plastics, polymers, metal or metallized substrates, crystalline or polycrystalline  
30 substrates, etc.). Of particular note are spatially addressable arrays (*i.e.*, gene chips, microtiter plates, etc.) of oligonucleotides and polynucleotides, or corresponding

oligopeptides and polypeptides, wherein at least one of the biopolymers present on the spatially addressable array comprises an oligonucleotide or polynucleotide sequence first disclosed in at least one of the sequences of SEQ ID NOS:1-4, or an amino acid sequence encoded thereby. Methods for attaching biopolymers to, or synthesizing biopolymers on, solid support matrices, and conducting binding studies thereon, are disclosed in, *inter alia*, U.S. Patent Nos. 5,700,637, 5,556,752, 5,744,305, 4,631,211, 5,445,934, 5,252,743, 4,713,326, 5,424,186, and 4,689,405, the disclosures of which are herein incorporated by reference in their entirety.

Addressable arrays comprising sequences first disclosed in SEQ ID NOS:1-4 can be used to identify and characterize the temporal and tissue specific expression of a gene. These addressable arrays incorporate oligonucleotide sequences of sufficient length to confer the required specificity, yet be within the limitations of the production technology. The length of these probes is usually within a range of between about 8 to about 2000 nucleotides. Preferably the probes consist of 60 nucleotides, and more preferably 25 nucleotides, from the sequences first disclosed in SEQ ID NOS:1-4.

For example, a series of the described oligonucleotide sequences, or the complements thereof, can be used in chip format to represent all or a portion of the described sequences. The oligonucleotides, typically between about 16 to about 40 (or any whole number within the stated range) nucleotides in length, can partially overlap each other, and/or the sequence may be represented using oligonucleotides that do not overlap. Accordingly, the described polynucleotide sequences shall typically comprise at least about two or three distinct oligonucleotide sequences of at least about 8 nucleotides in length that are each first disclosed in the described Sequence

Listing. Such oligonucleotide sequences can begin at any nucleotide present within a sequence in the Sequence Listing, and proceed in either a sense (5'-to-3') orientation vis-a-vis the described sequence or in an antisense orientation.

5 Microarray-based analysis allows the discovery of broad patterns of genetic activity, providing new understanding of gene functions, and generating novel and unexpected insight into transcriptional processes and biological mechanisms. The use of addressable arrays comprising sequences first disclosed in SEQ  
10 ID NOS:1-4 provides detailed information about transcriptional changes involved in a specific pathway, potentially leading to the identification of novel components, or gene functions that manifest themselves as novel phenotypes.

Probes consisting of sequences first disclosed in SEQ ID  
15 NOS:1-4 can also be used in the identification, selection, and validation of novel molecular targets for drug discovery. The use of these unique sequences permits the direct confirmation of drug targets, and recognition of drug dependent changes in gene expression that are modulated through pathways distinct from the  
20 intended target of the drug. These unique sequences therefore also have utility in defining and monitoring both drug action and toxicity.

As an example of utility, the sequences first disclosed in SEQ ID NOS:1-4 can be utilized in microarrays, or other assay  
25 formats, to screen collections of genetic material from patients who have a particular medical condition. These investigations can also be carried out using the sequences first disclosed in SEQ ID NOS:1-4 *in silico*, and by comparing previously collected genetic databases and the disclosed sequences using computer  
30 software known to those in the art.

Thus the sequences first disclosed in SEQ ID NOS:1-4 can be used to identify mutations associated with a particular disease, and also in diagnostic or prognostic assays.

Although the presently described sequences have been  
5 specifically described using nucleotide sequence, it should be appreciated that each of the sequences can uniquely be described using any of a wide variety of additional structural attributes, or combinations thereof. For example, a given sequence can be described by the net composition of the nucleotides present  
10 within a given region of the sequence, in conjunction with the presence of one or more specific oligonucleotide sequence(s) first disclosed in SEQ ID NOS:1-4. Alternatively, a restriction map specifying the relative positions of restriction endonuclease digestion sites, or various palindromic or other  
15 specific oligonucleotide sequences, can be used to structurally describe a given sequence. Such restriction maps, which are typically generated by widely available computer programs (e.g., the University of Wisconsin GCG sequence analysis package, SEQUENCHER 3.0, Gene Codes Corp., Ann Arbor, MI, etc.), can  
20 optionally be used in conjunction with one or more discrete nucleotide sequence(s) present in the sequence that can be described by the relative position of the sequence relative to one or more additional sequence(s) or one or more restriction sites present in the disclosed sequence.

25 For oligonucleotide probes, highly stringent conditions may refer, e.g., to washing in 6x SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligos), 48°C (for 17-base oligos), 55°C (for 20-base oligos), and 60°C (for 23-base oligos). These nucleic acid molecules may encode or act as NHP antisense molecules,  
30 useful, for example, in NHP gene regulation and/or as antisense primers in amplification reactions of NHP nucleic acid sequences. With respect to NHP gene regulation, such techniques

can be used to regulate biological functions. Further, such sequences can be used as part of ribozyme and/or triple helix sequences that are also useful for NHP gene regulation.

Inhibitory antisense or double stranded oligonucleotides  
5 can additionally comprise at least one modified base moiety that is selected from the group including, but not limited to, 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-  
10 2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine,  
15 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil,  
20 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

The antisense oligonucleotide can also comprise at least  
25 one modified sugar moiety selected from the group including, but not limited to, arabinose, 2-fluoroarabinose, xylulose, and hexose.

In yet another embodiment, the antisense oligonucleotide will comprise at least one modified phosphate backbone selected  
30 from the group including, but not limited to, a phosphorothioate, a phosphorodithioate, a phosphoramidothioate,

a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

In yet another embodiment, the antisense oligonucleotide is an  $\alpha$ -anomeric oligonucleotide. An  $\alpha$ -anomeric oligonucleotide  
5 forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual  $\beta$ -units, the strands run parallel to each other (Gautier *et al.*, 1987, Nucl. Acids Res. 15:6625-6641). The oligonucleotide is a 2'-O-methylribonucleotide (Inoue  
10 *et al.*, 1987, Nucl. Acids Res. 15:6131-6148), or a chimeric RNA-DNA analogue (Inoue *et al.*, 1987, FEBS Lett. 215:327-330). Alternatively, double stranded RNA can be used to disrupt the expression and function of a targeted NHP.

Oligonucleotides of the invention can be synthesized by standard methods known in the art, e.g., by use of an automated  
15 DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides can be synthesized (Stein *et al.*, 1988, Nucl. Acids Res. 16:3209), and methylphosphonate oligonucleotides can be prepared by use of controlled pore glass  
20 polymer supports (Sarin *et al.*, 1988, Proc. Natl. Acad. Sci. USA 85:7448-7451), etc.

Low stringency conditions are well-known to those of skill in the art, and will vary predictably depending on the specific organisms from which the library and the labeled sequences are  
25 derived. For guidance regarding such conditions see, for example, Sambrook *et al.*, 1989, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor, N.Y. (and periodic updates thereof); and Ausubel *et al.*, 1989, *supra*.

Alternatively, suitably labeled NHP nucleotide probes can  
30 be used to screen a human genomic library using appropriately stringent conditions or by PCR. The identification and characterization of human genomic clones is helpful for

identifying polymorphisms (including, but not limited to, nucleotide repeats, microsatellite alleles, single nucleotide polymorphisms, or coding single nucleotide polymorphisms), determining the genomic structure of a given locus/allele, and  
5 designing diagnostic tests. For example, sequences derived from regions adjacent to the intron/exon boundaries of the human gene can be used to design primers for use in amplification assays to detect mutations within the exons, introns, splice sites (e.g., splice acceptor and/or donor sites), etc., that can be used in  
10 diagnostics and pharmacogenomics.

For example, the present sequences can be used in restriction fragment length polymorphism (RFLP) analysis to identify specific individuals. In this technique, an individual's genomic DNA is digested with one or more  
15 restriction enzymes, and probed on a Southern blot to yield unique bands for identification (as generally described in U.S. Patent No. 5,272,057, incorporated herein by reference). In addition, the sequences of the present invention can be used to provide polynucleotide reagents, e.g., PCR primers, targeted to  
20 specific loci in the human genome, which can enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (i.e., another DNA sequence that is unique to a particular individual). Actual base sequence information can be used for identification  
25 as an accurate alternative to patterns formed by restriction enzyme generated fragments.

Further, a NHP gene homolog can be isolated from nucleic acids from an organism of interest by performing PCR using two degenerate or "wobble" oligonucleotide primer pools designed on  
30 the basis of amino acid sequences within the NHP products disclosed herein. The template for the reaction may be total RNA, mRNA, and/or cDNA obtained by reverse transcription of mRNA

prepared from, for example, human or non-human cell lines or tissue known to express, or suspected of expressing, an allele of a NHP gene.

The PCR product can be subcloned and sequenced to ensure  
5 that the amplified sequences represent the sequence of the desired NHP gene. The PCR fragment can then be used to isolate a full length cDNA clone by a variety of methods. For example, the amplified fragment can be labeled and used to screen a cDNA library, such as a bacteriophage cDNA library. Alternatively,  
10 the labeled fragment can be used to isolate genomic clones via the screening of a genomic library.

PCR technology can also be used to isolate full length cDNA sequences. For example, RNA can be isolated, following standard procedures, from an appropriate cellular or tissue source (*i.e.*,  
15 one known to express, or suspected of expressing, a NHP gene). A reverse transcription (RT) reaction can be performed on the RNA using an oligonucleotide primer specific for the most 5' end of the amplified fragment for the priming of first strand synthesis. The resulting RNA/DNA hybrid may then be "tailed"  
20 using a standard terminal transferase reaction, the hybrid may be digested with RNase H, and second strand synthesis may then be primed with a complementary primer. Thus, cDNA sequences upstream of the amplified fragment can be isolated. For a review of cloning strategies that can be used, see *e.g.*,  
25 Sambrook *et al.*, 1989, *supra*.

A cDNA encoding a mutant NHP sequence can be isolated, for example, by using PCR. In this case, the first cDNA strand may be synthesized by hybridizing an oligo-dT oligonucleotide to mRNA isolated from tissue known to express, or suspected of  
30 expressing, a NHP, in an individual putatively carrying a mutant NHP allele, and by extending the new strand with reverse transcriptase. The second strand of the cDNA is then

synthesized using an oligonucleotide that hybridizes specifically to the 5' end of the normal sequence. Using these two primers, the product is then amplified via PCR, optionally cloned into a suitable vector, and subjected to DNA sequence analysis through methods well-known to those of skill in the art. By comparing the DNA sequence of the mutant NHP allele to that of a corresponding normal NHP allele, the mutation(s) responsible for the loss or alteration of function of the mutant NHP gene product can be ascertained.

Alternatively, a genomic library can be constructed using DNA obtained from an individual suspected of carrying, or known to carry, a mutant NHP allele (e.g., a person manifesting a NHP-associated phenotype such as, for example, behavioral disorders, immune disorders, obesity, high blood pressure, etc.), or a cDNA library can be constructed using RNA from a tissue known to express, or suspected of expressing, a mutant NHP allele. A normal NHP gene, or any suitable fragment thereof, can then be labeled and used as a probe to identify the corresponding mutant NHP allele in such libraries. Clones containing mutant NHP sequences can then be purified and subjected to sequence analysis according to methods well-known to those skilled in the art.

Additionally, an expression library can be constructed utilizing cDNA synthesized from, for example, RNA isolated from a tissue known to express, or suspected of expressing, a mutant NHP allele in an individual suspected of carrying, or known to carry, such a mutant allele. In this manner, gene products made by the putatively mutant tissue may be expressed and screened using standard antibody screening techniques in conjunction with antibodies raised against a normal NHP product, as described below (for screening techniques, see, for example, Harlow and

Lane, eds., 1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor Press, Cold Spring Harbor, N.Y.).

Additionally, screening can be accomplished by screening with labeled NHP fusion proteins, such as, for example, alkaline phosphatase-NHP or NHP-alkaline phosphatase fusion proteins. In cases where a NHP mutation results in an expression product with altered function (e.g., as a result of a missense or a frameshift mutation), polyclonal antibodies to NHP are likely to cross-react with a corresponding mutant NHP expression product. Library clones detected via their reaction with such labeled antibodies can be purified and subjected to sequence analysis according to methods well-known in the art.

An additional application of the described novel human polynucleotide sequences is their use in the molecular mutagenesis/evolution of proteins that are at least partially encoded by the described novel sequences using, for example, polynucleotide shuffling or related methodologies. Such approaches are described in U.S. Patent Nos. 5,830,721, 5,837,458, 6,117,679, and 5,723,323, which are herein incorporated by reference in their entirety.

The invention also encompasses: (a) DNA vectors that contain any of the foregoing NHP coding sequences and/or their complements (*i.e.*, antisense); (b) DNA expression vectors that contain any of the foregoing NHP coding sequences operatively associated with a regulatory element that directs the expression of the coding sequences (for example, baculovirus as described in U.S. Patent No. 5,869,336, herein incorporated by reference); (c) genetically engineered host cells that contain any of the foregoing NHP coding sequences operatively associated with a regulatory element that directs the expression of the coding sequences in the host cell; and (d) genetically engineered host cells that express an endogenous NHP sequence under the control

of an exogenously introduced regulatory element (*i.e.*, gene activation). As used herein, regulatory elements include, but are not limited to, inducible and non-inducible promoters, enhancers, operators, and other elements known to those skilled  
5 in the art that drive and regulate expression. Such regulatory elements include, but are not limited to, the cytomegalovirus (hCMV) immediate early gene, regulatable, viral elements (particularly retroviral LTR promoters), the early or late promoters of SV40 or adenovirus, the *lac* system, the *trp* system,  
10 the *TAC* system, the *TRC* system, the major operator and promoter regions of phage lambda, the control regions of fd coat protein, the promoter for 3-phosphoglycerate kinase (PGK), the promoters of acid phosphatase, and the promoters of the yeast  $\alpha$ -mating factors.

15 Where, as in the present instance, some of the described NHP peptides or polypeptides are thought to be cytoplasmic or nuclear proteins (although processed forms or fragments can be secreted or membrane associated), expression systems can be engineered that produce soluble derivatives of a NHP  
20 (corresponding to a NHP extracellular and/or intracellular domains, or truncated polypeptides lacking one or more hydrophobic domains) and/or NHP fusion protein products (especially NHP-Ig fusion proteins, *i.e.*, fusions of a NHP domain to an IgFc), NHP antibodies, and anti-idiotypic  
25 antibodies (including Fab fragments) that can be used in therapeutic applications. Preferably, the above expression systems are engineered to allow the desired peptide or polypeptide to be recovered from the culture media.

The present invention also encompasses antibodies and anti-  
30 idiotypic antibodies (including Fab fragments), antagonists and agonists of a NHP, as well as compounds or nucleotide constructs that inhibit expression of a NHP sequence (transcription factor

inhibitors, antisense and ribozyme molecules, or open reading frame sequence or regulatory sequence replacement constructs), or promote the expression of a NHP (e.g., expression constructs in which NHP coding sequences are operatively associated with expression control elements such as promoters, promoter/enhancers, etc.).

The NHPs or NHP peptides, NHP fusion proteins, NHP nucleotide sequences, antibodies, antagonists and agonists can be useful for the detection of mutant NHPs or inappropriately expressed NHPs for the diagnosis of disease. The NHP proteins or peptides, NHP fusion proteins, NHP nucleotide sequences, host cell expression systems, antibodies, antagonists, agonists and genetically engineered cells and animals can be used for screening for drugs (or high throughput screening of combinatorial libraries) effective in the treatment of the symptomatic or phenotypic manifestations of perturbing the normal function of a NHP in the body. The use of engineered host cells and/or animals can offer an advantage in that such systems allow not only for the identification of compounds that bind to the endogenous receptor/ligand of a NHP, but can also identify compounds that trigger NHP-mediated activities or pathways.

Finally, the NHP products can be used as therapeutics. For example, soluble derivatives such as NHP peptides/domains corresponding to NHPs, NHP fusion protein products (especially NHP-Ig fusion proteins, i.e., fusions of a NHP, or a domain of a NHP, to an IgFc), NHP antibodies and anti-idiotypic antibodies (including Fab fragments), antagonists or agonists (including compounds that modulate or act on downstream targets in a NHP-mediated pathway) can be used to directly treat diseases or disorders. For instance, the administration of an effective amount of soluble NHP, or a NHP-IgFc fusion protein or an anti-

idiotypic antibody (or its Fab) that mimics the NHP could activate or effectively antagonize the endogenous NHP or a protein interactive therewith. Nucleotide constructs encoding such NHP products can be used to genetically engineer host cells  
5 to express such products *in vivo*; these genetically engineered cells function as "bioreactors" in the body delivering a continuous supply of a NHP, a NHP peptide, or a NHP fusion protein to the body. Nucleotide constructs encoding functional NHPs, mutant NHPs, as well as antisense and ribozyme molecules  
10 can also be used in "gene therapy" approaches for the modulation of NHP expression. Thus, the invention also encompasses pharmaceutical formulations and methods for treating biological disorders.

Various aspects of the invention are described in greater  
15 detail in the subsections below.

#### 5.1 THE NHP SEQUENCES

The cDNA sequences and corresponding deduced amino acid sequences of the described NHPs are presented in the Sequence  
20 Listing.

Expression analysis has provided evidence that the described NHPs can be expressed in a relatively narrow range of human tissues. In addition to serine-threonine kinases, the described NHPs also share significant similarity to a range of  
25 additional kinase families, including kinases associated with signal transduction, from a variety of phyla and species. Given the physiological importance of protein kinases, they have been subject to intense scrutiny as exemplified and discussed in U.S. Patent No. 5,817,479, herein incorporated by reference in its  
30 entirety.

The gene encoding the described NHPs is apparently encoded on human chromosome 11. As such, the described sequences are

useful, *inter alia*, for mapping the coding regions of the human genome, and particularly, chromosome 11 (see GENBANK accession no. AC074189).

An additional application of the described novel human  
5 polynucleotide sequences is their use in the molecular  
mutagenesis/evolution of proteins that are at least partially  
encoded by the described novel sequences using, for example,  
polynucleotide shuffling or related methodologies. Such  
approaches are described in U.S. Patent Nos. 5,830,721 and  
10 5,837,458, which are herein incorporated by reference in their  
entirety.

NHP gene products can also be expressed in transgenic  
animals. Animals of any species, including, but not limited to,  
worms, mice, rats, rabbits, guinea pigs, pigs, micro-pigs,  
15 birds, goats, and non-human primates, e.g., baboons, monkeys,  
and chimpanzees, may be used to generate NHP transgenic animals.

Any technique known in the art may be used to introduce a  
NHP transgene into animals to produce the founder lines of  
transgenic animals. Such techniques include, but are not  
20 limited to, pronuclear microinjection (Hoppe and Wagner, 1989,  
U.S. Patent No. 4,873,191); retrovirus-mediated gene transfer  
into germ lines (Van der Putten *et al.*, 1985, Proc. Natl. Acad.  
Sci. USA 82:6148-6152); gene targeting in embryonic stem cells  
(Thompson *et al.*, 1989, Cell 56:313-321); electroporation of  
25 embryos (Lo, 1983, Mol Cell. Biol. 3:1803-1814); and sperm-  
mediated gene transfer (Lavitrano *et al.*, 1989, Cell 57:717-  
723); *etc.* For a review of such techniques, see Gordon, 1989,  
Transgenic Animals, Intl. Rev. Cytol. 115:171-229, which is  
incorporated by reference herein in its entirety.

30 The present invention provides for transgenic animals that  
carry a NHP transgene in all their cells, as well as animals  
that carry a transgene in some, but not all their cells, *i.e.*,

mosaic animals or somatic cell transgenic animals. A transgene may be integrated as a single transgene, or in concatamers, e.g., head-to-head tandems or head-to-tail tandems. A transgene may also be selectively introduced into and activated in a particular cell-type by following, for example, the teaching of Lasko *et al.*, 1992, *Proc. Natl. Acad. Sci. USA* 89:6232-6236. The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell-type of interest, and will be apparent to those of skill in the art.

10       When it is desired that a NHP transgene be integrated into the chromosomal site of the endogenous NHP gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous NHP gene are designed for the purpose of integrating, 15 via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous NHP gene (*i.e.*, "knockout" animals).

      The transgene can also be selectively introduced into a particular cell-type, thus inactivating the endogenous NHP gene 20 in only that cell-type, by following, for example, the teaching of Gu *et al.*, 1994, *Science* 265:103-106. The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell-type of interest, and will be apparent to those of skill in the art.

25       Once transgenic animals have been generated, the expression of the recombinant NHP gene may be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to assay whether integration of the transgene has taken place. The 30 level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques that include, but are not limited to, Northern blot analysis of

tissue samples obtained from the animal, *in situ* hybridization analysis, and RT-PCR. Samples of NHP gene-expressing tissue may also be evaluated immunocytochemically using antibodies specific for the NHP transgene product.

5       The present invention also provides for "knockin" animals. Knockin animals are those having foreign gene or nonendogenous coding sequence inserted into a chromosome. For example, when a human coding sequence is used to replace its murine ortholog in the mouse. Such humanized knockin animals are useful for the *in*  
10 *vivo* study, testing and validation of, *intra alia*, human drug targets as well as for compounds that are directed at the same.

## 5.2 NHPS AND NHP POLYPEPTIDES

NHPs, NHP polypeptides, NHP peptide fragments, mutated,  
15 truncated, or deleted forms of the NHPs, and/or NHP fusion proteins can be prepared for a variety of uses. These uses include, but are not limited to, the generation of antibodies, as reagents in diagnostic assays, for the identification of other cellular gene products related to a NHP, and as reagents  
20 in assays for screening for compounds that can be used as pharmaceutical reagents useful in the therapeutic treatment of mental, biological, or medical disorders and diseases. Given the similarity information and expression data, the described NHPs can also be targeted (by drugs, oligos, antibodies, etc.)  
25 in order to treat disease, or to therapeutically augment the efficacy of therapeutic agents.

The Sequence Listing discloses the amino acid sequences encoded by the described NHP-encoding polynucleotides. The NHPs display initiator methionines that are present in DNA sequence  
30 contexts consistent with eucaryotic translation initiation sites. The NHPs do not display consensus signal sequences, which indicates that they may be cytoplasmic or possibly nuclear

proteins, although they may also be secreted or membrane associated.

The NHP amino acid sequences of the invention include the amino acid sequences presented in the Sequence Listing, as well as analogues and derivatives thereof. Further, corresponding NHP homologues from other species are encompassed by the invention. In fact, any NHP protein encoded by the NHP nucleotide sequences described above are within the scope of the invention, as are any novel polynucleotide sequences encoding all or any novel portion of an amino acid sequence presented in the Sequence Listing. The degenerate nature of the genetic code is well-known, and, accordingly, each amino acid presented in the Sequence Listing, is generically representative of the well-known nucleic acid "triplet" codon, or in many cases codons, that can encode the amino acid. As such, as contemplated herein, the amino acid sequences presented in the Sequence Listing, when taken together with the genetic code (see, for example, Table 4-1 at page 109 of "Molecular Cell Biology", 1986, J. Darnell et al., eds., Scientific American Books, New York, N.Y., herein incorporated by reference) are generically representative of all the various permutations and combinations of nucleic acid sequences that can encode such amino acid sequences.

The invention also encompasses proteins that are functionally equivalent to the NHPs encoded by the presently described nucleotide sequences as judged by any of a number of criteria, including, but not limited to, the ability to bind and modify a NHP substrate, or the ability to effect an identical or complementary downstream pathway, or a change in cellular metabolism (e.g., proteolytic activity, ion flux, tyrosine phosphorylation, etc.). Such functionally equivalent NHP proteins include, but are not limited to, additions or

substitutions of amino acid residues within the amino acid sequence encoded by the NHP nucleotide sequences described herein, but that result in a silent change, thus producing a functionally equivalent expression product. Amino acid  
5 substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan,  
10 and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid.

15 A variety of host-expression vector systems can be used to express the NHP nucleotide sequences of the invention. Where the NHP peptide or polypeptide can exist, or has been engineered to exist, as a soluble or secreted molecule, the soluble NHP peptide or polypeptide can be recovered from the culture media.  
20 Such expression systems also encompass engineered host cells that express a NHP, or functional equivalent, *in situ*. Purification or enrichment of a NHP from such expression systems can be accomplished using appropriate detergents and lipid micelles and methods well-known to those skilled in the art.  
25 However, such engineered host cells themselves may be used in situations where it is important not only to retain the structural and functional characteristics of a NHP, but to assess biological activity, e.g., in certain drug screening assays.

30 The expression systems that may be used for purposes of the invention include, but are not limited to, microorganisms such as bacteria (e.g., *E. coli*, *B. subtilis*) transformed with

recombinant bacteriophage DNA, plasmid DNA or cosmid DNA  
expression vectors containing NHP nucleotide sequences; yeast  
(e.g., *Saccharomyces*, *Pichia*) transformed with recombinant yeast  
expression vectors containing NHP nucleotide sequences; insect  
5 cell systems infected with recombinant virus expression vectors  
(e.g., baculovirus) containing NHP nucleotide sequences; plant  
cell systems infected with recombinant virus expression vectors  
(e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus,  
TMV) or transformed with recombinant plasmid expression vectors  
10 (e.g., Ti plasmid) containing NHP nucleotide sequences; or  
mammalian cell systems (e.g., COS, CHO, BHK, 293, 3T3) harboring  
recombinant expression constructs containing NHP nucleotide  
sequences and promoters derived from the genome of mammalian  
cells (e.g., metallothionein promoter) or from mammalian viruses  
15 (e.g., the adenovirus late promoter; the vaccinia virus 7.5K  
promoter).

In bacterial systems, a number of expression vectors may be  
advantageously selected depending upon the use intended for the  
NHP product being expressed. For example, when a large quantity  
20 of such a protein is to be produced for the generation of  
pharmaceutical compositions of or containing a NHP, or for  
raising antibodies to a NHP, vectors that direct the expression  
of high levels of fusion protein products that are readily  
purified may be desirable. Such vectors include, but are not  
25 limited to, the *E. coli* expression vector pUR278 (Ruther et al.,  
1983, EMBO J. 2:1791), in which a NHP coding sequence may be  
ligated individually into the vector in-frame with the *lacZ*  
coding region so that a fusion protein is produced; pIN vectors  
(Inouye and Inouye, 1985, Nucleic Acids Res. 13:3101-3109; Van  
30 Heeke and Schuster, 1989, J. Biol. Chem. 264:5503-5509); and the  
like. pGEX vectors may also be used to express foreign  
polypeptides as fusion proteins with glutathione S-transferase

(GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. The PGEX vectors are designed to include  
5 thrombin or factor Xa protease cleavage sites so that the cloned target expression product can be released from the GST moiety.

In an exemplary insect system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign polynucleotide sequences. The virus grows in  
10 *Spodoptera frugiperda* cells. A NHP coding sequence can be cloned individually into a non-essential region (for example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter). Successful insertion of a NHP coding sequence will result in  
15 inactivation of the polyhedrin gene and production of non-occluded recombinant virus (*i.e.*, virus lacking the proteinaceous coat coded for by the polyhedrin gene). These recombinant viruses are then used to infect *Spodoptera frugiperda* cells in which the inserted sequence is expressed  
20 (*e.g.*, see Smith *et al.*, 1983, J. Virol. 46:584; Smith, U.S. Patent No. 4,215,051).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the NHP nucleotide sequence of interest  
25 may be ligated to an adenovirus transcription/translation control complex, *e.g.*, the late promoter and tripartite leader sequence. This chimeric sequence may then be inserted in the adenovirus genome by *in vitro* or *in vivo* recombination. Insertion in a non-essential region of the viral genome (*e.g.*,  
30 region E1 or E3) will result in a recombinant virus that is viable and capable of expressing a NHP product in infected hosts (*e.g.*, see Logan and Shenk, 1984, Proc. Natl. Acad. Sci. USA

81:3655-3659). Specific initiation signals may also be required for efficient translation of inserted NHP nucleotide sequences. These signals include the ATG initiation codon and adjacent sequences. In cases where an entire NHP gene or cDNA, including  
5 its own initiation codon and adjacent sequences, is inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only a portion of a NHP coding sequence is inserted, exogenous translational control signals, including, perhaps, the  
10 ATG initiation codon, may be provided. Furthermore, the initiation codon should be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural  
15 and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements, transcription terminators, etc. (see Bitter et al., 1987, Methods in Enzymol. 153:516-544).

In addition, a host cell strain may be chosen that  
20 modulates the expression of the inserted sequences, or modifies and processes the expression product in the specific fashion desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may be important for the function of the protein. Different host cells have  
25 characteristic and specific mechanisms for the post-translational processing and modification of proteins and expression products. Appropriate cell lines or host systems can be chosen to ensure the desired modification and processing of the foreign protein expressed. To this end, eukaryotic host  
30 cells that possess the cellular machinery for the desired processing of the primary transcript, glycosylation, and phosphorylation of the expression product may be used. Such

mammalian host cells include, but are not limited to, CHO, VERO, BHK, HeLa, COS, MDCK, 293, 3T3, WI38, and in particular, human cell lines.

For long-term, high-yield production of recombinant  
5 proteins, stable expression is preferred. For example, cell lines that stably express the NHP sequences described herein can be engineered. Rather than using expression vectors that contain viral origins of replication, host cells can be transformed with DNA controlled by appropriate expression  
10 control elements (e.g., promoter, enhancer sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The  
15 selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci, which in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines that express  
20 the NHP product. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that affect the endogenous activity of the NHP product.

A number of selection systems may be used, including, but not limited to, the herpes simplex virus thymidine kinase  
25 (Wigler et al., 1977, Cell 11:223), hypoxanthine-guanine phosphoribosyltransferase (Szybalska and Szybalski, 1962, Proc. Natl. Acad. Sci. USA 48:2026), and adenine  
phosphoribosyltransferase (Lowy et al., 1980, Cell 22:817) genes, which can be employed in tk<sup>-</sup>, hgprt<sup>-</sup> or aprt<sup>-</sup> cells,  
30 respectively. Also, antimetabolite resistance can be used as the basis of selection for the following genes: dhfr, which confers resistance to methotrexate (Wigler et al., 1980, Proc.

Natl. Acad. Sci. USA 77:3567; O'Hare *et al.*, 1981, Proc. Natl. Acad. Sci. USA 78:1527); gpt, which confers resistance to mycophenolic acid (Mulligan and Berg, 1981, Proc. Natl. Acad. Sci. USA 78:2072); neo, which confers resistance to the  
5 aminoglycoside G-418 (Colberre-Garapin *et al.*, 1981, J. Mol. Biol. 150:1); and hygro, which confers resistance to hygromycin (Santerre *et al.*, 1984, Gene 30:147).

Alternatively, any fusion protein can be readily purified by utilizing an antibody specific for the fusion protein being  
10 expressed. An exemplary system allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht *et al.*, 1991, Proc. Natl. Acad. Sci. USA 88:8972-8976). In this system, the sequence of interest is subcloned into a vaccinia recombination plasmid such that the  
15 sequence's open reading frame is translationally fused to an amino-terminal tag consisting of six histidine residues. Extracts from cells infected with recombinant vaccinia virus are loaded onto Ni<sup>2+</sup>-nitriloacetic acid-agarose columns, and histidine-tagged proteins are selectively eluted with imidazole-  
20 containing buffers.

Also encompassed by the present invention are fusion proteins that direct a NHP to a target organ and/or facilitate transport across the membrane into the cytosol. Conjugation of NHPs to antibody molecules or their Fab fragments could be used  
25 to target cells bearing a particular epitope. Attaching an appropriate signal sequence to a NHP would also transport a NHP to a desired location within the cell. Alternatively, targeting of a NHP or its nucleic acid sequence might be achieved using liposome or lipid complex based delivery systems. Such  
30 technologies are described in "Liposomes: A Practical Approach", New, R.R.C., ed., Oxford University Press, N.Y., and in U.S. Patent Nos. 4,594,595, 5,459,127, 5,948,767 and 6,110,490 and

their respective disclosures, which are herein incorporated by reference in their entirety. Additionally embodied are novel protein constructs engineered in such a way that they facilitate transport of NHPs to a target site or desired organ, where they cross the cell membrane and/or the nucleus where the NHPs can exert their functional activity. This goal may be achieved by coupling of a NHP to a cytokine or other ligand that provides targeting specificity, and/or to a protein transducing domain (see generally U.S. Provisional Patent Application Ser. Nos. 60/111,701 and 60/056,713, both of which are herein incorporated by reference, for examples of such transducing sequences), to facilitate passage across cellular membranes, and can optionally be engineered to include nuclear localization signals.

### 5.3 ANTIBODIES TO NHP PRODUCTS

Antibodies that specifically recognize one or more epitopes of a NHP, epitopes of conserved variants of a NHP, or peptide fragments of a NHP, are also encompassed by the invention. Such antibodies include, but are not limited to, polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab')<sub>2</sub> fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above.

The antibodies of the invention can be used, for example, in the detection of a NHP in a biological sample and may, therefore, be utilized as part of a diagnostic or prognostic technique whereby patients may be tested for abnormal amounts of a NHP. Such antibodies may also be utilized in conjunction with, for example, compound screening schemes for the evaluation of the effect of test compounds on expression and/or activity of a NHP expression product. Additionally, such antibodies can be

used in conjunction with gene therapy to, for example, evaluate normal and/or engineered NHP-expressing cells prior to their introduction into a patient. Such antibodies may additionally be used in methods for the inhibition of abnormal NHP activity.

5 Thus, such antibodies may be utilized as a part of treatment methods.

For the production of antibodies, various host animals may be immunized by injection with a NHP, a NHP peptide (e.g., one corresponding to a functional domain of a NHP), truncated NHP  
10 polypeptides (NHP in which one or more domains have been deleted), functional equivalents of the NHPs or mutated variants of the NHPs. Such host animals may include, but are not limited to, pigs, rabbits, mice, goats, and rats, to name but a few. Various adjuvants may be used to increase the immunological  
15 response, depending on the host species, including, but not limited to, Freund's adjuvant (complete and incomplete), mineral salts such as aluminum hydroxide or aluminum phosphate, chitosan, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, and  
20 potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and *Corynebacterium parvum*. Alternatively, the immune response could be enhanced by combination and/or coupling with molecules such as keyhole limpet hemocyanin, tetanus toxoid, diphtheria toxoid, ovalbumin, cholera toxin, or  
25 fragments thereof. Polyclonal antibodies are heterogeneous populations of antibody molecules derived from the sera of the immunized animals.

Monoclonal antibodies, which are homogeneous populations of antibodies to a particular antigen, can be obtained by any  
30 technique that provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique (Kohler and Milstein, 1975,

Nature 256:495-497; and U.S. Patent No. 4,376,110), the human B-cell hybridoma technique (Kosbor *et al.*, 1983, Immunology Today 4:72; Cole *et al.*, 1983, Proc. Natl. Acad. Sci. USA 80:2026-2030), and the EBV-hybridoma technique (Cole *et al.*, 5 1985, Monoclonal Antibodies And Cancer Therapy, Alan R. Liss, Inc., pp. 77-96). Such antibodies may be of any immunoglobulin class, including IgG, IgM, IgE, IgA, or IgD, and any subclass thereof. The hybridomas producing the mAbs of this invention may be cultivated *in vitro* or *in vivo*. Production of high 10 titers of mAbs *in vivo* makes this the presently preferred method of production.

In addition, techniques developed for the production of "chimeric antibodies" (Morrison *et al.*, 1984, Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger *et al.*, 1984, Nature, 312:604- 15 608; Takeda *et al.*, 1985, Nature, 314:452-454), by splicing the genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity, can be used. A chimeric antibody is a molecule in which different portions are derived 20 from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region. Such technologies are described in U.S. Patent Nos. 6,114,598, 6,075,181 and 5,877,397 and their respective disclosures, which are herein incorporated by reference in their 25 entirety. Also encompassed by the present invention is the use of fully humanized monoclonal antibodies, as described in U.S. Patent No. 6,150,584 and respective disclosures, which are herein incorporated by reference in their entirety.

Alternatively, techniques described for the production of 30 single chain antibodies (U.S. Patent No. 4,946,778; Bird, 1988, Science 242:423-426; Huston *et al.*, 1988, Proc. Natl. Acad. Sci. USA 85:5879-5883; and Ward *et al.*, 1989, Nature 341:544-546) can

be adapted to produce single chain antibodies against NHP expression products. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide.

5       Antibody fragments that recognize specific epitopes may be generated by known techniques. For example, such fragments include, but are not limited to:  $F(ab')_2$  fragments, which can be produced by pepsin digestion of an antibody molecule; and Fab fragments, which can be generated by reducing the disulfide  
10 bridges of  $F(ab')_2$  fragments. Alternatively, Fab expression libraries may be constructed (Huse et al., 1989, Science, 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity.

Antibodies to a NHP can, in turn, be utilized to generate  
15 anti-idiotypic antibodies that "mimic" a given NHP, using techniques well-known to those skilled in the art (see, e.g., Greenspan and Bona, 1993, FASEB J. 7:437-444; and Nissinoff, 1991, J. Immunol. 147:2429-2438). For example, antibodies that bind to a NHP domain and competitively inhibit the binding of a  
20 NHP to its cognate receptor/ligand can be used to generate anti-idiotypes that "mimic" the NHP and, therefore, bind, activate, or neutralize a NHP, NHP receptor, or NHP ligand. Such anti-idiotypic antibodies or Fab fragments of such anti-idiotypes can be used in therapeutic regimens involving a NHP-mediated  
25 pathway.

Additionally given the high degree of relatedness of mammalian NHPs, the presently described knock-out mice (having never seen a NHP, and thus never been tolerized to a NHP) have a unique utility, as they can be advantageously applied to the  
30 generation of antibodies against the disclosed mammalian NHPs (i.e., a NHP will be immunogenic in NHP knock-out animals).

The present invention is not to be limited in scope by the specific embodiments described herein, which are intended as single illustrations of individual aspects of the invention, and functionally equivalent methods and components are within the  
5 scope of the invention. Indeed, various modifications of the invention, in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description. Such modifications are intended to fall within the scope of the appended claims. All cited publications, patents,  
10 and patent applications are herein incorporated by reference in their entirety.

## SEQUENCE LISTING

&lt;110&gt; LEXICON GENETICS INCORPORATED

&lt;120&gt; Novel Human Kinases and Polynucleotides Encoding the Same

&lt;130&gt; LEX-0300-PCT

&lt;150&gt; US 60/263,378

&lt;151&gt; 2001-01-23

&lt;160&gt; 4

&lt;170&gt; FastSEQ for Windows Version 4.0

&lt;210&gt; 1

&lt;211&gt; 2007

&lt;212&gt; DNA

&lt;213&gt; homo sapiens

&lt;400&gt; 1

```

atgacatcga cggggaagga cggcggcgcg cagcacgcgc agtatgttgg gccctaccgg 60
ctggagaaga cgctgggcaa ggggcagaca ggtctggtga agctgggggt tcaactgcgtc 120
acctgccaga aggtggccat caagatcgtc aaccgtgaga agctcagcga gtcggtgctg 180
atgaagggtg agcgggagat cgcgatactg aagctcattg agcaccacca cgtcctaaag 240
ctgcacgacg tttatgaaaa caaaaaatat ttgtacctgg tgctagaaca cgtgtcaggt 300
ggtgagctct tcgactacct ggtgaagaag gggaggctga cgcctaagga ggctcggaag 360
ttcttccggc agatcatctc tgcgctggac ttctgccaca gccactccat atgccacagg 420
gatctgaaac ctgaaaacct cctgctggac gagaagaaca acatccgcat cgcagacttt 480
ggcatggcgt ccctgcaggt tggcgacagc ctggttgaga ccagctgtgg gtccccccac 540
tacgcctgcc ccgaggtgat ccggggggag aagtatgacg gccggaaggc ggacgtgtgg 600
agctgcggcg tcactcctgt cgccttgctg gtgggggctc tgcccttcga cgatgacaac 660
ttgcgacagc tgctggagaa ggtgaagcgg ggcgtgttcc acatgccgca ctttatcccc 720
cccgactgcc agagtctgct acggggcatg atcgaggtgg acgccgcacg ccgcctcacg 780
ctagagcaca ttcagaaaca catatggtat atagggggca agaattgagcc cgaaccagag 840
cagcccatte ctgcgaaggt gcagatccgc tcgctgccca gcctggagga catcgacccc 900
gacgtgctgg acagcatgca ctcaactggc tgcttccgag accgcaacaa gctgctgcag 960
gacctgctgt ccgaggagga gaaccaggag aagatgatth acttcctcct cctggaccgg 1020
aaagaaaggt acccgagcca ggaggatgag gacctgcccc ccggaacga gatagaccct 1080
ccccggaagc gtgtggactc ccgatgctg aaccggcacg gcaagcggcg gccagaacgc 1140
aatccatgg aggtgctcag cgtgacggac ggcggctccc cgggtgcctgc gcggcggggc 1200
attgagatgg ccagcacagg ccagaggtct cgggtccatca gcggtgcctc ctcaggcctt 1260
tccaccagcc cactcagcag ccccggggtg acccctcacc cctcaccaag gggcagtccc 1320
ctccccaccc ccaaggggac acctgtccac acgccaagag agagcccggc tggcacgccc 1380
aaccacacgc cccggtccag cccagcgtc ggaggggtgc cctggagggc gcggctcaac 1440
tccatcaaga acagctttct gggctcacc cgcttccacc gccggaact gcaagttccg 1500
acgccggagg agatgtccaa cctgacacca gagtcgtccc cagagctggc gaagaagtcc 1560
tggtttggga acttcatcag cctggagaag gaggagcaga tcttcgtggt catcaaagac 1620
aaacctctga gctccatcaa ggctgacatc gtgcacgcct tcctgtcgat tcccagtctc 1680
agccacagcg tcacttccca aacgagcttc cgggccgagt acaaggccac gggggggcca 1740
gccgtgttcc agaagccggt caagttccag gttgatata cctacacgga ggggtggggag 1800
gcgcagaagg agaacggcat ctactccgtc accttcaccc tgctctcagg cccagccgt 1860
cgcttcaaga ggggtggtgga gaccatccag gccagctgc tgagcacaca cgaccgcct 1920
gcggcccagc acttgtcaga caccactaac tgtatggaaa tgatgacggg gcggctttcc 1980
aatgtggaa ttatcccga aagttaa

```

2007

WO 02/059287

PCT/US02/01818

<210> 2  
 <211> 668  
 <212> PRT  
 <213> homo sapiens

<400> 2  
 Met Thr Ser Thr Gly Lys Asp Gly Gly Ala Gln His Ala Gln Tyr Val  
 1 5 10 15  
 Gly Pro Tyr Arg Leu Glu Lys Thr Leu Gly Lys Gly Gln Thr Gly Leu  
 20 25 30  
 Val Lys Leu Gly Val His Cys Val Thr Cys Gln Lys Val Ala Ile Lys  
 35 40 45  
 Ile Val Asn Arg Glu Lys Leu Ser Glu Ser Val Leu Met Lys Val Glu  
 50 55 60  
 Arg Glu Ile Ala Ile Leu Lys Leu Ile Glu His Pro His Val Leu Lys  
 65 70 75 80  
 Leu His Asp Val Tyr Glu Asn Lys Lys Tyr Leu Tyr Leu Val Leu Glu  
 85 90 95  
 His Val Ser Gly Gly Glu Leu Phe Asp Tyr Leu Val Lys Lys Gly Arg  
 100 105 110  
 Leu Thr Pro Lys Glu Ala Arg Lys Phe Phe Arg Gln Ile Ile Ser Ala  
 115 120 125  
 Leu Asp Phe Cys His Ser His Ser Ile Cys His Arg Asp Leu Lys Pro  
 130 135 140  
 Glu Asn Leu Leu Leu Asp Glu Lys Asn Asn Ile Arg Ile Ala Asp Phe  
 145 150 155 160  
 Gly Met Ala Ser Leu Gln Val Gly Asp Ser Leu Leu Glu Thr Ser Cys  
 165 170 175  
 Gly Ser Pro His Tyr Ala Cys Pro Glu Val Ile Arg Gly Glu Lys Tyr  
 180 185 190  
 Asp Gly Arg Lys Ala Asp Val Trp Ser Cys Gly Val Ile Leu Phe Ala  
 195 200 205  
 Leu Leu Val Gly Ala Leu Pro Phe Asp Asp Asp Asn Leu Arg Gln Leu  
 210 215 220  
 Leu Glu Lys Val Lys Arg Gly Val Phe His Met Pro His Phe Ile Pro  
 225 230 235 240  
 Pro Asp Cys Gln Ser Leu Leu Arg Gly Met Ile Glu Val Asp Ala Ala  
 245 250 255  
 Arg Arg Leu Thr Leu Glu His Ile Gln Lys His Ile Trp Tyr Ile Gly  
 260 265 270  
 Gly Lys Asn Glu Pro Glu Pro Glu Gln Pro Ile Pro Arg Lys Val Gln  
 275 280 285  
 Ile Arg Ser Leu Pro Ser Leu Glu Asp Ile Asp Pro Asp Val Leu Asp  
 290 295 300  
 Ser Met His Ser Leu Gly Cys Phe Arg Asp Arg Asn Lys Leu Leu Gln  
 305 310 315 320  
 Asp Leu Leu Ser Glu Glu Glu Asn Gln Glu Lys Met Ile Tyr Phe Leu  
 325 330 335  
 Leu Leu Asp Arg Lys Glu Arg Tyr Pro Ser Gln Glu Asp Glu Asp Leu  
 340 345 350  
 Pro Pro Arg Asn Glu Ile Asp Pro Pro Arg Lys Arg Val Asp Ser Pro  
 355 360 365  
 Met Leu Asn Arg His Gly Lys Arg Arg Pro Glu Arg Lys Ser Met Glu  
 370 375 380  
 Val Leu Ser Val Thr Asp Gly Gly Ser Pro Val Pro Ala Arg Arg Ala  
 385 390 395 400

Ile Glu Met Ala Gln His Gly Gln Arg Ser Arg Ser Ile Ser Gly Ala  
 405 410 415  
 Ser Ser Gly Leu Ser Thr Ser Pro Leu Ser Ser Pro Arg Val Thr Pro  
 420 425 430  
 His Pro Ser Pro Arg Gly Ser Pro Leu Pro Thr Pro Lys Gly Thr Pro  
 435 440 445  
 Val His Thr Pro Lys Glu Ser Pro Ala Gly Thr Pro Asn Pro Thr Pro  
 450 455 460  
 Pro Ser Ser Pro Ser Val Gly Gly Val Pro Trp Arg Ala Arg Leu Asn  
 465 470 475 480  
 Ser Ile Lys Asn Ser Phe Leu Gly Ser Pro Arg Phe His Arg Arg Lys  
 485 490 495  
 Leu Gln Val Pro Thr Pro Glu Glu Met Ser Asn Leu Thr Pro Glu Ser  
 500 505 510  
 Ser Pro Glu Leu Ala Lys Lys Ser Trp Phe Gly Asn Phe Ile Ser Leu  
 515 520 525  
 Glu Lys Glu Glu Gln Ile Phe Val Val Ile Lys Asp Lys Pro Leu Ser  
 530 535 540  
 Ser Ile Lys Ala Asp Ile Val His Ala Phe Leu Ser Ile Pro Ser Leu  
 545 550 555 560  
 Ser His Ser Val Ile Ser Gln Thr Ser Phe Arg Ala Glu Tyr Lys Ala  
 565 570 575  
 Thr Gly Gly Pro Ala Val Phe Gln Lys Pro Val Lys Phe Gln Val Asp  
 580 585 590  
 Ile Thr Tyr Thr Glu Gly Gly Glu Ala Gln Lys Glu Asn Gly Ile Tyr  
 595 600 605  
 Ser Val Thr Phe Thr Leu Leu Ser Gly Pro Ser Arg Arg Phe Lys Arg  
 610 615 620  
 Val Val Glu Thr Ile Gln Ala Gln Leu Leu Ser Thr His Asp Pro Pro  
 625 630 635 640  
 Ala Ala Gln His Leu Ser Asp Thr Thr Asn Cys Met Glu Met Met Thr  
 645 650 655  
 Gly Arg Leu Ser Lys Cys Gly Ile Ile Pro Lys Ser  
 660 665

&lt;210&gt; 3

&lt;211&gt; 1827

&lt;212&gt; DNA

&lt;213&gt; homo sapiens

&lt;400&gt; 3

atgaagggtgg agcgggagat cgcgacacctg aagctcattg agcacccecca cgtcctaaag 60  
 ctgcacgacg tttatgaaaa caaaaaatat ttgtacctgg tgctagaaca cgtgtcaggt 120  
 ggtgagctct tgcactacct ggtgaagaag gggaggctga cgcctaagga ggctcggag 180  
 ttcttccggc agatcatctc tgcgctggac ttctgccaca gccactccat atgccacagg 240  
 gatctgaaac ctgaaaacct cctgctggac gagaagaaca acatccgcat cgcagacttt 300  
 ggcatggcgt ccctgcaggt tggcgacagc ctgttgaggaga ccagctgtgg gtccccccac 360  
 tacgcctgcc ccgaggtgat ccgggggggag aagtatgacg gccggaaggc ggacgtgtgg 420  
 agctgcggcg tcatcctgtt cgccttgctg gtgggggctc tgcccttcga cgatgacaac 480  
 ttgcgacagc tgctggagaa ggtgaagcgg ggcgtgttcc acatgccgca ctttatcccg 540  
 cccgactgcc agagtctgct acgggggcatg atcgaggtgg acgccgcacg ccgcctcacg 600  
 ctagagcaca ttcagaaaca catatggtat atagggggca agaagtgagcc cgaaccagag 660  
 cagcccatte ctcgcaaggt gcagatccgc tcgctgcccc gcctggagga catcgacccc 720  
 gacgtgctgg acagcatgca ctactgggc tgcttccgag accgcaacaa gctgctgcag 780  
 gacctgctgt ccgaggagga gaaccaggag aagatgattt acttctcct cctggaccgg 840

```

aaagaaaggt acccgagcca ggaggatgag gacctgcccc cccggaacga gatagaccct 900
ccccggaagc gtgtggactc cccgatgctg aaccggcacg gcaagcggcg gccagaacgc 960
aaatccatgg aggtgctcag cgtgacggac ggcgggtccc cggtgccctgc gcggcggggcc 1020
attgagatgg cccagcacgg ccagaggtct cgggccatca gcggtgcctc ctcaggcctt 1080
tccaccagcc cactcagcag cccccgggtg acccctcacc cctcaccaag gggcagtcct 1140
ctccccaccc ccaaggggac acctgtccac acgccaagag agagcccggc tggcacgccc 1200
aaccaccacgc ccccgctccag cccagcgcgc ggagggggtgc cctggagggc gcgggtcaac 1260
tccatcaaga acagctttct gggctcacc cgcttcacc gccggaaact gcaagttccg 1320
acgccggagg agatgtccaa cctgacacca gagtcgtccc cagagctggc gaagaagtcc 1380
tggtttggga acttcatcag cctggagaag gaggagcaga tcttcgtggc catcaaagac 1440
aaacctctga gctccatcaa ggctgacatc gtgcacgcct tcctgtcgat tcccagtctc 1500
agccacagcg tcatctccca aacgagcttc cgggccgagt acaaggccac gggggggcca 1560
gccgtgttcc agaagccggt caagttccag gttgatatca cctacacgga ggggtggggag 1620
gcgcagaagg agaacggcat ctactccgtc accttcacc tgctctcagg cccagccgt 1680
cgcttcaaga ggggtggtgga gaccatccag gccagctgc tgagcacaca cgacccgcct 1740
gcggcccagc acttgtcaga caccactaac tgtatggaaa tgatgacggg gcggctttcc 1800
aaatgtggaa ttatcccga aagttaa 1827

```

&lt;210&gt; 4

&lt;211&gt; 608

&lt;212&gt; PRT

&lt;213&gt; homo sapiens

&lt;400&gt; 4

```

Met Lys Val Glu Arg Glu Ile Ala Ile Leu Lys Leu Ile Glu His Pro
1          5          10          15
His Val Leu Lys Leu His Asp Val Tyr Glu Asn Lys Lys Tyr Leu Tyr
20          25          30
Leu Val Leu Glu His Val Ser Gly Gly Glu Leu Phe Asp Tyr Leu Val
35          40          45
Lys Lys Gly Arg Leu Thr Pro Lys Glu Ala Arg Lys Phe Phe Arg Gln
50          55          60
Ile Ile Ser Ala Leu Asp Phe Cys His Ser His Ser Ile Cys His Arg
65          70          75          80
Asp Leu Lys Pro Glu Asn Leu Leu Leu Asp Glu Lys Asn Asn Ile Arg
85          90          95
Ile Ala Asp Phe Gly Met Ala Ser Leu Gln Val Gly Asp Ser Leu Leu
100         105         110
Glu Thr Ser Cys Gly Ser Pro His Tyr Ala Cys Pro Glu Val Ile Arg
115         120         125
Gly Glu Lys Tyr Asp Gly Arg Lys Ala Asp Val Trp Ser Cys Gly Val
130         135         140
Ile Leu Phe Ala Leu Leu Val Gly Ala Leu Pro Phe Asp Asp Asp Asn
145         150         155         160
Leu Arg Gln Leu Leu Glu Lys Val Lys Arg Gly Val Phe His Met Pro
165         170         175
His Phe Ile Pro Pro Asp Cys Gln Ser Leu Leu Arg Gly Met Ile Glu
180         185         190
Val Asp Ala Ala Arg Arg Leu Thr Leu Glu His Ile Gln Lys His Ile
195         200         205
Trp Tyr Ile Gly Gly Lys Asn Glu Pro Glu Pro Glu Gln Pro Ile Pro
210         215         220
Arg Lys Val Gln Ile Arg Ser Leu Pro Ser Leu Glu Asp Ile Asp Pro
225         230         235         240
Asp Val Leu Asp Ser Met His Ser Leu Gly Cys Phe Arg Asp Arg Asn
245         250         255

```

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Lys | Leu | Leu | Gln | Asp | Leu | Leu | Ser | Glu | Glu | Glu | Asn | Gln | Glu | Lys | Met |
|     |     |     | 260 |     |     |     |     | 265 |     |     |     |     | 270 |     |     |
| Ile | Tyr | Phe | Leu | Leu | Leu | Asp | Arg | Lys | Glu | Arg | Tyr | Pro | Ser | Gln | Glu |
|     |     | 275 |     |     |     |     | 280 |     |     |     |     | 285 |     |     |     |
| Asp | Glu | Asp | Leu | Pro | Pro | Arg | Asn | Glu | Ile | Asp | Pro | Pro | Arg | Lys | Arg |
|     | 290 |     |     |     |     | 295 |     |     |     |     | 300 |     |     |     |     |
| Val | Asp | Ser | Pro | Met | Leu | Asn | Arg | His | Gly | Lys | Arg | Arg | Pro | Glu | Arg |
| 305 |     |     |     |     | 310 |     |     |     |     | 315 |     |     |     |     | 320 |
| Lys | Ser | Met | Glu | Val | Leu | Ser | Val | Thr | Asp | Gly | Gly | Ser | Pro | Val | Pro |
|     |     |     |     | 325 |     |     |     |     | 330 |     |     |     |     | 335 |     |
| Ala | Arg | Arg | Ala | Ile | Glu | Met | Ala | Gln | His | Gly | Gln | Arg | Ser | Arg | Ser |
|     |     |     | 340 |     |     |     |     | 345 |     |     |     |     | 350 |     |     |
| Ile | Ser | Gly | Ala | Ser | Ser | Gly | Leu | Ser | Thr | Ser | Pro | Leu | Ser | Ser | Pro |
|     |     | 355 |     |     |     |     | 360 |     |     |     |     | 365 |     |     |     |
| Arg | Val | Thr | Pro | His | Pro | Ser | Pro | Arg | Gly | Ser | Pro | Leu | Pro | Thr | Pro |
|     | 370 |     |     |     |     | 375 |     |     |     |     | 380 |     |     |     |     |
| Lys | Gly | Thr | Pro | Val | His | Thr | Pro | Lys | Glu | Ser | Pro | Ala | Gly | Thr | Pro |
| 385 |     |     |     |     | 390 |     |     |     |     | 395 |     |     |     |     | 400 |
| Asn | Pro | Thr | Pro | Pro | Ser | Ser | Pro | Ser | Val | Gly | Gly | Val | Pro | Trp | Arg |
|     |     |     |     | 405 |     |     |     |     | 410 |     |     |     |     | 415 |     |
| Ala | Arg | Leu | Asn | Ser | Ile | Lys | Asn | Ser | Phe | Leu | Gly | Ser | Pro | Arg | Phe |
|     |     |     | 420 |     |     |     |     | 425 |     |     |     |     | 430 |     |     |
| His | Arg | Arg | Lys | Leu | Gln | Val | Pro | Thr | Pro | Glu | Glu | Met | Ser | Asn | Leu |
|     |     | 435 |     |     |     |     | 440 |     |     |     |     | 445 |     |     |     |
| Thr | Pro | Glu | Ser | Ser | Pro | Glu | Leu | Ala | Lys | Lys | Ser | Trp | Phe | Gly | Asn |
|     | 450 |     |     |     |     | 455 |     |     |     |     |     | 460 |     |     |     |
| Phe | Ile | Ser | Leu | Glu | Lys | Glu | Glu | Gln | Ile | Phe | Val | Val | Ile | Lys | Asp |
| 465 |     |     |     |     | 470 |     |     |     |     | 475 |     |     |     |     | 480 |
| Lys | Pro | Leu | Ser | Ser | Ile | Lys | Ala | Asp | Ile | Val | His | Ala | Phe | Leu | Ser |
|     |     |     |     | 485 |     |     |     |     | 490 |     |     |     |     | 495 |     |
| Ile | Pro | Ser | Leu | Ser | His | Ser | Val | Ile | Ser | Gln | Thr | Ser | Phe | Arg | Ala |
|     |     |     | 500 |     |     |     |     | 505 |     |     |     |     | 510 |     |     |
| Glu | Tyr | Lys | Ala | Thr | Gly | Gly | Pro | Ala | Val | Phe | Gln | Lys | Pro | Val | Lys |
|     |     | 515 |     |     |     |     | 520 |     |     |     |     | 525 |     |     |     |
| Phe | Gln | Val | Asp | Ile | Thr | Tyr | Thr | Glu | Gly | Gly | Glu | Ala | Gln | Lys | Glu |
|     |     | 530 |     |     |     | 535 |     |     |     |     | 540 |     |     |     |     |
| Asn | Gly | Ile | Tyr | Ser | Val | Thr | Phe | Thr | Leu | Leu | Ser | Gly | Pro | Ser | Arg |
| 545 |     |     |     |     | 550 |     |     |     |     | 555 |     |     |     |     | 560 |
| Arg | Phe | Lys | Arg | Val | Val | Glu | Thr | Ile | Gln | Ala | Gln | Leu | Leu | Ser | Thr |
|     |     |     |     | 565 |     |     |     |     | 570 |     |     |     |     | 575 |     |
| His | Asp | Pro | Pro | Ala | Ala | Gln | His | Leu | Ser | Asp | Thr | Thr | Asn | Cys | Met |
|     |     |     | 580 |     |     |     |     | 585 |     |     |     |     | 590 |     |     |
| Glu | Met | Met | Thr | Gly | Arg | Leu | Ser | Lys | Cys | Gly | Ile | Ile | Pro | Lys | Ser |
|     |     | 595 |     |     |     |     | 600 |     |     |     |     | 605 |     |     |     |

## WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1.

5

2. An isolated nucleic acid molecule comprising a nucleotide sequence encoding the amino acid sequence shown in SEQ ID NO:2.

10

3. An isolated nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:3.

4. An isolated nucleic acid molecule comprising a nucleotide sequence encoding the amino acid sequence shown in  
15 SEQ ID NO:4.