

(19) **DANMARK**



Patent- og
Varemærkestyrelsen

(10) **DK/EP 1259628 T5**

(12) **Rettet oversættelse af
europæisk patentskrift**

-
- (51) Int.Cl.: **C 12 N 15/81 (2006.01)** **C 12 N 5/10 (2006.01)** **C 12 N 15/01 (2006.01)**
- (45) Oversættelsen bekendtgjort den: **2018-06-14**
- (80) Dato for Den Europæiske Patentmyndigheds bekendtgørelse om meddelelse af patentet: **2007-01-17**
- (86) Europæisk ansøgning nr.: **01911013.9**
- (86) Europæisk indleveringsdag: **2001-02-21**
- (87) Den europæiske ansøgnings publiceringsdag: **2002-11-27**
- (86) International ansøgning nr.: **US2001005447**
- (87) Internationalt publikationsnr.: **WO2001062945**
- (30) Prioritet: **2000-02-23 US 184336 P**
- (84) Designerede stater: **AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE TR**
- (73) Patenthaver: **The Johns Hopkins University, Suite 906, 111 Market Place, Baltimore, MD 21202, USA**
Morphotek, Inc., 210 Welsh Pool Road, Exton, PA 19341, USA
- (72) Opfinder: **GRASSO, Luigi, Apartment 816, 834 Chestnut Street, Philadelphia, PA 19107, USA**
VOGELSTEIN, Bert, 3700 Breton Way, Baltimore, MD 21208, USA
KINZLER, Kenneth, W., 1403 Halkirk Way, BelAir, MD 21015, USA
NICOLAIDES, Nicholas, C., 4 Cider Mill Court, Boothwyn, PA 19061, USA
SASS, Philip, M., 1903 Blackhawk Circle, Audubon, PA 19403, USA
- (74) Fuldmægtig i Danmark: **AWA Denmark A/S, Strandgade 56, 1401 København K, Danmark**
- (54) Benævnelse: **Fremgangsmåde til dannelse af hypermutabel gær**
- (56) Fremdragne publikationer:
CA-A- 2 240 609
NICOLAIDES NICHOLAS C ET AL: "A naturally occurring hPMS2 mutation can confer a dominant negative mutator phenotype" MOLECULAR AND CELLULAR BIOLOGY, AMERICAN SOCIETY FOR MICROBIOLOGY, WASHINGTON, US, vol. 18, no. 3, March 1998 (1998-03), pages 1635-1641, XP002165242 ISSN: 0270-7306 cited in the application
POLACZEK P ET AL: "Functional genetic tests of DNA mismatch repair protein activity in Saccharomyces cerevisiae" GENE: AN INTERNATIONAL JOURNAL ON GENES AND GENOMES, GB, ELSEVIER SCIENCE PUBLISHERS, BARKING, vol. 213, no. 1-2, 15 June 1998 (1998-06-15), pages 159-167, XP004125015 ISSN: 0378-1119
ALEXANDER ARONSHTAM ET AL.: "Dominant negative mutator mutations in the mutL gene of Escherichia coli" NUCLEIC ACIDS RESEARCH, vol. 24, no. 13, 1996, pages 2498-2504, XP002168605 cited in the application
MARILIA CASCALHO ET AL.: "Mismatch repair co-opted by hypermutation" SCIENCE, vol. 279, 20 February 1998 (1998-02-20), pages 1207-1210, XP002168607 DC

OPFINDELSSENS OMRÅDE

Opfindelsen angår området fejlparringsreparationsgener. Den angår især området *in situ*-mutagenese af encellede organismer.

OPFINDELSSENS BAGGRUND

- 5 Inden for de seneste fire år er den genetiske årsag til syndromet hereditær nonpolypøs kolorektal cancer (HNPCC), også kendt som Lynch syndrom II, blevet fastslået for størstedelen af de slægtninge, som lider af sygdommen (Liu, B., Parsons, R., Papadopoulos, N., Nicolaides, N.C., Lynch, H.T., Watson, P., Jass, J.R., Dunlop, M., Wyllie, A., Peltomaki, P., de la Chapelle, A., Hamilton, S.R., Vogelstein, B., og Kinzler, K.W. 1996. Analysis of mismatch repair genes in hereditary non-polyposis colorectal cancer patients. Nat. Med. 2:169-174). Den molekylære basis for HNPCC involverer genetisk ustabilitet, der er et resultat af defekt fejlparringsreparation (MMR). Patentansøgning CA-A-2240609 (The John Hopkins University) beskriver en fremgangsmåde til fremstilling af hypermutable celler eller dyr ved at indføre dominante negative alleler af et fejlparringsreparationsgen. Indtil nu er der blevet identificeret seks gener i mennesker, hvilke gener koder for proteiner og synes at deltage i MMR-processen, herunder *mutS*-homologerne *GTBP*, *hMSH2* og *hMSH3* og *mutL*-homologerne *hMLH1*, *hPMS1* og *hPMS2* (Bronner, C.E., Baker, S.M., Morrison, P.T., Warren, G., Smith, L.G., Lescoe, M.K., Kane, M., Earabino, C., Lipford, J., Lindblom, A., Tannergard, P., Bollag, R.J., Godwin, A., R., Ward, D.C., Nordenskjold, M., Fishel, R., Kolodner, R. og Liskay, R.M. 1994. Mutation in the DNA mismatch repair gene homologue *hMLH1* is associated with hereditary non-polyposis colon cancer. Nature 368:258-261; Fishel, R., Lescoe, M., Rao, M.R.S., Copeland, N.J., Jenkins, N.A., Garber, J., Kane, M. og Kolodner, R. 1993. The human mutator gene homolog *MSH2* and its association with hereditary nonpolyposis colon cancer. Cell 7:1027-1038; Leach, F.S., Nicolaides, N.C., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., Peltomaki, P., Sistonen, P., Aaltonen, L.A., Nystrom-Lahti, M., Guan, X.Y., Zhang, J., Meltzer, P.S., Yu, J.W., Kao, F.T., Chen, D.J., Cerosaletti, K.M., Fournier, R.E.K., Todd, S., Lewis, T., Leach R.J., Naylor, S.L., Weissenbach, J., Mecklin, J.P., Jarvinen, J.A., Petersen, G.M., Hamilton, S.R., Green, J., Jass, J., Watson, P., Lynch, H.T., Trent, J.M., de la Chapelle, A., Kinzler, K.W. og Vogelstein, B. 1993. Mutations of a *mutS* homolog in hereditary non-polyposis colorectal cancer. Cell 75:1215-1225; Nicolaides, N.C., Papadopoulos, N. Liu, B.,
- 10
- 15
- 20
- 25
- 30
- 35

- Wei, Y.F., Carter, K.C., Ruben, S.M., Rosen, C.A., Haseltine, W.A., Fleischmann, R.D., Fraser, C.M., Adams, M.D., Venter, C.J., Dunlop, M.G., Hamilton, S.R., Petersen, G.M., de la Chapelle, A., Vogelstein, B. og Kinzler, K.W. 1994. Mutations of two PMS homologs in hereditary nonpolyposis colon cancer. *Nature* 371:75-80; Nicolaides, N.C., Palombo, F., Kinzler, K.W., Vogelstein, B. og Jiricny, J. 1996. Molecular cloning of the N-terminus of GTBP. *Genomics* 31:395-397; Palombo, F., Hughes, M., Jiricny, J., Truong, O. og Hsuan, J. 1994. Mismatch repair and cancer. *Nature* 36:417; Palombo, F., Gallinari, P., Iaccarino, I., Lettieri, T., Hughes, M.A., Truong, O., Hsuan, J.J. og Jiricny, J.
- 10 1995. GTBP, a 160-kilodalton protein essential for mismatch-binding activity in human cells. *Science* 268:1912-1914; Papadopoulos, N., Nicolaides, N.C., Wei, Y.F., Carter, K.C., Ruben, S.M., Rosen, C.A., Haseltine, W.A., Fleischmann, R.D., Fraser, C.M., Adams, M.D., Venter, C.J., Dunlop, M.G., Hamilton, S.R., Petersen, G.M., de la Chapelle, A., Vogelstein, B. og Kinzler, K.W. 1994.
- 15 Mutation of a mutL homolog is associated with hereditary colon cancer. *Science* 263:1625-1629). Kimlinjemutationer i fire af disse gener (*hMSH2*, *hMLH1*, *hPMS1* og *hPMS2*) er blevet identificeret hos HNPCC-slægtninge (Bronner, C.E., Baker, S.M., Morrison, P.T., Warren, G., Smith, L.G., Lescoe, M.K., Kane, M., Earabino, C., Lipford, J., Lindblom, A., Tannergard, P., Bollag, R.J., Godwin, A., R., Ward, D.C., Nordenskjold, M., Fishel, R., Kolodner, R. og Liskay, R.M. 1994. Mutation in the DNA mismatch repair gene homologue *hMLH1* is associated with hereditary non-polyposis colon cancer. *Nature* 368:258-261; Leach, F.S., Nicolaides, N.C., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., Peltomaki, P., Sistonen, P., Aaltonen, L.A., Nystrom-Lahti, M.,
- 25 Guan, X.Y., Zhang, J., Meltzer, P.S., Yu, J.W., Kao, F.T., Chen, D.J., Cerosaletti, K.M., Fournier, R.E.K., Todd, S., Lewis, T., Leach R.J., Naylor, S.L., Weissenbach, J., Mecklin, J.P., Jarvinen, J.A., Petersen, G.M., Hamilton, S.R., Green, J., Jass, J., Watson, P., Lynch, H.T., Trent, J.M., de la Chapelle, A., Kinzler, K.W. og Vogelstein, B. 1993. Mutations of a *mutS* homolog in hereditary non-polyposis colorectal cancer. *Cell* 75:1215-1225; Liu, B., Parsons, R., Papadopoulos, N., Nicolaides, N.C., Lynch, H.T., Watson, P., Jass, J.R., Dunlop, M., Wyllie, A., Peltomaki, P., de la Chapelle, A., Hamilton, S.R., Vogelstein, B. og Kinzler, K.W. 1996. Analysis of mismatch repair genes in hereditary non-polyposis colorectal cancer patients. *Nat. Med.* 2:169-174; Nicolaides, N.C., Papadopoulos, N., Liu, B., Wei, Y.F., Carter, K.C., Ruben, S.M.,
- 35

Rosen, C.A., Haseltine, W.A., Fleischmann, R.D., Fraser, C.M., Adams, M.D., Venter, C.J., Dunlop, M.G., Hamilton, S.R., Petersen, G.M., de la Chapelle, A., Vogelstein, B. og Kinzler, K.W. 1994. Mutations of two PMS homologs in hereditary nonpolyposis colon cancer. *Nature* 371:75-80; Papadopoulos, N., Nicolaides, N.C., Wei, Y.F., Carter, K.C., Ruben, S.M., Rosen, C.A., Haseltine, W.A., Fleischmann, R.D., Fraser, C.M., Adams, M.D., Venter, C.J., Dunlop, M.G., Hamilton, S.R., Petersen, G.M., de la Chapelle, A., Vogelstein, B. og Kinzler, K.W. 1994. Mutation of a mutL homolog is associated with hereditary colon cancer. *Science* 263:1625-1629). Selvom mutatordefekten, som stammer fra MMR-defekten, kan påvirke en hvilken som helst DNA-sekvens, er mikrosatellitsekvenser særligt følsomme over for MMR-abnormiteter (Modrich, P. 1994. Mismatch repair, genetic stability, and cancer. *Science* 266:1959-1960). Mikrosatellitustabilitet (MI) er derfor en nyttig indikator for defekt MMR. Ud over dens forekomst i praktisk talt alle tumorer, der forekommer hos HNPCC-patienter, findes MI i en lille fraktion af sporadiske tumorer med karakteristiske molekulære og fænotypiske egenskaber (Perucho, M. 1996. Cancer of the microsatellite mutator phenotype. *Biol Chem.* 377:675-684).

HNPCC nedarves på en autosomal dominant måde, således at normale celler af familiemedlemmer, der lider af sygdommen, indeholder én mutantallel af det relevante MMR-gen (arvet fra en forælder, der lider af sygdommen) og én vildtype-allel (arvet fra forælderen, som ikke lider af sygdommen). Under de tidlige stadier af tumorudviklingen inaktiveres vildtype-allelen imidlertid gennem en somatisk mutation, hvilket efterlader cellen uden et funktionelt MMR-gen og resulterer i en udtalt defekt i MMR-aktivitet. Fordi en somatisk mutation, ud over en kimlinjemutation, er nødvendig for dannelse af defekt MMR i tumorcellerne, omtales denne mekanisme sædvanligvis som en mekanisme, der involverer to hits, som er analoge med den biallele inaktivering af tumorsuppressorgener, som initierer andre arvelige cancerformer (Leach, F.S., Nicolaides, N.C., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., Peltomäki, P., Sistonen, P., Aaltonen, L.A., Nystrom-Lahti, M., Guan, X.Y., Zhang, J., Meltzer, P.S., Yu, J.W., Kao, F.T., Chen, D.J., Cerosaletti, K.M., Fournier, R.E.K., Todd, S., Lewis, T., Leach, R.J., Naylor, S.L., Weissenbach, J., Mecklin, J.P., Jarvinen, J.A., Petersen, G.M., Hamilton, S.R., Green, J., Jass, J., Watson, P., Lynch, H.T., Trent, J.M., de la Chapelle, A., Kinzler, K.W. og Vogelstein, B. 1993. Mutations of a *mutS* homolog in hereditary non-polyposis colo-

rectal cancer. Cell 75:1215-1225; Liu, B., Parsons, R., Papadopoulos, N., Nicolaides, N.C., Lynch, H.T., Watson, P., Jass, J.R., Dunlop, M., Wyllie, A., Peltomaki, P., de la Chapelle, A., Hamilton, S.R., Vogelstein, B. og Kinzler, K.W. 1996. Analysis of mismatch repair genes in hereditary non-polyposis colorectal cancer patients. Nat. Med. 2:169-174; Parsons, R., Li, G.M., Longley, M.J., Fang, W.H., Papadopoulos, N., Jen, J., de la Chapelle, A., Kinzler, K.W., Vogelstein, B. og Modrich, P. 1993. Hypermutable and mismatch repair deficiency in RER⁺ tumor cells. Cell 75:1227-1236). På linje med denne to-hits-mekanisme opretholder de ikke-neoplastiske celler af HNPCC-patienter sædvanligvis næsten normale niveauer af MMR-aktivitet på grund af tilstedeværelsen af vildtype-allelen.

Evnen til at ændre signaltransduktionsvejene ved hjælp af manipulation af et genprodukts funktion, enten ved overekspression af vildtype-proteinet eller et fragment deraf eller ved indføring af mutationer i specifikke protein-domæner af proteinet, den såkaldte dominante negative inhiberende mutant, blev over et årti beskrevet i gærsystemet *Saccharomyces cerevisiae* af Herskowitz (Nature 329(6136):219-222, 1987). Det er blevet påvist, at overekspression af vildtype-genprodukter kan resultere i en lignende dominant negativ inhiberende fænotype, mest sandsynligt på grund af "udmætningen" af en faktor, såsom et protein, som er til stede ved lave niveauer og nødvendigt for aktivitet; fjernelse af proteinet ved binding til et højt niveau af dets beslægtede partner resulterer i den samme nettovirkning, hvilket fører til inaktivering af proteinet og den dermed forbundne signaltransduktionsvej. For nylig har arbejde udført af Nicolaides et al. (Nicolaides NC, Littman SJ, Modrich P, Kinzler KW og Vogelstein B. 1998. A naturally occurring hPMS2 mutation can confer a dominant negative mutator phenotype. Mol Cell Biol. 18:1635-1641) vist anvendeligheden af indførelse af dominante negative inhiberende fejlparingsreparationsmutanter i pattedyrceller for at bibringe overordnet DNA-hypermutableitet. Evnen til at manipulere MMR-processen og derfor forøge mutabiliteten af mål værtsgenomet efter ønske, i dette eksempel en pattedyrcelle, muliggør frembringelsen af innovative undertyper af celler eller varianter af de originale vildtype-celler. Disse varianter kan placeres under en specificeret, ønsket selektiv fremgangsmåde, hvilket resulterer i en hidtil ukendt organisme, som udtrykker ét eller flere ændrede biologiske molekyler og har et nyt træk. Begrebet dannelse og indføring af dominante nega-

tive alleler af et gen, herunder MMR-allelerne, i bakterieceller er blevet beskrevet som resulterende i genetisk ændrede prokaryote fejlparringsreparationsgener (Aronshtam A, Marinus MG. 1996. Dominant negative mutator mutations in the mutL gene of *Escherichia coli*. *Nucleic Acids Res* 24:2498-2504; Wu TH, Marinus MG. 1994. Dominant negative mutator mutations in the mutS gene of *Escherichia coli*. *J Bacteriol* 176:5393-400; Brosh RM Jr, Matson SW. 1995. Mutations in motif II of *Escherichia coli* DNA helicase II render the enzyme nonfunctional in both mismatch repair and excision repair with differential effects on the unwinding reaction. *J Bacteriol* 177:5612-5621). Endvidere er ændret MMR-aktivitet blevet påvist, når MMR-gener fra forskellige arter, herunder gær, pattedyrceller og planter, overudtrykkes (Fishel, R., Lescoe, M., Rao, M.R.S., Copeland, N.J., Jenkins, N.A., Garber, J., Kane, M. og Kolodner, R. 1993. The human mutator gene homolog *MSH2* and its association with hereditary nonpolyposis colon cancer. *Cell* 7:1027-1038; Studamire B, Quach T og Alani E. 1998. *Saccharomyces cerevisiae* Msh2p and Msh6p ATPase activities are both required during mismatch repair. *Mol Cell Biol* 18:7590-7601; Alani E, Sokolsky T, Studamire B, Miret JJ og Lahue RS. 1997. Genetic and biochemical analysis of Msh2p-Msh6p: role of ATP hydrolysis and Msh2p-Msh6p subunit interactions in mismatch base pair recognition. *Mol Cell Biol* 17:2436-2447; Lipkin SM, Wang V, Jacoby R, Banerjee-Basu S, Baxevanis AD, Lynch HT, Elliott RM og Collins FS. 2000. *MLH3*: a DNA mismatch repair gene associated with mammalian microsatellite instability. *Nat. Genet.* 24:27-35).

Der er et fortsat behov inden for teknikken for fremgangsmåder til genmanipulation af nyttige gærstammer for at øge deres adfærds karakteristika og -evner.

KORT BESKRIVELSE AF OPFINDELSEN

Det er et formål med den foreliggende opfindelse at tilvejebringe en fremgangsmåde til at gøre gærceller hypermutable.

Det er et andet formål med den foreliggende opfindelse at tilvejebringe hypermutable gærceller.

Det er et yderligere formål med opfindelsen at tilvejebringe en fremgangsmåde til mutation af et gen af interesse i en gær.

Det er endnu et yderligere formål med den foreliggende opfindelse at tilvejebringe en fremgangsmåde til frembringelse af gærtyper, som er hyper-

mutable.

Det er et formål med opfindelsen at tilvejebringe en fremgangsmåde til genoprettelse af normal fejlparringsreparationsaktivitet i hypermutable celler efter selektion af stamme.

- 5 Disse og andre formål med opfindelsen tilvejebringes ved hjælp af én eller flere af de følgende udførelsesformer. I én udførelsesform tilvejebringes en fremgangsmåde til dannelse af en hypermutabel gær. Et polynukleotid, der omfatter en dominant negativ allel af et fejlparringsreparationsgen, hvilken allel reguleres inducerbart, således at ekspression af den dominante negative
10 allel kan svækkes eller elimineres, indføres i en gær-celle. Cellen bliver således hypermutabel.

- I overensstemmelse med en anden udførelsesform tilvejebringes en homogen sammensætning af dyrkede, hypermutable gær-celler. Gær-cellerne omfatter en dominant negativ allel af et fejlparringsreparationsgen, hvilken allel reguleres inducerbart, således at ekspression af den dominante negative
15 allel kan svækkes eller elimineres.

- I overensstemmelse med endnu en anden udførelsesform ifølge opfindelsen tilvejebringes en fremgangsmåde til dannelse af en mutation i et gen af interesse. Der dyrkes en gær-cellekultur, der omfatter genet af interesse og en
20 dominant negativ allel af et fejlparringsreparationsgen, hvilken allel reguleres inducerbart, således at ekspression af den dominante negative allel kan svækkes eller elimineres. Gær-cellen er hypermutabel. Celler af kulturen undersøges for at bestemme, hvorvidt genet af interesse indeholder en mutation.

- 25 I endnu en anden udførelsesform ifølge opfindelsen tilvejebringes en fremgangsmåde til dannelse af en mutation i et gen af interesse. En gær-celle, der omfatter genet af interesse og et polynukleotid, der koder for en dominant negativ allel af et fejlparringsreparationsgen, hvilken allel reguleres inducerbart, således at ekspressionen af den dominante negative allel kan svækkes
30 eller elimineres, dyrkes for at danne en population af muterede, hypermutable gær-celler. Populationen af muterede, hypermutable gær-celler dyrkes under betingelser, hvor træk selekteres. Gær-celler, som vokser under betingelser, hvor træk selekteres, undersøges for at bestemme, hvorvidt genet af interesse indeholder en mutation.

- 35 Ved hjælp af den foreliggende opfindelse tilvejebringes der også en

fremgangsmåde til dannelse af forøget hypermutabel gær. En gærcelle eksponeres for et mutagen. Gærcellen er defekt i fejlparringsreparation (MMR) på grund af tilstedeværelsen af en dominant negativ allel af mindst ét MMR-gen, hvilken allel reguleres inducerbart, således at ekspresion af den dominante negative allel kan svækkes eller elimineres. Der opnås en forøget mutationsrate af gærcellen på grund af eksponering for mutagenet.

Ifølge endnu et andet aspekt af opfindelsen tilvejebringes en fremgangsmåde til dannelse af fejlparringsreparations (MMR)-kompetent gær med nye træk. En gærcelle, der omfatter et gen af interesse og et polynukleotid, som koder for en dominant negativ allel af et fejlparringsreparationsgen, hvilken allel reguleres inducerbart, således at ekspresion af den dominante negative allel kan svækkes eller elimineres, dyrkes for at danne en population af muterede, hypermutable gærceller. Populationen af muterede, hypermutable gærceller dyrkes under betingelser, hvor træk selekteres. Gærcellerne, som vokser under betingelser, hvor træk selekteres, undersøges for at bestemme, hvorvidt genet af interesse indeholder en mutation. Normal fejlparringsreparationsaktivitet genoprettes i gærcellerne.

Disse og andre udførelsesformer ifølge opfindelsen giver teknikken fremgangsmåder, som kan danne forøget mutabilitet i gær samt tilvejebringe encellede, eukaryote organismer, der indeholder potentielt nyttige mutationer, til dannelse af hidtil ukendte træk til kommerciel anvendelse.

DETALJERET BESKRIVELSE AF OPFINDELSEN

Det er en opdagelse ifølge den foreliggende opfindelse, at hypermutabel gær kan fremstilles ved at ændre aktiviteten af endogen fejlparringsreparationsaktivitet af værtsceller. Når dominante negative alleler af fejlparringsreparationsgener indføres og udtrykkes i gær, forøger de raten af spontane mutationer ved at reducere effektiviteten af endogen fejlparringsreparationsmediert DNA-reparationsaktivitet, og gør derved gæren særdeles modtagelig over for genetiske ændringer, dvs. hypermutabel. Hypermutabel gær kan derefter anvendes til at screene for mutationer i et gen eller et sæt gener i varierende datterceller, som udviser ét eller flere træk, der ikke findes i vildtype-cellerne.

Fremgangsmåden til fejlparringsreparation, også kaldet fejlparringskorrektur, er en udviklingsmæssigt højt konserveret fremgangsmåde, som udføres af proteinkomplekser beskrevet i celler så forskelligartede som prokaryote celler, såsom bakterier, til mere komplekse pattedyrceller (Modrich, P. 1994.

Mismatch repair, genetic stability, and cancer. *Science* 266:1959-1960; Parsons, R., Li, G.M., Longley, M., Modrich, P., Liu, B., Berk, T., Hamilton, S.R., Kinzler, K.W. og Vogelstein, B. 1995. Mismatch repair deficiency in phenotypically normal human cells. *Science* 268:738-740; Perucho, M. 1996. Cancer of the microsatellite mutator phenotype. *Biol Chem.* 377:675-684). Et fejlparringsreparationsgen er et gen, som koder for ét af proteinerne af et sådant fejlparringsreparationskompleks. Selvom det ikke ønskes at være bundet af nogen bestemt teori om virkningsmekanisme, menes et fejlparringsreparationskompleks at påvise forvridninger af DNA-helixen, hvilke forvridninger stammer fra ikke-komplementær parring af nukleotidbaser. Den ikke-komplementære base på den nyere DNA-streng skæres ud, og den udskårne base erstattes med den hensigtsmæssige base, der er komplementær med den ældre DNA-streng. På denne måde eliminerer celler mange mutationer, som forekommer som et resultat af fejl i DNA-replikation, hvilket resulterer i genetisk stabilitet af dattercellerne afledt af modercellen.

Nogle vildtype-alleler samt dominante negative alleler forårsager en fænotype med defekt fejlparringsreparation, selv i nærvær af en vildtype-allel i den samme celle. Et eksempel på en dominant negativ allel af et fejlparringsreparationsgen er det humane gen *hPMS2-134*, som bærer en trunkeringsmutation på kodon 134 (Parsons, R., Li, G.M., Longley, M., Modrich, P., Liu, B., Berk, T., Hamilton, S.R., Kinzler, K.W. og Vogelstein, B. 1995. Mismatch repair deficiency in phenotypically normal human cells. *Science* 268:738-740; Nicolaides NC, Littman SJ, Modrich P, Kinzler KW og Vogelstein B. 1998. A naturally occurring hPMS2 mutation can confer a dominant negative mutator phenotype. *Mol Cell Biol* 18:1635-1641). Mutationen får dette gens produkt til helt abnormt at terminere på stillingen for den 134. aminosyre, hvilket resulterer i et forkortet polypeptid, der indeholder de N-terminale 133 aminosyrer. En sådan mutation forårsager en forøgelse i raten af mutationer, hvilke mutationer akkumulerer i celler efter DNA-replikation. Ekspression af en dominant negativ allel af et fejlparringsreparationsgen resulterer i svækkelse af fejlparringsreparationsaktivitet, selv i nærvær af vildtype-allelen. En hvilken som helst fejlparringsreparationsallel, som frembringer en sådan virkning, kan anvendes i denne opfindelse, uanset om det er vildtype-allel eller en ændret allel, og uanset om den stammer fra pattedyr, gær, svampe, amfibier, insekter, planter eller bakterier. Anvendelsen af overudtrykte vildtype-MMR-genalleler

fra mennesker, mus, planter og gær i bakterier har derudover vist sig at forårsage en dominant negativ virkning på MMR-aktiviteten i bakterieværterne (Aronshtam A, Marinus MG. 1996. Dominant negative mutator mutations in the mutL gene of *Escherichia coli*. Nucleic Acids Res 24:2498-2504; Wu TH og Marinus MG. 1994. Dominant negative mutator mutations in the mutS gene of *Escherichia coli*. J Bacteriol 176:5393-400; Brosh RM Jr og Matson SW. 1995. Mutations in motif II of *Escherichia coli* DNA helicase II render the enzyme nonfunctional in both mismatch repair and excision repair with differential effects on the unwinding reaction. J Bacteriol 177:5612-5621; Lipkin SM, Wang V, Jacoby R, Banerjee-Basu S, Baxevanis AD, Lynch HT, Elliott RM og Collins FS. 2000. MLH3: a DNA mismatch repair gene associated with mammalian microsatellite instability. Nat Genet 24:27-35). Dette antyder, at perturbering af MMR-proteinkomplekset med flere komponenter kan opnås ved indføring af MMR-komponenter fra andre arter i gær.

15 Dominante negative alleler af et fejlparringsreparationsgen kan opnås fra cellerne af mennesker, dyr, gær, bakterier, planter eller andre organismer. Screeningsceller til defekt fejlparringsreparationsaktivitet kan identificere sådanne alleler. Fejlparringsreparationsgener kan være mutante eller vildtype. MMR-gærværter kan eventuelt muteres. Udtrykket gær anvendt i denne an-

20 søgning omfatter en hvilken som helst organisme fra eukaryot-riget, herunder, men ikke begrænset til, *Saccharomyces* sp., *Pichia* sp., *Schizosaccharomyces* sp., *Kluyveromyces* sp. og andre svampe (Gellissen, G. og Hollenberg, CP. Gen 190(1):87-97, 1997). Disse organismer kan eksponeres for f.eks. kemiske mutagener eller stråling og kan screenes for defekt fejlparringsrepa-

25 ration. Genomisk DNA, cDNA, mRNA eller protein fra en hvilken som helst celle, der koder for et fejlparringsreparationsprotein, kan analyseres for variationer af vildtype-sekvensen. Dominante negative alleler af et fejlparringsreparationsgen kan også dannes kunstigt, for eksempel ved at fremstille varianter af *hPMS2-134*-allelen eller andre fejlparringsreparationsgener (Nicolaidis NC, Littman SJ, Modrich P, Kinzler KW og Vogelstein B. 1998. A naturally occurring *hPMS2* mutation can confer a dominant negative mutator phenotype. Mol Cell Biol 18:1635-1641). Forskellige teknikker til stedorienteret mutagenese kan anvendes. Egnetheden af sådanne alleler, uanset om de er naturlige eller kunstige, til anvendelse i dannelsen af hypermutabel gær kan vurderes ved at

35 undersøge fejlparringsreparationsaktiviteten (under anvendelse af frem-

gangsmåder beskrevet i Nicolaides NC, Littman SJ, Modrich P, Kinzler KW og Vogelstein B. 1998. A naturally occurring hPMS2 mutation can confer a dominant negative mutator phenotype. *Mol Cell Biol* 18:1635-1641), som forårsages af allelen i nærvær af én eller flere vildtype-alleler, for at bestemme, 5 hvorvidt det er en dominant negativ allel.

En gær, som overudtrykker en vildtype-fejlparringsreparationsallel eller en dominant negativ allel af et fejlparringsreparationsgen, vil blive hypermutabel. Dette betyder, at den spontane mutationsrate af en sådan gær forøges sammenlignet med gær uden sådanne alleler. Graden af forøgelse af den 10 spontane mutationsrate kan være mindst 2, 5, 10, 20, 50, 100, 200, 500 eller 1000 gange højere end den for den normale gær, som målt som en funktion af gærfordobling/time.

Ifølge ét aspekt af opfindelsen indføres et polynukleotid, der koder for enten en vildtype-form eller en dominant negativ form af et fejlparringsreparationsprotein, i gær, hvor den dominante negative allel reguleres inducerbart, således at ekspression af den dominante negative allel kan svækkes eller elimineres. Genet kan være en hvilken som helst dominant negativ allel, 15 der koder for et protein, som er del af et fejlparringsreparationskompleks, f.eks. *mutS*-, *mutL*-, *mutH*- eller *mutY*-homologer af bakterie-, gær-, plante- eller pattedyrgener (Modrich, P. 1994. Mismatch repair, genetic stability, and cancer. *Science* 266:1959-1960; Prolla, T.A, Pang, Q., Alani, E., Kolodner, R.A og Liskay, R.M. 1994. MLH1, PMS1, and MSH2 Interaction during the initiation of DNA mismatch repair in yeast. *Science* 264:1091-1093). Den dominante negative allel kan være naturligt forekommende eller fremstillet i laboratoriet. 25 Polynukleotidet kan være i form af genomisk DNA, cDNA, RNA eller et kemisk syntetiseret polynukleotid eller polypeptid. Molekylet kan indføres i cellen ved transformation, elektroporering, parring, partikelbombardement eller andre fremgangsmåder beskrevet i litteraturen.

Transformation anvendes heri som en hvilken som helst fremgangsmåde, 30 de, hvorved et polynukleotid eller polypeptid indføres i en celle. Fremgangsmåden til transformation kan udføres i en gærkultur under anvendelse af en suspension af celler. Gæren kan være en hvilken som helst type, der er klassificeret under eukaryot-riget, i overensstemmelse med international konvention.

35 I almindelighed vil transformation udføres under anvendelse af en su-

- spension af celler, men andre fremgangsmåder kan også anvendes, når blot en tilstrækkelig fraktion af de behandlede celler inkorporerer polynukleotidet eller polypeptidet for at muliggøre vækst og anvendelse af transficerede celler. Proteinproduktet af polynukleotidet kan udtrykkes transient eller stabilt i
- 5 cellen. Teknikker til transformation er velkendte for fagmanden. Tilgængelige teknikker til indføring af et polynukleotid eller polypeptid i en gærcelle indbefatter, men er ikke begrænset til, elektroporering, virustransduktion, cellefusion, anvendelse af sfæroplaster eller kemisk kompetente celler (f.eks. calciumchlorid) og pakning af polynukleotidet sammen med lipid til fusion med
- 10 cellerne af interesse. Når en celle først er blevet omdannet med fejlparringsreparationsgenet eller -proteinet, kan cellen opformeres og manipuleres enten i flydende kultur eller på en fast agarmatrix, såsom en petriskål. Hvis den transficerede celle er stabil, vil genet blive udtrykt på et ensartet niveau i mange cellegenerationer, og en stabil, hypermutabel gærstamme følger.
- 15 En isoleret gærcelle kan opnås ud fra en gærkultur ved kemisk selektion af stammer under anvendelse af antibiotisk selektion af en ekspressionsvektor. Hvis gærcellen er afledt af en enkeltcelle, defineres den som en klon. Teknikker til enkeltcellekloning af mikroorganismer, såsom gær, er velkendte inden for teknikken.
- 20 Et polynukleotid, der koder for en dominant negativ form af et fejlparringsreparationsprotein, kan indføres i genomet af gær eller opformeres på et ekstra-kromosomalt plasmid, såsom 2-mikrometerplasmidet. Selektion af kloner, der indeholder en ekspressionsvektor for fejlparringsreparationsgenet, kan opnås ved at udplade celler på syntetisk, komplet medium uden den hen-
- 25 sigtsmæssige aminosyre eller andre væsentlige næringsstoffer, som beskrevet (J. C. Schneider og L. Guarente, *Methods in Enzymology* 194:373, 1991). Gæren kan være en hvilken som helst art, for hvilken egnede teknikker er tilgængelige, til fremstilling af transgene mikroorganismer, såsom, men ikke begrænset til, slægter, herunder *Saccharomyces*, *Schizosaccharomyces*, *Pi-*
- 30 *chia*, *Hansenula*, *Kluyveromyces* og andre.
- En hvilken som helst fremgangsmåde til fremstilling af transgen gær, hvilken fremgangsmåde er kendt inden for inden for teknikken, kan anvendes. Ifølge én fremgangsmåde til fremstilling af en transgen mikroorganisme indføres polynukleotidet i gæren ved hjælp af én af fremgangsmåderne, der er vel-
- 35 kendte for fagmanden. Derefter dyrkes gærkulturen under betingelser, som

selekterer for celler, hvor polynukleotidet, der koder for fejlparringsreparationsgenet, enten inkorporeres i værtsgenomet som en stabil enhed eller opformeres på et selvreplikerende ekstra-kromosomalt plasmid, og proteinet, der kodes af polynukleotidfragmentet, transkriberes og efterfølgende transla-
5 teres til et funktionelt protein i cellen. Når transgen gær først er manipuleret til at indeholde ekspressionskonstruktionen, opformeres den derefter for at danne og bibeholde en kultur af transgen gær på ubestemt tid.

Når en stabil, transgen gær celle først er blevet tilvejebragt til at udtrykke et defekt fejlparrings (MMR)-protein, kan gæren dyrkes for at danne hidtil
10 ukendte mutationer i ét eller flere målgener af interesse, hvilke målgener er indeholdt i samme gær celle. Et gen af interesse kan være et hvilket som helst gen, der naturligt er indeholdt i gæren, eller ét gen, der er indført i gærværten ved hjælp af rekombinant DNA-standardteknikker. Målgenerne kan eventuelt være kendte inden selektionen. Én fordel ved at anvende sådanne trans-
15 gene gær celler til frembringelse af mutationer i iboende eller ekstra-kromosomale gener i gæren er, at det er unødvendigt at eksponere cellerne for mutagen skade, uanset om det er kemisk stof eller stråling, for at frembringe en række vilkårlige genændringer i målgenerne. Dette skyldes den særdeles effektive beskaffenhed og spektret af naturligt forekommende muta-
20 tioner, som opstår som følge af den ændrede fremgangsmåde for fejlparrings-reparation. Det er imidlertid muligt at forøge spektret og frekvensen af mutationer ved hjælp af samtidig anvendelse af enten kemisk stof og/eller stråling sammen med MMR-defekte celler. Nettovirkningen af kombinationsbehandlingen er en forøgelse i mutationsraten i de genetisk ændrede gærtyper, der er
25 nyttige til at frembringe nye træk. Raten af kombinationsbehandlingen er højere end raten, hvor der kun anvendes MMR-defekte celler eller kun mutagenet med vildtype-MMR-celler.

MMR-defekt gær ifølge opfindelsen kan anvendes i genetiske screeninger til den direkte selektion af subklonvarianter, der udviser nye træk med kom-
30 mercielt ønskelige anvendelser. Dette muliggør, at de langsommelige og tidskrævende trin i genidentifikation, -isolering og -karakterisering undgås.

Mutationer kan påvises ved at undersøge den internt og/eller eksternt mutageniserede gær for ændringer i dens genotype og/eller fænotype. Gener, der frembringer ændrede fænotyper i MMR-defekte mikrobielle celler, kan ses
35 ved hjælp af en hvilken som helst af en række molekylære teknikker, der er

velkendte for fagmanden. Gærgenomet kan f.eks. isoleres, og et bibliotek af restriktionsfragmenter af gærgenomet kan klones i en plasmidvektor. Biblioteket kan indføres i en "normal" celle, og cellerne, som udviser den hidtil ukendte fænotype, kan screenes. Et plasmid kan isoleres fra de normale celler, der udviser den hidtil ukendte fænotype, og genet/generne kan karakteriseres ved DNA-sekvensanalyse. Der kan alternativt anvendes differentiell messenger-RNA-screening, hvori der anvendes driver- og tester-RNA (afledt af henholdsvis vildtype-mutant og hidtil ukendt mutant), efterfulgt af kloning af de differentielle transkripter og karakterisering deraf ved standardfremgangsmåder for molekylærbiologi, hvilke fremgangsmåder er velkendte for fagmanden. Hvis den søgte mutant endvidere kodes af et ekstra-kromosomalt plasmid, kan plasmidet, efter samtidig ekspression af det dominante negative MMR-gen og genet af interesse, og efter fænotypisk selektion, isoleres fra mutante kloner og analyseres ved hjælp af DNA-sekvensanalyse under anvendelse af fremgangsmåder, der er velkendte for fagmanden. Fænotypisk screening for træk i MMR-defekte mutanter kan ske ved hjælp af biokemisk aktivitet og/eller en let observerbar fænotype af det ændrede genprodukt. En mutant fænotype kan også detekteres ved at identificere ændringer i elektroforetisk mobilitet, DNA-binding i tilfælde af transkriptionsfaktorer, spektroskopiske egenskaber, såsom IR, CD, røntgenkrystallografi eller højfelts-NMR-analyse eller andre fysiske eller strukturelle karakteristika af et protein, der er kodet af et mutant gen. Det er også muligt at screene for ændret, hidtil ukendt funktion af et protein *in situ* i isoleret form eller i modelsystemer. Der kan screenes for ændring af en hvilken som helst egenskab af gæren, hvilken egenskab er forbundet med funktionen af genet af interesse, uanset om genet er kendt inden selektion eller ukendt.

Det er meningen, at de beskrevne fremgangsmåder for screening og selektion skal illustrere potentielle fremgangsmåder til tilvejebringelse af hidtil ukendte mutanter med kommercielt værdifulde træk, men det er ikke meningen, at de skal begrænse de mange mulige måder, hvorpå screening og selektion kan udføres af fagmanden.

Plasmidekspressionsvektorer, som indeholder et fejlparingsreparations (MMR)-geninsert, kan anvendes i kombination med et antal kommercielt tilgængelige, regulatoriske sekvenser til regulering af både det temporale og kvantitative biokemiske ekspressionsniveau af det dominante negative

MMR-protein. De regulatoriske sekvenser kan bestå af en promotor, en enhancer eller en promotor/enhancer-kombination og kan indsættes enten opstrøms eller nedstrøms for MMR-genet for at regulere ekspressionsniveauet. De regulatoriske sekvenser kan være en hvilken som helst regulatorisk sekvens, som er velkendt for fagmanden, herunder, men ikke begrænset til, AOX1-, GAP-, GAL1-, GAL10-, PHO5- og PGK-promotorer, der er indeholdt i ekstra-kromosomale ekspressionsvektorer med højt eller lavt kopital eller på konstruktioner, som er integreret i genomet via homolog rekombination. Disse typer af regulatoriske systemer er blevet beskrevet i videnskabelige publikationer og er kendte for fagmanden.

Når en mikroorganisme med et hidtil ukendt, ønsket træk af interesse først er dannet, svækkes eller elimineres aktiviteten af den afvigende MMR-aktivitet ønskeligt ved hjælp af en hvilken som helst fremgangsmåde, der er kendt inden for teknikken. Disse fremgangsmåder indbefatter, men er ikke begrænset til, fjernelse af en inducer fra dyrkningsmediet, hvilken inducer er ansvarlig for promotoraktivering, fjernelse af et plasmid fra en transformeret gær-celle og tilsætning af kemikalier, såsom 5-fluororotsyre, med henblik på et "loop-out" af genet af interesse.

I tilfælde af en inducerbart reguleret, dominant negativ MMR-allel vil ekspressionen af det dominante negative MMR-gen blive tændt (induceret) for at danne en population af hypermutable gærceller med nye træk. Ekspression af den dominante negative MMR-allel kan hurtigt slukkes for at rekonstituere en genetisk stabil stamme, som udviser et nyt træk af kommerciel interesse. Den resulterende gærstamme er nu nyttig som en stabil stamme, der kan anvendes til forskellige kommercielle anvendelser, afhængig af den selektionsfremgangsmåde, som den udsættes for.

I tilfælde hvor genetisk defekte fejlparingsreparationsgærtypen [stammer, såsom, men ikke begrænset til: M1 (mutS) og i EC2416 (mutS delta umuDC) og mutL- eller mutY-stammer] anvendes til at aflede nye træk, kan der anvendes transgene konstruktioner, som udtrykker vildtypefejlparingsreparationsgener, der er tilstrækkelige til at komplementere den genetiske defekt og derfor genoprette fejlparingsreparationsaktivitet af værten efter selektion af træk [Grzesiuk, E. et al. (Mutagenesis 13;127-132, 1998); Bridges, B.A., et al. (EMBO J. 16:3349-3356, 1997); LeClerc, J.E., Science 15:1208-1211, 1996); Jaworski, A. et al. (Proc. Natl. Acad. Sci USA

92:11019-11023, 1995)). Den resulterende gær er genetisk stabil og kan anvendes til forskellige kommercielle anvendelser.

Det er tidligere blevet påvist, at anvendelsen af overekspression af fremmede (eksogene, transgene) fejlparingsreparationsgener fra mennesker og gær, såsom MSH2, MLH1, MLH3, etc., frembringer en dominant negativ mutatorfænotype i gærværter (Shcherbakova, P.V., Hall, M.C., Lewis, M.S., Bennett, S.E., Martin, K.J., Bushel, P.R., Afshari, C.A. og Kunkel, T.A. Mol. Cell Biol. 21(3):940-951; Studamire B, Quach T og Alani E. 1998. Saccharomyces cerevisiae Msh2p and Msh6p ATPase activities are both required during mismatch repair. Mol Cell Biol 18:7590-7601; Alani E, Sokolsky T, Studamire B, Miret JJ og Lahue RS. 1997. Genetic and biochemical analysis of Msh2p-Msh6p: role of ATP hydrolysis and Msh2p-Msh6p subunit interactions in mismatch base pair recognition. Mol Cell Biol 17:2436-2447; Lipkin SM, Wang V, Jacoby R, Banerjee-Basu S, Baxevanis AD, Lynch HT, Elliott RM og Collins FS. 2000. MLH3: a DNA mismatch repair gene associated with mammalian microsatellite instability. Nat Genet 24:27-35). Endvidere har anvendelsen af gærstammer, der udtrykker prokaryote, dominante negative MMR-gener, samt værter, som har genomiske defekter i endogene MMR-proteiner, også tidligere vist sig at resultere i en dominant negativ mutatorfænotype (Evans, E., Sugawara, N., Haber, J.E. og Alani, E. Mol. Cell. 5(5):789-799, 2000; Aronshtam A og Marinus MG. 1996. Dominant negative mutator mutations in the mutL gene of Escherichia coli. Nucleic Acids Res 24:2498-2504; Wu TH og Marinus MG. 1994. Dominant negative mutator mutations in the mutS gene of Escherichia coli. J Bacteriol 176:5393-400; Brosh RM Jr og Matson SW. 1995. Mutations in motif II of Escherichia coli DNA helicase II render the enzyme nonfunctional in both mismatch repair and excision repair with differential effects on the unwinding reaction. J Bacteriol 177:5612-5621). Resultaterne beskrevet heri beskriver imidlertid anvendelsen af MMR-gener, herunder det humane PMSR2-gen (Nicolaidis, N.C., Carter, K.C., Shell, B.K., Papadopoulos, N., Vogelstein, B. og Kinzler, K.W. 1995. Genomic organization of the human PMS2 gene family. Genomics 30:195-206), det beslægtede PMS134-trunkerede MMR-gen (Nicolaidis N.C., Kinzler, K.W. og Vogelstein, B. 1995. Analysis of the 5' region of PMS2 reveals heterogenous transcripts and a novel overlapping gene. Genomics 29:329-334), de vegetabiliske fejlparingsreparationsgener (US-patentansøgning S.N. 09/749,601) og de gener,

som er homologe med de 134 N-terminale aminosyrer af PMS2-genet, til dannelse af hypermutabel gær.

DNA-mutagener kan anvendes sammen med MMR-defekte gærværter til forøgelse af den hypermutable produktion af genetiske ændringer. Dette reducerer endvidere MMR-aktivitet og er nyttigt til dannelse af mikroorganismer med kommercielt relevante træk.

Evnen til at danne hypermutable organismer under anvendelse af dominante negative alleler kan anvendes til at danne innovative gærstammer, som udviser nye træk, der er nyttige til en række anvendelser, herunder, men ikke begrænset til, industrien, til dannelse af nye biokemikalier, til afgiftning af skadelige kemikalier, enten biprodukter af fremstillingsprocesser eller dem, som anvendes som katalysatorer, samt til hjælp i remediering af toksiner, der er til stede i miljøet, herunder, men ikke begrænset til, polychlorbenzener (PCB'er), tungmetaller og andre miljømæssige farer. Hidtil ukendte gærstammer kan selekteres for forøget aktivitet for at frembringe enten øget mængde eller kvalitet af et terapeutisk molekyle med eller uden protein ved hjælp af biotransformation. Biotransformation er den enzymatiske omdannelse af ét kemisk mellemprodukt til det næste mellemprodukt eller produkt i en vej eller et skema ved hjælp af en mikrobe eller et ekstrakt, der er afledt af mikroben. Der er mange eksempler på biotransformation, der anvendes til kommerciel fremstilling af vigtige biologiske og kemiske produkter, herunder penicillin G, erythromycin og clavulansyre. Organismer, som er effektive ved omdannelse af "rå" materialer til avancerede mellemprodukter og/eller slutprodukter, kan også udføre biotransformation (Berry, A. Trends Biotechnol. 14(7):250-256). Evnen til at regulere DNA-hypermutabilitet i værtsgærstammer under anvendelse af en dominant negativ MMR (som beskrevet ovenfor) muliggør dannelsen af varierende undertyper, som kan selekteres for nye fænotyper af kommerciel interesse, herunder, men ikke begrænset til, organismer som er toksinresistente, som har evnen til nedbryde et toksin *in situ*, eller som har evnen til at omdanne et molekyle fra et mellemprodukt til enten et bearbejdet mellemprodukt eller et slutprodukt. Andre anvendelser, hvori der anvendes dominante negative MMR-gener til frembringelse af genetisk ændring af gærværter med henblik på opnåelse af nye træk, indbefatter, men er ikke begrænset til, rekombinante produktionsstammer, som frembringer større mængder af et rekombinant polypeptid, samt anvendelsen af ændrede

endogene gener, som kemisk kan omdanne eller katalysere nedstrøms-fremstillingsprocesser. En dominant negativ MMR-fænotype, som er regulerbar, kan anvendes til at fremstille en gærstamme med et kommercielt fordelagtigt træk. Under anvendelse af denne fremgangsmåde kan encellede gær-

5 celler, der udtrykker en dominant negativ MMR, selekteres direkte for fænotypen af interesse. Når en valgt gær med et specificeret træk først er isoleret, kan den hypermutable aktivitet af den dominante negative MMR-allel slukkes ved hjælp af adskillige fremgangsmåder, der er velkendte for fagmanden. Hvis for eksempel den dominante negative allel udtrykkes af et inducerbart

10 promotorsystem, kan induceren fjernes eller udtømmes. Sådanne systemer indbefatter, men er ikke begrænset til, promotorer, såsom: lactoseinducerbar GALi-GAL10-promotor (M. Johnston og R.W. Davis, *Mol. Cell Biol.* 4:1440, 1984); den fosfatinducerbare PHO5-promotor (A. Miyahara, A. Toh-e, C. Nosaki, F. Nosaki, F. Hamada, N. Ohtomo og K. Matsubara. *Proc. Natl. Acad.*

15 *Sci. U.S.A.* 80:1, 1983); alkoholdehydrogenase I (ADH)- og 3-phosphoglyceratkinase (PGK)-promotorer, som anses for at være konstitutive, men som kan undertrykkes/fremmes, når gærceller dyrkes i ikke-fermenteringsbare carbonkilder, såsom, men ikke begrænset til, lactat (G. Ammerer, *Methods in Enzymology* 194:192, 1991; J. Mellor, M.J. Dobson,

20 N.A. Roberts, M.F. Tuite, J.S. Emtage, S. White, D.A. Lowe, T. Patel, A.J. Kingsman og S.M. Kingsman, *Gene* 24:563, 1982); S. Hahn og L. Guarente, *Science* 240:317, 1988); Alkoholoxidase (AOX), i *Pichia pastoris* (Tschopp, JF, Brust, Pf, Cregg, JM, Stillman, CA og Gingeras, TR. *Nucleic Acids Res.* 15(9):3859-76, 1987; og den thiaminrepresserbare ekspressionspromotor

25 *mt1* i *Schizosaccharomyces pombe* (Moreno, MB, Duran, A. og Ribas, JC. *Yeast* 16(9):861-72, 2000). Gærceller kan omdannes ved hjælp af en hvilken som helst fremgangsmåde, der er kendt for fagmanden, herunder kemisk transformation med LiCl (Mount, R.C., Jordan, B.E. og Hadfield, C. *Methods Mol. Biol.* 53:139-145, 1996) og elektroporering (Thompson, JR, Register, E.,

30 Curotto, J., Kurtz, M. og Kelly, R. *Yeast* 14(6):565-71, 1998). Gærceller, der er blevet omdannet med DNA, kan selekteres for vækst ved hjælp af en række fremgangsmåder, herunder, men ikke begrænset til, selekterbare markører (URA3; Rose, M., Grisafi, P. og Botstein, D. *Gene* 29:113, 1984; LEU2; A. Andreadis, Y., Hsu, M., Hermodson, G., Kohlhaw og P. Schimmel. *J. Biol. Chem.*

35 259:8059, 1984; ARG4; G. Tschumper og J. Carbon. *Gene* 10:157, 1980; og

HIS3; K. Struhl, D.T. Stinchcomb, S., Scherer og R.W. Davis. Proc. Natl. Acad. Sci. U.S.A. 76:1035, 1979), og lægemidler, som inhiberer vækst af gærceller (tunicamycin, TUN; S. Hahn, J., Pinkham, R. Wei, R., Miller og L. Guarente. Mol. Cell Biol. 8:655, 1988). Rekombinant DNA kan indføres i gær, som beskrevet ovenfor, og gærvektorerne kan være indeholdt i gærcellen, enten ekstra-kromosomalt eller integreret i et specifikt locus. Ekstra-kromosomalt baserede gærekspressionsvektorer kan enten være centromeriske vektorer med højt kopital (såsom 2- μ m-vektoren Yep13; A.B. Rose og J.R. Broach, Methods in Enzymology 185:234, 1991), centromeriske vektorer med lavt kopital, hvilke vektorer indeholder autonomt replikerende sekvenser (ARS), såsom YRp7 (M. Fitzgerald-Hayes, L. Clarke og J. Carbon, Cell 29:235, 1982), samt integrationsvektorer, der muliggør, at genet af interesse indføres i et specificeret locus i værtsgenomet og opformeres på en stabil måde (R.J. Rothstein, Methods in Enzymology 101:202, 1991). Ektopisk ekspres-
 15 MMR-gener i gær kan svækkes eller elimineres fuldstændigt, efter ønske, ved hjælp af en række fremgangsmåder, herunder, men ikke begrænset til, fjernelse af den specifikke kemiske inducer fra mediet (f.eks. udtømt galactose, som driver ekspres-
 20 ekstra-kromosomalt replikerende plasmider kan "få fjernet" ekspres-sionsplasmid ved hjælp af cellevekst under ikke-selektive betingelser (f.eks. kan YEp13-holdige celler opformeres i nærvær af leucin), og celler, som har gener indsat i genomet, kan dyrkes med kemikalier, der tvinger det indsatte locus til et "loop-out" (f.eks. kan integranter med URA3 selekteres for tab af
 25 det indsatte gen ved vækst af integranter på 5-fluororotsyre) (J.D. Boeke, F. LaCroute og G.R. Fink. Mol. Gen. Genet. 197:345-346, 1984). Fjernelse af MMR-ekspres-
 30 vildtype-MMR-aktivitet i værtscellen fungerer normalt og reparerer DNA. De nyligt dannede, mutante gærstammer, som udviser hidtil ukendte, valgte træk, er egnede til en lang række kommercielle fremgangsmåder eller til gen/protein-opdagelse til at identificere nye biomolekyler, som er involveret i dannelsen af et særligt træk. Selvom det er blevet beskrevet, at MMR-defekt
 35 kan føre til så meget som en 1000-gange stigning i en værts endogene

DNA-mutationsrate, er det ikke sikkert, at MMR-defekten alene vil være tilstrækkelig til at ændre hvert gen i værtsbakteriens DNA til at danne ændrede biokemikalier med nye aktiviteter. Anvendelsen af kemiske mutagener og deres respektive analoger, såsom ethidiumbromid, EMS, MNNG, MNU, tamoxifen, 8-hydroxyguanin samt andre, såsom dem beskrevet i: Khromov-Borisov, N.N., et al. (Mutat. Res. 430:55-74, 1999); Ohe, T., et al. (Mutat. Res. 429:189-199, 1999); Hour, T.C., et al. (Food Chem. Toxicol. 37:569-579, 1999); Hrelia, P., et al. (Chem. Biol. Interact. 118:99-111, 1999); Garganta, F., et al. (Environ. Mol. Mutagen. 33:75-85, 1999); Ukawa-Ishikawa S., et al. (Mutat. Res. 412:99-107, 1998); www.ehs.utah.edu/ohh/mutagens, etc., kan derfor anvendes til at forøge spektret af mutationer yderligere og øge sandsynligheden for at opnå ændringer i ét eller flere gener, som igen kan danne værtsgær med ét eller flere ønskede nye træk. Fejlparringsreparationsdefekter fører til værter med en forøget resistens over for toksicitet af kemikalier med DNA-skadelig aktivitet. Dette træk muliggør dannelsen af yderligere genetisk forskelligartede værter, når fejlparringsdefekte gærtyper eksponeres for sådanne midler, hvilket ellers ville være umuligt på grund af de toksiske virkninger af sådanne kemiske mutagener [Colella, G., et al. (Br. J. Cancer 80:338-343, 1999); Moreland, N.J., et al. (Cancer Res. 59:2102-2106, 1999); Humbert, O., et al. (Carcinogenesis 20:205-214, 1999); Glaab, W.E., et al. (Mutat. Res. 398:197-207, 1998)]. Desuden er fejlparringsreparation ansvarlig for reparation af kemisk inducerede DNA-addukter, og derfor kunne blokering af denne fremgangsmåde teoretisk øge antallet, typerne, mutationsraten og de genomiske ændringer af en gær [Rasmussen, L.J. et al. (Carcinogenesis 17:2085-2088, 1996); Sledziewska-Gojska, E., et al. (Mutat. Res. 383:31-37, 1997) og Janion, C. et al. (Mutat. Res. 210:15-22, 1989)]. Ud over kemikalierne anført ovenfor kan andre typer af DNA-mutagener, herunder ioniserende stråling og UV-bestråling, hvilke DNA-mutagener er kendt for at forårsage DNA-mutagenese i gær, også anvendes til potentielt at forbedre denne fremgangsmåde (Lee CC, Lin HK, Lin JK. 1994. A reverse mutagenicity assay for alkylating agents based on a point mutation in the beta-lactamase gene at the active site serine codon. Mutagenesis 9:401-405; Vidal A, Abril N, Pueyo C. 1995. DNA repair by Ogt alkyltransferase influences EMS mutational specificity. Carcinogenesis 16:817-821). Disse midler, som er ekstremt toksiske over for værtsceller og derfor resulterer i en reduktion i den faktiske pool-størrelse

af ændrede gærceller, er mere tolererede i MMR-defekte værter og tillader igen et beriget spektrum og en beriget grad af genomisk mutagenese.

Beskrivelsen ovenfor beskriver i almindelighed den foreliggende opfindelse. En mere fuld forståelse kan opnås ved henvisning til følgende specifikke eksem-

- 5 pler, som udelukkende er til illustration og ikke har til hensigt at begrænse opfindelsens omfang.

EKSEMPLER

Eksempel 1: Dannelse af inducerbare MMR-dominante negative allelvektorer og gærceller, der indeholder ekspressionsvektorerne

- 10 Gærekspressionskonstrukter blev fremstillet for at bestemme, hvorvidt det humane PMS2-beslægtede gen (hPMSR2) (Nicolaidis et al. Genomics 30(2):195-206) og det humane PMS134-gen (Nicolaidis NC, Littman SJ, Modrich P, Kinzler KW og Vogelstein B. 1998. A naturally occurring hPMS2 mutation can confer a dominant negative mutator phenotype. Mol Cell Biol 18:1635-1641) kan inaktivere gær-MMR-aktiviteten og derved forøge den samlede frekvens af genomisk hypermutation, hvilket resulterer i dannelsen af varierende datterceller med hidtil ukendte træk efter værtsselektion. Til disse undersøgelser blev et plasmid, der koder for hPMS134-cDNA'et, ændret ved polymerasekædereaktion (PCR). 5'-oligonukleotidet har følgende struktur:
- 20 5'-ACG CAT ATG GAG CGA GCT GAG AGC TCG AGT-3', som indbefatter NdeI-restriktionsstedet CAT ATG. 3'-oligonukleotidet har følgende struktur: 5'-GAA TTC TTA TCA CGT AGA ATC GAG ACC GAG GAG AGG GTT AGG GAT AGG CTT ACC AGT TCC AAC CTT CGC CGA TGC-3', som indbefatter et EcoRI-sted GAA TTC og epitopen med 14 aminosyrer for V5-antistoffet. Oligonukleotiderne
- 25 blev anvendt til PCR under standardbetingelser, der indbefattede 25 cyklusser af PCR (95 °C i 1 minut, 55 °C i 1 minut, 72 °C i 1,5 minutter til 25 cyklusser, efterfulgt af 3 minutter ved 72 °C). PCR-fragmentet blev oprenset ved gelelektroforese og klonet i pTA2.1 (Invitrogen) ved standardfremgangsmåder for kloning (Sambrook et al., Molecular Cloning: A Laboratory Manual, tredje ud-
- 30 gave, 2001), hvilket dannede plasmidet pTA2.1-hPMS134. pTA2.1-hPMS134 blev spaltet med restriktionsenzymet EcoRI for at frigive insertet, som blev klonet i EcoRI-restriktionsstedet af pPIC3.5K (Invitrogen). Den følgende strategi, som ligner den, der er beskrevet ovenfor for kloning af humant PMS134, blev anvendt til fremstilling af en ekspressionsvektor for det humane beslægtede gen PMSR2. Først blev hPMSR2-fragmentet amplificeret ved hjælp af PCR
- 35

for at indføre to restriktionssteder, et NdeI-restriktionssted på 5'-enden og et EcoRI-sted på 3'-enden af fragmentet. 5'-oligonukleotidet, der blev anvendt til PCR, har følgende struktur: 5'-ACG CAT ATG TGT CCT TGG CGG CCT AGA-3', der indbefatter NdeI-restriktionsstedet CAT ATG. 3'-oligonukleotidet, der blev anvendt til PCR, har følgende struktur: 5'-GAA TTC TTA TTA CGT AGA ATC GAG ACC GAG GAG AGG GTT AGG GAT AGG CTT ACC CAT GTG TGA TGT TTC AGA GCT-3', der indbefatter et EcoRI-sted GAA TTC og V5-epitopen for at muliggøre detektion af antistof. Plasmidet, der indeholdt humant PMSR3 i pBluescript SK (Nicolaidis et al. Genomics 30 (2):195-206, 1995), blev anvendt som PCR-målet med de hPMS2-specifikke oligonukleotider ovenfor. Efter 25 cyklusser af PCR (95 °C i 1 minut, 55 °C i 1 minut og 72 °C i 1,5 minutter til 25 cyklusser, efterfulgt af 3 minutter ved 72 °C) blev PCR-fragmentet oprenset ved hjælp af gelelektroforese og klonet i pTA2.1 (Invitrogen) ved standardfremgangsmåder for kloning (Sambrook et al., Molecular Cloning: A Laboratory Manual, tredje udgave, 2001), hvilket dannede plasmidet pTA2.1-hR2. pTA2.1-hR2 blev derefter spaltet med restriktionsenzymet EcoRI for at frigive insertet (der er to EcoRI-restriktionssteder i det multiple kloningssted af pTA2.1, der flankerer insertet) og indført i gæreksepressionsvektoren pPIC3.5K (Invitrogen).

Pichia pastoris-gærceller blev omdannet med pPIC3.5K-vektor, pPIC3.5K-pms134 og pPIC3.5K-hR2 som følger. Først blev 5 ml YPD-medium (1 % gærekstrakt, 2 % bactopecton og 1 % dextrose) inokuleret med en enkelt koloni fra en YPD-plade (samme som ved YPD-væsken, men der var tilsat 2 % difco-agar til pladen) og inkuberet under omrystning natten over ved 30 °C. Overnatskulturen blev derefter anvendt til at inokulere 500 ml YPD-medium (200 µl overnatskultur), og kulturen blev inkuberet ved 30 °C, indtil den optiske densitet på 600 nm nåede 1,3 til 1,5. Cellerne blev derefter bundfældet ved centrifugering (4000 x g i 10 minutter) og derefter vasket 2 gange i sterilt vand (én volumen hver gang), hvorefter cellerne blev suspenderet i 20 ml 1M sorbitol. Sorbitol/cellesuspensionen blev bundfældet ved centrifugering (4.000 x g i 10 minutter) og suspenderet i 1 ml 1M sorbitol. 80 µl af celsuspensionen blev blandet med 5 til 10 µg lineariseret plasmid-DNA og anbragt i en kuvette på 0,2 cm, pulslængde 5 til 10 millisekunder ved feltstyrke på 7.500 V/cm. Derefter blev cellerne fortyndet i 1 ml af 1M sorbitol og overført til et rør på 15 ml og inkuberet ved 30 °C i 1 til 2 timer uden omrystning.

Derefter blev cellerne bundfældet ved centrifugering (4.000 x g i 10 minutter) og suspenderet i 100 µl sterilt vand, og 50 µl/plade blev spredt ud på den hensigtsmæssige, selektive dyrkningsplade. Pladerne blev inkuberet i 2 til 3 dage ved 30 °C, og kolonier blev lagt ud på YPD-plader til yderligere under-

5 søgelse.

Eksempel 2: Dannelse af hypermutabel gær med inducerbare dominante negative alleler af fejlparringsreparationsgener

Gærkloner, der udtrykker human PMS2-homolog PMS-R2 eller tom vektor, blev dyrket i flydende BMG (100mM kaliumphosphat, pH 6,0, 1,34 % YNB (gærnitrogenbase), 4 x 10⁻⁵ % biotin og 1 % glycerol)-kultur i 24 timer ved

10 30 °C. Den næste dag blev kulturerne fortyndet 1:100 i MM-medium (1,34 % YNB, 4 x 10⁻⁵% biotin og 0,5 % methanol) og inkuberet ved 30 °C under omrystning. Celler blev fjernet til mutantselektion 24 og 48 timer efter induktion af methanol, som beskrevet nedenfor (se eksempel 3).

15 **EKSEMPEL 3: Dominante negative MMR-gener kan frembringe nye genetiske varianter og kommercielt levedygtige træk i gær.**

Evnen til at udtrykke MMR-gener i gær, som beskrevet i eksempel 2, viser evnen til at danne genetiske ændringer og nye fænotyper i gær, der udtrykker dominante negative MMR-gener. I dette eksempel beskrives anvendeligheden af denne fremgangsmåde til dannelse af eukaryote stammer med

20 kommercielt relevante træk.

DANNELSE AF URACIL-AFHÆNGIG GÆRSTAMME

Ét eksempel på anvendelighed er dannelsen af en gærstamme, som er mutant for et specifikt metabolisk produkt, såsom en aminosyre eller et nukleotid. Fremstilling af en sådan gærstamme vil muliggøre rekombinant manipulation af gærstammen til indføringen af gener til skalérbare fremgangsmåde til rekombinant fremstilling. For at påvise, at MMR kan manipuleres i gær for at danne mutanter uden evnen til at frembringe specifikke molekulære byggeklodser, blev det følgende forsøg udført. Gærceller, som udtrykker en

25 methanolinducerbar human PMS2-homolog, hPMS2-R2 (som beskrevet i eksempel 1 ovenfor), blev dyrket i BMY-medium natten over, derefter fortyndet 1:100 og overført til MM-medium, hvilket resulterede i aktivering af AOX-promotoren og frembringelse af hPMS2-R2-MMR-genet, der er iboende i gærcellen. Kontrolceller blev behandlet på samme måde; disse celler indeholder

30 pPIC3.5-vektoren i gær og mangler et insert. Celler blev induceret i 24 og 48

timer og derefter selekteret for uracilkrævende mutationer som følger. Cellerne blev udpladet på 5-FOA-medium (Boeke, J.D., LaCroute, F. og Fink, G.R. Mol. Gen. Genet. 197:345-345, 1984). Pladerne fremstilles på følgende måde: (2 x koncentrat (filtersteriliseres): gærnitrogenbase 7 gram; 5-fluororotsyre 1 gram; uracil 50 milligram; glucose 20 gram og vand til 500 ml tilsættes til 500 ml 4 % agar (autoklaveret) og hældes på pladerne. Celler udplades på 5-FOA-plader (efter 0, 24 og 48 timer) og inkuberes ved 30 °C i mellem 3 og 5 dage. Data fra et typisk forsøg er vist i tabel 1. Der blev ikke observeret nogle uracilkrævende kloner i den ikke-inducerede eller inducerede kultur i gærceller, der indeholder den "tomme" vektor, hvorimod de celler, der indeholder MMR-genet, hPMS2-R2, har kloner, som kan vokse på selektionsmediet. Bemærk at den ikke-inducerede kultur af hPMS2-R2 ikke har nogle kolonier, som er resistente over for 5-FOA, hvilket viser, at genet skal induceres, for at den hidtil ukendte fænotype kan dannes. Det er blevet påvist, at mutagenerne (såsom ethylmethanolsulfonat) resulterer i et lavt antal af ura⁻-mutanter, og at den spontane mutationsrate til dannelse af denne klasse af mutanter er lav (Boeke, J.D., LaCroute, F. og Fink, G.R. Mol. Gen. Genet. 197:345-346, 1984).

Tabel 1: Dannelse af uracilkrævende mutante *Pichia pastoris*-gærceller. # betegner 24-timers methanolinduktion og @ 48-timers induktion. Til sammenligning er vist en behandlet/ubehandlet vildtype-gærcelle, (Galli, A. og Schiestl, R.H. Mutat. Res. 429(1):13-26, 1999).

Stamme	Podet	ura ⁻	URA ⁺	Frekvens (ura ⁻ -celler)
Vægt	100.000	0	~100.000	0
Tom	100.000	0	~100.000	0
pMOR ^{ye-1#}	100.000	14	~100.000	1/7.142
pMOR ^{ye2@}	100.000	123	~100.000	1/813
Vægt	100.000	1-0,1	100.000	1/10 ^{5-6*}
Mutagen	100.000	10	100.000	1/10.000

Ét eksempel på kommerciel anvendelighed er dannelsen af varmeresistente, rekombinante proteinproducentstammer. I den skalérbare fremgangsmåde til rekombinant fremstilling resulterer fermentering i stor skala af både prokaryoter og eukaryoter i dannelsen af meget kraftig varme i kulturen.

5 Denne varme skal spredes ved hjælp af fysiske midler, såsom under anvendelse af kølekapper, der omgiver kulturen, mens den aktivt vokser og frembringer produkt. Frembringelse af en gærcele, der effektivt kan modstå høj temperaturstigning, ville være fordelagtig til fremgangsmåder til rekombinant fremstilling i stor skala. Til dette formål kan gærstammen, som beskrevet i

10 eksempel 2, dyrkes i nærvær af methanol for at inducere det dominante negative MMR-gen, og cellerne, som dyrkes i forskellige tidsperioder (f.eks. 12, 24, 36 og 48 timer), anbringes derefter i plader og inkuberes ved forhøjede temperaturer for at selektere for mutanter, der modstår høj temperaturstigning (f.eks. 37 °C eller 42 °C). Disse stammer ville være nyttige til fermenteringsudvikling og "opskalering" af fremgangsmåder og skulle resultere i et

15 fald i fremstillingsomkostninger, da der mindre ofte vil være behov for at køle fermenteringen.

DANNELSE AF HØJE, REKOMBINANTE PROTEINPRODUCENT-STAMMER OG STAMMER MED MINDRE ENDOGEN PROTEASEAKTIVITET

20 Gær er en værdifuld organisme til rekombinant fremstilling, da den er en encellet organisme, der er billig at dyrke og let egner sig til fermentering i skala. Mange eukaryote proteiner, som ikke kan folde effektivt, når de udtrykkes i *Escherichia coli*-systemer, folder desuden med den rette konformation i gær og er strukturelt identiske med deres modstykker fra pattedyr. Der

25 er adskillige iboende begrænsninger ved mange proteiner, der udtrykkes i gær, herunder over- og/eller uhensigtsmæssig glycosylation af det rekombinante protein, proteolyse af endogene gærenzymer og utilstrækkelig sekretion af rekombinant protein fra det indre af gærcellen til mediet (som letter oprensning). Til dannelsen af gærceller med evnen til at oversecernere proteiner

30 eller med mindre endogen proteaseaktivitet og/eller mindre hyperglycosylationsaktivitet kan gærceller, som beskrevet i eksempel 1, dyrkes med methanol i 12, 24, 36 og 48 timer, og gærceller kan selekteres for evnen til at oversecernere proteinet af interesse, underglycosylere det eller en celle med svækket eller ingen proteaseaktivitet. En sådan stamme vil være nyttig til re-

35 kombinant fremstilling eller andre kommercielle formål og kan kombineres

med den varmeresistente stamme anført ovenfor.

For eksempel ville resultatet være en mutant gærcelle, der er resistent over for høj temperaturstigning og som kan udskille store mængder af protein i mediet.

- 5 Lignende resultater blev observeret med andre dominante negative mutanter, såsom PMSR2, PMSR3 og de humane MLH1-proteiner.

EKSEMPEL 4: Mutationer dannet i værtsgenomet af gær ved defekt MMR er genetisk stabile

- 10 Som beskrevet i eksempel 3 resulterer manipulation af MMR-vejen i gær i ændringer i værtsgenomet og evnen til at selekttere for et hidtil ukendt træk, f.eks. en gærcelles evne til at kræve et specifikt næringsstof. Det er vigtigt, at mutationerne indført af MMR-vejen er genetisk stabile og overføres til datterceller på en reproducerbar måde, når først vildtype-MMR-vejen er reableret.
- 15 For at bestemme den genetiske stabilitet af mutationer indført i gærgenomet blev følgende forsøg udført. Fem uafhængige kolonier fra pPIC3.5K-hPMS2-R2, som er *ura⁻*, fem vildtype-kontrolceller (*URA⁺*) og fem pPIC3.5K-omdannede celler ("tom vektor") blev dyrket natten over fra en isoleret koloni i 5 ml YPD (1 % gærekstrakt, 2 % bactopecton og 1 % dextrose) ved 30 °C
- 20 under omrystning. YPD-mediet indeholder alle de næringsstoffer, der er nødvendige for at få gær til vokse, herunder uracil. 1 µl af overnatskulturen, som var ved en optisk densitet (OD), som målt ved 600nm til >3,0, blev derefter fortyndet til en OD₆₀₀ på 0,01 i YPD, og kulturen blev inkuberet under omrystning ved 30 °C i yderligere 24 timer. Denne fremgangsmåde blev gentaget
- 25 endnu 3 gange i 5 overnatsinkubationer i alt. Dette er det ækvivalente af mere end 100 fordoblingsgenerationer (fra den første koloni på pladen til slutningen af den sidste overnatsinkubation). Celler (fem uafhængige kolonier, som var *ura⁻*, og fem, som var vildtype) blev derefter udpladet på YPD-plader ved en celledensitet på 300 til 1.000 celler/plade og inkuberet i to dage ved 30 °C.
- 30 Cellerne fra disse plader blev udpladet flere gange på de følgende plader og bedømt for vækst efter tre dages inkubation ved 30 °C; Synthetic Complete (SC) (Syntetisk Komplet) SC-ura (1,34 % gærnitrogenbase og ammoniumsulfat; 4×10^{-5} % biotin; suppleret med alle aminosyrer, ingen supplerende uracil; 2 % dextrose og 2 % agar); SC+URA (samme som SC-ura, men med supplerende
- 35 plader med 50 mg uracil/liter medium) og YPD-plader. De blev udpla-

det flere gange i følgende rækkefølge - SC-ura, SC-komplet og YPD. Hvis det hidtil ukendte træk, som er iboende i gærgenomet, som blev dannet ved ekspression af den mutante MMR (i dette eksempel den humane homolog af PMS2, hPMS2-R2), er ustabil, bør de uracil-afhængige celler "falde tilbage til" en uracil-uafhængig fænotype. Hvis fænotypen er stabil, bør vækst af de mutante celler under ikke-selektive betingelser resultere i gærceller, som bibeholder deres levedygtighedsafhængighed af eksogen supplering med uracil. Som det kan ses i dataene angivet i tabel 2, er den uracil-afhængige fænotype stabil, når gærcellerne dyrkes under ikke-selektive betingelser, hvilket viser, at den MMR-dannede fænotype afledt af mutation i én af de biosyntetiske uracilreaktionsvejgener er genetisk stabil.

Stamme	Podet	-ura	+URA	YPD
Vægt	650	650	650	650
Tom	560	560	560	560
pMOR ^{ye-1#}	730	0	730	730

Disse data viser anvendeligheden af at anvende et inducerbart ekspressionssystem og et dominant negativt MMR-gen i et eukaryot system for at danne genetisk ændrede stammer. Stammen udviklet i dette eksempel, en gærstamme, der nu kræver tilsætning af uracil for vækst, er potentielt nyttig som en stamme til rekombinant fremstilling; ved at konstruere en ekspressionsvektor, som indeholder vildtype-URA3-genet, på enten et integrationsplasmid eller en ekstra-kromosomal vektor, er det nu muligt at omdanne og danne hidtil ukendte celler, der udtrykker proteinet af interesse. Det er også muligt at modificere andre iboende gener i gærceller og selekttere for mutationer i gener, hvilke mutationer tilvejebringer andre nyttige fænotyper, såsom evnen til at udføre en hidtil ukendt biotransformation. Det er endvidere muligt at udtrykke et gen ekstra-kromosomalt i en gærcelle, der har ændret MMR-aktivitet, som beskrevet ovenfor, og selekttere for mutationer i det ekstra-kromosomale gen. På en måde, som ligner den beskrevet ovenfor, kan den mutante gærcelle derfor anbringes under specifikt selektivt tryk, og et hidtil ukendt protein med kommercielt vigtige biokemiske træk kan selekteres. Disse eksempler er kun ment som illustrationer og ikke som begrænsnin-

ger af omfanget af den foreliggende opfindelse. Endelig, som beskrevet ovenfor, svækkes eller elimineres MMR-aktiviteten fuldstændigt, når en mutation først er blevet indført i genet af interesse. Resultatet er en gærcelle, der indeholder en stabil mutation i målgenerne af interesse.

5 EKSEMPEL 5: Forøget dannelse af MMR-defekt gær og kemiske mutagener til dannelsen af nye træk

Det er tidligere blevet dokumenteret, at MMR-defekt giver forøget mutationsfrekvens og forøget resistens over for toksiske virkninger af kemiske mutagener (CM) og deres respektive analoger, såsom, men ikke begrænset til:

10 ethidiumbromid, EMS, MNNG, MNU, tamoxifen, 8-hydroxyguanin og andre angivet i, men ikke begrænset til, publikationer af: Khromov-Borisov, N.N., et al. *Mutat. Res.* 430:55-74, 1999; Ohe, T., et al. (*Mutat. Res.* 429:189-199, 1999; Hour, T.C. et al. *Food Chem. Toxicol.* 37:569-579, 1999; Hrelia, P., et al. *Chem. Biol. Interact.* 118:99-111, 1999; Garganta, F., et al. *Environ. Mol.*

15 *Mutagen.* 33:75-85, 1999; Ukawa-Ishikawa S., et al. *Mutat. Res.* 412:99-107, 1998; www.ehs.utah.edu/ohh/mutagens; Marcelino LA, Andre PC, Khrapko K, Coller HA, Griffith J og Thilly WG. Chemically induced mutations in mitochondrial DNA of human cells: mutational spectrum of N-methyl-N'-nitro-N-nitrosoguanidine. *Cancer Res.* 1998 Jul 1;58(13):2857-62; Koi M, Umar A,

20 Chauhan DP, Cherian SP, Carethers JM, Kunkel TA og Boland CR. Human chromosome 3 corrects mismatch repair deficiency and microsatellite instability and reduces N-methyl-N'-nitro-N-nitrosoguanidine tolerance in colon tumor cells with homozygous hMLH1 mutation. *Can res* 1994 54:4308-4312, 1994. Fejlparringsreparation fremkalder kromosomafvigelser i hamsterceller behandlet med methyleringsmidler eller 6-thioguanin, men ikke med ethyleringsmidler. For at vise CM'ers evne til at forøge mutationsfrekvensen i MMR-defekte gærceller blev det forudsagt, at eksponering af gærceller for CM'er i nærvær eller fravær af methanol (som inducerer ekspressionen af den iboende humane homolog til PMS2, hPMS2-R2) ville resultere i en forøgelse af mu-

25 tationer i gærcellen.

Gærceller, som udtrykker hPMS2-R2 (induceret eller ikke-induceret), og tomme vektorkontrolceller dyrkes (som beskrevet i eksempel 2 og 3) i 24 timer og fortyndes i MM-medium, som beskrevet ovenfor. Cellerne i MM inkuberes derefter enten med eller uden stigende mængder af ethylmethan-

35 sulfonat (EMS) fra 0, 1, 10, 50, 100 og 200µM. Aliquoter af 10 µl af kultur

(fortyndet i 300 µl MM) inkuberes i 30 minutter, 60 minutter og 120 minutter, efterfulgt af udpladning af celler på 5-FOA-plader, som beskrevet i eksempel 3 ovenfor. Mutanter selekteres og bedømmes som ovenfor. Det forudsiges, at der vil være en stigning i frekvensen af *ura⁻*-mutanter i PMS2-R2-kulturerne, 5 som induceres med methanol, i sammenligning med den ikke-inducerede udgangs- eller vildtype-stamme. I en yderligere udvidelse af dette eksempel vil humane PMS2-R2-holdige celler blive induceret i 24 og 48 timer og derefter mutageniseret med EMS. Dette vil muliggøre, at MMR-genet bliver fuldstændigt aktivt og udtrykt ved høje niveauer, hvilket resulterer i en stigning i an- 10 tallet af opnåede *ura⁻*-mutanter. Det forudsiges, at der ikke vil være nogen ændring i antallet af *ura⁻*-mutanter opnået i den ikke-inducerede udgangskontrol eller i vildtype-"tom vektor"-cellerne.

Dette eksempel viser anvendeligheden af at anvende et reguleret dominant negativt MMR-system plus kemiske mutagener til frembringelse af for- 15 øgede antal af genetisk ændrede gærstammer, som kan selekteres for nye træk. Denne fremgangsmåde er nyttig til at danne sådanne organismer til kommercielle anvendelser, såsom, men ikke begrænset til, rekombinant fremstilling, biotransformation og ændrede biokemikalier med forøgede aktiviteter. Den er også nyttig til at opnå ændringer af proteinaktivitet fra ektopisk ud- 20 trykte proteiner indeholdt på ekstra-kromosomale ekspressionsvektorer, der ligner dem beskrevet i eksempel 4 ovenfor.

EKSEMPEL 6: Alternative fremgangsmåder til inhibering af gær-MMR-aktivitet

Inhiberingen af MMR-aktivitet i en værtsorganisme kan opnås ved indfø- 25 ring af en dominant negativt allel, som vist i eksemplerne ovenfor. Denne ansøgning beskriver også anvendeligheden af at anvende regulerede systemer til at regulere MMR i gær for at frembringe genetisk diversitet og træk til kommercielle anvendelser. Yderligere fremgangsmåder til at regulere suppressionen af en værts MMR-aktivitet er ved anvendelse af genetisk rekombination 30 for at slå alleler af et MMR-gen i cellen af interesse ud. Dette kan opnås under anvendelse af homolog rekombination, som afbryder det endogene MMR-gen; 2) blokering af MMR-proteindimerisering med andre underenheder (hvilken proteindimerisering er nødvendig for aktivitet) ved at indføre polypeptider eller antistoffer i værten via fremgangsmåder til transfektion, hvilke frem- 35 gangsmåder anvendes rutinemæssigt af fagmanden (f.eks. elektroporering),

eller 3) forøgelse af ekspressionen af et MMR-gen under anvendelse af anti-sense-oligonukleotider.

MMR-gen-knockouts. Det er hensigten at frembringe afbrudte målrettede vektorer af et særligt MMR-gen og indføre det i gær-genomet under anvendelse af standardfremgangsmåder inden for teknikken. Gær, der udviser hypermutabilitet, vil være nyttig til at frembringe genetisk forskelligartet afkom til kommercielle anvendelser. Det vil blive bekræftet, at gær har mistet ekspressionen af MMR-genet under anvendelse af northern-standardteknikker og biokemiske standardteknikker (som beskrevet i reference 31). MMR-genloci kan slås ud, stammer kan selekteres for nye træk, og MMR kan gendannes ved at indføre et vildtype-MMR-gen til at supplere KO-locuset. Andre strategier indbefatter anvendelse af KO-vektorer, som kan målrette et MMR-genlocuset, selektere for værtstræk og derefter få KO-vektoren "splejset" fra genomet efter stammedannelse.

Blokerende peptider. MMR-underenheder (MutS- og MutL-proteiner) interagerer for at danne aktive MMR-komplekser. Peptider er i stand til specifikt at inhibere bindingen af to proteiner ved hjælp af kompetitiv inhibering. Indføring af peptider eller antistoffer i celler til konserverede domæner af et særligt MMR-gen for at afbryde aktivitet er simpelt for fagmanden. Gær vil blive kontrolleret for tab af ekspression af MMR-aktiviteten ved northern-standardteknikker og/eller biokemiske standardteknikker (som beskrevet i Nicolaides NC, Littman SJ, Modrich P, Kinzler KW og Vogelstein B. 1998. A naturally occurring hPMS2 mutation can confer a dominant negative mutator phenotype. Mol Cell Biol 18:1635-1641). Gær, der udviser hypermutabilitet, vil være nyttig til frembringelse af genetisk forskelligartet afkom til kommercielle anvendelser.

Diskussion

Resultaterne beskrevet ovenfor vil føre til flere konklusioner. For det første vil ekspression af dominante negative MMR-proteiner resultere i en stigning i mikrosatellit-ustabilitet og hypermutabilitet i gær. Gærcellens hypermutabilitet skyldes inhiberingen af den iboende, endogene MMR-biokemiske aktivitet i disse værter. Denne fremgangsmåde tilvejebringer et krav om anvendelse af MMR-gener og de produkter, der kodes deraf, til dannelse af hypermutabel gær til frembringelse af nye træk til kommercielle anvendelser.

EKSEMPLER PÅ MMR-GENER OG KODEDE POLYPEPTIDER

Gær MLH1-cDNA (deponeringsnummer U07187)

```

1 aaataggaat gtgatacctt ctattgcatg caaagatagt gtagggaggcg ctgctattgc
61 caaagacttt tgagaccgct tgctggttca ttatagttga ggagttctcg aagaccgagaa
121 attagcagtt ttcggtgttt agtaatcgcg ctagcatgct aggacaattt aactgcaaaa
181 ttttgatagc atagtgtatg taaatggaag gtaaaaaataa catagacctt tcaataagca
241 atgtctctca gaataaaagc acttgatgca tcagtgttta acaaaattgc tgcagggtgag
301 atcataatat ccccggtaaa tgctctcaaa gaaatgatgg agaattccat cgatgcgaat
361 gctacaatga ttgatatctt agtcaaggaa ggaggaatta aggtacttca aataacagat
421 aacggatctg gaattataaa agcagaccctg ccaatcttat gtgagcgatt caccagctcc
481 aaattacaaa aattcgaaga tttgagtcag attcaaacgt atggattccg aggagaagct
541 tttagccagta tctcacatgt ggcaagagtc acagtaacga caaaagttaa agaagacaga
601 tgtgcatgga gagtttcata tgcagaaggt aagatgttgg aaagccccaa acctgttgc
661 ggaaaagacg gtaccacgat cctagttgaa gacctttttt tcaatattcc ttctagatta
721 agggccttga ggtcccataa tgatgaatac tctaaaatat tagatgttgt cgggcgatac
781 gccattcatt ccaaggacat tggcttttct tgtaaaaagt tcggagactc taattattct
841 ttatcagtta aaccttcata tacagtcacg gataggatta ggactgtgtt caataaatct
901 gtggcttcga atttaattac ttttcatatc agcaaaagtag aagattttaa cctggaagc
961 gttgatggaa aggtgtgtaa tttgaatttc atatccaaaa agtccatttc attaatttt
1021 ttcatttaata atagactagt gacatgtgat cttctaagaa gagctttgaa cagcgtttac
1081 tccaattatc tgccaaaggg cttcagacct tttattttat tgggaattgt tatagatccg
1141 ggggtgtgtg atgttaacgt tcacccgaca aagagagagg ttcgtttcct gagccaagat
1201 gagatcatag agaaatcgc caatcaattg caccgcgaat tatctgccat tgcatactca
1261 cgtactttca aggtctcttc aatttcaaca aacaagccag agtcattgat accatttaat
1321 gacaccatag aaagtgtatg gaatagggaag agtctccgac aagcccaagt ggtagagaat
1381 tcatatacga cagccaatag tcaactaagg aaagcgaaaa gacaagagaa taaactagtc
1441 agaatagatg cttcacaaagc taaaattacg tcatttttat cctcaagtca acagtccaac
1501 tttgaaggat cgtctacaaa gcgacaactg agtgaaccca aggttaacaaa tgtaagccac
1561 toccaagagg cagaaagct gacactaaat gaaagcgaaac aaccgcgtga tgcacaatac
1621 atcaatgata atgacttgaa ggtcaaacct aagaagaaac aaaaagtggg ggattataaa
1681 gttccaagca ttgccgatga cgaagaagaat gcaactccga tttcaaaaaga ggggtatatt
1741 agagtaccta aggagcgagt taatgttaat cttacgagta tcaagaatgt gcgtgaaaaa
1801 gtatgtgatt cgatacatcg agaactaaca gacatttttg caaatttgaa ttactgtggg
1861 gttgtatagtg aggaagaag attagccgct attcagcatg acttaagctt tttttataa
1921 gattacggat ctgtgtgcta tgagctattc tatcagattg gtttgacaga cttcgcaaac
1981 ttttgtaaga taaacctaca gagtacaaat gtgtcagatg atatagttt gtataatctc
2041 ctatcagaat ttgacgagtt aaatgacgat gcttccaaag aaaaaataat tagtaaaaaa
2101 tgggacatga gcagtatgct aaatgagtac tattccatag aattggtgaa tgatggtcta
2161 gataatgact taaagtctgt gaagctaaaa tctctaccac tacttttaaa aggtacatt
2221 ccatctctgg tcaagttacc attttttata tatcgcttgg gtaaaagagt tgattgggag
2281 gatgaacaag agtgtctaga tggtaattta agagagattg cattactcta tatacctgat
2341 atggttccga aagtcgatac actcgatgca tcgttgtcag aagacgaaaa agcccagttt
2401 ataaatagaa aggaacacat atcctcatta ctagaacacg ttctcttccc ttgtatcaaa
2461 cgaaggttcc tggcccttag acacattctc aaggatgtcg tggaaatagc caaccttcca
2521 gatctataca aagtttttga gaggtgttaa ctttaaaacg ttttggtctg aataccaaa
2581 tttttgttta tttcctgagt gtgattgtgt ttcatttgaa agtgatgcc ctttccttta
2641 acgattcacc cgcgagattt caaaggatat gaaatatggt tgcagttagg aaagtatgtc
2701 agaaatgtat attcggattg aaactcttct aatagttctg aagtcacttg gttccgtatt
2761 gttttcgtcc tcttcctcaa gcaacgattc ttgtctaagc ttattcaacg gtaccaaaga
2821 cccgagtcct tttatgagag aaaacatttc atcatttttc aactcaatta tcttaatatc
2881 atttttagt attttgaaaa caggatggta aaacgaatca cctgaatcta gaagctgtac
2941 cttgtcccat aaaagtttta atttactgag cctttcggtc aagtaaacta gtttatctag
3001 ttttgaaccg aatattgtgg gcagatttgc agtaagttca gtttagatcta ctaaaagttg
3061 tttgacagca gccgattcca caaaaatttg gtaaaaggag atgaaagaga cctcgcgcgt
3121 aatggtttgc atcaccatcg gatgtctgtt gaaaaactca ctttttgcac ggaagttatt
3181 aacaataaga ctaatgatta ccttagaata atgtataa

```

Gaer-MLH1-protein (deponeringsnummer U07187)

MSLRKALDASVVNKIAAGEIIISPVNALKEMMENSIDANATMI
 DILVKEGGIKVLQITDNGSGINKADLPILCERFTTSKLQKFEDLSQIQTYGFRGEALA
 SISHVARVTVTTKVKEDRCANRVSYAEGKMLESPPKPVAGKDGTTILVEDLFFNIPSR
 RALRSHNDEYSKILDOVVGRYAIHSDIGFSCCKKFGDSNYSLSVKPSYTVQDRIRTVFN
 KSVASNLITFHISKVEDLNLESVDGKVCNLFISKKSISLIFFINNRLVTCOLLRRAL
 NSVYSNYLPGKFRPFYILGIVIDPAAVDVNVHPTKREVRFLSQDEIIEXIANQLHAEL
 SAIDTSRTFKASSISTNKPESLIPFNDTIESDRNRKSLROAQVVENSYTTANSQRLKA
 KRQENKLVRIDASQAKITSFLSSSQQFNFECSSTKRQLSEPKVTNVSHSQEAEKLTIN
 ESEQPRDANTINONDLKDQPKKKQKLGDKVPSTADDEKNALPISKDGYIRVPKERVN
 VNLTSIKKLREKVDSDSIHRELTDFANLNVGVVDEERRLAIIQHDLKFLIDYGSVC
 YELFYQIGLTFANFGKINLQSTNVSDDIVLYNLLSEFDELNDASKEKISKIWDMS
 SMLNEYYSIELVNDGLONDLKSVKLKSLPLLLKGYIPSLVKLPFFIYRLGKEVDWEDE
 QECLOGILREIALLYIPDMVPKVDTLDAASLSEDEKAQFINRKEHISLLEHVLFPCK
 RRFLAPRHILKOVVEIANLPDLYKVFERC

Muse- PMS2 protein

MEQTEGVSTE	CAKAIKPIDG	KSVHQICSGQ	VILSLSTAVK	ELIENSVDAG	ATTIDLRLKD	60
YGVDLIEVSD	NGCGVEEENF	EGLALKHHTS	KIQEFADLTQ	VETFGFRGEA	LSSLCALSDV	120
TISTCHGSAS	VGTRLVFDHN	GKITQKTPYP	RPKGTTVSVQ	HLFYTLPVRY	KEFQRNIKKE	180
YSKMVQVLQA	YCIISAGVRV	SCTNQLGQ GK	RHAVVCTSGT	SGMKENIGSV	FGQKQLQSLI	240
PFVQLPPSDA	VCEEYGLSTS	GRHKTFFSTF	ASFHSARTAP	GGVQQTGSFS	SSIRGPVTQQ	300
RSLSLSMRFY	HMYNRHQYPF	VVLNVSDVSE	CVDINVTPOK	RQILLQEEKL	LLAVLKTSLI	360
GMFDSANKL	NVNQQPLLDV	EGNLVKLHTA	ELEKPVPGKQ	DNSPSLKSTA	DEKRVASISR	420
LREAFSLHPT	KEIKSRGPET	AELTRSFPS	KRGVLSSYPS	DVISYRGLRG	SQDKLVSPTD	480
SPGDCMDREK	IEKDSGLSST	SAGSEEEFST	PEVASSFSSD	YNVSSLED RP	SQETINCGDL	540
DCRPPGTGQS	LKPEDHGYQC	KALPLARLSP	TNAKRFKTEE	RPSNVNISQR	LPGPQSTSAA	600
EVDVAIKMNK	RIVLLEFSL	SLAKRMKQLQ	HLKAQNKHEL	SYRKFRKIC	PGENQAAEDE	660
LRKEISKSMF	AEMEILGQFN	LGFIIVTKLKE	DLFLVDQHAA	DEKYNFEMLO	QHTVLQAQRL	720
ITPQTLNLTA	VNEAVLIENL	EIFRKNGFDF	VIDEDAPVTE	RAKLISLPTS	KNWTFGPQDI	780
DELIFMLSDS	PGVMCRPSRV	RQMFASRACR	KSVMIGTALN	ASEMKKLITH	MGEMDHPWNC	840
PHGRPTMRHV	ANLDVISQN					859

Muse- PMS2 cDNA

gaattccggt	gaaggctcctg	aagaatttcc	agattcctga	gtatcattgg	aggagacaga	60
taacctgtcg	tcaggtaacg	atggtgtata	tgcaacagaa	atgggtgttc	ctggagacgc	120
gtcttttccc	gagagcggca	ccgcaactct	cccgcggtga	ctgtgactgg	aggagtcctg	180
catccatgga	gcaaaccgaa	ggcgtgagta	cagaatgtgc	taaggccatc	aagcctattg	240

atgggaagtc agtccatcaa atttgttctg ggcaggtgat actcagttta agcaccgctg 300
 tgaaggagtt gatagaaaat agttagatg ctggtgctac tactattgat ctaaggctta 360
 aagactatgg ggtggacctc attgaagttt cagacaatgg atgtggggtg gaagaagaaa 420
 actttgaagg tctagctctg aaacatcaca catctaagat tcaagagttt gccgacctca 480
 cgcaggttga aactttcggc ttccgggggg aagctctgag ctctctgtgt gcactaagtg 540
 atgtcactat atctacctgc cacgggtctg caagcgttgg gactcgactg gtgtttgacc 600
 ataattggga aatcacccag aaaactccct acccccgacc taaagggaacc acagtcagt 660
 tgcagcactt attttataca ctacctgctg gtlacaaaga gtttcagagg aacattaaaa 720
 aggagtattc caaaatggtg cagggtctac aggcgtactg tatcatctca gcaggcgctc 780
 gtgtaagctg cactaatcag ctgggacagg ggaagcggca cgctgtgtgt tgcacaagcg 840
 gcacgtctgg catgaaggaa aatatcgggt ctgtgtttgg ccagaagcag ttgcaaagcc 900
 tcattccctt tgttcagctg cccctcagtg acgctgtgtg tgaagagtac ggcttgagca 960
 cttcaggacg ccacaaaacc ttttctacgt ttcgggcttc atttcacagt gcacgcacgg 1020
 cgccgggagg agtgcaacag acaggcagtt tttcttctac aatcagaggc cctgtgacct 1080
 agcaaaggct tctaagcttg tcaatgaggt tttatcacat gtataaccgg catcagtacc 1140
 catttgctgt ccttaacggt tccgttgact cagaatgtgt ggatattaat gtaactccag 1200
 ataaaaggca aattctacta caagaagaga agctattgct ggccgtttta aagacctcct 1260
 tgataggaat gtttgacagt gatgcaaca agcttaatgt caaccagcag ccactgctag 1320
 atgttgaaag taacttagta aagctgcata ctgcagaact agaaaagcct gtgccaggaa 1380
 agcaagataa ctctccttca ctgaagagca cagcagacga gaaaagggtg gcattccatct 1440
 ccaggtctgag agaggccttt tctcttctac ctactaaaga gatcaagtct aggggtccag 1500
 agactgctga actgacacgg agttttccaa gtgagaaaag gggcgtgtta tcctcttctc 1560
 cttcagacgt catctcttac agaggcctcc gtggctcgca ggacaaattg gtgagtccca 1620
 cggacagccc tggtagctgt atggacagag agaaaataga aaaagactca gggctcagca 1680
 gcacctcagc tggctctgag gaagagttca gcaccccgga agtgggcagt agcttttagca 1740
 gtgactataa cgtgagcttc ctagaagaca gaccttctca ggaaaccata aactgtgtgt 1800
 acctggactg ccgtcctcca ggtacaggac agtccttgaa gccagaagac catggatatc 1860
 aatgcaaagc tctacctcta gctcgtctgt caccacaaa tgccaagcgc ttcaagacag 1920
 aggaaagacc ctcaaatgtc aacatttctc aaagattgcc tggctcctcag agcacctcag 1980
 cagctgaggt cgatgtagcc ataaaaatga ataagagaat cgtgctcctc gaggctctctc 2040
 tgagttctct agctaagcga atgaagcagt tacagcacct aaaggcgagc aacaaacatg 2100
 aactgagttg cagaaaattt agggccaaga ttgcccctgg agaaaaccaa gcagcagaag 2160
 atgaactcag aaaagagatt agtaaatcga tgtttgcaga gatggagatc ttgggtcagt 2220
 ttaacctggg atttatagta accaaactga aagaggacct ctctctggtg gaccagcatg 2280
 ctgcggatga gaagtacaac tttagatgc tgcagcagca cacggtgctc caggcgagca 2340
 ggctcatcac accccagact ctgaacttaa ctgctgtcaa tgaagctgta ctgatagaaa 2400
 atctggaaat attcagaaag aatggctttg actttgtcat tgatgaggat gctccagtca 2460
 ctgaaagggc taaattgatt tccttaccac ctagtataaa ctggaccttt ggaccccaag 2520
 atatagatga actgatcttt atgttaagtg acagccctgg ggtcatgtgc cggccctcac 2580
 gagtcagaca gatgtttgct tccagagcct gtgggaagtc agtgatgatt ggaacygcgc 2640

tcaatgcgag cgagatgaag aagctcatca cccacatggg tgagatggac cacccttgga 2700
 actgccccca cggcaggcca accatgaggc acgttgccaa tctggatgtc atctctcaga 2760
 actgacacac cccttgtagc atagagttta ttacagattg ttcggtttgc aaagagaagg 2820
 ttttaagtaa tctgattatc gttgtacaaa aattagcatg ctgctttaat gtaactgagc 2880
 catttaaaag cagtgttaag gcaggcatga tggagtgttc ctctagctca gtaacttggg 2940
 tgatccggtg ggagctcatg tgagcccagg actttgagac cactccgagc cacattcatg 3000
 agactcaatt caaggacaaa aaaaaaaga tatttttgaa gccttttaaa aaaaaa 3056

human PMS2 protein

```

MKQLPAATVR LLSSSQIITS VVSVMKELIE NSLDAGATSV DVKLENYGFD KIEVRDNGEG 60
IKAVDAPVMA MKYYTSKINS HEDLENLTTY GFRGEALGSI CCIAEVLITT RTAADNFSTQ 120
YVLDGSGHIL SQKPSHLGQG TTVTALRLFK NLPVRKQFYS TAKKCKDEIK KIQDLLMSFG 180
ILKPDRLRIVF VHNKAVIWQK SRVSDHKMAL MSVLGTAVMN NMESFQYHSE ESQIYLSGFL 240
PKCDADHSFT SLSTPERSFI FINSRPVHOK DILKLIRHHY NLKCLKESTR LYPVFFLKID 300
VPTADVDVNL TPDKSQVLLQ NKESVLIALE NLMTTCYGPL PSTNSYENNK TDVSAADIVL 360
SKTAETDVLV NKVESGKKNY SNVDTSVIPF QNDMHNDSEG KNTDDCLNHQ ISIGDFGYGH 420
CSSEISNIDK NTKNAFQDIS MSNVSWENSQ TEYSKTCFIS SVKHTQSENG NKDHIDESGE 480
NEEEAGLENS SEISADEWSR GNILKNSVGE NIEPVKILVP EKSLPCKVSN NNYPIPEQMN 540
LNEDSCNKKS NVIDNKGSKV TAYDLLSNRV IKKPMASASAL FVQDHRPQFL IENPKTSLED 600
ATLQIEELWK TLSEEEKLKY EEKATKOLER YNSQMKRAIE QESQMSLKDG RKKIKPTSAA 660
NLAQKHKLKT SLSNQPKLDE LLQSQIEKRR SQNIKMVQIP FSMKNLKNF KKQNKVDLEE 720
KDEPCLIHNL RFPDAWLMTS KTEVMLNPNY RVEEALLFKR LLENHKLPAE PLEKPIMLTE 780
SLFNGSHYLD VLYKMTADDQ RYSGSTYLSQ PRLTANGFKI KLIPGVSITE NYLEIEGMAN 840
CLPFYGVADL KEILNAILNR NAKEVYECRP RKVISYLEGE AVRLSRQLPM YLSKEDIQDI 900
IYRMKHQFGN EIKECVHGRP FFHHLTYLPE TT 932

```

Human PMS2 cDNA

```

cgaggcgat cggtgtgtgc atccatggag cgagctgaga gctcgagtac agaacctgct 60
aaggccatca aacctattga tcggaagtca gtccatcaga ttgtctctgg gcagggtgga 120
ctgagtctaa gcactgcggt aaaggagtta gtagaaaaca gtctggatgc tgggtgccact 180
aatattgatc taaagcttaa ggactatgga gtggatctta ttgaagtttc agacaatgga 240
tgtgggtag aagaagaaaa cttcgaaggc ttaactctga aacatcacac atctaagatt 300
caagagtttg ccgacctaac tcaggttgaa acttttggct ttcgggggga agctctgagc 360
tcactttgtg cactgagcga tgtcaccatt tctacctgcc acgcatcggc gaaggttgga 420
actcgactga tgtttgatca caatgggaaa attatccaga aaaccccta ccccgcccc 480

```

```

agagggacca cagtcagcgt gcagcagtta ttttccacac tacctgtgcg ccataaggaa 540
tttcaaagga atattaagaa ggagtatgcc aaaatggtcc aggtcttaca tgcatactgt 600
atcatttcag caggcatccg tgtaagttgc accaatcagc ttggacaagg aaaacgacag 660
cctgtggtat gcacaggtgg aagccccagc ataaaggaaa atatcggtc tgtgtttggg 720
cagaagcagt tgcaaagcct cattcctttt gttcagctgc cccctagtga ctccgtgtgt 780
gaagagtacg gtttgagctg ttcggatgct ctgcataatc ttttttacat ctcaggtttc 840
atttcacaat gcacgcatgg agttggaagg agttcaacag acagacagtt tttctttatc 900
aaccggcggc cttgtgaccc agcaaaggtc tgcagactcg tgaatgaggt ctaccacatg 960
tataatcgac accagtatcc atttgttgtt cttaacattt ctgttgattc agaatgcgtt 1020
gatatcaatg ttactccaga taaaaggcaa attttgctac aagaggaaaa gcttttgttg 1080
gcagttttta agacctcttt gataggaatg tttgatagtg atgtcaacaa gctaaatgtc 1140
agtcagcagc cactgctgga tgttgaagggt aacttaataa aaatgcattc agcggatttg 1200
gaaaagccca tggtagaaaa gcaggatcaa tccccttcac taaggactgg agaagaaaaa 1260
aaagacgtgt ccatttccag actgcgagag gccttttctc ttcgtcacac aacagagaac 1320
aagcctcaca gcccaaagac tccagaacca agaaggagcc ctctaggaca gaaaaggggg 1380
atgctgtctt ctagcacttc aggtgccatc tctgacaaaag gcgtcctgag acctcagaaa 1440
gaggcagtga gttccagtca cggacccagt gaccctacgg acagagcggg ggtggagaag 1500
gactcggggc acggcagcac ttcctgggat tctgaggggt tcagcatccc agacacgggc 1560
agtcactgca gcagcgagta tgcggccagc tcccagggg acagggggtc gcaggaaacat 1620
gtggactctc aggagaaagc gcctgaaaact gacgactctt tttcagatgt ggactgccat 1680
tcaaaccagg aagataccgg atgtaaatth cgagtthtgc ctacagccaac taatctcgca 1740
accccaaaaca caaagcgttt taaaaaagaa gaaattcttt ccagttctga catttgtaa 1800
aagttagtaa atactcagga catgtcagcc tctcagggtt atgtagctgt gaaaattaat 1860
aagaaagttg tgccctgga cttttctatg agttctttag ctaaacgaat aaagcagtta 1920
catcatgaag cacagcaaag tgaaggggaa cagaattaca ggaagtttag ggcaaagatt 1980
tgtcctggag aaaatcaagc agccgaagat gaactaagaa aagagataag taaaacgatg 2040
tttgacagaa tggaaatcat tggtcagttt aacctgggat ttataataac caaactgaat 2100
gaggatatct tcatagtgga ccagcatgcc acggacgaga agtataactt cgagatgctg 2160
cagcagcaca ccgtgctcca ggggcagagg ctcatagcac ctacagactt caacttaact 2220
gctgttaatg aagctgttct gatagaaaat ctggaaatat ttagaaagaa tggctttgat 2280
tttgttatcg atgaaaatgc tccagtcact gaaagggtta aactgatttc cttgccaact 2340
agtaaaaaact ggaccttcgg accccaggac gtcgatgaac tgatcttcat gctgagcgac 2400
agccctgggg tcatgtgccg gccttcccga gtcaagcaga tgtttgcctc cagagcctgc 2460
cggaagtcgg tgatgattgg gactgctctt aacacaagcg agatgaagaa actgatcacc 2520
cacatggggg agatggacca cccctggaac tgtcccatg gaaggccaac catgagacac 2580
atcgccaacc tgggtgtcat ttctcagaac tgaccgtagt cactgtatgg aataattgg 2640
tttatcgagc atttttatgt ttgaaagac agagtcttca ctaacctttt ttgttttaa 2700
atgaaacctg ctacttaaaa aaaatacaca tcacacccat ttaaaagtga tcttgagaac 2760
cttttcaaac c
2771

```

human PMS1 protein

```

MKQLPAATVR LLSSSQIITS VVSVKELIE NSLDAGATSV DVKLENYGFD KIEVRDNGEG 60
IKAVDAPVMA MKYYTSKINS HEDLENLTTY GFRGEALGSI CCIAEVLITT RTAADNFSTQ 120
YVLDSGSHIL SQKPSHLGQG TTVTALRLFK NLPVRKQFYS TAKKCKDEIK KIQDLLMSFG 180
ILKPDLRIVF VHNKAVIWQK SRVSDHKMAL MSVLGTAVMN NMESFQYHSE ESQIYLSGFL 240
PKCDADHSFT SLSTPERSFI FINSRPVHOK DILKLIRHHY NLKCLKESTR LYPVFFLKID 300
VPTADVDVNL TPDKSQVLLQ NKESVLIALE NLMTTCYGPL PSTNSYENNK TDVSAADIVL 360
SKTAETDVLV NKVESSGKNY SNVDTSVIPF QNDMHNDESG KNTDDCLNHQ ISIGDFGYGH 420
CSSEISNIDK NTKNAFQDIS MSNVSWENSQ TEYSKTCFIS SVKHTQSENG NKDHIDESGE 480
NEEEAGLENS SEISADEWSR GNILKNSVGE NIEPVKILVP EKSLPCKVSN NNYPIPEQMN 540
LNEDSCNKKK NVIDNKSGKV TAYDLSNRV IKKPMASASAL FVQDHRPQFL IENPKTSLED 600
ATLQIEELWK TLSEEEKLKY EEKATKDLER YNSQMKRAIE QESQMSLKDG RKKIKPTS AW 660
NLAQKHKLKT SLSNQPKLDE LLQSQIEKRR SQNIKMQVIP FSMKNLKINF KQKNKVDLEE 720
KDEPCLIHNL RFPDAWLMTS KTEVMLLNPY RVEEALLFKR LLENHKLP AE PLEKPIMLTE 780
SLFNGSHYLD VLYKMTADDQ RYSGSTYLS D PRLTANGFKI KLIPGV SITE NYLEIEGMAN 840
CLPFYGVADL KEILNAILNR NAKEVYECRP RKVISYLEGE AVRLSRQLPM YLSKEDIQDI 900
IYRMKHQFGN EIKECVHGRP FFHHLTYLPE TT 932

```

Human PMS1 cDNA

```

ggcagcagtg gctgcttgcg gctagtggtat ggtaattgcc tgcctgcgcg tagcagcaag 60
ctgctctgtt aaaagcgaaa atgaaacaat tgcctgcggc aacagttcga ctcttttcaa 120
gttctcagat catcacttcg gtgggtcagtg ttgtaaaaga gcttattgaa aactccttgg 180
atgctggtgc cacaagcgta gatgttaaac tggagaacta tggatttgat aaaattgagg 240
tgcgagataa cggggagggg atcaaggctg ttgatgcacc tgaataggca atgaagtact 300
acacctcaaa aataaatagt catgaagatc ttgaaaattt gacaacttac ggttttcgtg 360
gagaagcctt ggggtcaatt tgttgtatag ctgaggtttt aattacaaca agaacggctg 420
ctgataattt tagcaccag tatgttttag atggcagtg ccacatactt tctcagaaac 480
cttcacatct tggtaagggt acaactgtaa ctgctttaag attatttaag aatctacctg 540
taagaaagca gttttactca actgcaaaaa aatgtaaaaga tgaaataaaa aagatccaag 600
atctcctcat gagctttggt atccttaaac ctgacttaag gattgtcttt gtacataaca 660
aggcagttat ttggcagaaa agcagagtat cagatcacaa gatggctctc atgtcagttc 720
tggggactgc tgttatgaac aatatggaat cctttcagta ccactctgaa gaatctcaga 780
tttatctcag tggatttctt ccaaagtgtg atgcagacca ctctttcact agtctttcaa 840
caccagaaag aagtttcac ttcataaaca gtcgaccagt acatcaaaaa gatattctaa 900
agttaatccg acatcattac aatctgaaat gcctaaagga atctactcgt ttgtatcctg 960
ttttctttct gaaaatcgat gtccctacag ctgatgttga tgtaaattta acaccagata 1020
aaagccaagt attattacaa aataaggaa ctgttttaat tgctcttgaa aatctgatga 1080
cgacttgta tggaccatta ctagtacaa attcttatga aaataataaa acagatgttt 1140
ccgcagctga catcgttctt agtaaaacag cagaaacaga tgtgcttttt aataaagtgg 1200

```

aatcatctgg aaagaattat tcaaattgtg atacttcagt cattccattc caaaatgata 1260
tgcataatga tgaatctgga aaaaacactg atgattgttt aaatcaccag ataagtattg 1320
gtgacttttg ttatggatcat ttagtagtg aaatttctaa cattgataaa aacactaaga 1380
atgcatttca ggacatttca atgagtaatg tatcatggga gaactctcag acggaatata 1440
gtaaaacttg ttttataagt tccgttaagc acacccagtc agaaaatggc aataaagacc 1500
atatagatga gagtggggaa aatgaggaag aagcaggtct tgaaaactct tcggaaattt 1560
ctgcagatga gtggagcagg ggaaatatac ttaaaaattc agtgggagag aatattgaac 1620
ctgtgaaaat tttagtgcct gaaaaaagtt taccatgtaa agtaagtaat aataattatc 1680
caatccctga acaaatgaat cttaatgaag attcatgtaa caaaaaatca aatgtaatag 1740
ataataaatc tggaaaagtt acagcttatg atttacttag caatcgagta atcaagaaac 1800
ccatgtcagc aagtgtctct tttgttcaag atcatcgctc tcagtttctc atagaaaatc 1860
ctaagactag ttttagaggat gcaacactac aaattgaaga actgtggaag acattgagtg 1920
aagaggaaaa actgaaatat gaagagaagg ctactaaaga cttggaacga tacaatagtc 1980
aatggaagag agccattgaa caggagtcac aaatgtcact aaaagatggc agaaaaaaga 2040
taaaacccac cagcgcatgg aatttgcccc agaagcacia gttaaaaacc tcattatcta 2100
atcaacccaa acttgatgaa ctcttcagt cccaaattga aaaaagaagg agtcaaaata 2160
ttaaaatggt acagatcccc ttttctatga aaaacttaaa aataaatttt aagaaacaaa 2220
acaaagttga cttagaagag aaggatgaac cttgcttgat ccacaatctc aggtttcctg 2280
atgcattgct aatgacatcc aaaacagagg taatgttatt aaatccatat agagtagaag 2340
aagccctgct atttaaaaga cttcttgaga atcataaact tcctgcagag ccactggaaa 2400
agccaattat gttaacagag agtcttttta atggatctca ttatttagac gttttatata 2460
aatgacagc agatgaccaa agatacagtg gatcaactta cctgtctgat cctcgtctta 2520
cagcgaatgg tttcaagata aaattgatac caggagtttc aattactgaa aattacttgg 2580
aatagaagg aatggctaatt tgtctcccat tctatggagt agcagattta aaagaaattc 2640
ttaatgctat attaaacaga aatgcaaagg aagtttatga atgtagacct cgcaaagtga 2700
taagtatttt agagggagaa gcagtcgctc tatccagaca attacccatg tacttatcaa 2760
aagaggacat ccaagacatt atctacagaa tgaagcacca gtttggaat gaaattaaag 2820
agtgtgttca tggctgcccc ttttttcac atttaacctc tcttcagaa actacatgat 2880
taaatatgtt taagaagatt agttaccatt gaaattggtt ctgtcataaa acagcatgag 2940
tctgttttta aattatcttt gtattatgtg tcacatggtt attttttaaa tgaggattca 3000
ctgacttggt tttatattga aaaaagttcc acgtattgta gaaaacgtaa ataaactaat 3060
aac

3063

human MSH2 protein

MAVQPKETLQ LESAAEVGFV RFFQGMPEKP TTTVRLFDRG DFYTAHGEDA LLAAREVFKT 60
QGVIKYMGPA GAKNLQSVVL SKMNFESFVK DLLLVRQYRV EVYKNRAGNK ASKENDWYLA 120
YKASPGNLSQ FEDILFGNND MSASIGVVGV KMSAVDQGRQ VGVGYVDSIQ RKLGLCEFPD 180
NDQFSNLEAL LIQIGPKECV LPGGETAGDM GKLRQIIQRG GILITERKKA DFSTKDIYQD 240
LNRLKGGKKG EQMNSAVLPE MENQVAVSSL SAVIKFLELL SDDSNFGQFE LTTDFDSQYM 300

KLDIAAVRAL NLFQGSVEDT TGSQSLAALL NKCKTPQQQR LVNQWIKQPL MDKNRIEERL 360
NLVEAFVEDA ELRQTLQEDL LRRFPDLNRL AKKFQRQAAN LQDCYRLYQG INQLPNVIQA 420
LEKHEGKHQK LLLAVFVTPL TDLRSDFSKF QEMIETTLDM DQVENHEFLV KPSFDPNLSE 480
LREIMNDLEK KMQSTLISAA RDLGLDPGKQ IKLDSSAQFG YYFRVTCKEE KVLRRNNKNS 540
TVDIQKNGVK FTNSKLTSLN EETKNKTEY EEAQDAIVKE IVNISSGYVE PMQTLNDVLA 600
QLDAVVVFAH VSNAPVPYV RPAILEKGQG RIILKASRHA CVEVQDEIAF IPNDVYFEKD 660
KQMFHIITGP NMGGKSTYIR QTVGVILMAQ IGCVPCESA EVSIVDCILA RVGAGDSQLK 720
GVSTFMAEML ETASILRSAT KDSLIIIDEL GRGTSTYDGF GLAWAISEYI ATKIGAFCMF 780
ATHFHELTAL ANQIPTVNNL HVTALTTEET LTMLYQVKKG VCDQSFGIHV AELANFPKHV 840
IECAKQKALE LEEFYIGES QGYDIMEPAA KKCYLEREQG EKIIQEFLSK VKQMPFTEMS 900
EENITIKLKQ LKAEVIKNN SFVNEIISRI KVT

934

Human MSH2 cDNA

```

ggcgggaaac agcttagtggt gtgtgggggtc gcgcattttc ttcaaccagg aggtgaggag 60
gtttcgacat ggcggtgcag ccgaaggaga cgtgcagtt ggagagcgcg gccgaggtcg 120
gcttcgtgcg cttctttcag ggcatgccgg agaagccgac caccacagtg cgccttttcg 180
accggggcga cttctatacg gcgcacggcg aggacgcgct gctggccgcc cgggaggtgt 240
tcaagaccca gggggtgatc aagtacatgg ggccggcagg agcaaagaat ctgcagagtg 300
ttgtgcttag taaaatgaat ttgaaatctt ttgtaaaaga tcttcttctg gttcgtcagt 360
atagagttga agtttataag aatagagctg gaaataaggc atccaaggag aatgattggt 420
atttggcata taaggcttct cctggcaatc tctctcagtt tgaagacatt ctctttggtg 480
acaatgatat gtcagcttcc atttgggttg tgggtgttaa aatgtccgca gttgatggcc 540
agagacaggt tggagttggg tatgtggatt ccatacagag gaaactagga ctgtgtgaat 600
tccctgataa tgatcagttc tccaatcttg aggctctcct catccagatt ggaccaaagg 660
aatgtgtttt acccgaggga gagactgctg gagacatggg gaaactgaga cagataattc 720
aaagaggagg aattctgac acagaaagaa aaaaagctga cttttccaca aaagacattt 780
atcaggacct caaccggttg ttgaaaggca aaaagggaga gcagatgaat agtgctgtat 840
tgccagaaat ggagaatcag gttgcagttt catcactgtc tgcggtaatc aagtttttag 900
aactcttata agatgattcc aactttggac agtttgaact gactactttt gacttcagcc 960
agtatatgaa attggatatt gcagcagtcg gagcccttaa cctttttcag ggttctgttg 1020
aagataccac tggctctcag tctctggctg ccttgctgaa taagtgtaaa accoctcaag 1080
gacaaagact tgtaaccag tggattaagc agcctctcat ggataagaac agaatagagg 1140
agagattgaa tttagtggaa gctttttag aagatgcaga attgaggcag actttacaag 1200
aagatttact tcgtcgattc ccagatctta accgacttgc caagaagttt caaagacaag 1260
cagcaaactt acaagattgt taccgactct atcagggtat aaatcaacta cctaattgta 1320
tacaggctct ggaaaaacat gaaggaaaac accagaaatt attgttggca gtttttgtga 1380
ctcctcttac tgatcttcgt tctgacttct ccaagtttca ggaaatgata gaaacaactt 1440
tagatatgga tcaggtggaa aaccatgaat tcttgtgaaa accttcattt gatcctaata 1500
tcagtgaatt aagagaaata atgaatgact tggaaaagaa gatgcagtcg acattaataa 1560
gtgcagccag agatcttggc ttggaccctg gcaaacagat taaactggat tccagtgcac 1620

```

```

agtttgata ttactttcgt gtaacctgta aggaagaaaa agtccttcgt aacaataaaa 1680
accttagtac tgtagatata cagaagaatg gtgttaaatt taccaacagc aaattgactt 1740
ctttaaatga agagtatacc aaaaataaaa cagaatatga agaagccag gatgccattg 1800
ttaaagaaat tgtcaatatt tcttcaggct atgtagaacc aatgcagaca ctcaatgatg 1860
tgtagctca gctagatgct gttgtcagct ttgctcacgt gtcaaatgga gcacctgttc 1920
catatgtacg accagccatt ttggagaaa gacaaggaa aattatatta aaagcatcca 1980
ggcatgcttg tgtgaagtt caagatgaaa ttgcatttat tcctaataac gtatactttg 2040
aaaaagataa acagatgttc cacatcatta ctggcccaa tatgggaggt aaatcaacat 2100
atattcgaca aactggggtg atagtaetca tggcccaa atgggtgttt gtgccatgtg 2160
agtcagcaga agtgtccatt gtggactgca tcttagcccg agtaggggt ggtgacagtc 2220
aattgaaagg agtctccacg ttcattgctg aaatgttga aactgcttct atcctcaggt 2280
ctgcaacca agattcatta ataatcatag atgaattggg aagaggaact tctacctacg 2340
atggatttgg gttagcatgg gctatatcag aatacattgc aacaaagatt ggtgcttttt 2400
gcatgtttgc aaccatttt catgaactta ctgccttggc caatcagata ccaactgtta 2460
ataatctaca tgtcacagca ctccactg aagagacctt aactatgctt tatcaggtga 2520
agaaagggtg ctgtgatcaa agttttggga ttcattgtgc agagcttctt aatttcccta 2580
agcatgtaat agagtgtgct aaacagaaa ccctggaact tgaggagttt cagtataattg 2640
gagaatcgca aggatatgat atcatggaac cagcagcaaa gaagtgtctat ctggaaaagag 2700
agcaagggtg aaaaattatt caggagtccc tgtccaagg gaaacaaatg ccctttactg 2760
aaatgtcaga agaaaacatc acaataaagt taaaacagct aaaagctgaa gtaatagcaa 2820
agaataatag ctttgtaaat gaaatcattt cacgaataaa agttactacg tgaaaaatcc 2880
cagtaatgga atgaaggtaa tattgataag ctattgtctg taatagtttt atattgtttt 2940
atattaaccc tttttccata gtgttaactg tcagtgtcca tgggctatca acttaataag 3000
atatttagta atattttact ttgaggacat tttcaaagat ttttattttg aaaaatgaga 3060
gctgtaactg aggactgttt gcaattgaca taggcaataa taagtgtatg gctgaatttt 3120
ataaataaaa tcatgtagtt tgtgg 3145

```

human MLH1 protein

```

MSFVAGVIRR LDETVVNRIR AGEVIQRPAN AIKEMIENCL DAKSTSIQVI VKEGGLKLIQ 60
IQDNGTGIRK EDLDIVCERF TTSKLQSFED LASISTYGFR GEALASISHV AHVTITTKTA 120
DGKCAIRASY SDGKLKAPPK PCAGNQGTQI TVEDLFYNIA TRRKALKNPS EEYKILEVV 180
GRYSVHNAGI SFSVKKQGET VADVRTLPA STVDNIRSIF GNAVSRELIE IGCEDKTALF 240
KMNGYISNAN YSVKKCIFLL FINHRLVEST SLRKAETVY AAYLPKNTHP FLYLSLEISP 300
QNVDVNVHPT KHEVHFLHEE SILERVQQHI ESKLLGSNSS RMYFTQTLPL GLAGPSGEMV 360
KSTTSLTSSS TSGSSDKVYA HQMVRTDSRE QKLDNFLQPL SKPLSSQPOA IVTEDKTDIS 420
SGRARQQDEE MLELPAPAEV AAKNQSLEGD TTKGTSEMSE KRGPTSSNPR KHRRESDVE 480
MVEDDSRKEM TAACTPRRRI INLTSLVLSLQ EEINEQGHEV LREMLHNHSF VGCVPQWAL 540
AQHQTKLYLL NTKLSEELF YQILIYDFAN FGVLRLEPA PLFDLAMLAL DSPESGWTEE 600
DGPKEGLAEY IVEFLKKKAE MLADYFSLEI DEEGNLIGLP LLIDNYVPPL EGLPIFILRL 660
ATEVNWDEEK ECFESLSKEC AMFYSIRKQY ISEESTLSGQ QSEVPGSIPN SWKWTVEHIV 720

```

YKALRSHILP PKHFTEDGNI LQLANLPDLY KVFERC

756

Human MLH1 cDNA

```

cttggctcct ctggcgccaa aatgtcgttc gtggcagggg ttattcggcg gctggacgag 60
acagtgggtga accgcatcgc ggcgggggaa gttatccagc ggccagctaa tgctatcaaa 120
gagatgattg agaactgttt agatgcaaaa tccacaagta ttcaagtgat tgttaaagag 180
ggaggcctga agttgattca gatccaagac aatggcaccg ggatcaggaa agaagatctg 240
gatattgtat gtgaaagggt cactactagt aaactgcagt cctttgagga tttagccagt 300
atttctacct atggctttcg aggtgaggct ttggccagca taagccatgt ggctcatgtt 360
actattacaa cgaaaacagc tgatggaaag tgtgcataca gagcaagtta ctacagatgga 420
aaactgaaag cccctcctaa accatgtgct ggcaatcaag ggacccagat cacgggtggag 480
gacctttttt acaacatagc cacgaggaga aaagctttta aaaatccaag tgaagaatat 540
gggaaaattt tggaggttgt tggcaggtat tcagtacaca atgcaggcat tagttttctca 600
gttaaaaaac aaggagagac agtagctgat gttaggacac taccacatgc ctcaaccgtg 660
gacaatatct gctccatctt tggaaatgct gttagtcgag aactgataga aattggatgt 720
gaggataaaa ccctagcctt caaaatgaat ggttacatat ccaatgcaaa ctactcagt 780
aagaagtga tcttcttact cttcatcaac catcgtcttg tagaatcaac ttccttgaga 840
aaagccatag aaacagtgtg tgcagcctat ttgccccaaa acacacaccc attcctgtac 900
ctcagtttag aaatcagtc cagaatgtg gatgttaatg tgcacccac aaagcatgaa 960
gttcacttcc tgcacgagga gagcatcctg gagcgggtgc agcagcacat cgagagcaag 1020
ctcctgggct ccaattcctc caggatgtac ttcacccaga ctttgctacc aggacttgct 1080
ggccctctg gggagatggt taaatccaca acaagtctga cctcgtcttc tacttctgga 1140
agtagtgata aggtctatgc ccaccagatg gttcgtacag attcccggga acagaagctt 1200
gatgcatttc tgcagcctct gagcaaacc cgtccagtc agccccaggc cattgtcaca 1260
gaggataaga cagatatctc tagtggcagg gctaggcagc aagatgagga gatgcttga 1320
ctcccagccc ctgctgaagt ggctgccaaa aatcagagct tggaggggga tacaacaaag 1380
gggacttcag aaatgtcaga gaagagagga cctacttcca gcaacccag aaagagacat 1440
cggaagatt ctgatgtgga aatggtgga gatgattccc gaaaggaaat gactgcagct 1500
tgtaccccc ggagaaggat cattaacctc actagtgttt tgagtctcca ggaagaaatt 1560
aatgagcagg gacatgaggt tctccgggag atgttgcata accactcctt cgtgggctgt 1620
gtgaatctc agtgggcctt ggcacagcat caaaccaagt tataccttct caacaccacc 1680
aagcttagtg aagaactgtt ctaccagata ctcatctatg attttgccaa ttttggtgtt 1740
ctcaggttat cggagccagc accgctcttt gaccttgcca tgcttgccct agatagttca 1800
gagagtggct ggacagagga agatgggtccc aaagaaggac ttgctgaata cattgttgag 1860
tttctgaaga agaaggctga gatgcttgca gactatttct ctttgaaat tgatgaggaa 1920
gggaacctga ttggattacc ccttctgatt gacaactatg tgcccccttt ggagggactg 1980
cctatcttca ttcttcgact agccactgag gtgaattggg acgaagaaaa ggaatgtttt 2040
gaaagcctca gtaaagaatg cgctatgttc tattccatcc ggaagcagta catatctgag 2100
gagtcgaccc tctcaggcca gcagagtga gtgcctggct ccattccaaa ctcttggaag 2160
tggactgtgg aacacattgt ctataaagcc ttgcgctcac acattctgcc tcctaaacat 2220

ttcacagaag atggaaatat cctgcagctt gctaacctgc ctgatctata caaagtcttt 2280
gagaggtggt aaatatggtt atttatgcac tgtgggatgt gttcttcttt ctctgtattc 2340
cgatacaaa gttgttatca aagtgtgata tacaaagtgt accaacataa gtgttggtag 2400
cacttaagac ttatacttgc ctctctgatg tattccttta tacacagtgg attgattata 2460
aataaataga tgtgtcttaa cata 2484

```

hPMS2-134 protein

```

MKQLPAATVR LLSSSQIITS VVSVKELIE NSLDAGATSV DVKLENYGFD KIEVRDNGEG 60
IKAVDAPVMA MKYYTSKINS HEDLENLTTY GFRGEALGSI CCIAEVLITT RTAADNFSTQ 120
YVLDGSGHIL SQK 133

```

hPMS2-134 cDNA

cgaggcggat cggtgtgtgc atccatggag cgagctgaga gctcgagtag agaacctgct 60
 aaggccatca aacctattga tcggaagtca gtccatcaga ttgctctggt gcaggtggta 120
 ctgagctctaa gcactgcggt aaaggagtta gtagaaaaca gtctggatgc tgggtgccact 180
 aatattgata taaagcttaa ggactatgga gtggatctta ttgaagtttc agacaatgga 240
 tgtggggtag aagaagaaaa cttcgaaggc ttaactctga aacatcacac atctaagatt 300
 caagagtttg ccgacctaac tcagggtgaa acttttggtt ttcgggggga agctctgagc 360
 tcactttgtg cactgagcga tgtcaccatt tctacctgcc acgcatcgcc gaagggttga 420
 acttga 426

hMSH6 (human cDNA) DEONERING U28946

MSRQSTLYSFFPKSPALSDANKASARASREGGRAAAAPGASPS
 GGDAWSEAGPGPRPLARSASPPKAKNLNGGLRRSVAPAAPTSCDFSPGDLVWAKMEG
 YPWWPCLVYNHPFDGTFIREKGKSVRVHVQFFDSDPTRGWVSKRLLKPYTGSKSKEAQ
 KGGHFYSAPPEILRAMQRADEALNKDKIKRLELAVCDEPSEPEEEEEMEVGTTYVTDK
 SEEDNEIESEEEVQPKTQGSRRSSRQIKRRRVIDSESDIGGSDVEFKPDTKEEGSSD
 EISSGVGDSESEGLNSPVKVARKRKRMVTGNGSLKRKSSRKETPSATKQATSISSETK
 NTLRAFSAPQNSAQAHVSGGGDDSSRPTVWYHETLEWLKEEKRRDEHRRRDPDHPDFD
 ASTLYVPEDFLNSCTPGMRKWWQIKSQNFDLVICYKVGKFYELYHMDALIGVSELGLV
 FMKGNWAHSGFPEIAFGRYSDSLVQKGYKVARVEQTETPEMMEARCRKMAHISKYDRV
 VRREICRIITKGTQTSVLEGDPSSENYSKYLLSLKEKEEDSSGHTRAYGVCVFDTSLG
 KFFIGQFSDDRHCSRFRFTLVAHYPPVQVLFEXGNLSKETTKTILKSSLSCSLQEGLI
 PG SQFWDASKTLRTLLEEEYFREKLSDGIGVMLPQVLKGMTSESDSIGLTPGEKSELALS
 ALGGCVFYLKKCLIDQELLSMANFEEYIPLDSDTVSTTRSGAIFTKAYQRMVLDVAVTL
 NNLEIFLNGTNGSTEGTLLEERVDTCHTPFGKRLKQWLCAPLCNHYAINDRLDAIEDL
 MVVPDKISEVVELLKKLPDLERLLSKIHNVGSPKLSQNHPSRAIMYEETTSYKKKII
 DFLSALEGFKVMCKIIGIMEEVADGFKSKI LKQVLSLQTKNPEGRFPDLTVELNRWD
 AFDHEKARKTGLITPKAGFDSYDQALADIRENEQSLLEYLEKQRNRIGCRTIVYWG
 GRNRYQLEIPENFTTRNLPEEYELKSTKKGCKRYWTKTIEKKLANLINAERDVS
 LK DCMRRLFYNFKNYKDWQSAVECIAVLVDVLLCLANYSRGGDGPMCRPVILLPEDTPPF

LELKGSRRPCITKTFFGDDFIPNDILIGCEEEQENGKAYCVLVTGPNMGGKSTLMRQ
 AGLLAVMAQMGCVPAEVCRLTPIDRVFTRLGASDRIMSGESTFFVELSETASILMHA
 TAHSLLVDELGRGTATFDGTAIANAVVKELAETIKCRTLFSTHYHSLVEDYSQNVAV
 RLGHMACMVENECEDPSETITFLYKFIKGACPKSYGFNAARLANLP EEVIQKGRKA
 REF EKMNQSLRLFREVCLASERSTVDAEAVHKLLTLIKEL"

hPMSR2 (human cDNA) DEPONERING U38964

```

1  ggcgcctccta cctgcaagtg gctagtgcca agtgctgggc cgccgctcct gccgtgcatg
61  ttggggagccc agtacatgca ggtgggctcc acacggagag gggcgagac ccggtgacag
121 ggctttacct ggtacatcgg catggcgcaa ccaaagcaag agaggggtgg ccgtgccaga
181 caccaacggt cggaaccgc cagacacca cggtcggaaa ccgccaagac accaacgctc
241 ggaaaccgcc agacaccaac gctcggaaac cgccagacac caaggctcgg aatccacgcc
301 aggccacgac ggagggcgac tacctccctt ctgaccctgc tgctggcgtt
cggaaaaaac
361 gcagtcgggt gtgctctgat tggtcaggc tctttgacgt caccgactcg acctttgaca
421 gagccactag gcgaaaagga gagacgggaa gtattttttc cgccccgccc ggaaagggtg
481 gagcacaacg tcgaaaagcag ccgttgggag ccagaggagg ggggcgctg tgggagccgt
541 ggagggaaact tcccagtc cggagggcga tccggtgttg catccttggg gcgagctgag
601 aactcgagta cagaacctgc taaggccatc aaacctattg atcgggaagtc agtccatcag
661 atttgctctg ggccgggtgt accgaagtcta aggccgaatg cggtagaagg gttagtagaa
721 aacagtctgg atgctgggtg cactaatgtt gatctaaagc ttaaggacta tggagtggat
781 ctcatgaaag ttccaggcaa tggatgtggg gtagaagaag aaaacttcga aggccttact
841 ctgaaacatc acacatgtaa gattcaagag ttgcccagc taactcaggt ggaaactttt
901 ggctttcggg gggaagctct gagctcactt tgtgcactga gtgatgtcac catttctacc
961 tgccgtgtat cagcgaaggt tgggactcga ctggtgtttg atcactatgg gaaaatcatc
1021 cagaaaaacc cctacccccg ccccagaggg atgacagtca gcgtgaagca gttattttct
1081 acgctacctg tgcaccataa agaatttcaa aggaatatta agaagaaacg tgcctgtctc
1141 ccttcgcct tctgccgtga ttgtcagttt cctgaggcct cccagccat gcttctgta
1201 cagcctgtag aactgactcc tagaagtacc ccacccacc cctgctcctt ggaggacaac
1261 gtgatcactg tattcagctc tgtcaagaat ggtccagggt ctctagatg atctgcacaa
1321 atggttcctc tctccttcc tgatgtctgc cattagcatt ggaataaagt tcttgcctgaa
1381 aatcaaaaaa aaaaaaaaaa aaaaaaaa

```

hPMSR2 (human protein) DEPONERING U38964

MAQPKQERVARARHQRSETARHQRSETAKTPTLGNROTPTLGNR

QTPRLGIHARPRRRATSSLTLLLAFGKNAVRICALIGPSLTSRTRPLTEPLGEKERR

EVFFPPRPERVEHNVESSRWEPRRRGACGSRGGNFPSPRGGSGVASLERAENSSTEPA

KAIKPIDRKSVMQICSGPVVPSLRPNAVKELVENS LDAGATNVDLKLKDYGVDLIEVS

GNGCGVEEENFEGFTLKHHTCKIQEFADLTQVETFGFRGEALSSLCALSDVTISTCRV

SAKVGTRLVFDHYGKIIQKTPYPRPGMTVSVKQLFSTLPVHHKEFQRNIIKKKRAFPP

FAFCRDCQFPEASPAMLPVQPVVELTPRSTPPHPCSLEDNVITVFSSVKNGPGSSR

5

hPMSR3 (human cDNA) DEPONERING U38979

```

1  ttttttagaaa ctgatgttta tttccatca accatttttc catgctgctt aagagaatat
61  gcaagaacag cttaagacca gtcagtgggt gctcctaccc attcagtggc ctgagcagtg
121 ggagagctgca gaccagtctt ccgtggcagg ctgagcgctc cagtcttcag tagggaattg

```

```

181 ctgaataggc acagaggcca cctgtacacc ttcagaccag tctgcaacct caggctgagt
241 agcagtgaac tcaggagcgg gagcagtcga ttcaccctga aattcctcct tggctactgc
301 cttctcagca gcagccctgct cttctttttc aatctcttca ggatctctgt agaagtacag
361 atcaggcatg acctcccatg ggtgttcacg ggaatggtg ccacgcatgc gcagaacttc
421 ccgagccagc atccaccaca ttaaacccac tgagtgaact ccttgttgtg tgcattgggt
481 ggcaatgtcc acatagcgca gaggagaatc tgtgttacac agcgcaatgg taggtagggt
541 aacataagat gcctccgtga gaggcgaagg ggcggcgagg cccgggcttg gccctgatgt
601 gtccctggcg gcctagacta ggccgtcgct gtatggtgag cccagggag gcggatctgg
661 gccccagaa ggacacccgc ctggatttgc cccgtagccc ggccggggcc cctcgggagc
721 agaacagcct tggtaggtg gacaggaggg gacctcgcga gcagacgcgc gcgccagcga
781 cagcagcccc gccccggcct ctcgggagcc ggggggcaga ggctgcggag cccagggagg
841 gtctatcagc cacagtctct gcatgtttcc aagagcaaca ggaatgaac acattgcagg
901 ggccagtgtc attcaaagat gtggctgtgg atttcacca ggaggagtgg cggcaactgg
961 accctgatga gaagatagca tacggggatg tgatgttggg gaactacagc catctagttt
1021 ctgtggggtg tgattatcac caagccaaac atcatcatgg agtggagggtg aaggaaagtgg
1081 agcaggggag ggagccgtgg ataattgga gtagaatttc atgtcaacat agtccagaac
1141 ctgctaaggc catcaaacct attgatcgga agtcagtcga tcagatttgc tctgggccag
1201 tggtagctag tctaagcact gcagtgaagg agttagtaga aaacagctcg gatgtgtgtg
1261 ccactaatat tgatctaaag ctttaaggact atggagtggg tctcattgaa gtttcagaca
1321 atggatgtgg ggtagaagaa gaaaactttg aaggcttaat ctctttcagc tctgaacat
1381 cacacatgta agattcaaga gtttgcgcac ctaactgaag ttgaaacttt cggttttcag
1441 ggggaagctc tgagctcact gtgtgcactg agcgatgtca ccattttctac ctgccacgcg
1501 ttggtgaagg ttgggactcg actggtgttt gatcacgatg ggaaatcat ccaggaaacc
1561 ccctaccccc accccagagg gaccacagtc agcgtgaagc agttattttc tacgtacct
1621 gtgcgcata aggaatttca aaggaatatt aagaagacgt gcctgtcttc ccttcgcctt
1681 ctgcccgtgt tgcagtttc ctgaggcctc cccagccatg cttcctgtac agcctgcaga
1741 actgtgagtc aattaaacct cttttcttca taaattaaaa aaaaa

```

hPMSR3 (human protein) DEONERING U38979

MCPWRPRLGRRCMVSPREADLGPQKDTRLDLPRSPARAPREQNS

LGEVDRRGPREQTRAPATAAPPRPLGSRGAEEAEPOEGLSATVSACFQEQQEMNTLQG

PVSFKDVAVDFTQEEWRQLDPDEKIAYGDMLENYSHLVSVGYDYHQAHHHGVEVKE

VEQGEFPWIMEGEFPCQHSPEPAKAIKPIDRKSVHQICSGPVVLSLSTAVKELVENS

DAGATNIDLKLDYGVDLIEVSDNGCGVEEENFEGLISFSSETSHM"

hPMSL9 (human cDNA) DEONERING NM_005395

```

1 atgtgtcctt ggccggcctag actaggccgt cgctgtatgg tgagcccccag ggaggcggat
61 ctgggcccc agaaggacac ccgcctggat ttgcccgtga gcccgcccg ggcccctcgg
121 gagcagaaca gccttggtga ggtggacagg aggggacctc gcgagcagac gcgcgcgcca
181 gcgacagcag ccccgcccc gcctctcggg agccgggggg cagaggtctg ggagccccag
241 gaggtctat cagccacagt ctctgcatgt ttccaagagc aacaggaaat gaacacattg
301 caggggccag tgcattcaa agatgtggct gtggatttca cccaggagga gtggcgccaa
361 ctggaccctg atgagaagat agcatacggg gatgtgatgt tggagaacta cagccatcta
421 gtttctgtgg ggtatgatta tcaccaagcc aaacatcatc atggagtggg ggtgaaggaa
481 gtggagcagg gagaggagcc gtggataatg gaaggtgaat ttccatgtca acatagtcca
541 gaacctgcta aggccatcaa acctattgat cggaagtcag tccatcagat ttgctctggg
601 ccagtggtag tgagtctaag cactgcagtg aaggagttag tagaaaacag tctggatgct

```

```

661 ggtgccacta atattgatct aaagcttaag gactatggag tggatctcat tgaagtttca
721 gacaatggat gtggggtaga agaagaaaac tttgaaggct taatctcttt cagctctgaa
781 acatcacaca tgtaa

```

hPMSL9 (human protein) DEONERING NM_005395

MCPWRPRLGRRCMVSPREADLGPQKDTRLDLPRSPARAPREQNS

LGEVDRRGPREQTRAPATAAPPRPLGSRGAEEAEPOEGLSATVSACFQEQQEMNTLQG

PVSFKDVAVDFTQEEWRQLDPDEKIAYGDMLENYSHLVSVGYDYHQAHHHGVEVKE

VEQGEFPWIMEGEFPCQHSPEPAKAIKPIDRKSVHQICSGPVVLSLSTAVKELVENS

DAGATNIDLKLDYGVDLIEVSDNGCGVEEENFEGLISFSSETSHM"

SEKVENSLISTE

- <110> Nicolaides, Nicholas
Sass, Philip
Kinzler, Kenneth
5 Grasso, Luigi
Vogelstein, Bert
- <120> Methods for generating hypermutable
yeast
10
- <130> 01107.00099
- <140> PCT/US01/05447
<141> 2001-02-21
15
- <150> 60/184,336
<151> 2000-02-23
- <160> 26
20
- <170> FastSEQ for Windows Version 3.0
- <210> 1
<211> 30
25 <212> DNA
<213> Artificial Sequence
- <220>
<223> PCR primer
30
- <400> 1
acgcatatgg agcgagctga gagctcgagt
30
- <210> 2
35 <211> 75

44

<212> DNA

<213> Artificial Sequence

<220>

5 <223> PCR primer

<400> 2

gaattcttat cacgtagaat cgagaccgag gagaggggta gggataggct

taccagttcc 60

10 aaccttcgcc gatgc

75

<210> 3

<211> 27

15 <212> DNA

<213> Artificial Sequence

<220>

<223> PCR primer

20

<400> 3

acgcatatgt gtccttggcg gcctaga

27

25 <210> 4

<211> 75

<212> DNA

<213> Artificial Sequence

30 <220>

<223> PCR primer

<400> 4

gaattcttat tacgtagaat cgagaccgag gagaggggta gggataggct

35 taccatgtg 60

tgatgtttca gagct

75

<210> 5

5 <211> 3218

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 5

```

aaataggaat gtgatacctt ctattgcatg caaagatagt gtaggaggcg
ctgctattgc      60
caaagacttt tgagaccgct tgctgtttca ttatagttga ggagttctcg
aagacgagaa      120
attagcagtt ttcggtgttt agtaatcgcg ctagcatgct aggacaattt
aactgcaaaa      180
ttttgatacg atagtgatag taaatggaag gtaaaaaataa catagaccta
tcaataagca      240
atgtctctca gaataaaagc acttgatgca tcagtgggta acaaaattgc
tgcaggtgag      300
atcataatat cccccgtaaa tgctctcaaa gaaatgatgg agaattccat
cgatcggaat      360
gctacaatga ttgatattct agtcaaggaa ggaggaatta aggtacttca
aataacagat      420
aacggatctg gaattaataa agcagacctg ccaatcttat gtgagcgatt
cacgacgtcc      480
aaattacaaa aattcgaaga tttgagtcag attcaaacgt atggattccg
aggagaagct      540
ttagccagta tctcacatgt ggcaagagtc acagtaacga caaaagttaa
agaagacaga      600
tgtgcatgga gagtttcata tgcagaaggt aagatgttgg aaagcccca
acctgttgct      660
ggaaaagacg gtaccacgat cctagttgaa gacctttttt tcaatattcc
ttctagatta      720
agggccttga ggtcccataa tgatgaatac tctaaaatat tagatgttgt
cgggcgatac      780
gccattcatt ccaaggacat tggcttttct tgtaaaaagt tcggagactc
taattattct      840
ttatcagtta aaccttcata tacagtccag gataggatta ggactgtgtt
caataaatct      900
gtggcttcga atttaattac ttttcatatc agcaaagtag aagattttaa
cctggaaagc      960
gttgatggaa aggtgtgtaa tttgaatttc atatccaaaa agtccatttc
attaattttt     1020
ttcattaata atagactagt gacatgtgat cttctaagaa gagctttgaa
cagcgtttac     1080

```

10

```

tccaattatc tgccaaaggg cttcagacct tttatttatt tgggaattgt
tatagatccg 1140
gcggtgtgtg atgttaacgt tcacccgaca aagagagagg ttcgtttcct
gagccaagat 1200
gagatcatag agaaaatcgc caatcaattg cagccgaat tatctgccat
tgatacttca 1260
cgtactttca aggtttcttc aatttcaaca aacaagccag agtcattgat
accatttaat 1320
gacaccatag aaagtgatag gaataggaag agtctccgac aagcccaagt
ggtagagaat 1380
tcataatcga cagccaatag tcaactaagg aaagcgaaaa gacaagagaa
taaactagtc 1440
agaatagatg cttcacaaagc taaaattacg tcatttttat cctcaagtca
acagttcaac 1500
tttgaaggat cgtctacaaa gcgacaactg agtgaacca aggtaacaaa
tgtaagccac 1560
tcccaagagg cagaaaagct gacactaaat gaaagcgaac aaccgcgtga
tgccaataca 1620
atcaatgata atgacttgaa ggatcaacct aagaagaaac aaaagttggg
ggattataaa 1680
gttccaagca ttgccgatga cgaaaagaat gcactcccg a tttcaaaaga
cgggtatatt 1740
agagtaccta aggagcaggt taatgttaat cttacgagta tcaagaaatt
gcgtgaaaaa 1800
gtagatgatt cgatacatcg agaactaaca gacatttttg caaatttgaa
ttacgttggg 1860
gttgtgatag aggaaagaag attagccgct attcagcatg acttaaagct
ttttttaata 1920
gattacggat ctgtgtgcta tgagctattc tatcagattg gtttgacaga
cttcgcaaac 1980
tttggttaaga taaacctaca gagtacaaat gtgtcagatg atatagtttt
gtataatctc 2040
ctatcagaat ttgacgagtt aaatgacgat gcttccaaag aaaaaataat
tagtaaaata 2100
tgggacatga gcagtatgct aaatgagtag tattccatag aattggtgaa
tgatgggtcta 2160
gataatgact taaagtctgt gaagctaaaa tctctaccac tactttttaa
aggctacatt 2220
ccatctctgg tcaagttacc attttttata tatcgcttgg gtaaagaagt
tgattgggag 2280
gatgaacaag agtgtctaga tgggtatttta agagagattg cattactcta
tatactgat 2340
atggttccga aagtcgatac actcgatgca tcgttgtcag aagacgaaaa
agcccagttt 2400
ataaatagaa aggaacacat atcctcatta ctagaacacg ttctcttccc
ttgtatcaaa 2460
cgaagggttc tggeccctag acacattctc aaggatgtcg tggaaatagc
caaccttcca 2520
gatctataca aagtttttga gaggtgttaa ctttaaaacg ttttggctgt
aataccaaag 2580
ttttgttta ttctctgagt gtgatttgtt ttcatttgaa agtgtatgcc
ctttccttta 2640
acgattcatc cgcgagattt caaaggatat gaaatatggt tgcagttagg
aaagtatgtc 2700
agaaatgtat attcggattg aaactcttct aatagttctg aagtcacttg
gttcogtatt 2760
gttttcgtcc tcttctcaa gcaacgattc ttgtctaagc ttattcaacg
gtaccaaaga 2820

```

```

cccgagtcct tttatgagag aaaacatttc atcatttttc aactcaatta
tcttaatatc 2880
atthttgtagt atthttgaaa caggatggta aaacgaatca cctgaatcta
gaagctgtac 2940
cttgteccat aaaagtttta atttactgag cctttcggtc aagtaaaacta
gtttatctag 3000
ttttgaaccg aatattgtgg gcagatttgc agtaagttca gttagatcta
ctaaaagtgt 3060
tttgacagca gccgattcca caaaaatttg gttaaaggag atgaaagaga
cctcgcgct 3120
aatggtttgc atcaccatcg gatgtctgtt gaaaaactca ctttttgcac
ggaagttatt 3180
aacaataaga ctaatgatta ccttagaata atgtataa
3218

```

<210> 6

<211> 3056

<212> DNA

5

<213> Mus musculus

<400> 6

```

gaattccggt gaaggctctg aagaatttcc agattcctga gtatcattgg
aggagacaga 60
taacctgtcg tcaggtaaacg atgggtgtata tgcaacagaa atgggtgttc
ctggagacgc 120
gtcttttccc gagagcggca ccgcaactct cccgcgggtga ctgtgactgg
aggagtcctg 180
catccatgga gcaaacggaa ggcgtgagta cagaatgtgc taaggccatc
aagcctattg 240
atgggaagtc agtccatcaa atttgttctg ggcagggtgat actcagttta
agcaccgctg 300
tgaaggagtt gatagaaaat agtgtagatg ctgggtgctac tactattgat
ctaaggctta 360
aagactatgg ggtggacctc attgaagttt cagacaatgg atgtggggta
gaagaagaaa 420
actttgaagg tctagctctg aaacatcaca catctaagat tcaagagttt
gccgacctca 480
cgcagggtga aactttcggc tttcgggggg aagctctgag ctctctgtgt
gcactaagtg 540
atgtcactat atctacctgc caccgggtctg caagcgttgg gactcgactg
gtgtttgacc 600
ataatgggaa aatcaccacg aaaactccct acccccgacc taaaggaacc
acagtcagtg 660
tgcagcactt attttataca ctaccctggt gttacaaaga gtttcagagg
aacattaaaa 720
aggagtattc caaatgggtg cagggtcttac aggcgtactg tatcatctca
gcaggcgtcc 780
gtgtaagctg cactaatcag ctccggacagg ggaagcggca cgctgtggtg
tgcacaagcg 840
gcacgtctgg catgaaggaa aatatcgggt ctgtgtttgg ccagaagcag
ttgcaaagcc 900
tcattccttt tgttcagctg cccctagtg acgtgtgtg tgaagagtac
ggcctgagca 960
cttcaggacg ccacaaaacc ttttctacgt ttcgggcttc atttcacagt
gcacgcacgg 1020
cgccgggagg agtgcaacag acaggcagtt tttcttcac aatcagaggc
cctgtgaccc 1080
agcaaaggtc tctaagcttg tcaatgaggt tttatcacat gtataaccgg
catcagtacc 1140

```

catttgctgt ccttaacgtt tccgttgact cagaatgtgt ggatattaat
 gtaactccag 1200
 ataaaaggca aattctacta caagaagaga agctattgct ggccgtttta
 aagacctcct 1260
 tgataggaat gtttgacagt gatgcaaaca agcttaatgt caaccagcag
 ccactgctag 1320
 atgttgaagg taacttagta aagctgcata ctgcagaact agaaaagcct
 gtgccaggaa 1380
 agcaagataa ctctccttca ctgaagagca cagcagacga gaaaagggta
 gcatccatct 1440
 ccaggctgag agaggccttt tctcttcctc ctactaaaga gatcaagtct
 aggggtccag 1500
 agactgctga actgacacgg agttttccaa gtgagaaaag gggcgtgtta
 tctctctatc 1560
 cttcagacgt catctcttac agaggcctcc gtggctcgca ggacaaattg
 gtgagtccca 1620
 cggacagccc tgggtgactgt atggacagag agaaaataga aaaagactca
 gggctcagca 1680
 gcacctcagc tggctctgag gaagagttca gacccccaga agtggccagt
 agcttttagca 1740
 gtgactataa cgtgagctcc ctagaagaca gaccttctca ggaaaccata
 aactgtgggtg 1800
 acctggactg ccgtctctca ggtacaggac agtccttgaa gccagaagac
 catggatatc 1860
 aatgcaaagc tctacctcta gctcgtctgt caccacaaaa tgccaagcgc
 ttcaagacag 1920
 agaaaagacc ctcaaagtgc aacatttctc aaagattgcc tggctcctcag
 agcacctcag 1980
 cagctgaggt cgatgtagcc ataaaaatga ataagagaat cgtgctcctc
 gagttctctc 2040
 tgagttctct agctaagcga atgaagcagt tacagcacct aaaggcgag
 aacaaacatg 2100
 aactgagtta cagaaaattt agggccaaga tttgccctgg agaaaacca
 gcagcagaag 2160
 atgaactcag aaaagagatt agtaaactga tgtttgcaga gatggagatc
 ttgggtcagt 2220
 ttaacctggg atttatagta accaaactga aagaggacct ctctctgggtg
 gaccagcatg 2280
 ctgcggatga gaagtacaac tttgagatgc tgcagcagca caccgtgctc
 caggcgagca 2340
 ggctcatcac accccagact ctgaacttaa ctgctgtcaa tgaagctgta
 ctgatagaaa 2400
 atctggaaat attcagaaaag aatggctttg actttgtcat tgatgaggat
 gctccagtca 2460
 ctgaaagggc taaattgatt tccctaccaa ctagtaaaaa ctggaccttt
 ggaccccaag 2520
 atatagatga actgatcttt atgttaagtg acagccctgg ggtcatgtgc
 cggccctcac 2580
 gagtcagaca gatgtttgct tccagagcct gtcggaagtc agtgatgatt
 ggaacggcgc 2640
 tcaatgcgag cgagatgaag aagctcatca cccacatggg tgagatggac
 caccctgga 2700
 actgccccca cggcaggcca accatgaggc acgttgccaa tctggatgtc
 atctctcaga 2760
 actgacacac cccttgtagc atagagttta ttacagattg ttcggtttgc
 aaagagaagg 2820
 ttttaagtaa tctgattatc gttgtacaaa aattagcatg ctgctttaat
 gtactggatc 2880

catttaaaag cagtgttaag gcaggcatga tggagtgttc ctctagctca
 gctacttggg 2940
 tgatccggtg ggagctcatg tgagcccagg actttgagac cactccgagc
 cacattcatg 3000
 agactcaatt caaggacaaa aaaaaaaga tatttttgaa gccttttaaa aaaaaa
 3056

<211> 2771

<212> DNA

<213> Homo sapiens

5

<400> 7

```

cgaggcggat cgggtgttgc atccatggag cgagctgaga gctcgagtac
agaacctgct      60
aaggccatca aacctattga tcggaagtca gtccatcaga tttgctctgg
gcagggtgta     120
ctgagttctaa gcactgcggt aaaggagtta gtagaaaaca gtctggatgc
tggtgccact     180
aatattgatc taaagcttaa ggactatgga gtggatctta ttgaagtttc
agacaatgga     240
tgtggggtag aagaagaaaa cttcgaaggc ttaactctga aacatcacac
atctaagatt     300
caagagtttg ccgacctaac tcaggttgaa acttttggct ttcgggggga
agctctgagc     360
tcactttgtg cactgagcga tgtcaccatt totacctgcc acgcatcggc
gaaggttgga     420
actcgactga tgtttgatca caatgggaaa attatccaga aaaccccta
ccccgcgcc     480
agagggaacca cagtcagcgt gcagcagtta tttccacac tacctgtgcg
ccataaggaa     540
tttcaaagga atattaagaa ggagtatgcc aaaatggtcc aggtcttaca
tgcatactgt     600
atcatttcag caggcatccg tgtaagttgc accaatcagc ttggacaagg
aaaacgacag     660
cctgtggtat gcacaggtgg aagccccagc ataaaggaaa atatcggttc
tgtgtttggg     720
cagaagcagt tgcaaagcct cattcctttt gttcagctgc cccctagtga
ctcgtgtgt     780
gaagagtacg gtttgagctg ttcggatgct ctgcataatc ttttttacat
ctcaggtttc     840
atttcacaat gcacgcatgg agttggaagg agttcaacag acagacagtt
tttctttatc     900
aaccggcggc cttgtgaccc agcaaaggtc tgcagactcg tgaatgaggt
ctaccacatg     960
tataatcgac accagtatcc atttgttgtt cttaacattt ctggtgatcc
agaatgcgtt    1020
gatataaatg ttactccaga taaaaggcaa attttgctac aagaggaaaa
gcttttgttg    1080
gcagttttta agacctcttt gataggaatg tttgatagtg atgtcaacaa
gctaaatgtc    1140
agtcagcagc cactgctgga tgttgaaggc aacttaataa aaatgcatgc
agcggatttg    1200
gaaaagccca tggtagaaaa gcaggatcaa tccccctcat taaggactgg
agaagaaaaa    1260
aaagacgtgt ccatttccag actgcgagag gccttttctc ttcgtcacac
aacagagaac    1320
aagcctcaca gcccaaagac tccagaacca agaaggagcc ctctaggaca
gaaaaggggt    1380

```

```

atgctgtctt ctagcacttc aggtgccatc tctgacaaag gcgtcctgag
acctcagaaa 1440
gaggcagtg gttccagtca cggacccagt gaccctacgg acagagcgga
ggtggagaag 1500
gactcggggc acggcagcac ttccgtggat tctgaggggt tcagcatccc
agacacgggc 1560
agtcactgca gcagcgagta tgccggccagc tccccagggg acagggggctc
gcaggaacat 1620
gtggactctc aggagaaagc gcctgaaact gacgactctt tttcagatgt
ggactgccat 1680
tcaaaccagg aagataccgg atgtaaattt cgagtcttgc ctcagccaac
taatctcgca 1740
accccaaaac caaagcgttt taaaaaagaa gaaattcttt ccagttctga
catttgtcaa 1800
aagttagtaa atactcagga catgtcagcc tctcagggtg atgtagctgt
gaaaattaat 1860
aagaaagttg tgcccctgga cttttctatg agttcttttag ctaaacgaat
aaagcagtta 1920
catcatgaag cacagcaaag tgaaggggaa cagaattaca ggaagtttag
ggcaaagatt 1980
tgtcctggag aaaatcaagc agccgaagat gaactaagaa aagagataag
taaacgatg 2040
tttgacagaa tggaaatcat tggtcagttt aacctgggat ttataataac
caaactgaat 2100
gaggatatct tcatagtggg ccagcatgcc acggacgaga agtataactt
cgagatgctg 2160
cagcagcaca ccgtgctcca ggggcagagg ctcatagcac ctcagactct
caacttaact 2220
gctgttaatg aagctgttct gatagaaaat ctggaaatat ttagaaagaa
tggtcttgat 2280
tttgttatcg atgaaaatgc tccagtcact gaaagggcta aactgatttc
cttgccaact 2340
agtaaaaact ggaccttcgg accccaggac gtcgatgaac tgatcttcat
gctgagcgac 2400
agccctgggg tcatgtgccg gccttcccga gtcaagcaga tgtttgccctc
cagagcctgc 2460
cggaagtccg tgatgattgg gactgctctt aacacaagcg agatgaagaa
actgatcacc 2520
cacatggggg agatggacca cccctgggaa tgtcccatg gaaggccaac
catgagacac 2580
atcgccaacc tgggtgtcat ttctcagaac tgaccgtagt cactgtatgg
aataattggt 2640
tttatcgag atttttatgt tttgaaagac agagtcttca ctaacctttt
ttgttttaaa 2700
atgaaacctg ctacttaaaa aaaatacaca tcacacccat ttaaaagtga
tcttgagaac 2760
cttttcaaac c
2771

```

<210> 8

<211> 3063

<212> DNA

5

<213> Homo sapiens

<400> 8

```

ggcacgagtg gctgcttgcg gctagtggat ggtaattgcc tgccctcgcg
tagcagcaag 60
ctgctctggt aaaagcgaaa atgaaacaat tgccctgcggc aacagttcga
ctcctttcaa 120

```

```

gtttctcagat catcacttcg gtggtcagtg ttgtaaaaga gcttattgaa
aactccttgg 180
atgctgggtgc cacaagcgta gatgttaaac tggagaacta tggatttgat
aaaattgagg 240
tgcgagataa cggggaggggt atcaaggctg ttgatgcacc tgtaatggca
atgaagtact 300
acacctcaaa aataaatagt catgaagatc ttgaaaattt gacaacttac
ggttttcgtg 360
gagaagcctt ggggtcaatt tgttgtatag ctgagggttt aattacaaca
agaacggctg 420
ctgataattt tagcaccag tatgttttag atggcagtg ccacatactt
tctcagaaac 480
cttcacatct tggccaaggt acaactgtaa ctgctttaag attatttaag
aatctacctg 540
taagaaagca gttttactca actgcaaaaa aatgtaaaga tgaaataaaa
aagatccaag 600
atctctcat gagctttgggt atccttaaac ctgacttaag gattgtcttt
gtacataaca 660
aggcagttat ttggcagaaa agcagagtat cagatcaca gatggctctc
atgtcagttc 720
tggggactgc tgttatgaac aatatggaat cctttcagta ccactctgaa
gaatctcaga 780
tttatctcag tggatttctt ccaaagtgtg atgcagacca ctctttcact
agtctttcaa 840
caccagaaag aagtttcatc ttcataaaca gtcgaccagt acatcaaaaa
gatatcttaa 900
agttaatccg acatcattac aatctgaaat gcctaaagga atctactcgt
ttgtatcctg 960
ttttctttct gaaaatcgat gttcctacag ctgatgttga tgtaaattta
acaccagata 1020
aaagccaagt attattacaa aataaggaat ctgttttaat tgctcttgaa
aatctgatga 1080
cgacttggtt tggaccatta cctagtacaa attcttatga aaataataaa
acagatgttt 1140
cgcgagctga catcgttctt agtaaaacag cagaaacaga tgtgcttttt
aataaagtg 1200
aatcatctgg aaagaattat tcaaagtgtg atacttcagt cattccattc
caaaatgata 1260
tgcataatga tgaatctgga aaaaacactg atgattgttt aaatcaccag
ataagtattg 1320
gtgactttgg ttatggtcac tgtagtagtg aaatttctaa cattgataaa
aacactaaga 1380
atgcatttca ggacatttca atgagtaatg tatcatggga gaactctcag
acggaatata 1440
gtaaaacttg ttttataagt tccgttaagc acaccagtc agaaaatggc
aataaagacc 1500
atatagatga gagtggggaa aatgaggaag aagcaggtct tgaaaactct
tcggaaattt 1560
gtgcagatga gtggagcagg ggaaatatac ttaaaaattc agtgggagag
aatattgaac 1620
ctgtgaaaat tttagtgcct gaaaaaagtt taccatgtaa agtaagtaat
aataattatc 1680
caatccctga acaaatgaat cttaatgaag attcatgtaa caaaaaatca
aatgtaatag 1740
ataataaatc tggaaaagtt acagcttatg atttacttag caatcgagta
atcaagaaac 1800
ccatgtcagc aagtgtcttt tttgttcaag atcatcgctc tcagtttctc
atagaaaatc 1860

```

```

ctaagactag ttttagaggat gcaacactac aaattgaaga actgtggaag
acattgagtg 1920
aagaggaaaa actgaaatat gaagagaagg ctactaaaga cttggaacga
tacaatagtc 1980
aaatgaagag agccattgaa caggagtcac aaatgtcact aaaagatggc
agaaaaaaga 2040
taaaacccac cagcgcatgg aatttggccc agaagcaca gttaaaaacc
tcattatcta 2100
atcaacccaa acttgatgaa ctcccttcagt cccaaattga aaaaagaagg
agtcaaaata 2160
ttaaaatggg acagatcccc ttttctatga aaaacttaaa aataaatttt
aagaaacaaa 2220
acaaagttag cttagaagag aaggatgaac cttgcttgat ccacaatctc
agggttcctg 2280
atgcatggct aatgacatcc aaaacagagg taatgttatt aaatccatat
agagtagaag 2340
aagccctgct atttaaaaga cttcttgaga atcataaaact tctgcagag
ccactggaaa 2400
agccaattat gttaacagag agtcttttta atggatctca ttatttagac
gttttatata 2460
aaatgacagc agatgaccaa agatacagtg gatcaactta cctgtctgat
cctcgtctta 2520
cagcgaatgg tttcaagata aaattgatac caggagtctc aattactgaa
aattacttgg 2580
aatagaagg aatggctaata tgtctcccat tctatggagt agcagattta
aaagaaattc 2640
ttaatgctat attaaacaga aatgcaaagg aagtttatga atgtagacct
cgcaaagtga 2700
taagtatttt agaggagaa gcagtgcgtc tatccagaca attacccatg
tacttatcaa 2760
aagaggacat ccaagacatt atctacagaa tgaagcacca gtttggaaat
gaaattaaag 2820
agtgtgttca tggtcgcca ttttttcac atttaaccta tcttcagaa
actacatgat 2880
taaataatgt taagaagatt agttaccatt gaaattgggt ctgtcataaa
acagcatgag 2940
tctggtttta aattatcttt gtattatgtg tcacatgggt attttttaaa
tgaggattca 3000
ctgacttgtt tttatattga aaaaagttcc acgtattgta gaaaacgtaa
ataaactaat 3060
aac
3063

```

<210> 9

<211> 3145

<212> DNA

5 <213> Homo sapiens

<400> 9

```

ggcgggaaac agcttagtg gtgtgggggc ggcattttc ttcaaccagg
aggtgaggag 60
gtttcgacat ggcgggtgcag ccgaaggaga cgctgcagtt ggagagcgcg
gccgaggtcg 120
gttcgtgctg cttctttcag ggcatgccgg agaagccgac caccacagtg
cgccttttgc 180
accggggcga cttctatacg gcgcacggcg aggacgcgct gctggccgcc
cgggaggtgt 240
tcaagaccca gggggtgatc aagtacatgg ggccggcagg agcaaagaat
ctgcagagtg 300

```

```

ttgtgcttag taaaatgaat ttggaatctt ttgtaaaaga tcttcttctg
gttcgtcagt 360
atagagttga agtttataag aatagagctg gaaataaggc atccaaggag
aatgattggt 420
atttggcata taaggcttct cctggcaatc tctctcagtt tgaagacatt
ctctttggta 480
acaatgatat gtcagcttcc attggtgttg tgggtgttaa aatgtccgca
gttgatggcc 540
agagacaggt tggagttggg tatgtggatt ccatacagag gaaactagga
ctgtgtgaat 600
tccctgataa tgatcagttc tccaatcttg aggcctctct catccagatt
ggaccaaaag 660
aatgtgtttt acccggagga gagactgctg gagacatggg gaaactgaga
cagataaattc 720
aaagaggagg aattctgata acagaaagaa aaaaagctga cttttccaca
aaagacattt 780
atcaggacct caaccggttg ttgaaaggca aaaagggaga gcagatgaat
agtgtgttat 840
tgccagaaat ggagaatcag gttgcagttt catcactgtc tgcggtaatc
aagtttttag 900
aactcttata agatgattcc aactttggac agtttgaact gactactttt
gacttcagcc 960
agtatatgaa attggatatt gcagcagtcg gagcccttaa cctttttcag
ggttctgttg 1020
aagataccac tggctctcag tctctggctg ccttgtctgaa taagtgtaaa
acccctcaag 1080
gacaaagact tgttaaccag tggattaagc agcctctcat ggataagaac
agaatagagg 1140
agagattgaa tttagtggaa gctttttaga aagatgcaga attgaggcag
actttacaag 1200
aagatttact tctgtcattc ccagatctta accgacttgc caagaagttt
caaagacaag 1260
cagcaaactt acaagattgt taccgactct atcaggggtat aaatcaacta
cctaattgta 1320
tacaggctct ggaaaaacat gaaggaaaac accagaaatt attgttggca
gtttttgtga 1380
ctcctcttac tgatcttcgt tctgacttct ccaagtttca ggaaatgata
gaaacaactt 1440
tagatatgga tcagggtggaa aacctatgaat tccttgtaaa accttcattt
gatcctaate 1500
tcagtgaatt aagagaaata atgaatgact tggaaaagaa gatgcagtcg
acattaataa 1560
gtgcagccag agatcttggc ttggaccctg gcaaacagat taaactggat
tccagtgcac 1620
agtttggata ttactttcgt gtaacctgta aggaagaaaa agtccttcgt
aacaataaaa 1680
acttttagtac tgtagatata cagaagaatg gtgttaaatt taccaacagc
aaattgactt 1740
ctttaaatga agagtatacc aaaaataaaa cagaatatga agaagcccag
gatgccattg 1800
ttaaagaaat tgtcaatatt tcttcaggct atgtagaacc aatgcagaca
ctcaatgatg 1860
tggttagctc gctagatgct gttgtcagct ttgctcacgt gtcaaagga
gcacctgttc 1920
catatgtacg accagccatt ttggagaaag gacaaggaag aattatatta
aaagcatcca 1980
ggcatgcttg tgttgaagtt caagatgaaa ttgcatttat tcctaattgac
gtatactttg 2040

```

```

aaaaagataa acagatgttc cacatcatta ctggccccaa tatgggaggt
aatcaacat 2100
atattcgaca aactgggggtg atagtactca tggcccaa at tgggtgtttt
tgccatgtg 2160
agtcagcaga agtgtccatt gtggactgca tcttagcccg agtaggggt
ggtgacagtc 2220
aattgaaagg agtctccacg ttcattggctg aaatgttgga aactgcttct
atcctcaggt 2280
ctgcaaccaa agattcatta ataatcatag atgaattggg aagaggaaact
tctacctacg 2340
atggatttgg gttagcatgg gctatatcag aatacattgc aacaagatt
ggtgcttttt 2400
gcatgtttgc aaccatttt catgaactta ctgccttggc caatcagata
ccaactgtta 2460
ataatctaca tgtcacagca ctcaccactg aagagacctt aactatgctt
taccaggtga 2520
agaaagggtg ctgtgatcaa agttttggga ttcattgttc agagcttgc
aatttcctta 2580
agcatgtaat agagtgtgct aaacagaaag ccctggaact tgaggagttt
cagtatatgt 2640
gagaatcgca aggatatgat atcatggaac cagcagcaaa gaagtgcctat
ctggaaagag 2700
agcaagggtga aaaaattatt caggagttcc tgtccaagg taaacaaatg
ccctttactg 2760
aaatgtcaga agaaaacac acaataaagt taaaacagct aaaagctgaa
gtaatagcaa 2820
agaataatag ctttgtaaat gaaatcattt caggaataaa agttactacg
tgaaaaatcc 2880
cagtaattgga atgaaggtaa tattgataag ctattgtctg taatagtttt
atattgtttt 2940
atattaaccc tttttccata gtgttaactg tcagtgccca tgggctatca
acttaataag 3000
atatttagta atattttact ttgaggacat tttcaaagat ttttattttg
aaaaatgaga 3060
gctgtaactg aggactgttt gcaattgaca taggcaataa taagtgatgt
gotgaatttt 3120
ataaataaaa tcatgtagtt tgtgg
3145

```

<210> 10

<211> 2484

<212> DNA

5

<213> Homo sapiens

<400> 10

```

cttggctctt ctggcgccaa aatgtcgttc gtggcagggg ttattcggcg
gctggacgag 60
acagtgtgga accgcacgac gccgggggaa gttatccagc ggccagctaa
tgctatcaaa 120
gagatgattg agaactgttt agatgcaaaa tccacaagta ttcaagtgat
tgtaaagag 180
ggaggcctga agttgattca gatccaagac aatggcaccg ggatcaggaa
agaagatctg 240
gatattgtat gtgaaagggt cactactagt aaactgcagt cctttgagga
ttagccagt 300
atttctacct atggctttcg aggtgaggct ttggccagca taagccatgt
ggctcatgtt 360
actattacaa cgaaaacagc tgatggaaag tgtgcataca gagcaagtta
ctcagatgga 420

```

```

aaactgaaag cccctcctaa accatgtgct ggcaatcaag ggacccagat
cacggtaggag 480
gacctttttt acaacatagc caccaggaga aaagctttta aaaatccaag
tgaagaatat 540
gggaaaattt tggaagttgt tggcaggat tcatcacaca atgcaggcat
tagtttctca 600
gttaaaaaac aaggagagac agtagctgat gttaggacac taccatgac
ctcaaccgtg 660
gacaatatcc gctccatctt tggaaatgct gttagtcgag aactgataga
aattggatgt 720
gaggataaaa ccctagcctt caaaatgaat ggttacatat ccaatgcaaa
ctactcagtg 780
aagaagtgca tcttcttact cttcatcaac catcgtctgg tagaatcaac
ttccttgaga 840
aaagccatag aaacagtgtg tgcagcctat ttgccaaaa acacacaccc
attcctgtac 900
ctcagtttag aatcagtcct ccagaatgtg gatgttaatg tgcacccac
aaagcatgaa 960
gttcacttcc tgcacgagga gagcatcctg gagcgggtgc agcagcacat
cgagagcaag 1020
ctcctgggct ccaattcctc caggatgtac ttcaccaga ctttgcctacc
aggacttgct 1080
ggccctctct gggagatggt taaatccaca acaagtctga cctcgtcttc
tacttctgga 1140
agtagtgata aggtctatgc ccaccagatg gttcgtacag attcccgga
acagaagctt 1200
gatgcatttc tgcagcctct gagcaaaccc ctgtccagtc agccccaggc
cattgtcaca 1260
gaggataaga cagatatctc tagtggcagg gctaggcagc aagatgagga
gatgcttgaa 1320
ctcccagccc ctgctgaagt ggctgccaaa aatcagagct tggaggggga
tacaacaaag 1380
gggacttcag aaatgtcaga gaagagagga cctacttcca gcaacccag
aaagagacat 1440
cggaagatt ctgatgtgga aatgggtgga gatgattccc gaaaggaaat
gactgcagct 1500
tgtaccccc ggagaaggat cattaacctc actagtgttt tgagtctcca
ggaagaaatt 1560
aatgagcagg gacatgaggt tctccgggag atgttgcata accactcct
cgtgggctgt 1620
gtgaatcctc agtgggcctt ggcacagcat caaaccaagt tataccttct
caacaccacc 1680
aagcttagtg aagaactgtt ctaccagata ctcatctatg attttgccaa
ttttggtgtt 1740
ctcaggttat cggagccagc accgctcttt gaccttgcca tgcctgcctt
agatagtcca 1800
gagagtggct ggacagagga agatgggtccc aaagaaggac ttgctgaata
cattgttgag 1860
tttctgaaga agaaggctga gatgcttgca gactatttct ctttggaaat
tgatgaggaa 1920
gggaacctga ttggattacc ccttctgatt gacaactatg tgccccctt
ggagggactg 1980
cctatcttca ttcttcgact agccactgag gtgaattggg acgaagaaaa
ggaatgtttt 2040
gaaagcctca gtaaagaatg cgctatgttc tattccatcc ggaagcagta
catatctgag 2100
gagtcgaccc tctcaggcca gcagagtga gtgcctggct ccattccaaa
ctcctggaag 2160

```

56

```

tggactgtgg aacacattgt ctataaagcc ttgcgctcac acattctgcc
tcctaaacat 2220
ttcacagaag atggaaatat cctgcagctt gctaacctgc ctgatctata
caaagtcctt 2280
gagagggtgtt aaatatgggtt atttatgcac tgtgggatgt gttcttcttt
ctctgtattc 2340
cgatacaaag tgttgtatca aagtgtgata tacaaagtgt accaacataa
gtgttggttag 2400
cacttaagac ttatacttgc cttctgatag tattccttta tacacagtgg
attgattata 2460
aataaataga tgtgtcttaa cata
2484

```

<210> 11

<211> 426

<212> DNA

5

<213> Homo sapiens

<400> 11

```

cgaggcggat cgggtgttgc atccatggag cgagctgaga gctcgagtac
agaacctgct 60
aaggccatca aacctattga tcggaagtca gtccatcaga ttgctcttgg
gcagggtggt 120
ctgagtctaa gcactgcggt aaaggagtta gtagaaaaca gtctggatgc
tggtgccact 180
aatattgatc taaagcttaa ggactatgga gtggatctta ttgaagtttc
agacaatgga 240
tgtggggttag aagaagaaaa cttcgaaggc ttaactctga aacatcacac
atctaagatt 300
caagagtttg ccgacctaac tcagggttgaa acttttggct ttcgggggga
agctctgagc 360
tcactttgtg cactgagcga tgtcaccatt tctacctgcc acgcacgggc
gaaggttgga 420
acttga
426

```

<210> 12

10

<211> 1408

<212> DNA

<213> Homo sapiens

<400> 12

```

ggcgctccta cctgcaagtg gctagtgcc aagtgtgggc cgccgctcct
gccgtgcatg      60
ttggggagcc agtacatgca ggtgggctcc acacggagag gggcgagac
ccggtgacag      120
ggctttacct ggtacatcgg catggcgcaa ccaaagcaag agaggggtggc
gcgtgccaga      180
caccaacggg cggaacccgc cagacaccaa cggtcggaaa ccgccaagac
accaacgctc      240
ggaaaccgcc agacaccaac gctcggaaac cgccagacac caaggctcgg
aatccacgcc      300
aggccacgac ggagggcgac tacctccctt ctgaccctgc tgctggcggt
cggaaaaaac      360
gcagtccggg gtgctctgat tgggtccaggc tctttgacgt cacggactcg
acctttgaca      420
gagccactag gcgaaaagga gagacgggaa gtattttttc cgccccgccc
ggaaagggtg      480

```

```

gagcacaacg tcgaaagcag ccgttgggag cccaggaggc ggggcgcctg
tgggagccgt      540
ggaggggaact ttcccagtc ccgaggcgga tccggtgttg catccttggg
gcgagctgag      600
aactcgagta cagaacctgc taaggccatc aaacctattg atcggaagtc
agtccatcag      660
atttgcctcg ggccggtggg accgagtcta aggccgaatg cggatgaagga
gttagtagaa      720
aacagtctgg atgctgggac cactaatggt gatctaaagc ttaaggacta
tgaggtggat      780
ctcattgaag ttccaggcaa tggatgtggg gtagaagaag aaaacttcga
aggctttact      840
ctgaaacatc acacatgtaa gattcaagag tttgccgacc taactcaggt
ggaaactttt      900
ggctttcggg gggaaagctc gagctcactt tgtgcactga gtgatgtcac
catttctacc      960
tgccgtgtat cagcgaaggt tgggactcga ctggtgtttg atcactatgg
gaaaatcacc     1020
cagaaaaccc cctacccccg ccccgaggag atgacagtca gcgtgaagca
gttattttct     1080
acgctacctg tgcaccataa agaatttcaa aggaatatta agaagaaacg
tgctgtcttc     1140
cccttcgcct tctgccgtga ttgtcagttt cctgaggcct ccccgccat
gcttcctgta     1200
cagcctgtag aactgactcc tagaagtacc ccacccacc cctgctcctt
ggaggacaac     1260
gtgatcactg tattcagctc tgtcaagaat ggtccagggt cttctagatg
atctgcacaa     1320
atggttcctc tcctccttcc tgatgtctgc cattagcatt ggaataaagt
tcctgctgaa     1380
aatccaaaaa aaaaaaaaaa aaaaaaaaaa
1408

```

<210> 13

<211> 1785

5

<212> DNA

<213> Homo sapiens

<400> 13

```

tttttagaaa ctgatgttta ttttccatca accatcttcc catgctgctt
aagagaatat 60
gcaagaacag cttaagacca gtcagtgggt gtcctaccc attcagtggc
ctgagcagtg 120
gggagctgca gaccagtctt ccgtggcagg ctgagcgctc cagtcttcag
tagggaattg 180
ctgaataggc acagagggca cctgtacacc ttcagaccag totgcaacct
caggctgagt 240
agcagtgaac tcaggagcgg gacgagtcca ttcacctga aattcctctt
tggtcactgc 300
cttctcagca gcagcctgct cttctttttc aatctcttca ggatctctgt
agaagtacag 360
atcaggcatg acctcccatg ggtgttcacg ggaaatggtg ccacgcatgc
gcagaacttc 420
ccgagccagc atccaccaca ttaaaccacac tgagttagct ccttctgtgt
tgcattgggt 480
ggcaatgtcc acatagcgca gaggagaatc tgtgttacac agcgcaatgg
taggtagggt 540
aacataagat gcctccgtga gaggcgaagg ggcggcggga cccgggcctg
gcccgtatgt 600

```

```

gtccttggcg gcctagacta ggccgtcgct gtatggtgag cccagggag
gcggatctgg 660
gccccagaa ggacacccgc ctggatttgc cccgtagccc ggccccggcc
cctcgggagc 720
agaacagcct tggtagagtg gacaggaggg gacctcgga gcagacgcgc
gcgccagcga 780
cagcagcccc gccccggcct ctccggagcc ggggggcaga ggctgcggag
ccccaggagg 840
gtctatcagc cacagtctct gcatgtttcc aagagcaaca ggaaatgaac
acattgcagg 900
ggccagtgtc attcaaagat gtggctgtgg atttcaccca ggaggagtgg
cggcaactgg 960
accctgatga gaagatagca tacggggatg tgatgttga gaactacagc
catctagttt 1020
ctgtggggta tgattatcac caagccaaac atcatcatgg agtggagggtg
aaggaaatgg 1080
agcagggaga ggagccgtgg ataatggaag gtgaatttcc atgtcaacat
agtccagaac 1140
ctgctaaggc catcaaacct attgatcgga agtcagtcca tcagatttgc
tctgggccag 1200
tggtactgag tctaagcact gcagtgaagg agttagtaga aaacagtctg
gatgctgggt 1260
ccactaatat tgatctaaag cttaaggact atggagtga tctcattgaa
gtttcagaca 1320
atggatgtgg ggtagaagaa gaaaactttg aaggcttaat ctctttcagc
tctgaaacat 1380
cacacatgta agattcaaga gtttgccgac ctaactgaag ttgaaacttt
cggttttcag 1440
ggggaagctc tgagctcact gtgtgcactg agegatgtca ccatttctac
ctgccacgcg 1500
ttggtgaagg ttgggactcg actggtgttt gatcacgatg ggaaaatcat
ccaggaaacc 1560
ccctaccccc accccagagg gaccacagtc agcgtgaagc agttattttc
tacgctacct 1620
gtgcgcata aggaatttca aaggaatatt aagaagacgt gcctgcttcc
ccttcgcctt 1680
ctgcgtgat tgcagtttc ctgaggcctc cccagccatg ctctctgtac
agcctgcaga 1740
actgtgagtc aattaaacct cttttcttca taaattaaaa aaaaa
1785

```

<210> 14

<211> 795

<212> DNA

<213> Homo sapiens

<400> 14

```

atgtgtcctt ggccggcctag actaggccgt cgctgtatgg tgagccccag
ggaggcggat      60
ctggggcccc agaaggacac ccgcctggat ttgccccgta gcccggccccg
ggccccctcgg      120
gagcagaaca gccttgggtga ggtggacagg aggggacctc gcgagcagac
gcgcgcgcga      180
gcgacagcag ccccgccccg gcctctcggg agccgggggg cagaggctgc
ggagccccag      240
gagggtctat cagccacagt ctctgcatgt ttccaagagc aacaggaaat
gaacacattg      300
cagggggccag tgtcattcaa agatgtggct gtggatttca cccaggagga
gtggcggcaa      360

```

```

ctggaccctg atgagaagat agcatacggg gatgtgatgt tggagaacta
cagccatcta      420
gtttctgtgg ggtatgatta tcaccaagcc aaacatcatc atggagtggg
ggtgaaggaa      480
gtggagcagg gagaggagcc gtggataatg gaaggatgaat ttccatgtca
acatagtcca      540
gaacctgcta aggccatcaa acctattgat cggaagtcag tccatcagat
ttgctctggg      600
ccagtgttac tgagtetaag cactgcagtg aaggagttag tagaaaacag
tctggatgct      660
ggtgccacta atattgatct aaagcttaag gactatggag tggatctcat
tgaagtttca      720
gacaatggat gtggggtaga agaagaaaac tttgaaggct taatctcttt
cagctctgaa      780
acatcacaca tgtaa
795

```

5

<210> 15

<211> 769

<212> PRT

<213> Saccharomyces cerevisiae

10

<400> 15

60

```

Met Ser Leu Arg Ile Lys Ala Leu Asp Ala Ser Val Val Asn Lys
Ile
1           5           10           15
Ala Ala Gly Glu Ile Ile Ile Ser Pro Val Asn Ala Leu Lys Glu
Met
           20           25           30
Met Glu Asn Ser Ile Asp Ala Asn Ala Thr Met Ile Asp Ile Leu
Val
           35           40           45
Lys Glu Gly Gly Ile Lys Val Leu Gln Ile Thr Asp Asn Gly Ser
Gly
50           55           60
Ile Asn Lys Ala Asp Leu Pro Ile Leu Cys Glu Arg Phe Thr Thr
Ser
65           70           75           80
Lys Leu Gln Lys Phe Glu Asp Leu Ser Gln Ile Gln Thr Tyr Gly
Phe
           85           90           95
Arg Gly Glu Ala Leu Ala Ser Ile Ser His Val Ala Arg Val Thr
Val
           100          105          110
Thr Thr Lys Val Lys Glu Asp Arg Cys Ala Trp Arg Val Ser Tyr
Ala
           115          120          125
Glu Gly Lys Met Leu Glu Ser Pro Lys Pro Val Ala Gly Lys Asp
Gly
130          135          140
Thr Thr Ile Leu Val Glu Asp Leu Phe Phe Asn Ile Pro Ser Arg
Leu
145          150          155
160
Arg Ala Leu Arg Ser His Asn Asp Glu Tyr Ser Lys Ile Leu Asp
Val
           165          170          175
Val Gly Arg Tyr Ala Ile His Ser Lys Asp Ile Gly Phe Ser Cys
Lys

```

```

      180              185              190
Lys Phe Gly Asp Ser Asn Tyr Ser Leu Ser Val Lys Pro Ser Tyr
Thr
      195              200              205
Val Gln Asp Arg Ile Arg Thr Val Phe Asn Lys Ser Val Ala Ser
Asn
      210              215              220
Leu Ile Thr Phe His Ile Ser Lys Val Glu Asp Leu Asn Leu Glu
Ser
      225              230              235
240
Val Asp Gly Lys Val Cys Asn Leu Asn Phe Ile Ser Lys Lys Ser
Ile
      245              250              255
Ser Leu Ile Phe Phe Ile Asn Asn Arg Leu Val Thr Cys Asp Leu
Leu
      260              265              270
Arg Arg Ala Leu Asn Ser Val Tyr Ser Asn Tyr Leu Pro Lys Gly
Phe
      275              280              285
Arg Pro Phe Ile Tyr Leu Gly Ile Val Ile Asp Pro Ala Ala Val
Asp
      290              295              300
Val Asn Val His Pro Thr Lys Arg Glu Val Arg Phe Leu Ser Gln
Asp
      305              310              315
320
Glu Ile Ile Glu Lys Ile Ala Asn Gln Leu His Ala Glu Leu Ser
Ala
      325              330              335
Ile Asp Thr Ser Arg Thr Phe Lys Ala Ser Ser Ile Ser Thr Asn
Lys
      340              345              350
Pro Glu Ser Leu Ile Pro Phe Asn Asp Thr Ile Glu Ser Asp Arg
Asn
      355              360              365
Arg Lys Ser Leu Arg Gln Ala Gln Val Val Glu Asn Ser Tyr Thr
Thr
      370              375              380
Ala Asn Ser Gln Leu Arg Lys Ala Lys Arg Gln Glu Asn Lys Leu
Val
      385              390              395
400
Arg Ile Asp Ala Ser Gln Ala Lys Ile Thr Ser Phe Leu Ser Ser
Ser
      405              410              415
Gln Gln Phe Asn Phe Glu Gly Ser Ser Thr Lys Arg Gln Leu Ser
Glu
      420              425              430
Pro Lys Val Thr Asn Val Ser His Ser Gln Glu Ala Glu Lys Leu
Thr
      435              440              445
Leu Asn Glu Ser Glu Gln Pro Arg Asp Ala Asn Thr Ile Asn Asp
Asn
      450              455              460
Asp Leu Lys Asp Gln Pro Lys Lys Lys Gln Lys Leu Gly Asp Tyr
Lys
      465              470              475
480

```

62

```

Val Pro Ser Ile Ala Asp Asp Glu Lys Asn Ala Leu Pro Ile Ser
Lys
                        485                        490                        495
Asp Gly Tyr Ile Arg Val Pro Lys Glu Arg Val Asn Val Asn Leu
Thr
                        500                        505                        510
Ser Ile Lys Lys Leu Arg Glu Lys Val Asp Asp Ser Ile His Arg
Glu
                        515                        520                        525
Leu Thr Asp Ile Phe Ala Asn Leu Asn Tyr Val Gly Val Val Asp
Glu
                        530                        535                        540
Glu Arg Arg Leu Ala Ala Ile Gln His Asp Leu Lys Leu Phe Leu
Ile
545                        550                        555
560
Asp Tyr Gly Ser Val Cys Tyr Glu Leu Phe Tyr Gln Ile Gly Leu
Thr
                        565                        570                        575
Asp Phe Ala Asn Phe Gly Lys Ile Asn Leu Gln Ser Thr Asn Val
Ser
                        580                        585                        590
Asp Asp Ile Val Leu Tyr Asn Leu Leu Ser Glu Phe Asp Glu Leu
Asn
                        595                        600                        605
Asp Asp Ala Ser Lys Glu Lys Ile Ile Ser Lys Ile Trp Asp Met
Ser
                        610                        615                        620
Ser Met Leu Asn Glu Tyr Tyr Ser Ile Glu Leu Val Asn Asp Gly
Leu
625                        630                        635
640
Asp Asn Asp Leu Lys Ser Val Lys Leu Lys Ser Leu Pro Leu Leu
Leu
                        645                        650                        655
Lys Gly Tyr Ile Pro Ser Leu Val Lys Leu Pro Phe Phe Ile Tyr
Arg
                        660                        665                        670
Leu Gly Lys Glu Val Asp Trp Glu Asp Glu Gln Glu Cys Leu Asp
Gly
                        675                        680                        685
Ile Leu Arg Glu Ile Ala Leu Leu Tyr Ile Pro Asp Met Val Pro
Lys
                        690                        695                        700
Val Asp Thr Leu Asp Ala Ser Leu Ser Glu Asp Glu Lys Ala Gln
Phe
705                        710                        715
720
Ile Asn Arg Lys Glu His Ile Ser Ser Leu Leu Glu His Val Leu
Phe
                        725                        730                        735
Pro Cys Ile Lys Arg Arg Phe Leu Ala Pro Arg His Ile Leu Lys
Asp
                        740                        745                        750
Val Val Glu Ile Ala Asn Leu Pro Asp Leu Tyr Lys Val Phe Glu
Arg
                        755                        760                        765
Cys

```

<210> 16

<211> 859

<212> PRT

5

<213> Mus musculus

<400> 16

```

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

```

```

                260                265                270
Phe His Ser Ala Arg Thr Ala Pro Gly Gly Val Gln Gln Thr Gly
Ser

                275                280                285
Phe Ser Ser Ser Ile Arg Gly Pro Val Thr Gln Gln Arg Ser Leu
Ser

                290                295                300
Leu Ser Met Arg Phe Tyr His Met Tyr Asn Arg His Gln Tyr Pro
Phe
305                310                315
320
Val Val Leu Asn Val Ser Val Asp Ser Glu Cys Val Asp Ile Asn
Val

                325                330                335
Thr Pro Asp Lys Arg Gln Ile Leu Leu Gln Glu Glu Lys Leu Leu
Leu

                340                345                350
Ala Val Leu Lys Thr Ser Leu Ile Gly Met Phe Asp Ser Asp Ala
Asn

                355                360                365
Lys Leu Asn Val Asn Gln Gln Pro Leu Leu Asp Val Glu Gly Asn
Leu

                370                375                380
Val Lys Leu His Thr Ala Glu Leu Glu Lys Pro Val Pro Gly Lys
Gln
385                390                395
400
Asp Asn Ser Pro Ser Leu Lys Ser Thr Ala Asp Glu Lys Arg Val
Ala

                405                410                415
Ser Ile Ser Arg Leu Arg Glu Ala Phe Ser Leu His Pro Thr Lys
Glu

                420                425                430
Ile Lys Ser Arg Gly Pro Glu Thr Ala Glu Leu Thr Arg Ser Phe
Pro

                435                440                445
Ser Glu Lys Arg Gly Val Leu Ser Ser Tyr Pro Ser Asp Val Ile
Ser

                450                455                460
Tyr Arg Gly Leu Arg Gly Ser Gln Asp Lys Leu Val Ser Pro Thr
Asp
465                470                475
480
Ser Pro Gly Asp Cys Met Asp Arg Glu Lys Ile Glu Lys Asp Ser
Gly

                485                490                495
Leu Ser Ser Thr Ser Ala Gly Ser Glu Glu Glu Phe Ser Thr Pro
Glu

                500                505                510
Val Ala Ser Ser Phe Ser Ser Asp Tyr Asn Val Ser Ser Leu Glu
Asp

                515                520                525
Arg Pro Ser Gln Glu Thr Ile Asn Cys Gly Asp Leu Asp Cys Arg
Pro

                530                535                540
Pro Gly Thr Gly Gln Ser Leu Lys Pro Glu Asp His Gly Tyr Gln
Cys
545                550                555
560

```

65

Lys	Ala	Leu	Pro	Leu	Ala	Arg	Leu	Ser	Pro	Thr	Asn	Ala	Lys	Arg
Phe														
				565					570					575
Lys	Thr	Glu	Glu	Arg	Pro	Ser	Asn	Val	Asn	Ile	Ser	Gln	Arg	Leu
Pro														
			580					585					590	
Gly	Pro	Gln	Ser	Thr	Ser	Ala	Ala	Glu	Val	Asp	Val	Ala	Ile	Lys
Met														
		595					600					605		
Asn	Lys	Arg	Ile	Val	Leu	Leu	Glu	Phe	Ser	Leu	Ser	Ser	Leu	Ala
Lys														
	610					615					620			
Arg	Met	Lys	Gln	Leu	Gln	His	Leu	Lys	Ala	Gln	Asn	Lys	His	Glu
Leu														
625				630						635				
640														
Ser	Tyr	Arg	Lys	Phe	Arg	Ala	Lys	Ile	Cys	Pro	Gly	Glu	Asn	Gln
Ala														
			645					650					655	
Ala	Glu	Asp	Glu	Leu	Arg	Lys	Glu	Ile	Ser	Lys	Ser	Met	Phe	Ala
Glu														
		660					665					670		
Met	Glu	Ile	Leu	Gly	Gln	Phe	Asn	Leu	Gly	Phe	Ile	Val	Thr	Lys
Leu														
		675				680					685			
Lys	Glu	Asp	Leu	Phe	Leu	Val	Asp	Gln	His	Ala	Ala	Asp	Glu	Lys
Tyr														
	690				695					700				
Asn	Phe	Glu	Met	Leu	Gln	Gln	His	Thr	Val	Leu	Gln	Ala	Gln	Arg
Leu														
705				710						715				
720														
Ile	Thr	Pro	Gln	Thr	Leu	Asn	Leu	Thr	Ala	Val	Asn	Glu	Ala	Val
Leu														
			725					730					735	
Ile	Glu	Asn	Leu	Glu	Ile	Phe	Arg	Lys	Asn	Gly	Phe	Asp	Phe	Val
Ile														
		740						745				750		
Asp	Glu	Asp	Ala	Pro	Val	Thr	Glu	Arg	Ala	Lys	Leu	Ile	Ser	Leu
Pro														
		755				760					765			
Thr	Ser	Lys	Asn	Trp	Thr	Phe	Gly	Pro	Gln	Asp	Ile	Asp	Glu	Leu
Ile														
	770				775					780				
Phe	Met	Leu	Ser	Asp	Ser	Pro	Gly	Val	Met	Cys	Arg	Pro	Ser	Arg
Val														
785				790						795				
800														
Arg	Gln	Met	Phe	Ala	Ser	Arg	Ala	Cys	Arg	Lys	Ser	Val	Met	Ile
Gly														
			805					810						

<210> 17

<211> 932

<212> PRT

<213> Homo sapiens

<400> 17

```

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

```

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

Thr Ala Tyr Asp Leu Leu Ser Asn Arg Val Ile Lys Lys Pro Met
 Ser
 565 570 575
 Ala Ser Ala Leu Phe Val Gln Asp His Arg Pro Gln Phe Leu Ile
 Glu
 580 585 590
 Asn Pro Lys Thr Ser Leu Glu Asp Ala Thr Leu Gln Ile Glu Glu
 Leu
 595 600 605
 Trp Lys Thr Leu Ser Glu Glu Glu Lys Leu Lys Tyr Glu Glu Lys
 Ala
 610 615 620
 Thr Lys Asp Leu Glu Arg Tyr Asn Ser Gln Met Lys Arg Ala Ile
 Glu
 625 630 635
 640
 Gln Glu Ser Gln Met Ser Leu Lys Asp Gly Arg Lys Lys Ile Lys
 Pro
 645 650 655
 Thr Ser Ala Trp Asn Leu Ala Gln Lys His Lys Leu Lys Thr Ser
 Leu
 660 665 670
 Ser Asn Gln Pro Lys Leu Asp Glu Leu Leu Gln Ser Gln Ile Glu
 Lys
 675 680 685
 Arg Arg Ser Gln Asn Ile Lys Met Val Gln Ile Pro Phe Ser Met
 Lys
 690 695 700
 Asn Leu Lys Ile Asn Phe Lys Lys Gln Asn Lys Val Asp Leu Glu
 Glu
 705 710 715
 720
 Lys Asp Glu Pro Cys Leu Ile His Asn Leu Arg Phe Pro Asp Ala
 Trp
 725 730 735
 Leu Met Thr Ser Lys Thr Glu Val Met Leu Leu Asn Pro Tyr Arg
 Val
 740 745 750
 Glu Glu Ala Leu Leu Phe Lys Arg Leu Leu Glu Asn His Lys Leu
 Pro
 755 760 765
 Ala Glu Pro Leu Glu Lys Pro Ile Met Leu Thr Glu Ser Leu Phe
 Asn
 770 775 780
 Gly Ser His Tyr Leu Asp Val Leu Tyr Lys Met Thr Ala Asp Asp
 Gln
 785 790 795
 800
 Arg Tyr Ser Gly Ser Thr Tyr Leu Ser Asp Pro Arg Leu Thr Ala
 Asn
 805 810 815
 Gly Phe Lys Ile Lys Leu Ile Pro Gly Val Ser Ile Thr Glu Asn
 Tyr
 820 825 830
 Leu Glu Ile Glu Gly Met Ala Asn Cys Leu Pro Phe Tyr Gly Val
 Ala
 835 840 845
 Asp Leu Lys Glu Ile Leu Asn Ala Ile Leu Asn Arg Asn Ala Lys
 Glu

69

```

      850      855      860
Val Tyr Glu Cys Arg Pro Arg Lys Val Ile Ser Tyr Leu Glu Gly
Glu
865      870      875
880
Ala Val Arg Leu Ser Arg Gln Leu Pro Met Tyr Leu Ser Lys Glu
Asp
      885      890      895
Ile Gln Asp Ile Ile Tyr Arg Met Lys His Gln Phe Gly Asn Glu
Ile
      900      905      910
Lys Glu Cys Val His Gly Arg Pro Phe Phe His His Leu Thr Tyr
Leu
      915      920      925
Pro Glu Thr Thr
930

```

<210> 18

<211> 932

<212> PRT

5 <213> Homo sapiens

<400> 18

```

Met Lys Gln Leu Pro Ala Ala Thr Val Arg Leu Leu Ser Ser Ser
Gln
1      5      10      15
Ile Ile Thr Ser Val Val Ser Val Val Lys Glu Leu Ile Glu Asn
Ser
      20      25      30
Leu Asp Ala Gly Ala Thr Ser Val Asp Val Lys Leu Glu Asn Tyr
Gly
      35      40      45
Phe Asp Lys Ile Glu Val Arg Asp Asn Gly Glu Gly Ile Lys Ala
Val
      50      55      60
Asp Ala Pro Val Met Ala Met Lys Tyr Tyr Thr Ser Lys Ile Asn
Ser
65      70      75      80
His Glu Asp Leu Glu Asn Leu Thr Thr Tyr Gly Phe Arg Gly Glu
Ala
      85      90      95
Leu Gly Ser Ile Cys Cys Ile Ala Glu Val Leu Ile Thr Thr Arg
Thr
      100      105      110
Ala Ala Asp Asn Phe Ser Thr Gln Tyr Val Leu Asp Gly Ser Gly
His
      115      120      125
Ile Leu Ser Gln Lys Pro Ser His Leu Gly Gln Gly Thr Thr Val
Thr
      130      135      140
Ala Leu Arg Leu Phe Lys Asn Leu Pro Val Arg Lys Gln Phe Tyr
Ser
145      150      155
160
Thr Ala Lys Lys Cys Lys Asp Glu Ile Lys Lys Ile Gln Asp Leu
Leu
      165      170      175
Met Ser Phe Gly Ile Leu Lys Pro Asp Leu Arg Ile Val Phe Val
His

```

```

      180      185      190
Asn Lys Ala Val Ile Trp Gln Lys Ser Arg Val Ser Asp His Lys
Met
      195      200      205
Ala Leu Met Ser Val Leu Gly Thr Ala Val Met Asn Asn Met Glu
Ser
      210      215      220
Phe Gln Tyr His Ser Glu Glu Ser Gln Ile Tyr Leu Ser Gly Phe
Leu
225      230      235
240
Pro Lys Cys Asp Ala Asp His Ser Phe Thr Ser Leu Ser Thr Pro
Glu
      245      250      255
Arg Ser Phe Ile Phe Ile Asn Ser Arg Pro Val His Gln Lys Asp
Ile
      260      265      270
Leu Lys Leu Ile Arg His His Tyr Asn Leu Lys Cys Leu Lys Glu
Ser
      275      280      285
Thr Arg Leu Tyr Pro Val Phe Phe Leu Lys Ile Asp Val Pro Thr
Ala
      290      295      300
Asp Val Asp Val Asn Leu Thr Pro Asp Lys Ser Gln Val Leu Leu
Gln
305      310      315
320
Asn Lys Glu Ser Val Leu Ile Ala Leu Glu Asn Leu Met Thr Thr
Cys
      325      330      335
Tyr Gly Pro Leu Pro Ser Thr Asn Ser Tyr Glu Asn Asn Lys Thr
Asp
      340      345      350
Val Ser Ala Ala Asp Ile Val Leu Ser Lys Thr Ala Glu Thr Asp
Val
      355      360      365
Leu Phe Asn Lys Val Glu Ser Ser Gly Lys Asn Tyr Ser Asn Val
Asp
      370      375      380
Thr Ser Val Ile Pro Phe Gln Asn Asp Met His Asn Asp Glu Ser
Gly
385      390      395
400
Lys Asn Thr Asp Asp Cys Leu Asn His Gln Ile Ser Ile Gly Asp
Phe
      405      410      415
Gly Tyr Gly His Cys Ser Ser Glu Ile Ser Asn Ile Asp Lys Asn
Thr
      420      425      430
Lys Asn Ala Phe Gln Asp Ile Ser Met Ser Asn Val Ser Trp Glu
Asn
      435      440      445
Ser Gln Thr Glu Tyr Ser Lys Thr Cys Phe Ile Ser Ser Val Lys
His
      450      455      460
Thr Gln Ser Glu Asn Gly Asn Lys Asp His Ile Asp Glu Ser Gly
Glu
465      470      475
480

```

Asn Glu Glu Glu Ala Gly Leu Glu Asn Ser Ser Glu Ile Ser Ala
 Asp
 485 490 495
 Glu Trp Ser Arg Gly Asn Ile Leu Lys Asn Ser Val Gly Glu Asn
 Ile
 500 505 510
 Glu Pro Val Lys Ile Leu Val Pro Glu Lys Ser Leu Pro Cys Lys
 Val
 515 520 525
 Ser Asn Asn Asn Tyr Pro Ile Pro Glu Gln Met Asn Leu Asn Glu
 Asp
 530 535 540
 Ser Cys Asn Lys Lys Ser Asn Val Ile Asp Asn Lys Ser Gly Lys
 Val
 545 550 555
 560
 Thr Ala Tyr Asp Leu Leu Ser Asn Arg Val Ile Lys Lys Pro Met
 Ser
 565 570 575
 Ala Ser Ala Leu Phe Val Gln Asp His Arg Pro Gln Phe Leu Ile
 Glu
 580 585 590
 Asn Pro Lys Thr Ser Leu Glu Asp Ala Thr Leu Gln Ile Glu Glu
 Leu
 595 600 605
 Trp Lys Thr Leu Ser Glu Glu Glu Lys Leu Lys Tyr Glu Glu Lys
 Ala
 610 615 620
 Thr Lys Asp Leu Glu Arg Tyr Asn Ser Gln Met Lys Arg Ala Ile
 Glu
 625 630 635
 640
 Gln Glu Ser Gln Met Ser Leu Lys Asp Gly Arg Lys Lys Ile Lys
 Pro
 645 650 655
 Thr Ser Ala Trp Asn Leu Ala Gln Lys His Lys Leu Lys Thr Ser
 Leu
 660 665 670
 Ser Asn Gln Pro Lys Leu Asp Glu Leu Leu Gln Ser Gln Ile Glu
 Lys
 675 680 685
 Arg Arg Ser Gln Asn Ile Lys Met Val Gln Ile Pro Phe Ser Met
 Lys
 690 695 700
 Asn Leu Lys Ile Asn Phe Lys Lys Gln Asn Lys Val Asp Leu Glu
 Glu
 705 710 715
 720
 Lys Asp Glu Pro Cys Leu Ile His Asn Leu Arg Phe Pro Asp Ala
 Trp
 725 730 735
 Leu Met Thr Ser Lys Thr Glu Val Met Leu Leu Asn Pro Tyr Arg
 Val
 740 745 750
 Glu Glu Ala Leu Leu Phe Lys Arg Leu Leu Glu Asn His Lys Leu
 Pro
 755 760 765
 Ala Glu Pro Leu Glu Lys Pro Ile Met Leu Thr Glu Ser Leu Phe
 Asn

```

      770              775              780
    Gly Ser His Tyr Leu Asp Val Leu Tyr Lys Met Thr Ala Asp Asp
    Gln
    785              790              795
    800
    Arg Tyr Ser Gly Ser Thr Tyr Leu Ser Asp Pro Arg Leu Thr Ala
    Asn
      805              810              815
    Gly Phe Lys Ile Lys Leu Ile Pro Gly Val Ser Ile Thr Glu Asn
    Tyr
      820              825              830
    Leu Glu Ile Glu Gly Met Ala Asn Cys Leu Pro Phe Tyr Gly Val
    Ala
      835              840              845
    Asp Leu Lys Glu Ile Leu Asn Ala Ile Leu Asn Arg Asn Ala Lys
    Glu
      850              855              860
    Val Tyr Glu Cys Arg Pro Arg Lys Val Ile Ser Tyr Leu Glu Gly
    Glu
    865              870              875
    880
    Ala Val Arg Leu Ser Arg Gln Leu Pro Met Tyr Leu Ser Lys Glu
    Asp
      885              890              895
    Ile Gln Asp Ile Ile Tyr Arg Met Lys His Gln Phe Gly Asn Glu
    Ile
      900              905              910
    Lys Glu Cys Val His Gly Arg Pro Phe Phe His His Leu Thr Tyr
    Leu
      915              920              925
    Pro Glu Thr Thr
    930

```

<210> 19

<211> 934

<212> PRT

5 <213> Homo sapiens

<400> 19

```

    Met Ala Val Gln Pro Lys Glu Thr Leu Gln Leu Glu Ser Ala Ala
    Glu
      1              5              10              15
    Val Gly Phe Val Arg Phe Phe Gln Gly Met Pro Glu Lys Pro Thr
    Thr
      20              25              30
    Thr Val Arg Leu Phe Asp Arg Gly Asp Phe Tyr Thr Ala His Gly
    Glu
      35              40              45
    Asp Ala Leu Leu Ala Ala Arg Glu Val Phe Lys Thr Gln Gly Val
    Ile
      50              55              60
    Lys Tyr Met Gly Pro Ala Gly Ala Lys Asn Leu Gln Ser Val Val
    Leu
      65              70              75              80
    Ser Lys Met Asn Phe Glu Ser Phe Val Lys Asp Leu Leu Leu Val
    Arg
      85              90              95
    Gln Tyr Arg Val Glu Val Tyr Lys Asn Arg Ala Gly Asn Lys Ala
    Ser

```

```

          100          105          110
Lys Glu Asn Asp Trp Tyr Leu Ala Tyr Lys Ala Ser Pro Gly Asn
Leu
          115          120          125
Ser Gln Phe Glu Asp Ile Leu Phe Gly Asn Asn Asp Met Ser Ala
Ser
          130          135          140
Ile Gly Val Val Gly Val Lys Met Ser Ala Val Asp Gly Gln Arg
Gln
          145          150          155
Val Gly Val Gly Tyr Val Asp Ser Ile Gln Arg Lys Leu Gly Leu
Cys
          165          170          175
Glu Phe Pro Asp Asn Asp Gln Phe Ser Asn Leu Glu Ala Leu Leu
Ile
          180          185          190
Gln Ile Gly Pro Lys Glu Cys Val Leu Pro Gly Gly Glu Thr Ala
Gly
          195          200          205
Asp Met Gly Lys Leu Arg Gln Ile Ile Gln Arg Gly Gly Ile Leu
Ile
          210          215          220
Thr Glu Arg Lys Lys Ala Asp Phe Ser Thr Lys Asp Ile Tyr Gln
Asp
          225          230          235
Leu Asn Arg Leu Leu Lys Gly Lys Lys Gly Glu Gln Met Asn Ser
Ala
          245          250          255
Val Leu Pro Glu Met Glu Asn Gln Val Ala Val Ser Ser Leu Ser
Ala
          260          265          270
Val Ile Lys Phe Leu Glu Leu Leu Ser Asp Asp Ser Asn Phe Gly
Gln
          275          280          285
Phe Glu Leu Thr Thr Phe Asp Phe Ser Gln Tyr Met Lys Leu Asp
Ile
          290          295          300
Ala Ala Val Arg Ala Leu Asn Leu Phe Gln Gly Ser Val Glu Asp
Thr
          305          310          315
Thr Gly Ser Gln Ser Leu Ala Ala Leu Leu Asn Lys Cys Lys Thr
Pro
          325          330          335
Gln Gly Gln Arg Leu Val Asn Gln Trp Ile Lys Gln Pro Leu Met
Asp
          340          345          350
Lys Asn Arg Ile Glu Glu Arg Leu Asn Leu Val Glu Ala Phe Val
Glu
          355          360          365
Asp Ala Glu Leu Arg Gln Thr Leu Gln Glu Asp Leu Leu Arg Arg
Phe
          370          375          380
Pro Asp Leu Asn Arg Leu Ala Lys Lys Phe Gln Arg Gln Ala Ala
Asn
          385          390          395
400

```

Leu Gln Asp Cys Tyr Arg Leu Tyr Gln Gly Ile Asn Gln Leu Pro
 Asn
 Val Ile Gln Ala Leu Glu Lys His Glu Gly Lys His Gln Lys Leu
 Leu
 Leu Ala Val Phe Val Thr Pro Leu Thr Asp Leu Arg Ser Asp Phe
 Ser
 Lys Phe Gln Glu Met Ile Glu Thr Thr Leu Asp Met Asp Gln Val
 Glu
 Asn His Glu Phe Leu Val Lys Pro Ser Phe Asp Pro Asn Leu Ser
 Glu
 465 470 475
 480
 Leu Arg Glu Ile Met Asn Asp Leu Glu Lys Lys Met Gln Ser Thr
 Leu
 Ile Ser Ala Ala Arg Asp Leu Gly Leu Asp Pro Gly Lys Gln Ile
 Lys
 Leu Asp Ser Ser Ala Gln Phe Gly Tyr Tyr Phe Arg Val Thr Cys
 Lys
 Glu Glu Lys Val Leu Arg Asn Asn Lys Asn Phe Ser Thr Val Asp
 Ile
 Gln Lys Asn Gly Val Lys Phe Thr Asn Ser Lys Leu Thr Ser Leu
 Asn
 545 550 555
 560
 Glu Glu Tyr Thr Lys Asn Lys Thr Glu Tyr Glu Glu Ala Gln Asp
 Ala
 Ile Val Lys Glu Ile Val Asn Ile Ser Ser Gly Tyr Val Glu Pro
 Met
 Gln Thr Leu Asn Asp Val Leu Ala Gln Leu Asp Ala Val Val Ser
 Phe
 Ala His Val Ser Asn Gly Ala Pro Val Pro Tyr Val Arg Pro Ala
 Ile
 Leu Glu Lys Gly Gln Gly Arg Ile Ile Leu Lys Ala Ser Arg His
 Ala
 625 630 635
 640
 Cys Val Glu Val Gln Asp Glu Ile Ala Phe Ile Pro Asn Asp Val
 Tyr
 Phe Glu Lys Asp Lys Gln Met Phe His Ile Ile Thr Gly Pro Asn
 Met
 Gly Gly Lys Ser Thr Tyr Ile Arg Gln Thr Gly Val Ile Val Leu
 Met
 Ala Gln Ile Gly Cys Phe Val Pro Cys Glu Ser Ala Glu Val Ser
 Ile

75

```

        690                695                700
Val Asp Cys Ile Leu Ala Arg Val Gly Ala Gly Asp Ser Gln Leu
Lys
705                710                715
720
Gly Val Ser Thr Phe Met Ala Glu Met Leu Glu Thr Ala Ser Ile
Leu
        725                730                735
Arg Ser Ala Thr Lys Asp Ser Leu Ile Ile Ile Asp Glu Leu Gly
Arg
        740                745                750
Gly Thr Ser Thr Tyr Asp Gly Phe Gly Leu Ala Trp Ala Ile Ser
Glu
        755                760                765
Tyr Ile Ala Thr Lys Ile Gly Ala Phe Cys Met Phe Ala Thr His
Phe
        770                775                780
His Glu Leu Thr Ala Leu Ala Asn Gln Ile Pro Thr Val Asn Asn
Leu
785                790                795
800
His Val Thr Ala Leu Thr Thr Glu Glu Thr Leu Thr Met Leu Tyr
Gln
        805                810                815
Val Lys Lys Gly Val Cys Asp Gln Ser Phe Gly Ile His Val Ala
Glu
        820                825                830
Leu Ala Asn Phe Pro Lys His Val Ile Glu Cys Ala Lys Gln Lys
Ala
        835                840                845
Leu Glu Leu Glu Glu Phe Gln Tyr Ile Gly Glu Ser Gln Gly Tyr
Asp
        850                855                860
Ile Met Glu Pro Ala Ala Lys Lys Cys Tyr Leu Glu Arg Glu Gln
Gly
865                870                875
880
Glu Lys Ile Ile Gln Glu Phe Leu Ser Lys Val Lys Gln Met Pro
Phe
        885                890                895
Thr Glu Met Ser Glu Glu Asn Ile Thr Ile Lys Leu Lys Gln Leu
Lys
        900                905                910
Ala Glu Val Ile Ala Lys Asn Asn Ser Phe Val Asn Glu Ile Ile
Ser
        915                920                925
Arg Ile Lys Val Thr Thr
930

```

<210> 20

<211> 756

<212> PRT

5

<213> Homo sapiens

<400> 20

```

Met Ser Phe Val Ala Gly Val Ile Arg Arg Leu Asp Glu Thr Val
Val
1                5                10                15

```

Asn Arg Ile Ala Ala Gly Glu Val Ile Gln Arg Pro Ala Asn Ala
 Ile
 Lys Glu Met 20 Ile Glu Asn Cys Leu Asp Ala Lys Ser Thr Ser Ile
 Gln 35 40 45
 Val Ile Val Lys Glu Gly Gly Leu Lys Leu Ile Gln Ile Gln Asp
 Asn 50 55 60
 Gly Thr Gly Ile Arg Lys Glu Asp Leu Asp Ile Val Cys Glu Arg
 Phe 65 70 75 80
 Thr Thr Ser Lys Leu Gln Ser Phe Glu Asp Leu Ala Ser Ile Ser
 Thr 85 90 95
 Tyr Gly Phe Arg Gly Glu Ala Leu Ala Ser Ile Ser His Val Ala
 His 100 105 110
 Val Thr Ile Thr Thr Lys Thr Ala Asp Gly Lys Cys Ala Tyr Arg
 Ala 115 120 125
 Ser Tyr Ser Asp Gly Lys Leu Lys Ala Pro Pro Lys Pro Cys Ala
 Gly 130 135 140
 Asn Gln Gly Thr Gln Ile Thr Val Glu Asp Leu Phe Tyr Asn Ile
 Ala 145 150 155
 160
 Thr Arg Arg Lys Ala Leu Lys Asn Pro Ser Glu Glu Tyr Gly Lys
 Ile 165 170 175
 Leu Glu Val Val Gly Arg Tyr Ser Val His Asn Ala Gly Ile Ser
 Phe 180 185 190
 Ser Val Lys Lys Gln Gly Glu Thr Val Ala Asp Val Arg Thr Leu
 Pro 195 200 205
 Asn Ala Ser Thr Val Asp Asn Ile Arg Ser Ile Phe Gly Asn Ala
 Val 210 215 220
 Ser Arg Glu Leu Ile Glu Ile Gly Cys Glu Asp Lys Thr Leu Ala
 Phe 225 230 235
 240
 Lys Met Asn Gly Tyr Ile Ser Asn Ala Asn Tyr Ser Val Lys Lys
 Cys 245 250 255
 Ile Phe Leu Leu Phe Ile Asn His Arg Leu Val Glu Ser Thr Ser
 Leu 260 265 270
 Arg Lys Ala Ile Glu Thr Val Tyr Ala Ala Tyr Leu Pro Lys Asn
 Thr 275 280 285
 His Pro Phe Leu Tyr Leu Ser Leu Glu Ile Ser Pro Gln Asn Val
 Asp 290 295 300
 Val Asn Val His Pro Thr Lys His Glu Val His Phe Leu His Glu
 Glu

```

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

```

78

Glu Tyr Ile Val Glu Phe Leu Lys Lys Lys Ala Glu Met Leu Ala
 Asp
 610 615 620
 Tyr Phe Ser Leu Glu Ile Asp Glu Glu Gly Asn Leu Ile Gly Leu
 Pro
 625 630 635
 640
 Leu Leu Ile Asp Asn Tyr Val Pro Pro Leu Glu Gly Leu Pro Ile
 Phe
 645 650 655
 Ile Leu Arg Leu Ala Thr Glu Val Asn Trp Asp Glu Glu Lys Glu
 Cys
 660 665 670
 Phe Glu Ser Leu Ser Lys Glu Cys Ala Met Phe Tyr Ser Ile Arg
 Lys
 675 680 685
 Gln Tyr Ile Ser Glu Glu Ser Thr Leu Ser Gly Gln Gln Ser Glu
 Val
 690 695 700
 Pro Gly Ser Ile Pro Asn Ser Trp Lys Trp Thr Val Glu His Ile
 Val
 705 710 715
 720
 Tyr Lys Ala Leu Arg Ser His Ile Leu Pro Pro Lys His Phe Thr
 Glu
 725 730 735
 Asp Gly Asn Ile Leu Gln Leu Ala Asn Leu Pro Asp Leu Tyr Lys
 Val
 740 745 750
 Phe Glu Arg Cys
 755

<210> 21

<211> 133

<212> PRT

5 <213> Homo sapiens

<400> 21

Met Lys Gln Leu Pro Ala Ala Thr Val Arg Leu Leu Ser Ser Ser
 Gln
 1 5 10 15
 Ile Ile Thr Ser Val Val Ser Val Val Lys Glu Leu Ile Glu Asn
 Ser
 20 25 30
 Leu Asp Ala Gly Ala Thr Ser Val Asp Val Lys Leu Glu Asn Tyr
 Gly
 35 40 45
 Phe Asp Lys Ile Glu Val Arg Asp Asn Gly Glu Gly Ile Lys Ala
 Val
 50 55 60
 Asp Ala Pro Val Met Ala Met Lys Tyr Tyr Thr Ser Lys Ile Asn
 Ser
 65 70 75 80
 His Glu Asp Leu Glu Asn Leu Thr Thr Tyr Gly Phe Arg Gly Glu
 Ala
 85 90 95
 Leu Gly Ser Ile Cys Cys Ile Ala Glu Val Leu Ile Thr Thr Arg
 Thr
 100 105 110

79

Ala Ala Asp Asn Phe Ser Thr Gln Tyr Val Leu Asp Gly Ser Gly
 His
 115 120 125
 Ile Leu Ser Gln Lys
 130

<210> 22

<211> 1360

<212> PRT

5 <213> Homo sapiens

<400> 22

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

Ser Ser Arg Gln Ile Lys Lys Arg Arg Val Ile Ser Asp Ser Glu
 Ser
 245 250 255
 Asp Ile Gly Gly Ser Asp Val Glu Phe Lys Pro Asp Thr Lys Glu
 Glu
 260 265 270
 Gly Ser Ser Asp Glu Ile Ser Ser Gly Val Gly Asp Ser Glu Ser
 Glu
 275 280 285
 Gly Leu Asn Ser Pro Val Lys Val Ala Arg Lys Arg Lys Arg Met
 Val
 290 295 300
 Thr Gly Asn Gly Ser Leu Lys Arg Lys Ser Ser Arg Lys Glu Thr
 Pro
 305 310 315
 320
 Ser Ala Thr Lys Gln Ala Thr Ser Ile Ser Ser Glu Thr Lys Asn
 Thr
 325 330 335
 Leu Arg Ala Phe Ser Ala Pro Gln Asn Ser Glu Ser Gln Ala His
 Val
 340 345 350
 Ser Gly Gly Gly Asp Asp Ser Ser Arg Pro Thr Val Trp Tyr His
 Glu
 355 360 365
 Thr Leu Glu Trp Leu Lys Glu Glu Lys Arg Arg Asp Glu His Arg
 Arg
 370 375 380
 Arg Pro Asp His Pro Asp Phe Asp Ala Ser Thr Leu Tyr Val Pro
 Glu
 385 390 395
 400
 Asp Phe Leu Asn Ser Cys Thr Pro Gly Met Arg Lys Trp Trp Gln
 Ile
 405 410 415
 Lys Ser Gln Asn Phe Asp Leu Val Ile Cys Tyr Lys Val Gly Lys
 Phe
 420 425 430
 Tyr Glu Leu Tyr His Met Asp Ala Leu Ile Gly Val Ser Glu Leu
 Gly
 435 440 445
 Leu Val Phe Met Lys Gly Asn Trp Ala His Ser Gly Phe Pro Glu
 Ile
 450 455 460
 Ala Phe Gly Arg Tyr Ser Asp Ser Leu Val Gln Lys Gly Tyr Lys
 Val
 465 470 475
 480
 Ala Arg Val Glu Gln Thr Glu Thr Pro Glu Met Met Glu Ala Arg
 Cys
 485 490 495
 Arg Lys Met Ala His Ile Ser Lys Tyr Asp Arg Val Val Arg Arg
 Glu
 500 505 510
 Ile Cys Arg Ile Ile Thr Lys Gly Thr Gln Thr Tyr Ser Val Leu
 Glu
 515 520 525
 Gly Asp Pro Ser Glu Asn Tyr Ser Lys Tyr Leu Leu Ser Leu Lys
 Glu

530 535 540
 Lys Glu Glu Asp Ser Ser Gly His Thr Arg Ala Tyr Gly Val Cys
 Phe
 545 550 555
 560
 Val Asp Thr Ser Leu Gly Lys Phe Phe Ile Gly Gln Phe Ser Asp
 Asp
 565 570 575
 Arg His Cys Ser Arg Phe Arg Thr Leu Val Ala His Tyr Pro Pro
 Val
 580 585 590
 Gln Val Leu Phe Glu Lys Gly Asn Leu Ser Lys Glu Thr Lys Thr
 Ile
 595 600 605
 Leu Lys Ser Ser Leu Ser Cys Ser Leu Gln Glu Gly Leu Ile Pro
 Gly
 610 615 620
 Ser Gln Phe Trp Asp Ala Ser Lys Thr Leu Arg Thr Leu Leu Glu
 Glu
 625 630 635
 640
 Glu Tyr Phe Arg Glu Lys Leu Ser Asp Gly Ile Gly Val Met Leu
 Pro
 645 650 655
 Gln Val Leu Lys Gly Met Thr Ser Glu Ser Asp Ser Ile Gly Leu
 Thr
 660 665 670
 Pro Gly Glu Lys Ser Glu Leu Ala Leu Ser Ala Leu Gly Gly Cys
 Val
 675 680 685
 Phe Tyr Leu Lys Lys Cys Leu Ile Asp Gln Glu Leu Leu Ser Met
 Ala
 690 695 700
 Asn Phe Glu Glu Tyr Ile Pro Leu Asp Ser Asp Thr Val Ser Thr
 Thr
 705 710 715
 720
 Arg Ser Gly Ala Ile Phe Thr Lys Ala Tyr Gln Arg Met Val Leu
 Asp
 725 730 735
 Ala Val Thr Leu Asn Asn Leu Glu Ile Phe Leu Asn Gly Thr Asn
 Gly
 740 745 750
 Ser Thr Glu Gly Thr Leu Leu Glu Arg Val Asp Thr Cys His Thr
 Pro
 755 760 765
 Phe Gly Lys Arg Leu Leu Lys Gln Trp Leu Cys Ala Pro Leu Cys
 Asn
 770 775 780
 His Tyr Ala Ile Asn Asp Arg Leu Asp Ala Ile Glu Asp Leu Met
 Val
 785 790 795
 800
 Val Pro Asp Lys Ile Ser Glu Val Val Glu Leu Leu Lys Lys Leu
 Pro
 805 810 815
 Asp Leu Glu Arg Leu Leu Ser Lys Ile His Asn Val Gly Ser Pro
 Leu
 820 825 830

Lys Ser Gln Asn His Pro Asp Ser Arg Ala Ile Met Tyr Glu Glu
 Thr
 835 840 845
 Thr Tyr Ser Lys Lys Lys Ile Ile Asp Phe Leu Ser Ala Leu Glu
 Gly
 850 855 860
 Phe Lys Val Met Cys Lys Ile Ile Gly Ile Met Glu Glu Val Ala
 Asp
 865 870 875
 880
 Gly Phe Lys Ser Lys Ile Leu Lys Gln Val Ile Ser Leu Gln Thr
 Lys
 885 890 895
 Asn Pro Glu Gly Arg Phe Pro Asp Leu Thr Val Glu Leu Asn Arg
 Trp
 900 905 910
 Asp Thr Ala Phe Asp His Glu Lys Ala Arg Lys Thr Gly Leu Ile
 Thr
 915 920 925
 Pro Lys Ala Gly Phe Asp Ser Asp Tyr Asp Gln Ala Leu Ala Asp
 Ile
 930 935 940
 Arg Glu Asn Glu Gln Ser Leu Leu Glu Tyr Leu Glu Lys Gln Arg
 Asn
 945 950 955
 960
 Arg Ile Gly Cys Arg Thr Ile Val Tyr Trp Gly Ile Gly Arg Asn
 Arg
 965 970 975
 Tyr Gln Leu Glu Ile Pro Glu Asn Phe Thr Thr Arg Asn Leu Pro
 Glu
 980 985 990
 Glu Tyr Glu Leu Lys Ser Thr Lys Lys Gly Cys Lys Arg Tyr Trp
 Thr
 995 1000 1005
 Lys Thr Ile Glu Lys Lys Leu Ala Asn Leu Ile Asn Ala Glu Glu
 Arg
 1010 1015 1020
 Arg Asp Val Ser Leu Lys Asp Cys Met Arg Arg Leu Phe Tyr Asn
 Phe
 1025 1030 1035
 1040
 Asp Lys Asn Tyr Lys Asp Trp Gln Ser Ala Val Glu Cys Ile Ala
 Val
 1045 1050 1055
 Leu Asp Val Leu Leu Cys Leu Ala Asn Tyr Ser Arg Gly Gly Asp
 Gly
 1060 1065 1070
 Pro Met Cys Arg Pro Val Ile Leu Leu Pro Glu Asp Thr Pro Pro
 Phe
 1075 1080 1085
 Leu Glu Leu Lys Gly Ser Arg His Pro Cys Ile Thr Lys Thr Phe
 Phe
 1090 1095 1100
 Gly Asp Asp Phe Ile Pro Asn Asp Ile Leu Ile Gly Cys Glu Glu
 Glu
 1105 1110 1115
 1120

83

Glu Gln Glu Asn Gly Lys Ala Tyr Cys Val Leu Val Thr Gly Pro
 Asn
 1125 1130 1135
 Met Gly Gly Lys Ser Thr Leu Met Arg Gln Ala Gly Leu Leu Ala
 Val
 1140 1145 1150
 Met Ala Gln Met Gly Cys Tyr Val Pro Ala Glu Val Cys Arg Leu
 Thr
 1155 1160 1165
 Pro Ile Asp Arg Val Phe Thr Arg Leu Gly Ala Ser Asp Arg Ile
 Met
 1170 1175 1180
 Ser Gly Glu Ser Thr Phe Phe Val Glu Leu Ser Glu Thr Ala Ser
 Ile
 1185 1190 1195
 1200
 Leu Met His Ala Thr Ala His Ser Leu Val Leu Val Asp Glu Leu
 Gly
 1205 1210 1215
 Arg Gly Thr Ala Thr Phe Asp Gly Thr Ala Ile Ala Asn Ala Val
 Val
 1220 1225 1230
 Lys Glu Leu Ala Glu Thr Ile Lys Cys Arg Thr Leu Phe Ser Thr
 His
 1235 1240 1245
 Tyr His Ser Leu Val Glu Asp Tyr Ser Gln Asn Val Ala Val Arg
 Leu
 1250 1255 1260
 Gly His Met Ala Cys Met Val Glu Asn Glu Cys Glu Asp Pro Ser
 Gln
 1265 1270 1275
 1280
 Glu Thr Ile Thr Phe Leu Tyr Lys Phe Ile Lys Gly Ala Cys Pro
 Lys
 1285 1290 1295
 Ser Tyr Gly Phe Asn Ala Ala Arg Leu Ala Asn Leu Pro Glu Glu
 Val
 1300 1305 1310
 Ile Gln Lys Gly His Arg Lys Ala Arg Glu Phe Glu Lys Met Asn
 Gln
 1315 1320 1325
 Ser Leu Arg Leu Phe Arg Glu Val Cys Leu Ala Ser Glu Arg Ser
 Thr
 1330 1335 1340
 Val Asp Ala Glu Ala Val His Lys Leu Leu Thr Leu Ile Lys Glu
 Leu
 1345 1350 1355
 1360

<210> 23

<211> 389

<212> PRT

5

<213> Homo sapiens

<400> 23

Met Ala Gln Pro Lys Gln Glu Arg Val Ala Arg Ala Arg His Gln
 Arg
 1 5 10 15

Ser Glu Thr Ala Arg His Gln Arg Ser Glu Thr Ala Lys Thr Pro
 Thr
 20 25 30
 Leu Gly Asn Arg Gln Thr Pro Thr Leu Gly Asn Arg Gln Thr Pro
 Arg
 35 40 45
 Leu Gly Ile His Ala Arg Pro Arg Arg Arg Ala Thr Thr Ser Leu
 Leu
 50 55 60
 Thr Leu Leu Leu Ala Phe Gly Lys Asn Ala Val Arg Cys Ala Leu
 Ile
 65 70 75 80
 Gly Pro Gly Ser Leu Thr Ser Arg Thr Arg Pro Leu Thr Glu Pro
 Leu
 85 90 95
 Gly Glu Lys Glu Arg Arg Glu Val Phe Phe Pro Pro Arg Pro Glu
 Arg
 100 105 110
 Val Glu His Asn Val Glu Ser Ser Arg Trp Glu Pro Arg Arg
 Gly
 115 120 125
 Ala Cys Gly Ser Arg Gly Gly Asn Phe Pro Ser Pro Arg Gly Gly
 Ser
 130 135 140
 Gly Val Ala Ser Leu Glu Arg Ala Glu Asn Ser Ser Thr Glu Pro
 Ala
 145 150 155
 160
 Lys Ala Ile Lys Pro Ile Asp Arg Lys Ser Val His Gln Ile Cys
 Ser
 165 170 175
 Gly Pro Val Val Pro Ser Leu Arg Pro Asn Ala Val Lys Glu Leu
 Val
 180 185 190
 Glu Asn Ser Leu Asp Ala Gly Ala Thr Asn Val Asp Leu Lys Leu
 Lys
 195 200 205
 Asp Tyr Gly Val Asp Leu Ile Glu Val Ser Gly Asn Gly Cys Gly
 Val
 210 215 220
 Glu Glu Glu Asn Phe Glu Gly Phe Thr Leu Lys His His Thr Cys
 Lys
 225 230 235
 240
 Ile Gln Glu Phe Ala Asp Leu Thr Gln Val Glu Thr Phe Gly Phe
 Arg
 245 250 255
 Gly Glu Ala Leu Ser Ser Leu Cys Ala Leu Ser Asp Val Thr Ile
 Ser
 260 265 270
 Thr Cys Arg Val Ser Ala Lys Val Gly Thr Arg Leu Val Phe Asp
 His
 275 280 285
 Tyr Gly Lys Ile Ile Gln Lys Thr Pro Tyr Pro Arg Pro Arg Gly
 Met
 290 295 300
 Thr Val Ser Val Lys Gln Leu Phe Ser Thr Leu Pro Val His His
 Lys

85

```

305          310          315
320
Glu Phe Gln Arg Asn Ile Lys Lys Lys Arg Ala Cys Phe Pro Phe
Ala
          325          330          335
Phe Cys Arg Asp Cys Gln Phe Pro Glu Ala Ser Pro Ala Met Leu
Pro
          340          345          350
Val Gln Pro Val Glu Leu Thr Pro Arg Ser Thr Pro Pro His Pro
Cys
          355          360          365
Ser Leu Glu Asp Asn Val Ile Thr Val Phe Ser Ser Val Lys Asn
Gly
          370          375          380
Pro Gly Ser Ser Arg
385

```

<210> 24

<211> 264

<212> PRT

5 <213> Homo sapiens

<400> 24

```

Met Cys Pro Trp Arg Pro Arg Leu Gly Arg Arg Cys Met Val Ser
Pro
1          5          10          15
Arg Glu Ala Asp Leu Gly Pro Gln Lys Asp Thr Arg Leu Asp Leu
Pro
          20          25          30
Arg Ser Pro Ala Arg Ala Pro Arg Glu Gln Asn Ser Leu Gly Glu
Val
          35          40          45
Asp Arg Arg Gly Pro Arg Glu Gln Thr Arg Ala Pro Ala Thr Ala
Ala
          50          55          60
Pro Pro Arg Pro Leu Gly Ser Arg Gly Ala Glu Ala Ala Glu Pro
Gln
65          70          75          80
Glu Gly Leu Ser Ala Thr Val Ser Ala Cys Phe Gln Glu Gln Gln
Glu
          85          90          95
Met Asn Thr Leu Gln Gly Pro Val Ser Phe Lys Asp Val Ala Val
Asp
          100          105          110
Phe Thr Gln Glu Glu Trp Arg Gln Leu Asp Pro Asp Glu Lys Ile
Ala
          115          120          125
Tyr Gly Asp Val Met Leu Glu Asn Tyr Ser His Leu Val Ser Val
Gly
          130          135          140
Tyr Asp Tyr His Gln Ala Lys His His His Gly Val Glu Val Lys
Glu
145          150          155
Val Glu Gln Gly Glu Glu Pro Trp Ile Met Glu Gly Glu Phe Pro
Cys
          165          170          175
Gln His Ser Pro Glu Pro Ala Lys Ala Ile Lys Pro Ile Asp Arg
Lys

```

86

```

      180      185      190
Ser Val His Gln Ile Cys Ser Gly Pro Val Val Leu Ser Leu Ser
Thr
      195      200      205
Ala Val Lys Glu Leu Val Glu Asn Ser Leu Asp Ala Gly Ala Thr
Asn
      210      215      220
Ile Asp Leu Lys Leu Lys Asp Tyr Gly Val Asp Leu Ile Glu Val
Ser
      225      230      235
Asp Asn Gly Cys Gly Val Glu Glu Glu Asn Phe Glu Gly Leu Ile
Ser
      245      250      255
Phe Ser Ser Glu Thr Ser His Met
      260

```

<210> 25

<211> 264

<212> PRT

5 <213> Homo sapiens

<400> 25

```

Met Cys Pro Trp Arg Pro Arg Leu Gly Arg Arg Cys Met Val Ser
Pro
  1      5      10      15
Arg Glu Ala Asp Leu Gly Pro Gln Lys Asp Thr Arg Leu Asp Leu
Pro
      20      25      30
Arg Ser Pro Ala Arg Ala Pro Arg Glu Gln Asn Ser Leu Gly Glu
Val
      35      40      45
Asp Arg Arg Gly Pro Arg Glu Gln Thr Arg Ala Pro Ala Thr Ala
Ala
      50      55      60
Pro Pro Arg Pro Leu Gly Ser Arg Gly Ala Glu Ala Ala Glu Pro
Gln
      65      70      75      80
Glu Gly Leu Ser Ala Thr Val Ser Ala Cys Phe Gln Glu Gln Gln
Glu
      85      90      95
Met Asn Thr Leu Gln Gly Pro Val Ser Phe Lys Asp Val Ala Val
Asp
      100      105      110
Phe Thr Gln Glu Glu Trp Arg Gln Leu Asp Pro Asp Glu Lys Ile
Ala
      115      120      125
Tyr Gly Asp Val Met Leu Glu Asn Tyr Ser His Leu Val Ser Val
Gly
      130      135      140
Tyr Asp Tyr His Gln Ala Lys His His His Gly Val Glu Val Lys
Glu
      145      150      155
Val Glu Gln Gly Glu Glu Pro Trp Ile Met Glu Gly Glu Phe Pro
Cys
      165      170      175
Gln His Ser Pro Glu Pro Ala Lys Ala Ile Lys Pro Ile Asp Arg
Lys

```

87

```

                180                185                190
Ser Val His Gln Ile Cys Ser Gly Pro Val Val Leu Ser Leu Ser
Thr
                195                200                205
Ala Val Lys Glu Leu Val Glu Asn Ser Leu Asp Ala Gly Ala Thr
Asn
                210                215                220
Ile Asp Leu Lys Leu Lys Asp Tyr Gly Val Asp Leu Ile Glu Val
Ser
225                230                235
240
Asp Asn Gly Cys Gly Val Glu Glu Glu Asn Phe Glu Gly Leu Ile
Ser
                245                250                255
Phe Ser Ser Glu Thr Ser His Met
260

```

<210> 26

<400> 26

5

000

P A T E N T K R A V

1. Fremgangsmåde til fremstilling af en hypermutabel gær, hvilken fremgangsmåde omfatter trinnene:

indføring af et polynukleotid i en gær, hvilket polynukleotid omfatter en
 5 dominant negativ allel af et fejlparringsreparationsgen, hvor den dominante negative allel reguleres inducerbart, således at ekspressionen af den dominante negative allel kan svækkes eller elimineres, hvorved cellen bliver hypermutabel.

2. Fremgangsmåde ifølge krav 1, hvor fejlparringsreparationsgenet er en
 10 MutH-homolog.

3. Fremgangsmåde ifølge krav 1, hvor fejlparringsreparationsgenet er en MutS-homolog.

4. Fremgangsmåde ifølge krav 1, hvor fejlparringsreparationsgenet er en MutL-homolog.

15 5. Fremgangsmåde ifølge krav 1, hvor fejlparringsreparationsgenet er en MutY-homolog.

6. Fremgangsmåde ifølge krav 1, hvor fejlparringsreparationsgenet er PMS2.

7. Fremgangsmåde ifølge krav 1, hvor fejlparringsreparationsgenet er
 20 vegetabilsk PMS2.

8. Fremgangsmåde ifølge krav 1, hvor fejlparringsreparationsgenet er MLH1.

9. Fremgangsmåde ifølge krav 1, hvor fejlparringsreparationsgenet er MLH3.

25 10. Fremgangsmåde ifølge krav 1, hvor fejlparringsreparationsgenet er MSH2.

11. Fremgangsmåde ifølge krav 1, hvor fejlparringsreparationsgenet er en PMSR- eller PMSL-homolog.

12. Fremgangsmåde ifølge krav 4, hvor allelen omfatter en trunke-
 30 ringsmutation.

13. Fremgangsmåde ifølge krav 6, hvor allelen omfatter en trunke-
 ringsmutation.

14. Fremgangsmåde ifølge krav 7, hvor allelen omfatter en trunkerings-
 mutation.

35 15. Fremgangsmåde ifølge krav 3, hvor allelen omfatter en trunke-

ringsmutation.

16. Fremgangsmåde ifølge krav 3, hvor allelen omfatter en trunke-
ringsmutation i kodon 134.

17. Fremgangsmåde ifølge krav 4, hvor allelen omfatter en trunke-
5 ringsmutation i kodon 134.

18. Fremgangsmåde ifølge krav 6, hvor allelen omfatter en trunke-
ringsmutation i kodon 134.

19. Fremgangsmåde ifølge krav 1, hvor polynukleotidet indføres i en
gær ved parring.

10 20. Fremgangsmåde ifølge krav 6, hvor fejlparingsreparationsgenet er
pattedyr-*PMS2*.

21. Fremgangsmåde ifølge krav 14, hvor fejlparingsreparationsgenet er
vegetabilsk *PMS2*.

22. Fremgangsmåde ifølge krav 12, hvor fejlparingsreparationsgenet er
15 *MLH1*.

23. Fremgangsmåde ifølge krav 12, hvor fejlparingsreparationsgenet er
MLH3.

24. Fremgangsmåde ifølge krav 15, hvor fejlparingsreparationsgenet er
MSH2.

20 25. Fremgangsmåde ifølge krav 15, hvor fejlparingsreparationsgenet er
MSH3.

26. Fremgangsmåde ifølge krav 15, hvor fejlparingsreparationsgenet er
MSH6.

27. Fremgangsmåde ifølge krav 12, hvor fejlparingsreparationsgenet er
25 en vegetabilsk MutL-homolog.

28. Sammensætning af dyrket, hypermutabel gær, hvilken sammen-
sætning omfatter en dominant negativ allel af et fejlparingsreparationsgen,
hvor den dominante negative allel reguleres inducerbart, således at ekspres-
sion af den dominante negative allel kan svækkes eller elimineres.

30 29. Isoleret, hypermutabel gær ifølge krav 28, hvor fejlparingsrepara-
tionsgenet er et mutL-gen eller en mutL-homolog.

30. Isoleret, hypermutabel gær ifølge krav 28, hvor fejlparingsrepara-
tionsgenet er et *PMS2*-gen eller en *PMS2*-homolog.

31. Isoleret, hypermutabel gær ifølge krav 28, hvor fejlparingsrepara-
35 tionsgenet er et *MLH1* eller en homolog.

32. Isoleret, hypermutabel gær ifølge krav 28, hvor fejlparringsreparationsgenet er en *PMSR*-homolog.

33. Isoleret, hypermutabel gær ifølge krav 28, hvor fejlparringsreparationsgenet er *mutS* eller en homolog.

5 34. Isoleret, hypermutabel gær ifølge krav 28, hvor fejlparringsreparationsgenet er eukaryot.

35. Isoleret, hypermutabel gær ifølge krav 28, hvor fejlparringsreparationsgenet er prokaryot.

10 36. Isoleret, hypermutabel gær ifølge krav 30, hvor cellerne udtrykker et protein, som består af de første 133 aminosyrer af PMS2.

37. Isoleret, hypermutabel gær ifølge krav 30, hvilken gær omfatter et protein, der består af de første 133 aminosyrer af PMS2.

38. Isoleret, hypermutabel gær ifølge krav 33, hvilken gær omfatter et pattedyr-MutS-protein.

15 39. Isoleret, hypermutabel gær ifølge krav 31, hvilken gær omfatter et protein, der består af et pattedyr-MutL-protein.

40. Isoleret, hypermutabel gær ifølge krav 28, hvilken gær omfatter et eukaryot MutL-protein.

20 41. Isoleret, hypermutabel gær ifølge krav 28, hvilken gær omfatter et eukaryot MutS-protein.

42. Fremgangsmåde til dannelsen af en mutation i et gen af interesse, hvilken fremgangsmåde omfatter trinnene med dyrkning af en gærcelle, som omfatter genet af interesse og en dominant negativ allel af et fejlparringsreparationsgen, hvilken dominant negativ allel reguleres inducerbart, således at
25 ekspression af den dominante negative allel kan svækkes eller elimineres, hvor cellen er hypermutabel; undersøgelse af cellen til bestemmelse af, hvorvidt genet af interesse indeholder en mutation; og genoprettelse af normal fejlparringsreparationsaktivitet i gærcellen.

43. Fremgangsmåde ifølge krav 42, hvor trinnet med undersøgelse omfatter analyse af en nukleotidsekvens af genet af interesse.
30

44. Fremgangsmåde ifølge krav 42, hvor trinnet med undersøgelse omfatter analyse af mRNA, der er transkriberet fra genet af interesse.

45. Fremgangsmåde ifølge krav 42, hvor trinnet med undersøgelse omfatter analyse af et protein, der er kodet af genet af interesse.

35 46. Fremgangsmåde ifølge krav 42, hvor trinnet med undersøgelse om-

fatter analyse af en fænotype, der er forbundet med genet af interesse.

47. Fremgangsmåde ifølge krav 42, hvor gæren fremstilles ved hjælp af den fremgangsmåde at indføre et polynukleotid, som omfatter en dominant negativ allel af et fejlparringsreparationsgen, i en gærce5lle, hvorved gærce5llen bliver hypermutabel.

48. Fremgangsmåde ifølge krav 47, hvor trinnet med undersøgelse omfatter analyse af nukleotidsekvensen fra genet af interesse.

49. Fremgangsmåde ifølge krav 47, hvor trinnet med undersøgelse omfatter analyse af et protein, der er kodet af genet af interesse.

- 10 50. Fremgangsmåde ifølge krav 47, hvor trinnet med undersøgelse omfatter analyse af fænotypen af genet af interesse.

51. Fremgangsmåde til dannelse af en mutation i et gen af interesse i en hypermutabel celle og efterfølgende stabilisering af genomet af cellen, hvilken fremgangsmåde omfatter trinnene:

- 15 dyrkning af en hypermutabel gærce15lle, som omfatter genet af interesse og et polynukleotid, der koder for en dominant negativ allel af et fejlparringsreparationsgen under regulering af et inducerbart transkriptionsregulatorisk element, til dannelse af en population af muterede, hypermutable gærce15ller;

- 20 dyrkning af populationen af muterede, hypermutable gærce20ller under betingelser, hvor træk selekteres;

undersøgelse af gærce20llerne, som vokser under betingelser, hvor træk selekteres, for at bestemme, hvorvidt genet af interesse indeholder en mutation; og

- 25 genoprettelse af fejlparringsreparationsaktivitet i cellen ved faldende ekspression af den dominante negative allel, hvorved der dannes en mutation i genet af interesse, og cellens genom stabiliseres.

52. Fremgangsmåde ifølge krav 51, hvor trinnet med undersøgelse omfatter analyse af en nukleotidsekvens af genet af interesse.

- 30 53. Fremgangsmåde ifølge krav 51, hvor trinnet med undersøgelse omfatter analyse af mRNA, der er transkriberet fra genet af interesse.

54. Fremgangsmåde ifølge krav 51, hvor trinnet med undersøgelse omfatter analyse af et protein, der er kodet af genet af interesse.

55. Fremgangsmåde ifølge krav 51, hvor trinnet med undersøgelse omfatter analyse af en fænotype, der er forbundet med genet af interesse.

- 35 56. Fremgangsmåde ifølge krav 51, hvilken fremgangsmåde yderligere

omfatter trinnet med anvendelse af gærcellerne, som indeholder en mutation i genet af interesse, til frembringelse af et rekombinant produkt.

57. Fremgangsmåde ifølge krav 51, hvilken fremgangsmåde yderligere omfatter trinnet med anvendelse af gærcellerne, som indeholder en mutation i genet af interesse, til udførelse af en biotransformation.

58. Fremgangsmåde ifølge krav 51, hvilken fremgangsmåde yderligere omfatter trinnet med anvendelse af gærcellerne, som indeholder en mutation i genet af interesse, til udførelse af bioremediering.

59. Fremgangsmåde ifølge krav 51, hvilken fremgangsmåde yderligere omfatter trinnet med anvendelse af gærcellerne, som indeholder en mutation i genet af interesse, til identificering af gener, der koder for virusantigener.

60. Fremgangsmåde ifølge krav 51, hvilken fremgangsmåde yderligere omfatter trinnet med anvendelse af gærcellerne, som indeholder en mutation i genet af interesse, til identificering af gærantigener.

61. Fremgangsmåde ifølge krav 51, hvilken fremgangsmåde yderligere omfatter trinnet med anvendelse af gærcellerne, som indeholder en mutation i genet af interesse, til identificering af farmaceutiske mål.

62. Fremgangsmåde ifølge krav 51, hvor mutationen i genet af interesse forårsager antibiotisk resistens, og genet klones.

63. Fremgangsmåde ifølge krav 51, hvilken fremgangsmåde yderligere omfatter trinnet med anvendelse af gærcellerne, som indeholder en mutation i genet af interesse, til screening af sammensatte biblioteker.

64. Fremgangsmåde til dannelse af hypermutationsfremmet gær, hvilken fremgangsmåde omfatter trinnene: eksponering af en gærcelle for et mutagen, hvor gærcellen er defekt med hensyn til fejlparringsreparation (MMR) på grund af tilstedeværelsen af en dominant negativ allel af mindst ét MMR-gen, som reguleres inducerbart, således at ekspresion af den dominante negative allel kan svækkes eller elimineres, hvorved der opnås en forøget mutationsrate af gærcellen.

65. Fremgangsmåde ifølge krav 64, hvor mutagenet er et DNA-alkyleringsmiddel.

66. Fremgangsmåde ifølge krav 64, hvor mutagenet er et DNA-interkalerende middel.

67. Fremgangsmåde ifølge krav 64, hvor mutagenet er et DNA-oxideringsmiddel.

68. Fremgangsmåde ifølge krav 64, hvor mutagenet er ioniserende stråling.

69. Fremgangsmåde ifølge krav 64, hvor mutagenet er ultraviolet bestråling.

5 70. Fremgangsmåde ifølge krav 64, hvor den dominante negative allel reguleres inducerbart.

71. Fremgangsmåde til dannelsen af fejlparringsreparation (MMR)-kompetent gær med nye træk, hvilken fremgangsmåde omfatter trinnene:

10 dyrkning af en gærcelle, som omfatter et gen af interesse og et polynukleotid, der koder for en dominant negativ allel af et fejlparringsreparationsgen, som reguleres inducerbart, således at ekspresionen af den dominante negative allel kan svækkes eller elimineres, til fremstilling af en population af muterede, hypermutable gærceller;

15 dyrkning af populationen af muterede, hypermutable gærceller under betingelser, hvor træk selekteres;

undersøgelse af gærcellerne, som vokser under betingelser, hvor træk selekteres, for at bestemme, hvorvidt genet af interesse indeholder en mutation;

genoprettelse af normal fejlparringsreparationsaktivitet i gærcellerne.

20 72. Fremgangsmåde ifølge krav 71, hvor gærcellen eksponeres for et mutagen til forøgelse af mutationsraten inden dyrkningstrinnet.

73. Fremgangsmåde ifølge krav 71, hvor trinnet med genoprettelse af normal fejlparringsreparationsaktivitet omfatter fjernelse af en inducer, som regulerer transkription af den dominante negative allel fra gærcellerne.

25 74. Fremgangsmåde ifølge krav 72, hvor trinnet med genoprettelse af normal fejlparringsreparationsaktivitet omfatter fjernelse af en inducer, som regulerer transkription af den dominante negative allel fra gærcellerne.

75. Fremgangsmåde ifølge krav 73, hvor induceren er methanol.

76. Fremgangsmåde ifølge krav 73, hvor induceren er galactose.

30 77. Fremgangsmåde ifølge krav 71, hvor trinnet med genoprettelse af normal fejlparringsreparationsaktivitet omfatter udskæring af den dominante negative allel ved homolog rekombination.

78. Fremgangsmåde ifølge krav 71, hvor trinnet med genoprettelse af normal fejlparringsreparationsaktivitet involverer inaktivering af den dominante negative allel.

35

79. Fremgangsmåde ifølge krav 71, hvor trinnet med genoprettelse af normal fejlparringsreparationsaktivitet omfatter anvendelse af selektionsbetingelser på gærcellerne, hvorunder celler, der har mistet den dominante negative allel, kan vokse, men hvorunder celler, som indeholder den dominante negative allel, ikke kan vokse.

80. Fremgangsmåde ifølge krav 71, hvor trinnet med genoprettelse af normal fejlparringsreparationsaktivitet udføres efter trinnet med dyrkning under betingelser, hvor træk selekteres.

81. Fremgangsmåde ifølge krav 72, hvor trinnet med genoprettelse af normal fejlparringsreparationsaktivitet udføres efter trinnet med eksponering for et mutagen og efter trinnet med dyrkning under betingelser, hvor træk selekteres.

82. Fremgangsmåde ifølge krav 72, hvor mutagenet er ioniserende stråling.

83. Fremgangsmåde ifølge krav 72, hvor mutagenet er ultraviolet (UV) bestråling.

84. Fremgangsmåde ifølge krav 71, hvor normal fejlparringsreparationsaktivitet genoprettes ved hjælp af komplementering med en vildtype-fejlparringsreparationsallel.