

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2005-511002

(P2005-511002A)

(43) 公表日 平成17年4月28日(2005.4.28)

(51) Int. Cl. ⁷	F I	テーマコード (参考)
C 1 2 N 15/09	C 1 2 N 15/00	Z N A A
A O 1 K 67/027	A O 1 K 67/027	2 G O 4 5
A 6 1 K 45/00	A 6 1 K 45/00	4 B O 2 4
A 6 1 P 19/00	A 6 1 P 19/00	4 B O 6 3
C 1 2 N 5/10	C 1 2 Q 1/02	4 B O 6 5
		4 C O 8 4
	審査請求 未請求 予備審査請求 有	(全 484 頁) 最終頁に続く

(21) 出願番号	特願2002-589632 (P2002-589632)	(71) 出願人	397012945
(86) (22) 出願日	平成14年5月13日 (2002.5.13)		ワイス
(85) 翻訳文提出日	平成15年11月11日 (2003.11.11)		アメリカ合衆国 ニュージャージー州 マ
(86) 国際出願番号	PCT/US2002/014876		ジソン, ファイブ ジラルダ ファームズ
(87) 国際公開番号	W02002/092764		(番地なし)
(87) 国際公開日	平成14年11月21日 (2002.11.21)	(74) 代理人	100099623
(31) 優先権主張番号	60/290, 071		弁理士 奥山 尚一
(32) 優先日	平成13年5月11日 (2001.5.11)	(74) 代理人	100096769
(33) 優先権主張国	米国 (US)		弁理士 有原 幸一
(31) 優先権主張番号	60/291, 311	(74) 代理人	100107319
(32) 優先日	平成13年5月17日 (2001.5.17)		弁理士 松島 鉄男
(33) 優先権主張国	米国 (US)	(72) 発明者	バビィ, フィリップ
(31) 優先権主張番号	60/353, 058		アメリカ合衆国マサチューセッツ州018
(32) 優先日	平成14年2月1日 (2002.2.1)		27, ダンスタブル, アダムズ・ストリー
(33) 優先権主張国	米国 (US)		ト 75
			最終頁に続く

(54) 【発明の名称】 骨量変調のトランスジェニック動物モデル

(57) 【要約】

本発明は動物細胞およびトランスジェニック動物においてHBMタンパク質を発現させるのに用いられる方法および材料に関する。本発明また高骨量遺伝子、対応する野生型遺伝子、およびその変異体を発現するトランスジェニック動物にも関する。本発明はコード化配列などの核酸、オリゴヌクレオチドプライマーおよびプローブ、タンパク質、クローニングベクター、発現ベクター、形質転換された宿主、医薬用組成物の開発方法、骨発達に関与する分子の同定方法、ならびに骨発達に関与する疾患の診断および処置方法を提供する。好ましい実施形態では、本発明は骨粗鬆症を処置、診断および予防するための方法を志向する。

【特許請求の範囲】

【請求項 1】

配列番号 2 の位置 5 8 2 におけるチミンを少なくとも含み、配列番号 2 の少なくとも 1 5 個の連続するヌクレオチドを含む配列に機能的に連結され、動物細胞および／またはヒト細胞におけるタンパク質の発現を導くことができるプロモーター領域を含む核酸を含む体細胞および／または生殖細胞を有するトランスジェニック動物。

【請求項 2】

配列番号 4 をコードする配列を含み、配列番号 4 の位置 1 7 1 のバリンに相当するバリンのコドンを含み、核酸であって、動物細胞および／またはヒト細胞におけるタンパク質の発現を導くことができる、機能的に連結されたプロモーター領域をさらに含む核酸を含む体細胞および／または生殖細胞を有するトランスジェニック動物。 10

【請求項 3】

核酸が配列番号 2 を含む請求項 1 または 2 に記載のトランスジェニック動物。

【請求項 4】

配列番号 1 の少なくとも 1 5 個の連続するヌクレオチドを含む配列に機能的に連結された、動物細胞および／またはヒト細胞におけるタンパク質の発現を導くプロモーター領域を含む核酸を含む体細胞および／または生殖細胞を有する、骨密度変調の研究のためのトランスジェニック動物であって、骨量が、骨密度、骨力、骨梁数、骨の大きさおよび骨組織結合性の中から選択される 1 つ以上のパラメーターについて、同一種の非トランスジェニック動物と比較して、変調されているトランスジェニック動物。 20

【請求項 5】

プロモーター領域が、CMV、RSV、SV40 および EF-1 α 、CMV β アクチン、ヒストン、I 型コラーゲン、TGF β 1、SX2、c fos/c jun、Cbfa1、Fra/Jun、Dlx5、オステオカルシン、オステオポンチン、骨シアロタンパク質、またはコラゲナーゼプロモーター領域である請求項 1 から 4 のいずれかに記載のトランスジェニック動物。

【請求項 6】

プロモーター領域が、骨特異的プロモーター領域である請求項 1 から 4 のいずれかに記載のトランスジェニック動物。

【請求項 7】

プロモーター領域が、CMV β アクチンプロモーター領域である請求項 1 から 4 のいずれかに記載のトランスジェニック動物。 30

【請求項 8】

プロモーター領域が、I 型コラーゲンプロモーター領域である請求項 1 から 4 のいずれかに記載のトランスジェニック動物。

【請求項 9】

請求項 1 から 8 のいずれかに記載のトランスジェニック動物の胚。

【請求項 10】

トランスジェニック動物が、ヒト HBM タンパク質を発現する請求項 1 から 3 のいずれかに記載のトランスジェニック動物。 40

【請求項 11】

ヒト HBM タンパク質が、骨組織において最も多く発現される請求項 10 に記載のトランスジェニック動物。

【請求項 12】

HBM 表現型を表す請求項 1 から 3 のいずれかに記載のトランスジェニック動物。

【請求項 13】

骨量が、骨密度、骨力、骨梁数、骨の大きさおよび骨組織結合性の中から選択される 1 つ以上のパラメーターについて、同一種の非トランスジェニック動物と比較して、変調されている請求項 1 から 3 のいずれかに記載のトランスジェニック動物。

【請求項 14】

ヒトHBMタンパク質が発現され、繁殖力があり、ヒトHBM遺伝子をその子孫に伝える請求項1から3のいずれかに記載のトランスジェニック動物。

【請求項15】

ヒトLRP5タンパク質が発現され、繁殖力があり、ヒトHBM遺伝子をその子孫に伝える請求項4に記載のトランスジェニック動物。

【請求項16】

請求項1から15のいずれかに記載のトランスジェニック動物またはその子孫から産生されたトランスジェニック動物。

【請求項17】

LRP5をコードする遺伝子に変化を加えたゲノムを有するトランスジェニックマウスであって、該変化が、相同組換えによるジーンターゲティングのための核酸であって、マウスLRP5遺伝子に相同である第1セクションと、マウスLRP5遺伝子の別のセクションに相同である第2セクションと、第1および第2セクション間におけるLRP5遺伝子の一部の操作された欠失、核酸配列の変化または核酸の挿入を含む中央セクションとを含む核酸の、胚性幹細胞または多能性細胞への導入により引き起こされ、該核酸が内因性遺伝子と相同組換えできるトランスジェニックマウス。 10

【請求項18】

ジーンターゲティングのための核酸の中央セクションが、ATG開始コドンの操作された欠失、操作されたフレームシフト変異、操作された停止コドン、ネオマイシン抵抗性配列、loxP組換え部位、または合成転写中止配列を含む請求項17に記載のトランスジェニックマウス。 20

【請求項19】

ジーンターゲティングのための核酸が、マウスLRP5遺伝子のイントロン配列およびエクソン配列の双方をさらに含む請求項17に記載のトランスジェニックマウス。

【請求項20】

ジーンターゲティングのための核酸が、マウスLRP5遺伝子のアミノ酸配列の位置170においてグリシンからバリンへの変化をコードするコドンを含み、変化した遺伝子がHBMタンパク質をコードする請求項17に記載のトランスジェニックマウス。

【請求項21】

前記変化が、LRP5遺伝子が発現しないようにするLRP5遺伝子の破壊である請求項17に記載のトランスジェニックマウス。 30

【請求項22】

LRP6をコードする遺伝子に変化を加えたゲノムを有するトランスジェニックマウスであって、該変化が、相同組換えによるジーンターゲティングのための核酸であって、マウスLRP6遺伝子に相同である第1セクションと、マウスLRP6遺伝子の別のセクションに相同である第2セクションと、第1および第2セクション間におけるLRP6遺伝子の一部の操作された欠失、核酸配列の変化または核酸の挿入を含む中央セクションとを含む核酸の、胚性幹細胞または多能性細胞への導入により引き起こされ、該核酸が内因性遺伝子と相同組換えでき、該トランスジェニック動物が変調されたWnt活性、Dkk活性、脂質レベルまたは骨量を有するトランスジェニックマウス。 40

【請求項23】

前記変化が、LRP6遺伝子が発現しないようにするLRP6遺伝子の破壊である請求項22に記載のトランスジェニックマウス。

【請求項24】

繁殖力があり、前記変化した遺伝子をその子孫に伝える請求項22に記載のトランスジェニックマウスまたはその子孫。

【請求項25】

繁殖力があり、前記変化した遺伝子をその子孫に伝える請求項17から24のいずれかに記載のトランスジェニックマウス。

【請求項26】

繁殖力があり、前記変化した遺伝子をその子孫に伝えるトランスジェニックマウスであって、子孫が、野生型と比較して、骨密度、骨力、骨梁数、骨の大きさおよび骨組織結合性の中から選択される少なくとも3つのパラメーターにより示される変調された骨量の表現型を表す請求項17から24のいずれかに記載のトランスジェニックマウス。

【請求項27】

マウス胚性幹細胞をマウス胚盤胞へ導入することで産生される請求項17から26のいずれかに記載のトランスジェニックマウス。

【請求項28】

請求項17から27のいずれかに記載のトランスジェニックマウスから産生されたトランスジェニックマウス。

【請求項29】

請求項1から28のいずれかに記載のトランスジェニック動物からなる動物の第1群および対照動物の第2群を含む骨密度変調の研究のための動物モデル。

【請求項30】

対照動物群が、ヒトLRP5をコードする核酸を含む細胞を有するトランスジェニック動物を含む請求項29に記載の動物モデル。

【請求項31】

骨量決定要素の研究方法であって、

(a) 請求項1から28のいずれかに記載のトランスジェニック動物の第1群を供するステップと、

(b) 該トランスジェニック動物における少なくとも1つの骨発達パラメーターを測定するステップと、
を含む方法。

【請求項32】

骨量のモデレーターの研究方法であって、

(a) 請求項1から28のいずれかに記載のトランスジェニック動物の第1群を供するステップと、

(b) 被験化合物を投与するステップと、

(c) 被験化合物を投与されたトランスジェニック動物における少なくとも1つの発達パラメーターを測定するステップと、
を含む方法。

【請求項33】

骨量の研究方法であって、

(a) 請求項1から28のいずれかに記載のトランスジェニック動物の第1群を供するステップと、

(b) 実験手順を施すステップと、

(c) 実験手順を施された動物における少なくとも1つの発達パラメーターを測定するステップと、
を含む方法。

【請求項34】

被験化合物が、注射投与、経口投与、坐薬投与、移植片投与、または局所的投与される請求項32に記載の方法。

【請求項35】

実験手順が、卵巣切除、骨負荷の制限、および骨負荷の増強の中から選択される請求項33に記載の方法。

【請求項36】

トランスジェニック動物群がHBMを発現し、HBMを発現し、被験化合物を投与されたトランスジェニック動物において骨量密度が変調されている請求項32に記載の方法。

【請求項37】

トランスジェニックマウス群がHBMを発現し、HBMを発現し、被験化合物を投与さ

10

20

30

40

50

れたトランスジェニックマウスにおいて骨量密度が上昇している請求項 36 に記載の方法。

【請求項 38】

被験化合物が、ホルモン、成長因子、ペプチド、RNA、DNA、ミネラル、ビタミン、天然生成物、または合成有機化合物を含む請求項 32 に記載の方法。

【請求項 39】

実験手順が、外科的手順、遺伝子治療手順、薬物治療手順、食事療法、または運動を含む請求項 33 に記載の方法。

【請求項 40】

HBM が骨障害に及ぼす影響を研究する方法であって、

10

(a) 骨障害の表現型を有する動物の胚を供するステップと、

(b) 請求項 1 から 4 のいずれかに記載の核酸を該胚の第 1 の群に導入するステップと、

(c) 胚を偽妊娠マウスに移すステップと、

(d) 得られたマウスの少なくとも 1 つの発達パラメーターを測定するステップと、を含む方法。

【請求項 41】

骨形成／再吸収の代理マーカーを同定する方法であって、

(a) 請求項 29 に記載の骨発達の動物モデルを供するステップと、

(b) 該動物における代理マーカー候補を定量的に測定するステップと、

20

(c) トランスジェニック動物群の測定値を対照動物群の測定値と比較するステップと、

を含む方法。

【請求項 42】

HBM が心臓障害に及ぼす影響を研究する方法であって、

(a) 請求項 1 から 28 のいずれかに記載のトランスジェニック動物の第 1 群を供するステップと、

(b) 被験化合物を投与されたトランスジェニック動物の少なくとも 1 つの心臓健康パラメーターを測定するステップと、

を含む方法。

30

【請求項 43】

HBM が心臓障害に及ぼす影響を研究する方法であって、

(a) 請求項 1 から 28 のいずれかに記載のトランスジェニック動物の第 1 群を供するステップと、

(b) 被験化合物を投与するステップと、

(c) 被験化合物を投与されたトランスジェニック動物の少なくとも 1 つの心臓健康パラメーターを測定するステップと、

を含む方法であって、被験化合物が、ホルモン、成長因子、ペプチド、RNA、DNA、ミネラル、ビタミン、天然生成物、または合成有機化合物を含む方法。

【請求項 44】

40

心臓健康パラメーターが、血清脂質濃度である請求項 43 に記載の方法。

【請求項 45】

心臓保護処置から骨量変調効果をスクリーニングする方法であって、

請求項 31 に記載の方法を含み、心臓保護処置を前記動物の第 1 群のサブグループに投与するステップをさらに含む方法。

【請求項 46】

請求項 1 から 28 のいずれかに記載のトランスジェニック動物またはその子孫に由来する単離された細胞。

【請求項 47】

骨芽細胞または破骨細胞である請求項 46 に記載の単離された細胞。

50

【請求項 48】

内因性 L R P 5 の発現が、相同組換えまたは非相同組換えにより導入された変化した遺伝子の調節配列により変調されるトランスジェニック動物。

【請求項 49】

骨密度を変調する方法であって、

L R P 5、H B M または L R P 6 の内因性の発現を変調する化合物を投与するステップを含む方法。

【請求項 50】

心臓保護処置から骨量変調効果をスクリーニングする方法であって、

請求項 1 から 28 のいずれかに記載のトランスジェニック動物の第 1 群を供するステップと、

心臓保護処置を該動物の第 1 群のサブグループに投与するステップと、

少なくとも処置されたマウスにおいて少なくとも 1 つの骨変調パラメーターを測定するステップと、

を含む方法。

【請求項 51】

前記核酸が、配列番号 4 の位置 1 3 3 0 で少なくともアラニンからバリンへの置換をさらにコードする請求項 1 から 16 のいずれかに記載のトランスジェニック動物。

【請求項 52】

相同組換えによるジーンターゲティングのための核酸であって、マウス L R P 5 遺伝子に相同である第 1 セクションと、マウス L R P 5 遺伝子の別のセクションに相同である第 2 セクションと、第 1 および第 2 セクション間における L R P 5 遺伝子の一部の操作された欠失、核酸配列の変化または核酸の挿入を含む中央セクションとを含み、内因性遺伝子と相同組換えできる核酸。

【請求項 53】

中央セクションが、A T G 開始コドンの操作された欠失、操作されたフレームシフト変異、操作された停止コドン、ネオマイシン抵抗性配列、l o x P 組換え部位、または合成転写中止配列を含む請求項 52 に記載の核酸。

【請求項 54】

マウス L R P 5 遺伝子のイントロン配列およびエクソン配列の双方をさらに含む請求項 52 に記載の核酸。

【請求項 55】

マウス L R P 5 遺伝子のアミノ酸配列の位置 1 7 0 でグリシンからバリンへの変化をコードするコドンを含み、請求項 52 に記載の単離された核酸。

【請求項 56】

L R P 5 をコードする遺伝子に変化を加えたゲノムを有するトランスジェニックマウスを産生する方法であって、

(a) 請求項 52 に記載の核酸を供するステップと、

(b) 該核酸をマウス胚性幹細胞に導入するステップと、

(c) 該核酸を含む胚性幹細胞を選択するステップと、

(d) ステップ (c) の胚性幹細胞をマウス胚盤胞に導入するステップと、

(e) ステップ (d) の胚盤胞を偽妊娠マウスに移すステップと、

(f) 移した胚盤胞を、該核酸に関してキメラであるマウスに発達させるステップと、を含む方法。

【請求項 57】

胚性幹細胞の導入が、マイクロインジェクションによるものである請求項 56 に記載の方法。

【請求項 58】

(a) キメラマウスと野生型マウスと交配させ、前記変化に関してヘテロ接合性であるマウスを得るステップと、

(b) 該ヘテロ接合性マウスを交配させ、前記変化に関してホモ接合性であるマウスを産生するステップと、

をさらに含む請求項 57 に記載の方法。

【請求項 59】

骨量に関連する遺伝子を同定する方法であって、

(a) 請求項 29 に記載の骨発達の動物モデルを供するステップと、

(b) 該動物における遺伝子発現のプロファイルを測定するステップと、

(c) 該動物の第 1 の群の測定値を動物の対照群の測定値と比較するステップと、を含む方法。

【発明の詳細な説明】

10

【技術分野】

【0001】

本発明は一般に遺伝的、ゲノムおよび分子生物学の分野に関する。本発明は高骨量遺伝子 (high bone mass gene) および対応する野生型遺伝子、ならびにその変異体の単離、検出および配列に関する。本発明はまた高骨量 (HBM) 遺伝子、対応する野生型遺伝子およびその変異体にも関する。本発明で同定された遺伝子は骨発達のオントロロジーおよび生理学に関係している。本発明はまた核酸、タンパク質、クローニングベクター、発現ベクター、形質転換された宿主、医薬用組成物を開発する方法、骨発達に関与する分子を同定する方法、ならびに骨発達に関与する疾患の診断および処置方法をも提供する。本発明は、さらに HBM 表現型を研究するためのトランスジェニック動物、HBM 遺伝子の作用のメカニズム、ならびに正常および異常な骨状態に影響する因子および処置に関する。好ましい実施形態では、本発明は代謝性骨疾患、例えば骨粗鬆症などの骨の正常および異常な状態を処置、診断、予防およびスクリーニングする方法を志向する。

20

【背景技術】

【0002】

骨粗鬆症の最も一般的な型の 2 つは閉経後および老年性骨粗鬆症である。骨粗鬆症は男性および女性を襲い、その他の骨の異常と共に、高齢者にとってかつてない健康上のリスクを提示している。骨粗鬆症の最も一般的な型は閉経期に関連する。大抵の女性は月経の停止後 3 から 6 年以内に骨の骨梁区画の骨量の 20 から 60 % を喪失する。この急速な喪失は一般に骨吸収および形成の増加に関連する。しかしながら、再吸収のサイクルはより優勢であり、結果的には骨量の純損失となる。骨粗鬆症は閉経後の女性の間では一般的で深刻な疾患である。米国内でおよそ 25000000 人の女性がこの疾患に苦しんでいる。骨粗鬆症の結果は双方共に個人にとって有害であり、また疾患の続発症から慢性化ならびに過度および長期の援助 (入院および老人ホームでの介護) の必要性のために多大な経済的損失の原因となる。これはとりわけより高齢の患者において真実である。さらに、骨粗鬆症は一般に生命を脅かす症状としては考えられていないが、死亡率の 20 から 30 % は高齢の女性の股関節部の骨折に関連している。この大きなパーセンテージの死亡率が直接的に閉経後骨粗鬆症に関連し得る。

30

【0003】

閉経後骨粗鬆症の影響に対して最も脆弱な骨の組織は骨梁骨である。この組織はしばしば海綿状骨と称され、とりわけ関節近くの骨の末端近くおよび脊椎の椎骨に集中している。骨梁組織は互いに相互接続している小型の構造ならびに骨の外側表面および中央の骨幹を構成するより堅固で、密な皮質組織を特徴とする。この骨梁の交差する網目は外側の皮質構造に対して側壁保持を提供し、全体の構造の生物機械的強度にとって非常に重要である。閉経後骨粗鬆症では、網目の再吸収および骨梁の喪失が第 1 次的であり、これは骨不全および骨折に至る。閉経後の女性の骨梁の喪失に鑑みて、最も一般的な骨折が骨梁支持に高度に依存する骨に関連するもの、例えば椎骨、大腿骨頸部および前腕であることは驚くべきことではない。実際に、股関節部骨折、コーレス骨折、および椎骨圧挫が閉経後骨粗鬆症の徴候である。

40

【0004】

50

最も早くに認められた閉経後骨粗鬆症の処置の方法はエストロゲン置換療法である。この治療は高頻度で成功するが、長期にわたるエストロゲン処置の望ましくない副作用のために患者のコンプライアンスが低い。前記したエストロゲン置換療法の高頻度の副作用には月経の再発、胃内ガス貯留、抑うつ、および乳癌または子宮癌の恐れなどがある。子宮摘出を行っていないその女性の子宮癌の既知の脅威を制限するために、エストロゲンおよびプロゲステンの周期的治療法がしばしば用いられる。このプロトコルは受胎調節計画に用いられるものに類似しており、しばしばプロゲステンの副作用特性のために女性に耐容されない。さらに最近では、元来乳癌の処置のために開発された特定の抗エストロゲンが閉経後骨粗鬆症の実験モデルで有効であることが示されている。これらの作用物質のなかにラロキシフェン(米国特許第5,393,763号およびBlackら、J.Clin.Invest.93:63-69(1994))がある。加えて、乳癌の処置に広く用いられている臨床薬であるタモキシフェンは乳癌を患う閉経後の女性の骨塩量を増加させることが示されている(Loveら、N.Engl.J.Med.326:852-856(1992))。 10

【0005】

閉経後骨粗鬆症の処置のための別の治療はカルシトニンの使用である。カルシトニンは骨再吸収を阻止し、多くの国でこの使用が承認されている天然発生ペプチドである(Overgaardら、Br.Med.J.305:556-561(1992))。しかしながらカルシトニンの使用は幾分制限されている。その効果は骨塩量の増加においてまさに中程度であり、処置には非常に経費がかかる。閉経後骨粗鬆症の処置のための別の治療はビスホスフォネートの使用である。これらの化合物は元来パジェット病および悪性高カルシウム血症で使用するために開発されたものである。これらは骨再吸収を阻止することが示されている。アレンドロネートはこのクラスの化合物の1つであり、閉経後骨粗鬆症の処置用に承認されている。これらの作用物質は骨粗鬆症の処置の助けとなり得るが、これらの作用物質は骨軟化症、骨における極度に長い半減期(2年以上)、および「凍結骨症候群(frozen bone syndrome)」、例えば正常な骨リモデリングの休止の可能性を含む潜在的な傾向性をも有している。 20

【0006】

老年性骨粗鬆症は骨塩量の喪失により特徴づけられ、骨折率、罹患率および関連する死亡率の増加に至るという点で閉経後骨粗鬆症に類似している。これは一般に晩年、すなわち70歳後に起こる。組織化学的には、老年性骨粗鬆症は女性においてより一般的になっているが、より高齢の男性集団の出現で、この疾患は双方の性別における主要な因子になりつつある。ホルモン例えばテストステロンまたはエストロゲンがこの疾患で、あるとすれば何の役割を果たしているのかは明らかではなく、その病因は不明確である。この疾患の処置はあまり満足のいくものではない。ホルモン治療、女性ではエストロゲン、男性ではテストステロンは不明瞭な結果を示しており；カルシトニンおよびビスホスフォネートは幾分役立つ。 30

【0007】

成熟時の骨格の骨量のピークは主として遺伝子調節下にある。双生児の研究により成人の一卵性双生児間の骨量の不一致は二卵性双生児間の不一致よりも小さいことが示されており(Slemendaら、J.Bone Miner.Res.6:561-567(1991);Youngら、J.Bone Miner.Res.6:561-567(1995);Pocockら、J.Clin.Invest.80:706-710(1987);Kellyら、J.Bone Miner.Res.8:11-17(1993))、骨格骨量の不一致の60%またはそれ以上までが遺伝されていると推定されている(Krallら、J.Bone Miner.Res.10:S367(1993))。骨格骨量のピークは老齢期の骨量の最も強力な決定要因であるが(Huiら、Ann.Int.Med.111:355-361(1989))、成人および晩年の加齢関連の骨損失の速度もまた強力な決定要素(determinant)である(Huiら、Osteoporosis Int.1:30-34(1995))。骨量は原則的に骨折リスクの決定要素であるので、遺伝した成熟時に達成した骨格骨量のピークは晩年の骨折の個々のリスクの重要な決定因子である。従って、骨量の遺伝的基盤の研究は骨粗鬆症による骨折の病因学において相当興味深いものである。 40

【0008】

最近では、骨粗鬆症の分野で骨量ピークの遺伝的調節に強く興味をもたれるようになって 50

てきている。興味は主に、正常範囲内での骨量の変動にとの関連性を試験するための適当な多型性を有する候補遺伝子に注がれているか、または骨粗鬆症の患者に見出される範囲の骨量低下に関連する遺伝子および遺伝子座の試験に注がれている。ビタミンDレセプター遺伝子座(VDR)(Morrisonら、Nature,367:284-287(1994))、PTH遺伝子(Howardら、J.Clin.Endocrinol.Metab.80:2800-2805(1995);Johnsonら、J.Bone Miner.Res.8:11-17(1995);Gongら、J.Bone Miner.Res.10:S462(1995))、およびエストロゲンレセプター遺伝子(Hosoiら、J.Bone Miner.Res.10:S170(1995);Morrisonら、Nature 367:284-287(1994))がこの仕事で最も傑出している。骨量(表現型)は連続的、量的、多遺伝子的特質であり、環境因子、例えば栄養、同時罹患している疾患、年齢、身体活動度、およびその他の因子により混乱されるので、これらの研究は困難である。また、この型の研究設計には多くの対象が必要とされる。とりわけ今日までのVDR研究の結果は解り難く、矛盾している(Garneroら、J.Bone Miner.Res.10:1283-1288(1995);Eismanら、J.Bone Miner.Res.10:1289-1293(1995);Peacock、J.Bone Miner.Res.10:1294-1297(1995))。さらに、この仕事はこれまでは、遺伝的影響が骨量において効果を奏するメカニズムにそれほど光を注いでいない。

【0009】

骨量ピークは環境因子よりもむしろ遺伝的因子により主に決定されることは周知であるが、骨量の変動に関連する遺伝子座(および究極的には遺伝子)を決定する研究は困難で、経費がかかる。連鎖解析、例えば同胞対または拡大家族の力を利用する研究設計は単純な関連性研究よりも一般に有益であるが、後者も価値がある。しかしながら、骨量に関する遺伝連鎖研究は2つの主要な問題により妨害される。第1の問題は前記で簡単に論じた表現型である。骨量は連続的な、量的な特質であり、別個の表現型を確立するのは困難である。測定のための各々の解剖学的な部位はいくつかの遺伝子により影響を受け、その多くは部位間で異なっているかもしれない。第2の問題は表現型の年齢要素である。個体を骨量が低いと同定できるまでに、その親またはその他の前世代の構成員が死亡している可能性が高く、従って研究に利用できず、若い世代は未だ骨量ピークに達していないかもしれない。その表現型の遺伝子分析を不正確にしている。

【0010】

それにもかかわらず、連鎖解析を用いて遺伝的「障害」を引き起こす遺伝子の位置を見出すことができ、障害の生化学的性質の知識を何ら必要としない、すなわち障害を引き起こすと考えられている変異タンパク質を知っている必要はない。伝統的な研究法は公知のタンパク質を評価すべき候補として関連づける、疾患の過程に関係する仮説に依存する。連鎖解析を用いる遺伝子局在研究法を用いて最初に欠損遺伝子が位置する一般的な染色体領域を見出し、次にできるだけ正確に具体的な変異遺伝子の位置を決定するために領域の大きさを徐々に狭めることができる。候補領域内で遺伝子そのものを見出した後、メッセンジャーRNAおよびタンパク質を同定し、DNAと一緒に変異に関して検査する。

【0011】

疾患を引き起こす変化した遺伝子を見出す前でも疾患の位置を出生前診断に用いることができるので、遺伝子局在研究法は現実的な意味を有している。連鎖解析により、その多くが病気の子供を有してもいない家族が、疾患遺伝子のキャリアであるかどうかを知り、分子診断を通して生まれていない子供の状態を評価することが可能になる。次いで、家族内の疾患の伝達を用いて欠陥遺伝子を見出すことができる。本明細書で用いるように、「高骨量」(HBM)に対する言及は病態への言及と類似しているが、現実的な観点から、対象が骨粗鬆症として知られている疾患を回避するのを、高骨量が実際に助けることができる。

【0012】

親から子孫への染色体の遺伝の性質のために連鎖解析は可能である。減数分裂の間に、2つの親の相同対は娘細胞への適切な分離に導かれる。それらは並び、対を形成しているが、2つの相同体は染色体片を交換し、その事象を「クロスオーバー」または「組換え」と称される。得られた染色体はキメラであり、すなわちそれらは双方の親相同対に由来す

る部分を含有している。2つの配列が染色体上ですぐ近くであるほど、組換え事象がその間で生じる可能性は低くなり、より近くで連鎖する。連鎖解析実験では、染色体上の2つの位置は、それらの間の組換えの頻度を決定するために、1世代から次の世代に追跡される。遺伝される疾患の研究では、染色体位置の1つを疾患遺伝子またはその正常な対応物により標識する、すなわち染色体領域の遺伝を、個体が障害の症状を表示するかどうかを試験することにより決定することができる。標識する2つの相同体が、それらが保持する「マーカー」配列のコピーに基づいて区別され得るように、その他の位置を集団の天然の変動を示すDNA配列により標識する。どの家族でも、遺伝的マーカー配列の遺伝を病態の遺伝と比較する。常染色体優性の障害、例えば高骨量を担持する家族内で、罹患する全ての個体がマーカーの同一の形態を担持し、罹患していない全ての個体は少なくとも1つのマーカーとは異なる形態を担持する場合、疾患遺伝子およびマーカーが互いに接近して位置している可能性が大きい。このように、染色体を公知のマーカーで系統的に調べ、病態と比較することができる。異なる家族から得られたデータを組み合わせて、統計学的方法を用いてコンピューターにより一緒に分析する。結果は、それらの間で異なる距離が可能である遺伝子マーカーおよび疾患の間の連鎖の可能性を示す情報である。陽性の結果は、疾患がマーカーに非常に接近しており、一方陰性の結果は染色体上で非常に離れているか、または全く異なる染色体上であることを意味し得る。

【0013】

連鎖解析は、規定のマーカー遺伝子座で罹患家族の全員を分類し、特定の病態の同時遺伝をマーカープローブで評価し、それによりどのような頻度でそのうちの2つが同時に遺伝されたかを決定することにより実施される。組換え頻度を2つの遺伝子座の遺伝子距離の測定として用いることができる。組換え頻度1%は1地図単位または1センチモルガン(cM)と等価であり、これはDNA1000kbとほぼ等価である。この関係は約20%または20cMまで保持される。

【0014】

ヒトの全ゲノムは3300cMの長さである。5から10cMのマーカー遺伝子座内の未知の疾患遺伝子を見出すために、ヒト全ゲノムをおよそ10cM間隔で有益なマーカー遺伝子座およそ330を検査することができる(Botsteinら、Am.J.Hum.Genet.32:314-331(1980))。連鎖結果の信頼性は多くの統計学的方法を用いることにより確立される。ヒトにおける連鎖解析に最も一般的に用いられる方法は、Ott、Am.J.Hum.Genet.28:528-529(1976)によりコンピュータープログラムLIPEDに組み込まれたLODスコア法(Morton, Prog.Clin.Biol.Res.147:245-265(1984);Mortonら、Am.J.Hum.Genet.38:868-883(1986))である。LODスコアは2つの遺伝子座が所定の距離で連鎖する可能性の、連鎖しない可能性に対する比率の対数である(>50cM離れている)。対数値を用いる利点は同一の疾患を有する家族間で和を取ることができる点である。これは比較的小型のヒト家族で行うことが必要となる。

【0015】

慣例により、+3.0以上の全LODスコア(すなわち規定の組換え頻度での連鎖の可能性は連鎖していない可能性よりも1000倍大きくなる)はその特定の組換え頻度での連鎖の有意な証拠であると考えられる。-2.0未満の全LODスコア(すなわち連鎖していない可能性は規定の頻度での連鎖の可能性よりも100倍大きくなる)は検討中の2つの遺伝子座が特定の組換え頻度で連鎖されていないという強力な証拠であると考えられる。最近までは、たいていの連鎖解析が2点データを基にして実施されており、これは検討中の障害と特定の遺伝子マーカー間の関係である。しかしながら、ここ数年のヒトゲノムのマッピングにおける急速な進歩およびコンピューター方法学における付随的な改良の結果として、多点データを用いる連鎖解析を実行することが可能になっている。マーカー間の組換え距離が解っている場合、多点解析により疾患といくつかの連結された遺伝子マーカーの間の連鎖の同時解析が提供される。

【0016】

多点解析は2つの理由から有利である。第1に通常家系の有益性が増す。各々の家系は

、マーカー遺伝子座に関してヘテロ接合性である親の数および家族内の罹患した個体数に依存して、特定量の潜在性の情報を有している。しかしながら、全てのそれらの個体において有益であるほど十分に多型性であるマーカーは少ししかない。複数のマーカーが同時に検討される場合、次いで少なくとも1つのマーカーに関してヘテロ接合性である個体の可能性は大きく増加する。第2に、マーカー間で疾患遺伝子の位置の提示を決定することができる。これによりフランキングマーカーの同定が可能になり、従って疾患遺伝子が属する小型の領域の単離が実際に可能になる。Lathropら、Proc.Natl.Acad.Sci.USA 81:3443-3446(1984)は、多点解析に関して最も広く用いられているコンピューターパッケージ、LINKAGEについて記載している。

【0017】

10

当分野では高骨量表現型に関連する遺伝子を同定する必要性がある。高骨量遺伝子および表現型の研究のための手段に関する必要性もある。より一般的には、診断手段および処置の開発のための必要性がある。本発明はこれら、およびその他の重要な目的を志向する。

【発明の開示】

【0018】

本発明は遺伝子連鎖および変異分析による染色体11q13.3上のLRP5遺伝子およびLRP5遺伝子のHBM遺伝子座の同定について記載する。LRP5遺伝子およびそれがコードするLRP5タンパク質を発明者および共同研究者は以前にZmax1遺伝子およびZmax1（またZmaxとも）と称した。遺伝子およびその産物はその他の研究者により別の命名法によりLR3とも称されている。Zmax、Zmax1、LRP5およびLR3は同義語である。遺伝子に連結した遺伝子マーカーの使用はその発見の助けとなった。連鎖解析および変異分析を用いることにより、HBMの素因を有するヒトを容易に同定することができる。バクテリア人工染色体を用いるクローニング法により発明者が11q13.3の染色体領域に着目して、常染色体優性遺伝子のシーケンシングを促進することが可能になった。加えて、本発明はLRP5遺伝子およびHBM遺伝子を同定し、HBM遺伝子およびHBM表現型を生成するLRP5遺伝子の位置582でのグアニンからチミンへの多型性を同定する。

20

【0019】

本発明によると、LRP5遺伝子およびHBM遺伝子を同定し、これを用いてヒトがHBMに対する素因を有しているかどうかを決定することができ、従って骨密度低下を特徴とする疾患、例えば骨粗鬆症などになりにくいか、または異常な高骨密度を特徴とする疾患、例えば大理石骨病に対する素因を有し、なりやすいかどうかを決定することができる。HBM遺伝子を有する高齢の個体はHBMタンパク質を発現し、従って骨粗鬆症を進展させない。換言すれば、HBM遺伝子は骨粗鬆症のサプレッサーである。このインビボ観察は、正常な個体のHBM遺伝子もしくはタンパク質、またはそのフラグメントでの処置が骨粗鬆症を改善するという強力な証拠である。

30

【0020】

本発明は、動物モデルにおける骨密度変調の研究に有用であるLRP5およびHBMのための発現ベクターを提供する。発現ベクターは動物組織、とりわけ骨組織において広範にLRP5およびHBMの発現を誘導するプロモーターを含む。発現ベクターはまたトランスジェニックおよびその他の遺伝的に修飾された動物の産生のための直線状核酸配列を提供するのにも役立つ。

40

【0021】

1つの実施形態によると、LRP5を発現しないノックアウトマウスや相同なマウスHBMタンパク質を発現するノックインマウスを産生するために、マウスおよびその他の動物においてジーンターゲティングするためのベクターが提供される。予め融合した接合体においてノックアウトカセットをインビトロで欠失させることができる条件的ノックアウト／ノックインベクターが提供される。本発明はジーンターゲティングベクターの組換えDNAを含む動物胚幹性細胞を提供する。

50

【0022】

別の実施形態では、LRP5およびHBMを発現する動物細胞が、骨密度のモデュレーターの調査に使用するために提供される。

【0023】

さらに別の実施形態によると、LRP5遺伝子およびHBM遺伝子を発現するトランスジェニック動物または一般的なプロモーターおよび骨特異的プロモーターの調節下のその他の関連する変異体が提供される。内因性LRP5遺伝子または異種性LRP5もしくはHBM遺伝子が、誘導可能なまたは条件的なプロモーター、例えばGENESWITCH（登録商標）系の調節下であるトランスジェニック動物もまた提供される。本発明は、HBM表現型およびその分子メカニズムの研究、診断およびスクリーニング手段の開発、ならびに処置および治療用化合物の試験および開発のためのこれらの動物の使用方法を提供

10

【0024】

さらに、このような処置は骨損傷、とりわけ骨折の処置において、このような損傷の治療における骨リモデリングのために指示される。本発明の手段により例えばストレス骨折由来の人が病気にかかりやすいまたは苦痛（すなわち圧により誘導される微小骨折、実際に骨皮質を通る真性骨折に至る）を同定および／または処置できる。さらに、本発明の方法および組成物は、治療の過程が骨リモデリングに関与する2次骨粗鬆症、例えば副腎皮質ステロイド投与を伴う内分泌条件、甲状腺機能亢進症、性腺機能低下症、血液学的悪性腫瘍、吸収不良およびアルコール依存症、およびビタミンDおよび／またはリン酸塩代謝物に伴う障害、例えば骨軟化症およびくる病、および骨リモデリング異常または障害を特徴とする疾患、例えばパジェット病の処置ならびに良性または悪性であり得る骨の新生物に有用である。

20

【0025】

種々の実施形態では、本発明はHBMおよびLRP5を発現する核酸、タンパク質、ベクター、形質転換された宿主、およびヒトHBMおよびLRP5遺伝子および関連する変異体を担持するトランスジェニック動物、HBM相同体に関するノックイン動物またはこれらの遺伝子に関するノックアウト動物を志向する。

【0026】

加えて、本発明は例えば骨発達障害のための遺伝子治療、医薬品開発、および診断アッセイを含む本発明の前記の実施形態の適用を志向する。好ましい実施形態では、本発明は骨粗鬆症の処置、診断、予防およびスクリーニングのための方法を志向する。

30

【0027】

本発明の別の態様によると、配列番号2の位置582でチミンを少なくとも含み、配列番号2の少なくとも15個の連続するヌクレオチドを含む配列に機能的に連結され(operably linked)、動物細胞および／またはヒト細胞においてタンパク質の発現を導く(direct)プロモーター領域を含む核酸を含む体細胞および／または生殖細胞を有するトランスジェニック動物が提供される。

【0028】

意図される別の実施形態は、配列番号4をコードし、配列番号4の位置171のバリンに相当するバリンのコドンを含み、配列を含む核酸であって、動物細胞および／またはヒト細胞においてタンパク質の発現を導く、機能的に連結したプロモーター領域をさらに含む核酸を含む体細胞および／または生殖細胞を有するトランスジェニック動物；配列番号1の少なくとも15個の連続するヌクレオチドを含む配列に機能的に連結された、動物細胞および／またはヒト細胞においてタンパク質の発現を導くプロモーター領域を含む核酸を含む体細胞および／または生殖細胞を有する骨密度変調(bone density modulation)の研究のためのトランスジェニック動物を含む。このような動物の子孫もまた意図される。動物は好ましくはマウスであるが、いずれかの非ヒト動物（例えば霊長類、イヌ、ネコ、げっ歯類、ヒツジ、ウシ等）を含み得る。

40

【0029】

50

このような動物は、骨密度または骨量(bone mass)の変調の研究および骨密度または骨量の変調に影響する方法および処置の開発に有用である。骨密度および／または骨量の変調を、1つまたはそれ以上のパラメーター、例えば骨塩量(bone mineral density)、骨強度、骨梁数(trabecular number)、骨の大きさ、および骨組織結合性(bone tissue connectivity)の変化により評価することができる。

【0030】

本発明の別の目的は、配列番号2の位置582でチミンを少なくとも含み、配列番号2の少なくとも15個の連続するヌクレオチドを含む配列に機能的に連結され、動物細胞および／またはヒト細胞におけるタンパク質の発現を導くプロモーター領域を含む核酸を含む動物の胚を提供することである。

10

【0031】

相同組換えによるジーンターゲティングのための核酸であって、マウスLRP5遺伝子に相同である第1セクションと、マウスLRP5遺伝子の別のセクションに相同である第2セクションと、第1および第2セクション間におけるLRP5遺伝子の一部の操作された欠失、核酸配列の変化または核酸の挿入を含む中央セクションとを含む核酸を提供することも本発明のさらなる目的である。

【0032】

本発明の別の目的は、そのゲノムがLRP5をコードする遺伝子に変化(alteration of the gene)を有するトランスジェニック動物、好ましくはトランスジェニックマウスを産生する方法を提供することである。この方法は、

20

- (a) LRP5またはHBMをコードする核酸を供するステップと；
 - (b) 該核酸をマウス胚幹性細胞に導入するステップと；
 - (c) 該核酸を含む胚性幹細胞を選択するステップと；
 - (d) ステップ(c)の胚性幹細胞をマウス胚盤胞に導入するステップと；
 - (e) ステップ(d)の胚盤胞を偽妊娠マウスに移すステップと；
 - (f) 移した胚盤胞を、該核酸に関してキメラであるマウスに発達させるステップと；
- を含むことができる。

【0033】

本発明の別の実施形態では、記載されたように得られた動物は、次いで、例えば、キメラマウスと野生型マウスと交配させ、前記変化に関してヘテロ接合性であるマウスを得るステップと；該ヘテロ接合性マウスを交配させ、前記変化に関してホモ接合性であるマウスを産生するステップと；により、さらに交配させることができる。

30

【0034】

本発明の別の態様によると、HBM発現を変調する作用物質を同定するための方法であって、(a)請求項13に記載の細胞を供するステップと；(b)該細胞を被験化合物に暴露するステップと；(c)HBMの発現を測定するステップと；を含む方法が提供される。

【0035】

本発明の別の目的は、前記トランスジェニック動物の第1の群を供するステップと；(b)被験化合物を投与するステップと；(c)被験化合物を投与された動物における少なくとも1つの発達のパラメーターを測定するステップとにより骨量モデレーターを研究することである。特に限定されるものではないが、被験化合物は、ホルモン、成長因子、ペプチド、RNA、DNA、ミネラル、ビタミン、天然生成物、または合成有機化合物などを含む。

40

【0036】

別の態様では、骨量変調および骨発達を、前記トランスジェニック動物の1つの群を利用し、動物に実験手順を施し、発達のパラメーターを測定する方法により研究することができる。実験手順には、例えば、卵巣切除、骨負荷の制限(restricted bone loading)および骨負荷の増強(increased bone loading)などがある。

【0037】

50

本発明の別の態様によると、単離された核酸配列（配列番号 689～697）：

(1)

5' -GTCAGCCTGGAGGAGTTCTCA-3' ;

5' -TCACCCTTGGCAATACAGATGT-3' ;および

6-FAM-5' -CCCACCCATGTGCCCCGTGACA-3' ;または

(2)

5' -CGTGATTGCCGACGATCTC-3' ;

5' -TTCCGGCCGCTAGTCTTGT-3' ;および

6-FAM-5' -CGCACCCGTTCTGGTCTGACGCAGTAC-3' ;

を含む、ヒトLRP5 mRNAまたはHBM mRNAを定量するための試薬セットが提供される。 10

【0038】

別の試薬セットは、

5' -CTTCCCCACGAGTATGTTGGT-3' ;および

5' -AAGGGACCGTGCTGTGAGC-3' ;および

6-FAM-5' -AGCCCCTCATGTGCCTCTCAACTTCATAG-3' ;

を含む。

【0039】

本発明の別の態様によると、1つまたはそれ以上の以下のアミノ酸の置換：G171V、A214V、E128V、A65V、G199V、M282、G479V、G781V、Q1087V、G171K、G171F、G171I、G171Qを有する配列番号3の変異体が提供される。 20

【0040】

本発明の別の目的はLRP6の対応する変異体を提供することである。LRP6の好ましい変異体はG158Vを含む。

【0041】

さらに本発明の実施の形態によると、HBMが他の骨障害に及ぼす影響を研究する方法であって、（a）骨障害の表現型を有する動物の胚を供するステップと；（b）LRP5、HBM、その変異体をコードする核酸を該胚の第1の群に導入するステップと；（c）胚を偽妊娠マウスに移すステップと；（d）得られたマウスの少なくとも1つの発達パラメーターを測定するステップと；の工程を含む方法が提供される。核酸は、いずれの動物に由来してもよく、ヒトLRP5またはヒトHBMに限定されない。 30

【0042】

本発明のさらなる態様によると、LRP5またはHBMに関連する心臓障害を研究する方法であって、（a）前記したトランスジェニック動物の第1群を供するステップと；（b）被験化合物を投与されたトランスジェニック動物の少なくとも1つの心臓健康パラメーターを測定するステップと；を含む方法が提供される。別の方法では、推定される心臓病薬から有効なものをスクリーニングするのにこれらの動物を用いることができる。

【0043】

本発明のさらに別の実施形態によると、心臓保護処置の骨量変調効果を評価する方法であって、請求項9に記載の動物の第1群を提供するステップと；心臓保護処置(cardio-protective treatment)を該動物の第1群のサブグループに施すステップと；少なくとも処置されたマウスにおいて少なくとも1つの骨変調パラメーターを測定するステップとを含む方法が提供される。 40

【0044】

本発明の別の態様によると、骨量のモデュレーターの研究方法であって、（a）配列番号1の少なくとも15個の連続するヌクレオチドを含む配列に機能的に連結された、動物細胞および／またはヒト細胞におけるタンパク質の発現を導くプロモーター領域を含む核酸を含む体細胞および／または生殖細胞を有する動物の第1群を供するステップと；（b）被験化合物を投与するステップと；（c）被験化合物を投与された動物における少なく 50

とも1つの骨発達パラメーターを測定するステップとを含む方法が提供される。

【0045】

別の態様によると、骨量に及ぼす実験手順の影響を研究する方法であって、(a)配列番号1の少なくとも15個の連続するヌクレオチドを含む配列に機能的に連結された、動物細胞および／またはヒト細胞におけるタンパク質の発現を導くプロモーター領域を含む核酸を含む体細胞および／または生殖細胞を有する動物の第1群を供するステップと；(b)実験手順を施すステップと；(c)少なくとも1つの骨量のパラメーターを測定し、実験手順を施された動物における骨量を評価するステップとを含む方法が提供される。

【0046】

本発明のこれらのおよびその他の態様を以下にさらに詳細に記載する。

10

【発明を実施するための最良の形態】

【0047】

本明細書および請求の範囲の理解を助けるために、以下の定義を提供する。

【0048】

「遺伝子」は、その鋳型またはメッセンジャーRNAを介して特定のペプチドに特徴的なアミノ酸配列をコードするDNA配列を意味する。「遺伝子」なる用語は介在性、非コード化領域および制御領域を含み、5'および3'末端を含むことができる。

【0049】

「核酸」は、一本鎖および二本鎖核酸、DNA、RNA（例えばmRNA、tRNA）、cDNA、組換えDNA（rDNA）、rRNA、アンチセンス核酸、オリゴヌクレオチド、およびオリゴマー、ならびにポリヌクレオチドを含むことを意味する。ハイブリッド、例えばRNAおよび／またはDNAの二本鎖領域または二本鎖RNA：DNAハイブリッドを含んでもよい。修飾塩基、例えばビオチン化、トリチル化、フルオロフォア、イノシン等を含んでもよい。

20

【0050】

「遺伝子配列」は、遺伝子を含む、非転写または非翻訳配列を含有するDNAを含む核酸分子を意味する。該用語はまた同一DNA分子上に存在する（複数の）遺伝子、（複数の）遺伝子フラグメント、（複数の）非転写配列または（複数の）非翻訳配列のいずれかの組み合わせを含むことをも意図する。

【0051】

本発明の核酸配列をDNA、cDNA、合成DNA、合成RNAまたはその組み合わせを含む種々の供給源から誘導することができる。このような配列は天然発生イントロンを含んでも含まなくてもよいゲノムDNAを含むことができる。さらにこのようなゲノムDNAをプロモーター領域および／またはポリ(A)配列に随伴されて得ることができる。配列、ゲノムDNAまたはcDNAをいずれかのいくつかの方法で得ることができる。当分野で周知の手段によりゲノムDNAを適当な細胞から抽出し、精製することができる。また別に、mRNAを細胞から単離し、これを用いて逆転写またはその他の手段によりcDNAを生成することができる。

30

【0052】

「cDNA」は、RNA依存性DNAポリメラーゼ（逆転写酵素）の作用によりRNA鋳型から生成された相補性またはコピーDNAを意味する。従って、「cDNAクローン」は1つの鎖が、クローニングベクターまたはPCR増幅で運び込まれた目的のRNA分子に相補的である二本鎖DNA配列を意味する。cDNAはまた逆転写酵素による一本鎖合成の後の一本鎖でもよい。この形態では、これはPCR鋳型に有用であり、クローニングベクターに運び込まれる必要はない。この用語は介在配列が除去されている遺伝子を含む。従って「遺伝子」なる用語はしばしば一般的に用いられるように、cDNAおよびcDNAクローンを含む核酸分子をも含む。

40

【0053】

「組換えDNA」は、インビトロでcDNAまたはゲノムDNA配列をスプライシングするか、または方法例えばPCR変異誘発により配列を変化させることにより操作されて

50

いる分子を意味する。

【0054】

「クローニング」は特定の遺伝子またはその他のDNA配列をベクター分子に挿入するためのインビトロ組換え技術の使用を意味する。望ましい遺伝子のクローニングに成功するために、DNAフラグメントを作製するための方法、フラグメントをベクター分子に結合させるための方法、それが複製できる宿主に複合性のDNA分子を導入するための方法、および受容者宿主細胞の中から標的遺伝子を有するクローンを選択するための方法を用いることが必要である。

【0055】

「cDNAライブラリー」は、特定の組織または細胞供給源で発現される遺伝子の全体的または部分的レパートリーを含むcDNAインサートを含有する組換えDNA分子を意味する。当業者に既知の、ならびに、例えばCowellおよびAustin、「cDNAライブラリープロトコル」Methods in Molecular Biology(1997)に記載される方法によりこのようなcDNAライブラリーを調製することができる。

【0056】

「クローニングベヒクル」は、宿主細胞内で複製できるプラスミドまたはファージDNAまたはその他のDNA配列を意味する。この用語は人工染色体、例えばBACおよびYACを含むこともできる。クローニングベヒクルは、このようなDNA配列がDNAの本質的な生物学的機能を喪失せずに確定できる様式で切断され得る1つまたはそれ以上のエンドヌクレアーゼ認識部位を特徴とし、これは形質転換細胞の同定において使用するのに適したマーカーを含有できる。

【0057】

「発現」は、遺伝子の最終産物を生成するための、遺伝子配列の転写およびそれに続くプロセッシングステップ、例えば得られたmRNAの翻訳を含む過程を意味する。最終産物はタンパク質（例えば酵素またはレセプター）または核酸（例えばtRNA、アンチセンスRNA、またはその他の制御配列）でよい。「発現調節配列(expression control sequence)」なる用語は、構造遺伝子に機能的に連結された場合、その遺伝子の発現を調節または制御するヌクレオチド配列を意味する。これには例えばlacシステム、trpシステム、λファージの主要オペレーターおよびプロモーター領域、fdコートタンパク質の調節領域および原核または真核細胞で遺伝子の発現を調節することが解っているその他の配列などがある。発現調節配列は、ベクターが原核または真核細胞宿主において機能的に連結された遺伝子を発現するように設計されているかどうかに依存して変化し、転写エレメント、例えばエンハンサーエレメント、終止配列、組織特異性エレメントならびに／または翻訳開始および終止部位を含有することができる。

【0058】

「発現ベヒクル」は、クローニングベヒクルに類似するベヒクルまたはベクターを意味するが、これは宿主への形質転換の後、クローン化されている遺伝子を発現することができる。クローン化遺伝子は通常発現調節配列の調節下にある（すなわちそれに機能的に連結されている）。

【0059】

「オペレーター」は、特異的レプレッサーと相互作用できるDNA配列を意味し、それにより（複数の）隣接する遺伝子の転写を調節する。

【0060】

「プロモーター」は、RNAポリメラーゼにより認識され得るDNA配列を意味する。このような配列の存在によりRNAポリメラーゼを結合させ、機能的に連結された遺伝子配列の転写を開始することができる。

【0061】

「プロモーター領域」は、転写の開始に必要であるプロモーターおよびその他の遺伝子配列を含むことを意図する。プロモーター領域の存在は機能的に連結された遺伝子配列の発現を引き起こすのに十分である。「プロモーター」なる用語はしばしば当分野で一般的

10

20

30

40

50

にプロモーター領域を示すために用いられる。特定の細胞型で遺伝子の発現を導く多くの異なるプロモーターが当分野で公知である。組織特異的プロモーターは特定の組織型の細胞でより高レベル（または低レベル）の発現を引き起こす核酸配列を含むことができる。

【0062】

「機能的に連結された」は、プロモーターが遺伝子の発現の開始を調節することを意味する。宿主細胞への導入時にプロモーターが（複数の）近接したDNA配列の1つまたはそれ以上の種のRNAへの転写を決定する場合、プロモーターは近接したDNAの配列に機能的に連結されている。プロモーターがそのDNA配列の転写を開始できる場合、プロモーターはDNA配列に機能的に連結されている。

【0063】

10

「原核生物」は、細菌などの真核を有さない全ての生物を意味する。

【0064】

「真核生物」は、哺乳動物細胞などの真核を有する生物および細胞を意味する。

【0065】

「宿主」には原核生物および真核生物、例えば酵母および糸状菌類、ならびに植物および動物細胞などがある。該用語は複製可能な発現ベヒクルの受容者である生物または細胞を含む。

【0066】

本明細書で用いられる「動物」なる用語は、ヒト以外の全ての脊椎動物を含む。これはまた胚および胎仔段階を含む全ての発達段階の個々の動物を含む。

20

【0067】

「トランスジェニック動物」は、直接的または間接的に、慎重な遺伝子操作によるかまたは細胞下レベルで操作された先祖からの遺伝により、例えば組換えウイルスベクター（例えばアデノウイルス、レトロウイルス、ヘルペスウイルス、アデノ随伴ウイルス、レンチウイルス）でのマイクロインジェクションまたは感染により受け取った遺伝情報を担持する1つまたはそれ以上の細胞を含有する動物である。この導入されたDNA分子を染色体に組み込むことができるか、またはこれは染色体外複製DNAでよい。

【0068】

本明細書で用いる「胚性幹細胞」または「ES細胞」は、未分化細胞であることを意味する、多能性の通常胚由来の細胞または細胞株である。これらの細胞はまた、宿主胚に組み込まれた場合、相同組換えおよびそれに続く体内のいずれかの組織への発達により外来性DNAを組み込むことが可能である。当分野でよく理解されている方法により胚組織以外の供給源から多能性細胞を単離することが可能である。

30

【0069】

マウスの胚性幹細胞により、トランスジェニック細胞を選択し、ジーンターゲティングを実施するための研究が可能になった。別のトランスジェニック動物トランスジェニック技術で可能であるよりもさらに多くの遺伝子操作がこれにより可能になる。例えば、マウスES細胞をインビトロでコロニーとして成長させるのが比較的容易である。標準的な手順により細胞をトランスフェクトでき、抗生物質抵抗性によりトランスジェニック細胞をクローンで選択することができる。例えばDoetschmanら、胚性幹細胞での遺伝子移入 (gene transfer in embryonic stem cells)、Pinkert編、Transgenic Animal Technology: A Laboratory Handbook (1994)、Academic Press Inc.、ニューヨーク115-146頁参照。さらに、この方法の効率は、十分なトランスジェニックコロニー（数百から数千）を生成して相同組換えのための第2の選択を可能にできるようなものである。次いでマウスES細胞を、生殖細胞を含む正常な宿主胚と組み合わせて、その強度を保持するために、得られたキメラ動物において全て組織まで発達させることができる。次いでトランスジェニック修飾を次世代まで伝えることができる。

40

【0070】

初期移植前マウス胚から胚性幹（ES）細胞株をインビトロで誘導するための方法は周知である。例えばEvansら、Nature 29:154-156(1981)およびMartin、Proc.Nat.Acad.Sci.

50

USA 78:7634-7638(1981)参照。繊維芽細胞の支持細胞層または分化阻止供給源が存在すれば、未分化状態でES細胞を継代することができる。

【0071】

「体細胞」なる用語は、精子もしくは卵細胞ではないか、または精子もしくは卵細胞になることができるいずれかの動物またはヒト細胞を示す。「生殖細胞」または「生殖系細胞」なる用語は、精子もしくは卵細胞のいずれかであるか、または精子もしくは卵細胞に発達することができ、従って子孫にその遺伝情報を渡すことができるいずれかの細胞を意味する。「生殖細胞株トランスジェニック動物」なる用語は、遺伝情報が生殖細胞株に組み込まれ、それにより遺伝情報を子孫に移す能力を付与するトランスジェニック動物を意味する。このような子孫が実際にいくつかまたは全てのその情報を有している場合、次いでそれもまたトランスジェニック動物である。

10

【0072】

遺伝情報の遺伝子変化は、受容者が属する動物の種に対して外来性であるか、または特定の個々の受容者に対してのみ外来性であり得る。後者の場合、変化したまたは導入された遺伝子は、内在性の遺伝子とは異なって発現されるかもしれない。

【0073】

遺伝子の「フラグメント」は、遺伝子配列のいずれかの部分を意味する。生物学的に活性なフラグメントとは、遺伝子の少なくとも1つの生物学的活性を保持する遺伝子のいずれかの部分を意味する。

【0074】

「生物学的に活性」とは、保存的に置換された変異体、タンパク質をコードする遺伝子の対立遺伝子、またはタンパク質のポリペプチドフラグメントを含むタンパク質およびポリペプチドの形態であって、野生型タンパク質またはポリペプチドの生物学的および／または免疫学的活性を保持するものを意味する。好ましくは、活性は、骨量の発達または表現型における変化を誘起する活性である。生物学的に活性とはまたLRP5 (Zmax1)、LPR6、およびHBM、例えばWnt経路に関連するシグナル発生経路を直接的または間接的であろうとなかろうと、およびインビボまたはインビトロアッセイであろうとなかろうと、変調する能力を意味する。

20

【0075】

「有効量」または「投与有効量」または「治療上有効量」は、本発明のポリペプチドの生物学的活性を変調する作用物質の量を意味する。

30

【0076】

「変異体」は、構造および生物学的活性または免疫学的特性が遺伝子全体または遺伝子のフラグメントのいずれかに実質的に類似している遺伝子を意味する。2つの遺伝子が類似の活性を有している場合、たとえコードされるアミノ酸残基の配列が同一でなくても、これらは本明細書でその用語が用いられるような変異体と考えられる。

【0077】

「核酸の増幅」は、例えばポリメラーゼ連鎖反応(PCR)、ライゲーション増幅(またはリガーゼ連鎖反応、LCR)およびQ-βレプリカーゼの使用に基づく増幅方法のような方法を意味する。これらの方法は当分野で周知であり、例えば米国特許第4,683,195号および第4,683,202号に記載されている。PCRを実施するための試薬およびハードウェアは市販により入手可能である。HBMから配列を増幅するのに有用なプライマーは、好ましくはHBM領域またはその標的領域の側面に位置する領域の配列に相補的であり、それに特異的にハイブリダイズする。増幅により作製されたHBM配列を直接シーケンシングすることができる。また別に、増幅された(複数の)配列を配列分析の前にクローン化することができる。

40

【0078】

「抗体」は、非限定例としてはポリクローナル、モノクローナル、キメラ、ヒト、ヒト化、二重特異性、多重特異性、primatized(商標)抗体を含むことを意味する。「抗体」なる用語は、好ましくはHBMタンパク質およびそのフラグメントに、またはHBM領域

50

、とりわけHBM遺伝子座またはその部分に由来する核酸配列に結合できるポリクローナルおよび／またはモノクローナル抗体およびそのフラグメント、ならびにその免疫学的結合等価物を意味する。抗体なる用語は共に同種の分子体、または混合物、例えば複数の異なる分子体から構成される血清生成物を意味するために用いられる。タンパク質をタンパク質合成器で合成により調製し、キャリア分子に結合させ、数ヶ月にわたってウサギに注射することができる。HBMタンパク質またはフラグメントに対する免疫反応性に関してウサギ血清を試験する。マウスにタンパク質またはそのフラグメントを注射することによりモノクローナル抗体を作ることができる。モノクローナル抗体をELISAによりスクリーニングし、EBMタンパク質またはそのフラグメントで特異的免疫反応性に関して試験する。Harlowら、Antibodies:A Laboratory Manual、Cold Spring Harbor Laboratory 10、Cold Spring Harbor、ニューヨーク州(1988)。これらの抗体はアッセイおよび医薬品において有用である。

【0079】

L R P 5 遺伝子およびそれがコードするL R P 5 タンパク質は以前から発明者によりZmax1 遺伝子およびZmax1 と称されている。遺伝子およびその生成物はまた当分野ではL R 3 とも称されている。Zmax、Zmax1、L R P 5 およびL R 3 は同義語である。「HBMタンパク質」とは、それがグリシン171のバリンへの変化を有する以外はL R P 5 と同一であるタンパク質を意味している。HBMタンパク質はL R P 5 真性相同体をコードするいずれかの生物に関して定義されている。例えば、マウスHBMタンパク質とはグリシン171のバリンへの置換体を有しているマウスL R P 5 を意味する。 20

【0080】

本発明の1つの実施形態では、「HBM遺伝子」とはHBM特性または表現型を示す個体において見出されるゲノムDNA配列を意味し、その配列は配列番号4で示されるタンパク質をコードする。HBM遺伝子およびL R P 5 遺伝子是对立遺伝子である。HBM遺伝子によりコードされるタンパク質は骨量上昇を引き起こす特性を有しているが、一方L R P 5 遺伝子によりコードされるタンパク質は有していない。HBM遺伝子およびL R P 5 遺伝子は、HBM遺伝子は位置582でチミンを有しているが、一方L R P 5 遺伝子は位置582でグアニンを有しているという点で異なっている。HBM遺伝子は配列番号2に示す核酸配列を含む。HBM遺伝子を「HBM多型性」とも称することができる。

【0081】

本発明のまた別の実施形態では、「HBM遺伝子」はまた、HBM表現型に至るL R P 5 (Zmax1) またはL R P 6 のいずれかの対立遺伝子変異体をも意味することができる。このような変異体は、本明細書で記載する野生型タンパク質コード化配列からの変化、および／またはL R P 5 の発現調節配列における変化を含んでよい。このような変異体の好ましい実例はL R P 5 タンパク質の発現増加に至る内因性L R P 5 プロモーター領域の変化である。 30

【0082】

「正常」、「野生型」、「影響されない」および「L R P 5」の全ては、配列番号3に示すタンパク質をコードするゲノムDNA配列を意味する。L R P 5 遺伝子は位置582でグアニンを有する。L R P 5 (Zmax1) 遺伝子は配列番号1に示す核酸配列を含む 40。「正常」、「野生型」、「影響されない」および「L R P 5」はまた骨量増加に寄与しないタンパク質をコードするゲノム配列の対立遺伝子変異体をも意味する。L R P 5 遺伝子はヒト集団において一般的であり、一方HBM遺伝子は希少である。

【0083】

「5 Y W T D + E G F」は、5個のY W T D 反復とそれに続くE G F 反復からなる、L R P 5 タンパク質で見出される反復単位を意味する。

【0084】

「骨発達」は、一般に、例えば正常な発達、疾病状態の経過中に生じる変化、および加齢と共に生じる変化などの長期にわたる骨の変化に関わるいずれかの過程を意味する。これは構造変化および動的な速度変化、例えば成長速度、吸収速度、骨修復速度等を意味し得 50

る。「骨発達障害」は、とりわけ、例えば疾病状態の経過中に生じる変化、および加齢と共に生じる変化などの骨発達におけるいずれかの障害を意味する。骨発達は本質的に進行性かまたは周期性でよい。骨発達中に変化し得る骨の態様には、例えばミネラル化、特定の解剖学的特長の形成、および種々細胞型の相対数または絶対数などがある。

【0085】

「骨変調」または「骨形成の変調」は、当業者に理解されるように、例えばとりわけ破骨および造骨活性による骨再吸収および骨付加成長などの骨リモデリングに関与するいずれかの生理学的過程に影響する能力を意味し、本明細書で用いるいくつかまたは全ての骨形成および発達を含み得る。

【0086】

骨は、破骨細胞により古いまたは不必要な骨を除去し、骨芽細胞により新たな骨を再構築することにより継続的にそれ自体を改造および更新している動的組織である。これらの過程間の共役の本質は、骨成長中のモデリングならびにリモデリングおよび修復による成人の骨格完全性の維持の双方に寄与して機械的使用の毎日の必要性に合致する。骨再吸収および形成の間の均衡の非共役による引き起こされる多くの骨疾患がある。加齢と共に骨ターンオーバーにおける漸進的な「生理学的」不均衡があり、これはとりわけ閉経によるエストロゲン補助の喪失のために女性において悪化し、進行性の骨喪失に至る。骨量の低下および骨構築の劣化は骨脆弱性および特発性骨折の増加に至る。喪失された骨の10%ごとに骨折の危険性は倍化する。脊椎または近接する大腿骨において、正常な骨量ピーク以下の骨ミネラル密度(BMD) 2.5またはそれ以上の標準偏差を有する個体を骨粗鬆症と分類する。しかしながら、正常以下の、1と2.5の間の標準偏差のBMDを有するオステオペニア個体は明らかに骨喪失関連障害を患う危険性がある。

【0087】

パラメーターの測定、例えばpDXA-X線法による骨ミネラル密度(BMD)および骨ミネラル含量(BMC)、X線による骨の大きさ、厚さまたは体積、例えばカルシエン標識により測定される骨形成速度pQCTおよび/またはCT法により測定される全体、骨梁および中央シャフト密度、 μ CT法により測定される結合性およびその他の組織学的パラメーター、好ましくは各々大腿骨および脊椎で測定される機械的屈曲および圧縮力により骨変調を評価することができる。当業者に理解されるように、これらの測定値の性質により、各々は所定の状況に関して多かれ少なかれ適当であり得る。さらに、骨折の有無の臨床履歴、骨形状、骨形態学、結合性、正常な履歴、骨折修復速度、およびその他の骨品質のパラメーターなどのパラメーターおよび方法論が当分野で公知であり、使用されている。最も好ましくは、このような測定が適当である場合、骨品質を脊椎の圧縮力により評価できる。骨変調を種々パラメーターの変化速度により評価することもできる。最も好ましくは、骨変調を1つ以上の年齢で評価する。

【0088】

「正常な骨密度」は、HBM関連研究の文脈では、2つの標準偏差内でZスコア0の骨密度を意味する。一般的な文脈では、正常な骨密度パラメーターの範囲は通常統計学的方法により決定される。正常なパラメーターは年齢および性別で正規化されたパラメーターの標準偏差約1または2内にあり、好ましくは標準偏差約2である。有意な統計的測定は関連する測定値が平均と有意に異なる可能性を提示できるP値である。有意なP値は $P < 0.05$ 、 0.01 、 0.005 および 0.001 であり、好ましくは少なくとも $P < 0.01$ である。

【0089】

「HBM」は、この用語が骨密度、ミネラル含量および大きさに関しても表現し得るが、高骨量を意味する。

【0090】

「HBM表現型」は、骨変調の1、2、3、4、5またはそれ以上の定量パラメーター、好ましくは骨密度およびミネラル含量および骨力パラメーターにおいて、そのパラメーターの年齢および性別の標準を超えて標準偏差約2またはそれ以上、好ましくは2、2.

10

20

30

40

50

5、3またはそれ以上の増加を特徴とし得る。HBM表現型は、少なくとも1つのパラメーター、好ましくは少なくとも2つのパラメーター、および最も好ましくは少なくとも3つまたはそれ以上のパラメーターにおける統計学的に有意な増加を特徴とする。HBM表現型はまた、1つまたはそれ以上の骨品質パラメーターにおける増加をも特徴とし、最も好ましくは増加しているパラメーターがいずれかの骨品質パラメーターの低下により随伴される。最も好ましくは、骨変調パラメーターおよび／または骨品質測定値における増加が1つ以上の年齢で観察される。

【0091】

「LRP5系」は、精製されたタンパク質、細胞抽出物、細胞、動物、ヒトまたは、LRP5が正常または変異形態で提示されている物のいずれかその他の組成物を意味する。 10

【0092】

「単離された」なる用語は、ヒトの手により天然の環境から変化されている物質を意味する。単離されたペプチドは例えば実質的に純粋な形態であるか、またはそうでなければ例えば単離された細胞株またはトランスジェニック動物における発現により元来の環境から移動されている。単離された配列は、当該単離された配列の少なくとも1つの末端が天然の状態において隣接する配列と隣接しないように、例えば実質的に純粋な形態の分子である、または、その元来の環境から移動されている。

【0093】

「代理マーカー」は、HBM遺伝子または骨量増加とまたはその双方に関連しているが、骨密度よりも測定が容易である細胞、組織、ヒトまたは動物において観察できる診断用指標、症状、徴候またはその他の特徴を意味する。一般的な代理マーカーの概念は診断医学において十分承認されている。 20

【0094】

本発明は各々配列番号1および3により同定された形態のLRP5遺伝子およびLRP5タンパク質、およびその他の密接に関連する変異体、ならびにその正確な発現に必要なLRP5の隣接する染色体領域を包含する。好ましい実施形態では、本発明は配列番号1の核酸配列の少なくとも15個の連続するヌクレオチドを志向する。

【0095】

I. 序論

本発明はまた各々配列番号2および4により示される形態のHBM遺伝子およびHBMタンパク質、およびその他の密接に関連する変異体ならびにその正確な発現に必要なHBM遺伝子の隣接する染色体領域をも包含する。好ましい実施形態では、本発明は配列番号2の核酸配列の少なくとも15個の連続するヌクレオチドを志向する。より好ましくは、本発明は配列番号2の核酸配列の少なくとも15個の連続するヌクレオチドであって、その15個の連続するヌクレオチドの1つはヌクレオチド582のチミンであるヌクレオチドを志向する。 30

【0096】

本発明はまたLRP5遺伝子領域のヌクレオチド配列およびHBM遺伝子領域のヌクレオチド配列にも関連する。さらにとりわけ、好ましい実施形態はLRP5遺伝子領域B200E21HおよびB527D2-Hのセグメントを含有するBACクローンである。好ましい実施形態は配列番号5～12からなるBACクローンのヌクレオチド配列である。 40

【0097】

本発明はまたLRP5遺伝子およびHBM遺伝子のためのDNAプローブ、LRP5遺伝子およびHBM遺伝子を増幅するためのPCRプライマー、LRP5遺伝子およびHBM遺伝子におけるヌクレオチド多型性、およびLRP5遺伝子およびHBM遺伝子の制御エレメントを同定するためのヌクレオチド配列の使用にも関する。

【0098】

本発明はさらに染色体11q13.3上の遺伝子マーカーD11S987とSNP CONTIG033-6の間のLRP5遺伝子およびHBM遺伝子の染色体位置の位置決定、ならびにLRP5遺伝子およびHBM遺伝子のDNA配列について記載する。染色体位 50

置は遺伝子のマッピングに用いられるマッピングパネルにさらなる遺伝子マーカーを付加することにより、およびさらなる個体を含むように家系を拡張することにより染色体位置を絞り込んだ。遺伝子型を同定された新たな個体が領域を狭める重大な組換え事象を宿すので、家系の拡張は非常に重要である。11q13.3の領域で遺伝子を同定するために、この染色体領域を含有するBACクロンのセットを同定した。ゲノムDNAシーケンシングのための鋳型として、および直接cDNA選択によりコード化配列を同定するための試薬としてもBACクロンを供した。ゲノムシーケンシングおよび直接cDNA選択を用いて11q13.3から1500000個のDNAの塩基対特徴付けをした。この領域内でLRP5遺伝子を同定し、次いで罹患および非罹患個体の変異分析の後、HBM遺伝子を見出した。

10

【0099】

遺伝子が特定の染色体領域に遺伝子的に位置決定されている場合、重複クロンのセットにおけるDNAの領域全体のクローニング（物理的マッピング）、直接cDNA選択の組み合わせによるこれらのクロンによりコードされる遺伝子の特徴付け、エクソントラッピングおよびDNAシーケンシング（遺伝子同定）、ならびにHBM家系の罹患および非罹患メンバーの比較DNAシーケンシングによるこれらの遺伝子における変異の同定（変異分析）を含む一連のステップにより分子レベルでこの領域の遺伝子の特徴付けすることができる。

【0100】

物理的マッピングは、目的の染色体領域における独特の分子標識を増幅するように設計されたPCRアッセイを用いて大腸菌またはサッカロマイセス・セレビシエにおいて増殖したベクターにクローン化されたヒトDNAのライブラリーをスクリーニングすることにより達成される。HBM候補領域の物理的マッピングを作製するために、バクテリア人工染色体（BAC: Bacterial Artificial Chromosome）にクローン化されたヒトDNAのライブラリーを、ヒトゲノムプロジェクトの努力により予め染色体11q12-q13にマッピングされた配列標識部位（STS: Sequence Tagged Site）マーカーのセットでスクリーニングした。

20

【0101】

STSはPCRにより検定することができるヒトゲノムにおける独特な分子標識である。ヒトゲノムプロジェクトの努力を合わせることで、22個の常染色体および2個の性染色体における数千のSTSの位置が決定された。位置クローニングの努力のために、遺伝子マッピングに用いられたマーカーを物理的マッピングのためのSTSとしても使用できるので、物理的地図が遺伝子地図に結びつけられる。BACライブラリーを遺伝子マーカー、遺伝子、およびランダムDNAフラグメントから誘導されるSTSの組み合わせでスクリーニングすることにより、目的の染色体領域における全てのDNAを表す重複クロンからなる物理的地図を組み立てることができる。

30

【0102】

BACは大腸菌で増殖させたヒトまたはその他のDNAの大きな（80キロ塩基から200キロ塩基）セグメントのためのクローニングベクターである。BACを用いて物理地図を構築するために、BACクロンのライブラリーをスクリーニングして、所定のSTSまたはSTSのセットに相当するDNA配列を宿す個々のクロンを同定する。たいていのヒトゲノムの至るところでSTSマーカーがおよそ20から50キロ塩基離れて配置されており、個々のBACクロンは典型的に少なくとも2つのSTSマーカーを含有するようになる。加えて、スクリーニングされたBACライブラリーはヒトゲノムを6回被覆するのに十分なクローン化DNAを含有する。従って、個々のSTSは典型的には1個以上のBACクロンを同定する。およそ50キロ塩基離れて位置する一連のSTSマーカーで6倍被覆するBACライブラリーをスクリーニングすることにより、一連の重複BACクロンからなる物理地図、すなわちBACコンティグをヒトゲノムのいずれかの領域に関して組み立てることができる。物理地図を調製するために用いられた多くのSTSマーカーもまた遺伝子マーカーであるので、この地図は遺伝子地図に密接に結びつけられ

40

50

ている。

【0103】

物理地図を構築する場合、ゲノムのSTS地図に、所定の位置で重複するBACクローンを同定することが不可能になるギャップが存在することがしばしば生じる。典型的には、公に利用可能な文献およびワールド・ワイド・ウェブ・リソースにより同定されたSTSのセットから物理地図を最初に構築する。最初の地図は未知の分子距離のギャップにより分けられたいくつかの別個のBACコンティグから成る。これらのギャップを埋めるBACクローンを同定するために、ギャップのいずれかの側のクローンの末端から新規のSTSマーカを開発する必要がある。これはギャップをフランキングする200から300の末端塩基対をシーケンシングし、100またはそれ以上の塩基対の配列を増幅するPCRアッセイを開発することにより行われる。末端配列がヒトゲノム内で独特であることが実証される場合、次いで新規のSTSを用いてBASライブラリーをスクリーニングし、物理地図のギャップからのDNAを含有するさらなるBACを同定することができる。HBM候補領域の大きさ(2, 000, 000またはそれ以上の塩基対)の領域を被覆するBACコンティグを組み立てるために、しばしばいくつかのクローンの末端から新規のSTSマーカを開発することが必要である。

【0104】

BACコンティグを構築した後、重複クローンのこのセットを、染色体領域をコードする遺伝子を同定するための鋳型として提供する。多くの方法により遺伝子同定を達成することができる。一般的に3つの方法が用いられている。(1)染色体領域全体を表すBACコンティグから選択されたBACのセットをシーケンシングでき、コンピューターによる方法を用いて全ての遺伝子を同定することができる、(2)BACコンティグからのBACを試薬として用いて、直接cDNA選択と称される方法により、領域においてコードされる遺伝子に相当するcDNAクローニングすることができる、または(3)エクソントラッピングと称される手順において特異的DNA配列モチーフを選択することにより、BACコンティグからのBACを用いてコード化配列を同定することができる。本発明は最初の2つの方法により同定された遺伝子を含む。

【0105】

HBM候補領域を表す全BACコンティグをシーケンシングするために、プラスミドベクターへのサブクローニングおよびそれに続くこれらのサブクローンのDNAシーケンシングのためのBACのセットを選択した。BACにクローニングされたDNAはゲノムDNAを表すので、このシーケンシングをcDNAシーケンシングと区別するためにゲノムシーケンシングと称する。目的の染色体領域のゲノムシーケンシングを開始するために、いくつかの非重複BACクローンを選択した。各々のBACクローンに関するDNAを調製し、クローンを無作為の小型のフラグメントに切断し、これを経て標準的なプラスミドベクター、例えばpUC18にクローニングする。次いでプラスミドクローンを成長させて小型のフラグメントを増殖させ、これはシーケンシングの鋳型になる。BAC DNA配列のための十分な被覆および配列の品質を確実にするために、十分なプラスミドクローンをシーケンシングしてBACクローンの6倍の被覆を生じる。例えば、BACが100キロ塩基対である場合、次いでファゲミドをシーケンシングして600キロ塩基の配列を生じる。ファゲミドベクターでのクローニングの前にBAC DNAを無作為に切断するので、600キロ塩基対の未処置のDNA配列をコンピューターによる方法により配列コンティグと称する重複DNA配列に組み立てることができる。コンピューターによる方法で最初の遺伝子を同定する目的のために、各々のBAC 6倍の被覆が1000塩基対から20000塩基対の10個から20個の配列コンティグを生じるのに十分である。

【0106】

本発明で用いるシーケンシング計画は、最初にHBM候補領域のBACコンティグから「種」BACをシーケンシングすることである。次いで「種」BACの配列を用いてコンティグから最低限重複するBACを同定し、続いてこれらをシーケンシングした。

この方法で、各 B A C に残るいくつかの小型の配列ギャップを伴う候補領域全体をシーケンシングした。この配列をコンピューターによる遺伝子同定のための鋳型として提供した。コンピューターによる遺伝子同定の 1 つの方法は B A C の配列を公に利用可能な c D N A のデータベースおよびゲノム配列、例えばユニジーン、d b E S T、G e n b a n k と比較することである。これらの比較は典型的にはコンピューターアルゴリズムおよびプログラムの B L A S T ファミリーを用いて行う (Altschul ら、J.Mol.Biol.215:403-410(1990))。B A C 配列をタンパク質配列に翻訳し、そしてタンパク質配列を用いて、タンパク質配列を分析するために設計された B L A S T のバージョンを用いて公に利用可能なタンパク質データベースを検索することができる (Altschul ら、Nucl.Acids.Res.25:3389-3402 (1997))。別の方法はコンピューターアルゴリズム、例えば M Z E F (Zhang、Proc.Natl.Acad.Sci.94:565-568(1997)) および G R A I L (Uberbacher ら、Methods Enzymol.266:259-281(1996)) を用いることであり、これは全てのエクソンに一般的である特異的 D N A 配列モチーフの存在、およびヒトタンパク質コード化配列に典型的なコドン利用の存在に基づいて配列におけるエクソンの位置を推定する。

【0107】

コンピューターによる方法での遺伝子の同定に加えて、直接 c D N A 選択 (Del Mastro ら、Genome Res.5(2):185-194(1995)) によっても遺伝子を同定した。直接 c D N A 選択では、目的の組織からの c D N A プールを調製し、候補領域からの B A C を液体ハイブリダイゼーションアッセイにおいて使用して、B A C のコード化領域に塩基対形成する c D N A を捕捉する。本明細書にて記載した方法では、ポリ A R N A から c D N A の第 1 の鎖をランダムプライミングし、標準的な方法により c D N A の第 2 の鎖を合成し、c D N A フラグメントの末端にリンカーを加えることにより、いくつかの異なる組織から c D N A プールを作製した。リンカーを用いて c D N A プールを増幅する。B A C クローンをインビトロ D N A 合成のための鋳型として用いて B A C D N A のビオチン標識コピーを作製する。次いで B A C D N A のビオチン標識コピーを変性し、これもまた変性されている、P C R 増幅し、リンカーを付けた過剰の c D N A プールと共にインキュベートする。B A C D N A および c D N A を溶液中にアニーリングさせ、B A C と c D N A の間のヘテロ二本鎖を、ストレプトアビジンコーティング磁性ビーズを用いて単離する。次いで B A C により捕捉された c D N A をリンカー配列に相補的なプライマーを用いて増幅し、ハイブリダイゼーション/選択方法を第 2 ラウンドで繰り返す。直接 c D N A 選択の 2 回目のラウンドの後、c D N A フラグメントをクローニングし、これらの直接選択されたフラグメントのライブラリーを作製する。

【0108】

直接選択により単離された c D N A クローンを 2 つの方法により分析する。H B M 候補領域からの B A C のプールはゲノム D N A 配列を提供するために用いられるので、c D N A は個々の B A C にマッピングされなければならない。これはマイクロタイター皿で B A C を整列させ、その D N A を高密度グリッドで複製することにより達成する。次いで個々の c D N A クローンをグリッドにハイブリダイズして、直接選択に用いたセットからの個々の B A C に対する配列同一性を有することを確認し、その B A C の特異的同一性を決定する。個々の B A C に対応することが確認された c D N A クローンをシーケンシングする。直接選択により単離された c D N A クローンが以前に同定された遺伝子と配列同一性または類似性を共有するかどうかを決定するために、D N A およびタンパク質コード化配列をプログラムの B L A S T ファミリーを用いて公に利用可能なデータベースと比較する。

【0109】

B A C シーケンシングおよび直接 c D N A 選択により提供されるゲノム D N A 配列および c D N A 配列の組み合わせにより、その領域における推定遺伝子の最初のリストが得られる。その領域の遺伝子は全て H B M 遺伝子座に関する候補であった。各遺伝子をさらに特徴付けるために、ノーザンブロットを実施して各遺伝子に対応する転写物の大きさを決定し、推定エクソンと一緒に転写されて個々の遺伝子を作製することを決定した。各遺

伝子のノーザンブロット分析のために、直接選択された cDNA クローンから、またはゲノム DNA からもしくは目的の推定遺伝子をコードする BAC から特異的フラグメントを PCR 増幅することによりプローブを調製した。ノーザンブロットにより転写物の大きさおよびそれが発現された組織に関する情報が得られた。高度には発現されない転写物に関しては、反応の鋳型として目的の組織からの RNA を用いて逆転写 PCR アッセイを実施する必要があるときもあった。

【0110】

コンピューターによる方法によるおよび直接 cDNA 選択による遺伝子同定により、染色体の領域における遺伝子についての独特な情報が提供される。遺伝子が同定された場合、次いで各々の遺伝子における変異に関して異なる個体を試験することが可能である。

10

【0111】

本発明はまた各々配列番号 2 および 4 により示される形態の HBM 遺伝子および HBM タンパク質、およびその他の密接に関連する変異体、ならびにその正確な発現に必要な HBM 遺伝子の隣接する染色体領域をも包含する。好ましい実施形態では、本発明は配列番号 2 の単離された核酸配列およびその変異体を志向する。配列番号 2 の変異体には、配列番号 2 の核酸配列またはそのフラグメントに対して少なくとも約 90%、好ましくは 95%、またはより好ましくは 98% の類似性または同一性を有しているポリヌクレオチドが含まれる。従って、配列番号 2 またはそのフラグメントに 96%、97%、および 99% 類似している配列もまた本明細書にて意図される。

【0112】

BLAST または FASTA または標準的なデフォルトパラメーターを用いるその他の適当なアルゴリズムを用いて高骨量 (HBM) 変異体間の変異体の程度の決定を実施することができる。好ましくは、配列番号 2 のコード化領域に関して同一性を決定するが、非コード化ドメインを含めることもできる。加えて、アラインメントプログラムを用いて、保存配列または異なる動物種にわたる潜在モチーフを同定することができる。アラインメントプログラムを用いて、関連する遺伝子およびそれがコードするタンパク質の核酸および／またはタンパク質配列を整列させることができる。好ましいアラインメントプログラムには CLUSTALW、PILEUP および GAP などがあり、好ましくはこれをデフォルトパラメーターと共に用いる。例えば、このようなプログラムを用いて LRP5 (Zmax1 としても公知)、HBM、LDL レセプター関連タンパク質 6 (LRP6) および関連する配列の配列を整列させることができる。

20

30

【0113】

ポリペプチドをコードする参照ヌクレオチド配列に少なくとも例えば 90% 「類似する」ヌクレオチド配列を有するポリヌクレオチドにより、ポリヌクレオチド配列が、参照ヌクレオチド配列の各 100 ヌクレオチドあたり 10 個までの点変異を含み得ること以外は、ポリヌクレオチドのヌクレオチド配列が参照配列と同一であることが意図される。参照配列のこれらの変異は配列番号 2 のいずれの位置で生じてもよく、サイレントでもよく、または、アミノ酸置換をコードしても、しなくてもよい。

【0114】

別の実施形態は、配列番号 2 のこのようなポリヌクレオチド変異体が、配列番号 2 の少なくとも 15、20、25、30、35、40、45、50、60、70、80、90、100、150、200、300、400 または 500 個の連続するヌクレオチドである核酸配列を含むことを意図する。さらに好ましくは、このようなポリヌクレオチド変異体は、配列番号 2 または配列番号 2 の別の変異体のヌクレオチド 582 において多型性 (G → T 置換) に対応する連続する核酸配列を有し、それによりコードされるポリペプチドが対象に投与された場合、これは骨量を変調する変異を含む。意図される配列番号 2 の全ての変異体は、対象に投与された場合、骨変調を誘起するタンパク質またはポリペプチドをコードする特性を有する。対象に投与された場合に骨量の変調に寄与し得る別の変異体は、HBM 多型性を含有することが解っているドメイン内にあり、これは β プロペラドメイン (YWTDMotif) をコードする。また別に、骨量を変調し、および／または対象に

40

50

におけるHBM表現型に至るLRP5の別の変異体は、LRP5の別の保存されたドメイン、例えば図4（例えばRGD細胞外付着部位、LDLおよびカルシウムの結合部位、システイン富化成長因子反復、理想PEST領域、および内部移行ドメイン）に示すもののいずれかをコードする核酸配列における変異に起因し得る。

【0115】

意図されるHBMポリヌクレオチドにはストリンジェント条件下で配列番号2にハイブリダイズするものが含まれる。特に限定されるものではないが、当分野で公知のハイブリダイゼーション法には、(a) 50℃で0.1xSSPE(0.62MのNaCl、0.06MのNaH₂PO₄・H₂O、0.075MのEDTA、pH7.4)および0.1%ドデシル硫酸ナトリウム(SDS)で洗浄；(b) 42℃で50%ホルムアミド、5xSSC(0.75MのNaCl、0.075Mのクエン酸ナトリウム)、50mMのリン酸ナトリウム(pH6から8)、0.1%ピロリン酸ナトリウム、5xデンハート溶液、超音波処理サケ精子DNA(50μ/ml)、0.1%SDSおよび10%デキストラン硫酸で洗浄、続いて42℃で0.2xSSCおよび0.1%SDSで洗浄；ならびに(c) 65℃で0.5MのNaPO₄、7%SDSで洗浄、続いて60℃で0.5xSSCおよび0.1%SDSで洗浄などがある。配列番号2またはその核酸フラグメントに対するハイブリダイゼーションによりHBM変異体を単離することができる別の条件を、ハイブリダイゼーション温度を変化させることにより実行できる。高度にストリンジェントなハイブリダイゼーション条件は配列番号2またはそのフラグメントの融点(T_m)以下の約20℃で実施される。好ましいストリンジェンシーは配列番号2またはそのフラグメントのT_m以下の約5から10℃で実施される。さらなるハイブリダイゼーション条件をSambrookら、Molecular Cloning: A Laboratory Manual(1989)の第11章に記載されるよう、または当業者に知られているように調製することができる。

【0116】

また別に、縮重プライマーおよびポリメラーゼ連鎖反応(PCR)技術を用いて、哺乳動物ライブラリー(例えばウマ、霊長類、ヤギ、ウシ、ヒツジ、ネコ、ブタ、およびイヌ)をプロービングして配列番号2またはそのフラグメントの変異体を同定することができる。好ましくはストリンジェント条件下で配列番号2のオープンリーディングフレームに、またはその配列の非コード化部分にハイブリダイズするプライマーを利用する。さらに好ましくは、このようなプライマーは配列番号2内の保存されたドメインにハイブリダイズする。例えば、保存されたドメインはYWT Dβプロペラドメインまたはその他のドメインをコードするもの、例えば図4に列挙するものを含む。好ましいプライマーは典型的には15ヌクレオチドの長さであるが、少なくとも約16、17、18、19、20、21、22、23、24、25、30、35または40ヌクレオチドの長さに変化し得る。異種性ハイブリダイゼーションは、縮重PCRプライマーを用いて標的遺伝子または核酸ハイブリダイゼーションを増幅することである。対応するアミノ酸配列(例えばその配列番号4フラグメント)をコードできる可能な全てのヌクレオチド配列を表す、10以上、好ましくは15またはそれ以上のヌクレオチドの長さの混合オリゴヌクレオチドを調製することにより、配列番号2およびそれによりコードされるポリペプチドの変異体に関するプローブを得ることができる。この方法はGouldら、Proc.Natl.Acad.Sci.USA 86(6):1934-1938(1989)により明確に文献化されている。

【0117】

別の実施形態には、配列番号4およびそのフラグメントに少なくとも約90%類似し、対象に投与された場合にその対象において骨量を変調するHBMポリペプチドをコードする核酸などがある。このようなHBMポリペプチドには配列番号4の位置171またはマウス相同体の位置170に対応するバリン(GlyのValへの置換)を有する変異体などがある。その他の好ましい実施形態には、配列番号4の少なくとも約10、15、20、25、30、35、40、45、50、60、70、80、90、100、150、200、250、300、500またはそれ以上の連続するアミノ酸を有する高骨量ポリペプチドなどがある。このような意図される連続する配列は好ましくは高骨量、例えば

配列番号4のバリン171に対応する多型性と重複する。配列番号4およびそのフラグメントに少なくとも約95%、96%、97%、98%および99%類似するポリペプチドをコードするポリヌクレオチドもまた意図される。

【0118】

別の実施形態では、配列番号4をコードする合成核酸であって、配列番号4におけるアミノ酸は変化しないが、恐らく該配列番号4をコードする得られた合成ポリヌクレオチドは配列番号2に少なくとも約50%類似するものであるように、コードの縮重に基づいて保存置換(conservatively substituted)されている核酸が意図される。

【0119】

配列番号4またはそのフラグメントの参照アミノ酸配列に対して少なくとも、例えば95%「同一」であるアミノ酸配列を有するポリペプチドとは、ポリペプチド配列が配列番号4の参照アミノ酸配列の各100個のアミノ酸あたり5個までのアミノ酸変化を含んでよい以外は、ポリペプチドのアミノ酸配列が参照配列と同一であることを意図している。換言すれば、参照アミノ酸配列に95%同一であるアミノ酸を有するポリペプチドを得るために、参照配列の5%までのアミノ酸残基が欠失または別のアミノ酸と置換され得るか、または参照配列の全アミノ酸残基の5%までのアミノ酸の数が参照配列に挿入され得る。これらの参照配列の変化は参照アミノ酸配列のアミノもしくはカルボキシ末端位置で、またはこれらの末端位置の間のどこかで、参照配列において別個に残基の間かまたは参照配列内の1つまたはそれ以上の連続するグループのいずれかで散在して生じ得る。

【0120】

アミノ酸残基が保存置換されているさらなるEBMポリペプチドおよび該HBMポリペプチドをコードする核酸が意図される。例えば、表現型としてサイレントアミノ酸置換を作製する方法に関する手引きはBowieら、「タンパク質配列におけるメッセージの解説：アミノ酸置換に対する認容性(Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions)」Science 247:1306-1310(1990)にて提供され、ここで著者はアミノ酸配列の変化に対する認容性の研究に関して2つの主要な研究方法があることを示している。第1の方法は、変異が天然の選択により許容されるか、または拒絶される進化の過程に依存する。第2の研究法は遺伝子操作を用いて、クローン化された遺伝子の特定の位置でのアミノ酸変化、および機能を維持する配列を同定するための選択またはスクリーニングを導入する。これらの研究により、タンパク質が驚くほどにアミノ酸置換に認容性があることを示している。著者はさらに、アミノ酸変化がタンパク質の特定の位置で許容状態である可能性があることを示している。Bowieら(前出)では多くの表現型置換が記載されており、そこに引用された参照は、参照によりその全てを本明細書に組み込む。あまり保存されていない種にわたるドメインであって、構造的にまたは機能的に重要なドメイン(例えば結合部位、触媒部位、もしくはβプロペラまたはその他の図4に記載されるドメイン)に相当しないドメインにおける置換が好ましい。

【0121】

LRP5/HBMの変異体

Springerら、J.Molecular Biology 283:837-862(1998)に記載するように、2個のYWTD-βプロペラ含有分子(ニワトリLRP1およびヒトニドゲン、各々Swiss-PROT参照LRP1 CH7およびNIDO HU1を用いる)からのモデル予測に基づいて、LRP5第1βプロペラモジュールの構造モデルを作製した。モデルに基づいて、特定のアミノ酸残基をHBM/LRP5の重要な変異体として同定した。以下の3つのカテゴリーがこのような変異体の事例を提供する。

【0122】

βプロペラの形状は側面が内側に傾斜し、中央に孔があいているディスクに類似している。残基171はプロペラモジュール1の翼4のドメインの外側または上部表面のループにある。従って、変異体は別の翼の構造的に同等の位置にある残基において変化した残基を含み；および結合ポケットに対してさらにわずかに内部にあるが、依然表面に接近可能な残基が、HBMによる骨量変調の研究に、骨量障害の医薬品および処置の開発に、なら

10

20

30

40

50

びに本発明の別の目的に関する本発明の重要な実施形態である。このような変異体の実例は以下のとおりである。

【0123】

A214V（翼5の171の等価な位置；アラニンは別のプロペラで保存されていない）、

E128V（翼3の171の等価な位置；グルタミン酸は別のプロペラで保存されていない）、

A65V（翼2の171の等価な位置；アラニンは1から3のプロペラで保存されているが4では保存されていない）、

G199V（翼5の接近可能な内部の位置；グリシンは1から3のプロペラで保存されているが4では保存されていない）；および

M282V（翼1の接近可能な内部の位置；メチオニンは1から3のプロペラで保存されているが4では保存されていない）。

【0124】

LRP5は4つのβプロペラ構造を有している。1／3のβプロペラモジュールはヒトLRP5の残基171に相当する位置でグリシンを保存している。従って、別のプロペラの等価の位置でバリンを担持する変異体は本発明の重要な実施形態である。以下の変異体はHBMによる骨量変調の研究に、骨量障害の医薬品および処置の開発に、ならびに本発明の別の目的に有用である。G479V、G781VおよびQ1087V。

【0125】

G171V HBM多型性により、開いた結合ポケットに突出し、リガンド／タンパク質相互作用を変化させる可能性があるバリン残基からの側鎖を伴うβプロペラ1の「占有空間」に至る。グリシン残基はLRP5プロペラ1、2および3で保存されるが、プロペラ4ではグルタミン酸が保存される。従って、HBMの以下の変異体がHBMによる骨量変調の研究に、骨量障害の医薬品および処置の開発に、ならびに本発明の別の目的に関する本発明の重要な実施形態である。

【0126】

G171K（これは荷電した側鎖を導入する）、

G171F（これは環状の側鎖を導入する）、

G171I（これは分岐した側鎖を導入する）、および

G171Q（これはプロペラ4残基を導入する）。

【0127】

さらに、LRP6はLRP5の最も密接な相同体である。従って、LRP6により骨密度を変調させることもできる。LRP6は、LRP5と同一ではないが、類似していると予測されるβプロペラ構造を有している。ヒトLRP5のグリシン171に対応する位置はヒトLRP6のグリシン158である。従って、LRP6の対応する変異体はLRP5対その関連するフェミリーメンバーの特異性の研究に、骨量障害の医薬品および処置の開発に、ならびに本発明の別の目的に関する本発明の重要な実施形態である。とりわけ、例えば、ヒトLRP6およびその他の種のLRP6相同体の類似の変異体の構造的に等価な位置、残基158でのグリシンのバリンへの置換は重要な研究用手段を提示している。

【0128】

当業者はこれらが、本発明をよりうまく記載するために提示される極わずかな説明用の実例であり、多くのその他の変異体は本発明の範囲内になることを意図し得ることを認識している。

【0129】

ポリペプチドまたはポリペプチドをコードする核酸の骨量変調活性を決定する方法を、骨量を研究するための異なる動物モデルを用いて実施することができる。例えば、卵巣切除されたネズミモデルまたは自然発生骨粗鬆症マウス株を利用して、LRP5変調物質が動物モデルにおいて相応して骨量を変調するかどうかを決定することができる。このようなマウスのインビボ分析に関しては、Kaluら、J.Bone Miner.Res.14:593-601(1999)、お

10

20

30

40

50

よびShimizuら、Mamm.Genome 10:81-87(1999)を参照されたい。

【0130】

使用することができるさらなるインビボアッセイは、LRP5の発現が変化しているか、またはHBMをコードする核酸が導入されているトランスジェニック動物およびノックアウト動物である。次いでこれらの動物を利用して骨量を変調する化合物または組成物を同定することができる。

【0131】

本発明のポリペプチドは好ましくは、単離された形態で提供される。「単離されたポリペプチド」とは元来の環境から除去されたポリペプチドを企図する。従って、組換え宿主細胞内で生成および含有されるポリペプチドは、本発明の目的に関して「単離された」と考えることができる。「単離されたポリペプチド」とは部分的または実質的に組換え宿主から精製されているポリペプチドであるとも意図される。同様に、「単離された核酸」または「単離されたポリヌクレオチド」とは別の核酸およびタンパク質夾雑物から精製されている核酸配列を意味する。

10

【0132】

本発明はまた各々配列番号1および3に示される形態のLRP5遺伝子およびLRP5タンパク質、ならびにその他の密接に関連する変異体をも包含する。本発明はまたその正確な発現に必要なLRP5の隣接する染色体領域をも包含する。

【0133】

好ましい実施形態では、本発明は配列番号1の核酸配列の少なくとも15個の連続するヌクレオチドを志向する。配列番号1および3で示される形態のLRP5遺伝子およびLRP5タンパク質の変異体を、各々G171V HBM多型性を有さないHBM遺伝子およびHBMタンパク質に関して前記したように一般的に同定することができる。

20

【0134】

本発明はまたLRP5遺伝子領域のヌクレオチド配列およびHBM遺伝子領域のヌクレオチド配列にも関する。さらにとりわけ、好ましい実施形態はLRP5遺伝子領域B200E21-HおよびB527D12-Hのセグメントを含有するBACクローンである。好ましい実施形態は配列番号5~12からなるBACクローンのヌクレオチド配列である。

【0135】

本発明はまたLRP5遺伝子およびHBM遺伝子に関するDNAプローブを同定するためのヌクレオチド配列、LRP5遺伝子およびHBM遺伝子を増幅するためのPCRプライマー、LRP5遺伝子およびHBM遺伝子におけるヌクレオチド多型性、ならびにLRP5遺伝子およびHBM遺伝子の制御エレメントの使用に関する。

30

【0136】

II. DXA測定を用いる表現型分類

Creighton大学(Omaha, Nebraska)で実施した脊椎骨ミネラル含量(BMC)および骨ミネラル密度(BMD)測定はNorland Instrumentsデンシトメーター(Norland XR2600デンシトメーター、二重エネルギーX線吸光光度分析法、DXA)を用いて行った。別の位置での脊椎BMCおよびBMDは利用できる機械を用いた。現在米国で操作されているDXA機械は800と推定される。最も大きな都市にはDXA能力、通常LunarまたはHologic機械を有するオフィスまたは撮影所がある。脊椎BMCおよびBMDデータを提供する各々の場所には機械からのプリントアウトのコピーを含み、BMDの測定のための目的の領域が適当に選択されていることの証明が提供される。完全な臨床履歴および骨格放射線写真が得られた。

40

【0137】

ヒトおよび動物対象、好ましくはヒトにおけるHBM表現型を診断基準：例えば非常に高度な脊椎BMD；いずれかの公知の高骨量症候群を欠く臨床履歴；ならびに四肢骨格の正常な形状を示すラジオグラフィー放射線写真；を用いて記載することができる。

【0138】

50

I I I . マイクロサテライトマーカーの遺伝子型解析

遺伝子間隔をJohnsonら、Am.J.Hum.Genet.60:1326-1332(1997)により元来報告されたものより小さな領域まで狭めるために、染色体11q12-13上のさらなるマイクロサテライトマーカーを分類した。新規のマーカーには、D11S4191、D11S1883、D11S1785、D11S4113、D11S4136、D11S4139(Dibら、Nature 380:152-154(1996)、FGF3(Polymeropolousら、Nucl.Acids Res.18:7468(1990))、ならびにGTC HBM マーカー1、GTC HBM マーカー2、GTC HBM マーカー3、GTC HBM マーカー4、GTC HBM マーカー5、GTC HBM マーカー6、およびGTC HBM マーカー7などがある(図2参照)。

【0139】

10

資格のある瀉血専門医が血液(20ml)を、薄紫色のキャップのチューブ(EDTA含有)に抜き取って入れた。DNA抽出まで血液を冷蔵保存した。品質または収量が低下しないで冷蔵庫で7日間保存した血液からDNAを抽出した。離れた場所で血液を抜き取られたこれらの対象に関しては、多くの場合、配送プロトコルを用いて成功した。血液サンプルを冷却するための冷凍パックと共に発泡スチロール容器に入れて、翌日配達で配送した。薄紫色のキャップのチューブを個々のプラスチック配送用チューブ上に、そして次いで「ジップ・ロック」バイオハザードバッグに内に置いた。翌日にサンプルが到着したとき、これを即座に処理してDNAを抽出した。

【0140】

20

DNA抽出手順はGentra Systems,Inc.(Minneapolis、ミネソタ州)から購入したキットを使用した。簡単には、手順は3容量の赤血球溶解バッファーを全血に添加することを含む。室温で10分間インキュベートした後、Beckmanテーブルトップ遠心で溶液を2000Xgで10分間遠心した。白血球ペレットを細胞溶解バッファーに再懸濁した。一度ペレットが完全に再懸濁され、細胞の塊がなくなったら、溶液をRNアーゼAで、37℃で15分間消化した。提供されたタンパク質沈殿溶液の添加によりタンパク質を沈殿させ、遠心により除去した。イソプロパノールの添加により上澄からDNAを沈殿させた。この方法は単純および迅速であり、1から2時間を要するのみであり、同時に多量のサンプルの処理が可能であった。DNA収量は通常全血20mlに関して>8mgであり、MW>50kbであった。保存コード化された50μgアリコートで、-80℃でエタノール沈殿としてDNAが得られた。

30

【0141】

蛍光標識オリゴヌクレオチドプライマーおよび未標識オリゴヌクレオチドプライマーを用いてDNAを遺伝子型解析した。Integrated DNA Technologies,Inc(Coralville、アイオワ州)から標識および未標識オリゴヌクレオチドを入手した。マイクロサテライト遺伝子型解析のためのその他の全ての試薬はPerkin Elmer-Applied Biosystems,Inc.(「PE-ABI」)(Norwalk、コネチカット州)から購入した。個々のPCR反応を、PE-ABIに記載されるように、AmpliTaq(商標)DNAポリメラーゼを用いて各々のマーカーに関して実施した。反応物を脱イオン化ホルムアミド、ブルー・デキストランおよびTAMRA 350サイズ・スタンダード(PE-ABI)を含有する負荷バッファー3.5μlに加えた。95℃で5分間加熱してDNAを変性させた後、モデル377DNAシーケンサー(PE-ABI、Foster City、カリフォルニア州)の操作マニュアルに記載されているように、サンプルを負荷し、電気泳動した。ゲル電気泳動の後、PE-ABI GENESCAN(登録商標)およびGENOTYPER(商標)ソフトウェアを用いてデータを分析した。最初に、GENESCAN(商標)ソフトウェア内で、分析の第1ステップの前にレーントラッキングを手作業で最適化した。ゲルレーンデータを抽出した後、各レーンの標準的な曲線プロファイルを試験し、直線性およびサイズコーリングを確認した。これらのパラメーターのいずれかで問題を有したレーンを再トラッキングし、確認した。一度全レーンをトラッキングし、サイズスタンダードを正確に同定すると、対立遺伝子同定のためにデータをGENOTYPER(商標)にインポートした。対立遺伝子の呼び出し(ピニング)を迅速にするために、Dr.Guy Van Campのインターネットウェブサイト(<<http://alt.www.uia.ac.be/u/dnalab/ld.html>>)から

40

50

のプログラム・Linkage Designerを使用した。このプログラムはGENOTYPER（商標）により作製したデータの家系ドローイングプログラムCyrillic(バージョン2.0、Cherwell Scientific Publishing Limited, Oxford、英国)へのインポート、およびそれに続くLINKAGE(Lathropら、Am.J.Hum.Genet.37:482-498(1985))プログラムを用いる連鎖解析を大いに促進する。

【0142】

I V. 連鎖解析

図1は本発明の遺伝子連鎖研究において使用した個体の家系を示す。とりわけ、LINKAGE(Lathropら、Am.J.Hum.Genet.37:482-498(1985))プログラムのMLINKおよびLINKMAPコンポーネントを用いて2点連鎖解析を実施した。家系／マーカーデータをMakepedプログラムへのpreファイルとしてCyrillicからエクスポートし、連鎖解析用に適当なpedファイルに変換した。

【0143】

元来の連鎖解析は3つのモデルを用いて実施した。(i)常染色体優性、十分に浸透したモデル、(ii)浸透度が低下した常染色体優性モデル、および(iii)量的形質モデル。広く拡大した親族の22人のメンバーからのリンクマーカーに関してDNAを分析することにより、HBM遺伝子座は染色体11q12-13にマッピングされた。6から22cmの範囲の間隔で22個の常染色体にわたる345個の蛍光マーカーのパネルを伴う高度に自動化された技術を用いた。染色体11のこの領域からのマーカーのみが連鎖の証拠を示した(LODスコア~3.0)。最も高いLODスコア(5.74)が2点により得られ、多点解析はD11S987(図2の地図位置55)であった。95%信頼区間によりマーカーD11S905とD11S937の間にHBM遺伝子座が配置された(図2の地図位置41-71)。ハプロタイプ分析でもLRP5遺伝子がこの同一の領域に配置される。マーカーD11S987、D11S905およびD11S937のさらなる記載をGyapayら、Nature Genetics 7巻(1994)に見出すことができる。

【0144】

本発明では、発明者はHBM間隔をマーカーD11S987とGTC HBM マーカー5の間の領域に狭めることを報告している。これらの2つのマーカーは元来の分析(D11S905およびD11S937)からデリミッティング・マーカーの間にあり、そして互いにおよそ3cmである。マーカーD11S4191、D11S1883、D11S1785、D11S4113、D11S4136、D11S4139(Dibら、Nature 380:152-154(1996)、FGF3(Polymeropoulosら、Nucl.Acids Res.18:7468(1990)) (遺伝子マーカーに関する情報はゲノムデータベースのインターネットサイト、<http://gdwww.gdv.org/> <<http://gdvwww.gdv.org/>>で見出すことができる)、ならびにマーカー、GTC HBM マーカー1、GTC HBM マーカー2、GTC HBM マーカー3、GTC HBM マーカー4、GTC HBM マーカー5、GTC HBM マーカー6、およびGTC HBM マーカー7からの遺伝子型データを用いて間隔を狭めることができる。

【0145】

図1で示すように、前記の遺伝子マーカーを用いるハロタイプ解析により、LRP5遺伝子が局在する染色体11の間隔を有意に絞り込む個体9019および9020の組換え事象(交差)を同定する。個体9019はHBM遺伝子を有する母方の染色体からの染色体11の部分、および染色体11相同体からの部分を遺伝しているHBMを罹患する個体である。HBM遺伝子担持染色体から遺伝された部分は、マーカーD11S935、D11S1313、GTC HBM マーカー4、D11S987、D11S1296、GTC HBM マーカー6、GTC HBM マーカー2、D11S970、GTC HBM マーカー3、D11S113、GTC HBM マーカー1、GTC HBM マーカー7およびGTC HBM マーカー5を含む。D11S4136からおよびテロメア方向へ続く部分は非HBM染色体に由来する。このデータによりLRP5遺伝子はマーカーGTC HBM マーカー5に対してセントロメアよりの位置に配置される。個体9020もまた重要な組換え事象を呈する非罹患個体である。この個体はHBM遺伝子を担持

する彼女の父（個体 0 1 1 5）の染色体 1 1 相同体からのマーカー D 1 1 S 9 3 5、D 1 1 S 1 3 1 3、G T C H B M マーカー 4、D 1 1 S 9 8 7、D 1 1 S 1 2 9 6 および G T C H B M マーカー 6、ならびに H B M 遺伝子を担持しない彼女の父の染色体 1 1 からのマーカー G T C H B M マーカー 2、D 1 1 S 9 7 0、G T C H B M マーカー 3、G T C H B M マーカー 1、G T C H B M マーカー 7、G T C H B M マーカー 5、D 1 1 S 4 1 3 6、D 1 1 S 4 1 3 9、D 1 1 S 1 3 1 4、および D 1 1 S 9 3 7 を含む組換え父方染色体 1 1 を遺伝する。マーカー D 1 1 S 1 1 3 は個体 0 1 1 5 のそのホモ接合体特性のため情報価値がない。この組換え事象により H B M 領域のセントロメア境界はマーカー D 1 1 S 1 2 9 6 と D 1 1 S 9 8 7 の間に配置される。

【0 1 4 6】

10

また 2 点連鎖解析を用いて染色体 1 1 上の L R P 5 遺伝子の位置を確認した。十分な浸透度のモデルでの 2 点連鎖解析の連鎖結果を以下の表 1 に提示する。この表は第 1 コラムで遺伝子マーカー、および表の上から全体にわたって組換え分画を列挙する。列の各セルは第 1 列で示す組換え分画で L R P 5 遺伝子に対する連鎖に関して試験した個々のマーカーの L O D スコアを示す。例えば、マーカー D 1 1 S 9 7 0 でのピーク L O D スコアは 7 . 6 6 であり、これはハロタイプ解析により定義された間隔内にある。

【0 1 4 7】

【表 1】

表 1

マーカー	0.0	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4
D11S935	—無限大	0.39	0.49	0.47	0.41	0.33	0.25	0.17	0.10
D11S1313	—無限大	2.64	2.86	2.80	2.59	2.30	1.93	1.49	1.00
D11S987	—無限大	5.49	5.18	4.70	4.13	3.49	2.79	2.03	1.26
D11S4113	4.35	3.99	3.62	3.24	2.83	2.40	1.94	1.46	0.97
D11S1337	2.29	2.06	1.81	1.55	1.27	0.99	0.70	0.42	0.18
D11S970	7.66	6.99	6.29	5.56	4.79	3.99	3.15	2.30	1.44
D11S4136	6.34	5.79	5.22	4.61	3.98	3.30	2.59	1.85	1.11
D11S4139	6.80	6.28	5.73	5.13	4.50	3.84	3.13	2.38	1.59
FGF3	0.59	3.23	3.15	2.91	2.61	2.25	1.84	1.40	0.92
D11S1314	6.96	6.49	5.94	5.34	4.69	4.01	3.27	2.49	1.67
D11S937	—無限大	4.98	4.86	4.52	4.06	3.51	2.88	2.20	1.47

20

30

【0 1 4 8】

単一塩基多型（S N P）がさらに H B M 領域を定義する。この S N P は S N P コンテイング 0 3 3 - 6 と称され、そして遺伝子マーカー G T C H B M マーカー 5 に対して 2 5 k b セントロメアより位置する。この S N P は遺伝子マーカー G T C H B M マーカー 7 に対してテロメアよりである。S N P コンテイング 0 3 3 - 6 は H B M 罹患個体 0 1 1 3 に存在する。しかしながら、0 1 1 3 の息子である H B M 罹患個体 9 0 1 9 はこの S N P を担持しない。従って、これは交差がこの S N P に対してセントロメアよりであることを示している。遺伝子マーカー G T C H B M マーカー 5 および G T C H B M マーカー 7 のプライマー配列を以下の表 2 に示す。

40

【0 1 4 9】

【表 2】

表 2

マーカー	プライマー (順行性)	プライマー (逆行性)
GTC_HBM_ マーカー_5	TTTTGGGTACACAATTGAGTCG (配列番号: 63)	AAAACTGTGGGTGCTTCTGG (配列番号: 65)
GTC_HBM_ マーカー_7	GTGATTGAGCCAATCCTGAGA (配列番号: 64)	TGAGCCAAATAAACCCCTTCT (配列番号: 66)

【0150】

記載した親族は非常に興味深いいくつかの特徴を有し、顕著なものでは彼らの骨は、非常に高密度であるが、絶対的に正常な形態を有している。HBMに罹患した個体の骨格の外寸は正常であり、髄腔は存在するが、造血の干渉はない。HBMに罹患したメンバーは骨折に対して抵抗するようであり、試験したメンバーには神経学的な症状はなく、いずれかの器官や系における機能の障害の症状はない。親族のHBMに罹患したメンバーはひどい病気または身体障害もなく、高齢まで生存する。さらに、HBM表現型はその他の骨障害、例えば骨粗鬆症、骨粗鬆症偽性神経膠腫、エンゲルマン病、リビング病、高ホスファターゼ血症、ファン・ブッヘム病、メロレオストーシス、大理石骨病、濃化異骨症、硬化狭窄症、オステオポイキローシス、先端巨大症、パジェット病、繊維性骨異形成症、管状狭窄症、骨形成不全症、副甲状腺機能低下症、偽性副甲状腺機能低下症、偽性偽性副甲状腺機能低下症、1次および2次性副甲状腺機能亢進症および随伴される症候群、高カルシウム尿症、甲状腺の髄様癌、骨軟化症ならびにその他の疾患に適合しない。明らかに、このファミリーのHBM遺伝子座は骨密度の制御において非常に強力で、実質的な役割を有しており、その同定は骨密度を制御する(複数の)経路および骨粗鬆症のような疾患の病因を理解するのに重要なステップである。

【0151】

加えて、HBM遺伝子を担持する、従ってHBMタンパク質を発現する年輩の個体は正常な個体の骨量特性の喪失を示さない。換言すれば、HBM遺伝子は骨粗鬆症のサプレッサーである。本質的には、HBM遺伝子を担持する個体はHBMタンパク質を与えられ、結果的に骨粗鬆症を進行しない。このインビボ観察は、正常な個体のHBM遺伝子もしくはタンパク質、またはそのフラグメントでの処置が骨粗鬆症を改善するという強い証拠である。

【0152】

V. 物理マッピング

HBM遺伝子座のクローニングおよび特徴付けのための試薬を提供するために、前記した遺伝子マッピングデータを用いて染色体11q13.3上にLRP5遺伝子を含む領域の物理地図を構築した。物理地図は分子ランドマークの順序付けされたセットおよび染色体11q13.3からのLRP5遺伝子領域を含むBACクローンのセットから成る。

【0153】

種々の公に利用可能なマッピングリソースを利用してHBM領域の既存のSTSマーカー(Olsonら、Science 245:1434-1435(1989))を同定した。リソースにはGDB、the Whitehead Institute Genome Center、dbSTSおよびdbEST(NCBI)、11db、Texas Southwestern大学 GSETEC、the Stanford Human Genome Center、ならびにいくつかの参考文献(Courseauxら、Genomics 40:13-23(1997)、Courseauxら、Genomics 37:354-365(1996)、Guruら、Genomics 42:436-445(1997)、Hosodaら、Genes Cells 2:345-357(1997)、Jamesら、Nat.Genet. 8:70-76(1994)、Kitamuraら、DNA Reseach 4:281-289(1997)、Lemmensら、Genomics 44:94-100(1997)、Smithら、Genome Res.7:835-842(1997))などがある。地図を手作業で組み込み、LRP5を含む領域にマッピングするマーカーを同定した。

【0154】

GDBまたは参考文献から入手した既存のSTSに関するプライマーを以下の表3に列

挙する。従って、表3はLRP5遺伝子領域の物理地図を調製するために用いたSTSマーカーを示す。

【0155】

【表3A】

TABLE 3: HBM STS Table

STS Name	Locus Name	Type	Gene Access.	Size (kb)	Seq ID No.	Forward Primer	Reverse Primer	Gene Name
ACTN3		Gene	GDB:197568	0.164	67	CTGAGTCTGGCTCTCTC	66: TTGAGAGTCTGGCTCTC	Actinin, alpha 3 - skeletal muscle
PC-BPC-Y		Gene	GDB:197884	0.125	69	CTCAGTGCATGAGTATGGA	70: CAGATCACTGATGCTGAGG	Pyruvate Carboxylase
D11S2181E		Gene		0.322	71	GTTTCAGAGAGTCTGAGAGT	72: TCTGCGAGTCTGCTGTTGAG	Adenosine Receptor (A2) Gene
ADRBK1		Gene	GDB:459017E	0.259	73	TTATGTGATTCGGTGGC	74: GCGCTGTCTGCTGCTTCCAG	Beta-adrenergic receptor kinase
RPL12/PP1(22)		Gene	GDB:197568	0.208	75	GAGAGCAATATAGGGAGCC	76: TGGTCTTGTAAGAGCTGAGA	sim. to Human endogenous retrovirus mRNA long terminal repeat
CSPT1-PCRT		Gene	GDB:270068	0.197	77	GAGATACGGGAGGAGCTGAGAGCTACGT	78: ATGACCAAGTCTGACACCCGCTTCACA	Protein phosphatase 1, catalytic subunit, alpha isoform
NDUFR1		Gene		0.521	81	CATGTGCGCCACAGAGCTGAGAGCTATCAT	82: CAGATCTGTGATGCTGCTGG	Glutathione S-transferase pi
PSANK1		Gene		0.157	83	CAGAGAGTCAAGGAGCTT	84: ATCTCTCTGATGCTGCTGCT	NADH dehydrogenase (ubiquinone) flavoprotein 1 (F1ND)
PSANK2		Gene		0.3	85	CAGAGTCAAGGAGCTT	86: CAGAGTCAAGGAGCTT	Aldehyde Dehydrogenase 8 (ALDH8)
U5520		EST	GDB:314521	0.211	87	AAGTCCAGGCTGCAAGGAG	88: CCGTGTGCTGCTTTCATGTA	Human ribosomal protein L37 (PSANK1) pseudogene.
AFM289va9		MSAT	GDB:199805	0.287	89	AAGGTGTGAGGATCACTGG	90: AGCTCATGCGGGCTTAT	Preprogalactin (GAL1)
GALN1		Gene		0.322	91	GTTCTCTCGAGTGTATCAAG	92: ATGCGACAGGAGCTGAGACA	B-cell CLL/lymphoma 1 - Cyclin D1 (PRAD1 gene)
BCL11(BCL12)		Gene	GDB:177850	0.205	93	GATCAGGAGCTTCTGCTGGCTC	94: TCCACATTGAGGAGCTGAGGAGG	Cyclin D1
CCND1		Gene	GDB:459014	0.248	97	GACAGCTGTATGAGGCTTCTAGGC	98: TTGCACTGTCTGATGAGA	Fibroblast growth factor 4
FGF4		Gene	GDB:459013	0.549	99	CACGATGTATGAGGCTTCTAGGC	99: CAGGCGCAAGGAGCTGACACAGGC	Fibroblast growth factor 3
FGF3-PCR1		Gene	GDB:188927	0.161	101	TTCTGGGTGTGCTGAAT	102: ACACAGTCTGCTTAAAGGGT	Choline Kinase
AFM164Zf12		MSAT	GDB:188151	0.22	103	CATTGGGAAATCCAGAAGA	104: TAGGTGTCTTATTTTGTGCTTC	
SHGC-15295		MSAT	GDB:122324E	0.275	105	GACATACATGAGATGATAGAGG	106: GAGCCATACAGGAGTAAG	
SHGC-3084		STS	GDB:740600	0.147	107	GACACAGGAGGTAGTGGC	108: TGAGGACACATACATGATGGG	
SHGC-14407		STS	GDB:740102	0.167	109	GAGGTGTCTGCTTAAATCTTTTG	110: GAGCATATGTATGTATGATGAGAG	
SHGC-10946		STS	GDB:674522	0.311	113	ACTCTACACCCAGTCCG	112: TCTTGGCTGCTGATGTTCTCGG	
S1516		STS	GDB:196290	0.166	115	TTGAGACAGGATAGCTGAGG	114: TCGTGTGCTGCTGCTGCTGAG	
AFM147XD10		MSAT	GDB:307895	0.163	117	AGCTGAGCTCTACAGATG	118: CAGAGGCTGGTAGAGAGTG	
D11S1987		MSAT	GDB:195002	0.082	119	GATCCAGTGTGGGCAATAAAGG	120: GGTGCGACATGACCTCTTAAG	
D11S417E		MSAT	GDB:611922	0.237	121	CAGGCCAGTCTCTTG	122: CGTGCTCAGATGAAGTG	
AFM272/2/5		EST	GDB:458164	0.216	123	ACCTACGGTGTATCCG	124: CTGAGGCCGCTCTTGG	
WI-17803		EST	GDB:457860	0.128	127	ACTTATGTCAGCTGGGCG	128: ACTCCCTGCTGCTTGGC	
SGC31923		EST	GDB:677852	0.324	129	GAGCAGGAGAGAGAGC	130: CCAAGTGGCTGTTTATG	Transformion-sensitization protein IEF SSP 3521
SGC35223		EST	GDB:459258	0.13	131	AGCCACTTATGTTATTTATGTC	132: AAGAGTGAACAAAGCAACATACG	ZNF162 - splicing factor 1
WI-16754		EST	GDB:678904	0.224	135	GTGGAGTGTGGATTTGG	134: TAGTGTCTGTATGATGATGTCGGG	
WI-16915		EST	GDB:678904	0.125	137	GTGGTGTCTGCTGCTGCTG	138: TCGCCGCAAAAGATGTAAAG	
SGC30608		EST	GDB:458405E	0.128	139	TCAGAGGAGAGCTGTTTATGA	140: CTCTGCTTGAAGTCTCTAGAGCC	Mitogen inducible gene (MIG-2)
WI-17683		EST	GDB:458334E	0.128	141	CAGCCGGGTTTATTTGAA	142: GATCGCAGGAGCCTAGGAC	
SGC31567		Gene	GDB:1222237	0.199	143	GCAATATGAACATTTATTTGGG	144: CTCTGAGGAGGAGGAGCAGAG	Human tel. interactive protein (TIP60)
SGC30658		EST	GDB:4578432	0.207	145	CTACACACACACAGAGC	146: CAGCGCAAGGCTGCTTC	Calcium activated neutral protease large subunit, muCAMP, calpain
SGC34590		EST	GDB:4584037	0.15	147	GTTGTCTGACCTCAGGCTGTC	148: TTTCTGCTCAACATCACTACTCC	
SGC33927		EST		0.13	149	GCGTGGGATATAGAGTCA	150: TACGTGCGCCAGAGAGTAGG	
WI-8071		EST	GDB:4582382	0.15	151	TATATATGCCAGTCAAGGCAT	152: AGCTTGCAGATGAGGCC	
WI-12334		EST	GDB:1222235	0.124	153	TGGTTTTAAAGCTTTATGAGAAA	154: TGTGTATGATACCTGTTCTCG	
WI-18402		EST	GDB:1222237	0.127	155	AATATTTAAAGAGAGAGAGGCA	156: TGGGTGTGAGCTGCTCTGA	
WI-18071		EST	GDB:4581874	0.13	157	GTTACAGAAAACATTTAGAGAT	158: TGAGCTTAGTTCCTCTCTCTG	HLA-B
WI-12858		EST	GDB:4584947	0.131	159	TGAAAGCAATTTATTTACCG	160: TGTGCGGCTGTTGGATTT	HLA-B
SGC33787		EST	GDB:4576606	0.209	161	TTGAAAGCAATTTATTTACCG	162: TGTCTCTCTCTCTCTCTCTG	
AFM343YB5		MSAT	GDB:4581106	0.181	165	AACACAGACAGTGTGCA	164: TGTGCAATTCCTCCCAAAA	
SGC33744		EST	GDB:4575826	0.181	165	AACACAGACAGTGTGCA	166: CCGTGAAGAGTGAATGCT	
SGC32272		EST	GDB:4581592	0.135	169	CAGAGAGTGTATGAGGCG	168: TATCTGTGTATGCTCAATGT	
SGC34148		EST	GDB:4583084	0.171	171	CAGATAAAGAGTCACTATGGCTCA	170: ACTGAGAGCTGTGCTCT	Human 1.1 kb mRNA upregulated in retinoid acid treated HL-60 neutrophilic cells
WI-18546		EST	GDB:4574598	0.133	173	TATGTATAGCATATGAGACCC	172: CAGTCTGCGCATTTGTCGG	
SGC31103		EST	GDB:4587285	0.173	175	CTATGCCAGAGTGAACAGG	174: TGCCAGTTAGGACAGTCA	
SGC30028		Gene	GDB:4580505	0.128	177	GCCAGCTTTATGATTAAGTCC	176: TGCCAGTTAGGAGCTATGTCG	Human pyruvate carboxylase precursor
WI-2875		STS	GDB:678546	0.125	178	CATCCCAACCATCACTGAT	178: CACTGGAAGTACAGTGTGG	
SGC38985		Gene	GDB:4571182	0.223	181	AGATACATTTGCAACAGTGG	180: GGGGAGTACTTACAGATTGA	
GCT16B07		STS	GDB:628245	0.137	183	GAGGTTTGTGCTGCTGATC	182: TGAAGGATATTATAGCTGGA	LAR-interacting protein 1b
D11S4270		STS					184: TGAGGCTGGGAGAGATATA	

【表 3 B】

TABLE 3: HBM STS Table

STS Name	Locus Name	Type	Gene Access	Size (kb)	SEQ ID No.	Forward Primer	Reverse Primer	Gene Name
WIL4504	D11S1974	EST	GDH:588142	0.174	185	CGTGTATGACGACACCG	186: CAGCTAAGTGTGACATGCCA	
SGC31049		EST	GDH:459093	0.15	187	TGTTAGTGTGTGACATGTTCT	188: CAGAGTGCAGTGGATGCG	NDUFV1
TIGR-A002117		EST	GDH:1222182	0.189	189	AGTACAGCAGCAGATAG	190: CATTCCACATGCGATAGC	amplaxin (EMS1)
WI-5998	D11S2382	EST	GDH:458683	0.1	191	CATGCTATGAGGTGTGCTACAGG	192: GATTTTGTGCTGCTGCTGCG	Nuclear mitotic apparatus protein 1, NUMA
WI-19987		EST	GDH:4575846	0.15	193	TACAGCAGCAGAGTTTCC	194: AGGTGTGTGTGCTGCTGCTGCG	
SGC31812		EST	GDH:4587688	0.101	195	CAGTGTATGCTATGCTGCTGAGG	196: TTGATTTTGTGCTGCTGCTGCG	
WI-13500		EST	GDH:4577883	0.18	197	CCGACGCTGCTGCTGCTGCTG	198: CAGGACGCTTGTGCTGCTGCTG	
CHLC GAAT1B01.P7933	D11S971	MSAT	GDH:684755	0.103	199	AGCAGCAGCAGCTGCTGCTG	200: ACAGGACATGCGCTGCTGCTG	LAR-interacting protein 1a mRNA
SGC35519		Gene	GDH:4577180	0.134	201	GATGGGTGCTGCTGCTGCTGCTG	202: ACATTTATTTGCTGCTGCTGCTG	Camiline palmitoyl transferase 1
WI-11974		EST	GDH:1222255	0.108	203	AGCCTTTTATGCTGCTGCTGCTG	204: ATGTGCTGCTGCTGCTGCTGCTG	Beta-adrenergic receptor kinase 1, ADRB1
WI-15244		Gene	GDH:4574740	0.108	205	TGACATCAAAATGCGGCA	206: TGCTGCTGCTGCTGCTGCTGCTG	
WI-17486		EST	GDH:4583336	0.131	207	TGTTATTTGCTGCTGCTGCTGCTG	208: GATGCTGCTGCTGCTGCTGCTG	FGF4
WI-9159	D11S4381	EST	GDH:678144	0.171	209	CCAGCAATTTATTTATGTTGCTGCG	210: GTAGATGCTGCTGCTGCTGCTG	
WI-4232		STS	GDH:1222250	0.175	211	CTTATTTGCTGCTGCTGCTGCTG	212: AGTGTGCTGCTGCTGCTGCTG	
SHGC-1167		EST	GDH:4586788	0.181	213	CAGTGTGCTGCTGCTGCTGCTG	214: CAGCTGCTGCTGCTGCTGCTG	Human DNA helicase gen (SNBP2)
WI-14303		EST	GDH:4576938	0.15	215	CTGCAATTTATTTATGCTGCTGCTG	216: TCGTGTGCTGCTGCTGCTGCTG	
WI-16597		EST	GDH:4585666	0.13	217	CAGGCGCTGCTGCTGCTGCTGCTG	218: AGCTGCTGCTGCTGCTGCTGCTG	
RC2B5ICAT1FORRC29S1CAT1	D11S970	MSAT	GDH:191084	0.15	219	ACACATGCTGCTGCTGCTGCTG	220: TGACCTGCTGCTGCTGCTGCTG	
UT979	D11S1296	MSAT	GDH:198523	0.382	221	CATTCCCGAGTTTCCAGAC	222: GTGCTGGGATTTAGAGGTGT	
12B11282	D11S1959E	EST	GDH:335218	0.07	223	GCAGAGAGTCTGCTGCTGCTG	224: CCACTGTAGAGACACACAC	
D11S468		STS		0.096	225	AGTGGGCGGACGCTGCTGCTG	226: CAGACAGATAGCCTGCTGCTG	
D11S668		STS		0.143	227	TCCCTCATCCCTGCTGCTGCTG	228: AGCCCTGCTGCTGCTGCTGCTG	
RH18048		Gene	GDH:179349	0.188	229	GATGCTTACCTTACCTGCTGCTG	230: AGATGCTGCTGCTGCTGCTGCTG	Aldelyde dehydrogenase (ALDH8)
IGHMBP2		Gene	GDH:4572853	0.699	231	TGCGAGAGCATGCTGCTGCTG	232: GAGAGCGCGGAGGCTGCTG	Human DNA helicase gen (SNBP2)
NUMA		Gene	GDH:4590244	0.277	233	CTGCTATGCAACAGATTTGCTGCTG	234: GGTGTGCTGCTGCTGCTGCTG	Nuclear mitotic apparatus protein 1, NUMA
KRN1		Gene	GDH:4590232	0.278	235	AGTGGGAGAGCTGCTGCTGCTGCTG	236: CAGTGTGCTGCTGCTGCTGCTG	High sulphur keratin, KRN
Cda1106	D11S2302E	EST	GDH:445887	0.081	237	CATATGATGCTGCTGCTGCTGCTG	238: CAAAGCAGATGCTGCTGCTG	
RH10753		Gene	GDH:4593586	0.194	239	GGAGTACAGCATGATTTGCTG	240: CATGCTGCTGCTGCTGCTGCTG	protein phosphatase 2A, PP2A
EMS1		Gene	GDH:4590116	0.84	241	CGCTGCTGCTGCTGCTGCTGCTG	242: GGGCATGAGGCTGCTGCTGCTG	Amplaxin
SHGC-11098		Gene	GDH:6737674	0.137	243	GCTCTATGCTGCTGCTGCTGCTG	244: CCGTGGCATGCTGCTGCTGCTG	Androgen Receptor
RH18051		Gene	GDH:4590093	0.392	245	CTGCTATGCTGCTGCTGCTGCTG	246: ATGCAAGTACGCTGCTGCTGCTG	51C protein, host cell polyphosphate phosphatase-like 1
Cda11c11		EST	GDH:4572858	0.195	247	TGAGTGTGCTGCTGCTGCTGCTG	248: CAGAGCTGCTGCTGCTGCTG	NOF1
12A91260	D11S2297E	EST	GDH:445699	0.1	249	ACAGAGCTGCTGCTGCTGCTGCTG	250: GATGAGCAGCTGCTGCTGCTG	
D11S1857E		EST	GDH:335210	0.247	251	TTTCCCATATGCTGCTGCTG	252: CATTCCCGAGCTGCTGCTGCTG	NDUFV1
D11S2245E		EST	GDH:445695	0.19	253	CTTGTGCTGCTGCTGCTGCTG	254: CAGGTGCTGCTGCTGCTGCTG	
D11S4136		MSAT	GDH:609546	0.18	255	GATGCTGCTGCTGCTGCTGCTG	256: TATGCTGCTGCTGCTGCTG	
D11S2288E		EST	GDH:614025	0.2	257	GAGCTGCTGCTGCTGCTGCTG	258: GCACTGCTGCTGCTGCTGCTG	
D11S351E		EST	GDH:4456842	0.158	259	AGGGAATGCTGCTGCTGCTGCTG	260: CCACTGCTGCTGCTGCTGCTG	
SHGC-1364		EST	GDH:4562765	0.137	261	AGTGGCAATGCTGCTGCTGCTGCTG	262: CCACTGCTGCTGCTGCTGCTG	
RH17410		EST	GDH:4571587	0.126	263	TGACATCTTGTGCTGCTGCTGCTG	264: AGTATCCCTGCTGCTGCTGCTG	
RH17414		EST	GDH:4571595	0.121	265	AGCTGCTGCTGCTGCTGCTGCTG	266: CAAAGTGTGCTGCTGCTGCTG	
RH17770		EST	GDH:4572301	0.267	267	CTGCTGCTGCTGCTGCTGCTGCTG	268: TGTGCTGCTGCTGCTGCTGCTG	
SEA		EST	GDH:4590169	0.13	269	CTCAGGCGGAGCTGCTGCTGCTG	270: GGAAGCTGCTGCTGCTGCTGCTG	S13 avian erythroblastosis oncogene homolog
RH10989		EST	GDH:4563460	0.107	271	AATGATGCTGCTGCTGCTGCTG	272: ACTGAGAACTGCTGCTGCTG	
TIGR-A006P20		EST	GDH:4587692	0.236	273	GACATGCTGCTGCTGCTGCTGCTG	274: GGTACAGTGTGCTGCTGCTG	
TIGR-A007D15		Gene	GDH:4588596	0.24	275	CTATGCTGCTGCTGCTGCTGCTG	276: ATGCTGCTGCTGCTGCTGCTG	Menin gene (MEN1)
TIGR-A008B14		EST	GDH:4588892	0.141	277	GTAATGAGAAACAGCAGCAATGA	278: CATTGAGTGTGCTGCTGCTG	
TIGR-A008K11		EST	GDH:4589094	0.203	279	AAGTAAACAAATGAGGAGGAC	280: CCGTCCCAAGCTGCTGCTG	
TIGR-A008P15		EST	GDH:4589862	0.182	281	AGTCTGCTGCTGCTGCTGCTGCTG	282: GAGGAGCTGCTGCTGCTGCTG	
TIGR-A008T11		EST	GDH:4589978	0.138	283	CATGCTGCTGCTGCTGCTGCTG	284: GAATGCTGCTGCTGCTGCTGCTG	
TIGR-A008U48		EST	GDH:45899364	0.107	285	GTGTGAGGAGAAAGGCACT	286: CTGCGAGTGTGCTGCTGCTG	
TIGR-A008X45		EST	GDH:45899364	0.242	287	CAAGTGTGCTGCTGCTGCTGCTG	288: CAGGCGCTGCTGCTGCTGCTG	
SHGC-11839	D11S4611	Gene	GDH:740339	0.131	289	TTTCTGCTGCTGCTGCTGCTG	290: GAGTGTGCTGCTGCTGCTGCTG	Feline receptor 2 (FBR2)
NIB1242	D11S4929E	EST	GDH:3888276	0.149	291	TTTCTGCTGCTGCTGCTGCTG	292: CCACTGCTGCTGCTGCTGCTG	cGMP-stimulated 3',5'-cyclic nucleotide phosphodiesterase PDE2A3 (PDE2A)
SHGC-13599	D22S1553	Gene	GDH:737558	0.147	293	CACCAAGAGTGTGCTGCTGCTG	294: ACTATGAGCTGCTGCTGCTGCTG	Macrophage Migration Inhibitory factor
SHGC-11867	D11S4331	Gene	GDH:674684	0.14	295	CTGCTGCTGCTGCTGCTGCTGCTG	296: TTGCTGCTGCTGCTGCTGCTGCTG	P2U Putnoceptor
SHGC-15349	D12S1214	EST	GDH:740819	0.141	297	TACAGCGCTGCTGCTGCTGCTG	298: ACATGCTGCTGCTGCTGCTGCTG	
Bda84a05	D11S2235E	EST	GDH:445682	0.095	299	CCTGAGCTGCTGCTGCTGCTG	300: CCGTGTGCTGCTGCTGCTGCTG	
Bda99a07	D11S2238E	EST	GDH:445674	0.091	301	TACAGCTGCTGCTGCTGCTGCTG	302: CAAATGAGCTGCTGCTGCTGCTG	

TABLE 3: HBM STS Table

STS Name	Locus Name	Type	GB Access	Size (kb)	SEQ ID No.	Forward Primer	Reverse Primer	Gene Name
to11								
D11S4284	EST	GDB:197840	0.173	303: CGGCACTTTCATCCAGGAC	303: GGGTGAGGAGTCCGACAAAT			Folate receptor2 (FBP2)
D11S4284	EST	GDB:632620	0.173	305: TTCCATTATTGACGACCTG	306: CTTAAGCCACTGTGTTTGG			
D11S4433	Gene	GDB:878143	0.324	307: CCTCTACACCTGCMAAGC	308: TGGAGAACCCACAGAGAC			Folate receptor3 (FBP3)
WI-14325	EST	GDB:457850	0.132	309: AAGCACAAGAGTACAGCACA	310: GTGTGTGGGCGACAGATTG			
WI-15192	EST	GDB:457860	0.15	311: AGAGACCTTCTCTCAGAC	312: AGAATCTCATCAGGGGCGC			
WI-17872	EST	GDB:4577492	0.141	313: AAAAAAGCAGTGTCTAAATTTGA	314: AATTGTTTGTGTTTGTGTTGAT			
SHFGC-30732	EST	GDB:4577830	0.105	315: GATTAGGGAGTACAGCTGCGG	316: GGGGACAAATATAGTATTACGG			
ts1SG4286	EST	GDB:4566057	0.123	317: CCACTCATATATGTTGTGACG	318: TGCGTGGCCAGAGAAGAG			DRES9
WI-13814	EST	GDB:4579260	0.15	319: TAAAGTGCCTAAATCATGAC	320: CCAAGAGGATGACCAAGTGG			Human VEGF related factor isoform VRF186 precursor (VRF)
WI-14222	Gene	GDB:581614	0.128	321: CCACTGTCTTATCAAGGTTGG	322: CTCTGTGCGCATGACATCTATACA			
WI-17872/30	EST	GDB:596509	0.118	323: CCACTGTCTTATCTTCCACAG	324: AGAAGGCCATATCAATGACAG			
SHFGC-31329	EST	GDB:4567386	0.15	325: AGCTTAAGTAGGACCAACCTAGG	326: GGATGCTGCTACTCCAGAAAG			
SGC-33858	EST	GDB:4578600	0.127	327: TGTTTATTTCACAGCTGCC	328: AGAGTGGGTGCGAGGCCAG			
WI-12181	EST	GDB:1222208	0.15	329: TTTTATTTTACAGCAATTGAGG	330: TGAAGAAAGTAAACACAGTCTATC			
WI-13701	EST	GDB:4574892	0.15	331: ATGAATCTTAAGCAGATCCCA	332: CACAGAGTCCGAGGGTCTGT			
WI-14069	EST	GDB:4584373	0.15	333: AAGGCCCTTATTATCTGTCTGTG	334: GCGTCAAGAGTGGTGGGT			
WI-14272	EST	GDB:4578525	0.125	335: GCTTCTAAGTCTTAGACTCAGCTGG	336: AGCCACACAGTGCAGCTACG			
WI-17347	EST	GDB:4578523	0.127	337: TTGTTTAATGATGCCAGA	338: TGGTCCCAATGCATCATCCG			
ts1SG1597	EST	GDB:4564419	0.127	339: ACACAGCATGCAGGGAGAG	340: ATCCCTGGTGTAGGTTGG			
ts1SG1938	EST	GDB:4564568	0.137	341: GATGGAAGTAGCTCTGCTGG	342: GGAAGGCCACGAGTACTACG			
ts1SG2759	EST	GDB:4565137	0.141	343: CCGGTGCTTGGMAAGT	344: GAGTGTCTGTGTTGGGGGA			
RH87	EST	GDB:4559690	0.17	345: TTAGAGGCATGAGTCACTAGCG	346: ACCAGTCTCACAGCCGCTTAA			
ts1SG4784	EST	GDB:4573113	0.141	347: CCGTGGCTGCACACACAG	348: GCTCATGACCTTTCAGGGC			
ts1SG4957	EST	GDB:4569055	0.171	349: AGATACGGCCAAACACCTGG	350: GTTGATATATAGCAGGGGCCCC			
ts1SG4974	EST	GDB:4569063	0.168	351: TCTGAGGTCAAGGGGCTGTCT	352: GCTCTGGAATAATGCTGTGTGA			
ts1SG8144	EST	GDB:4573137	0.17	353: ACTGATCCGCTCCGACCC	354: TCCTCTCAACTCTGTCACAGA			
ts1SG3975	EST	GDB:4569989	0.19	355: GATGATCAAGGCTCAACCTG	356: TGGAGAGTGGCTGTGAGCC			Human protein kinase (MLK-3)
SHFGC-10667	Gene	GDB:740246	0.377	357: CTCAGCTGGCTCAAGTCT	358: TCAAAAGTGGTGGTGACAGC			FOF3
SHFGC-11930	Gene	GDB:1231223	0.21	359: ATTCCAGAGCCAGCTCCAA	360: GTTAAATGTGTGATGACACAAAGC			
SHGC-32788	EST	GDB:4567878	0.125	361: GATCATGAGCTGTGACAG	362: TACATTTGAACATTTAAACCTGA			FK506-Binding Protein Precursor (FKBP-13)
FKBP2	EST	GDB:4565096	0.064	363: AACTGAGCTGACCAAGCTGGGA	364: TGGACAGTGTGCTGTGATGG			
WI-13116	EST	GDB:4565096	0.202	365: TTATCCGCTTATTTTCTCTTTG	366: TGGTCAAGCTGTGCTGTGATGG			4F2 Cell- Surface Antigen Heavy Chain (4F2HC)
MDU1	Gene	GDB:4590068	0.859	367: TCTCAAGGCTGTGCGAGTACC	368: GTGATCTCCAACTGTCTGAACC			
D11S5779	STS	GDB:196278	0.108	369: GTGGCTGCAGCTATGTAAGACAG	370: CAGCAGACAGATGCGCTAAGTCC			
D11S3866	STS	GDB:547681	0.135	371: CTGATGAGACACAGAACAG	372: TAAAGCCCTAATGCTGCTCC			
D11S3830	STS	GTC:547609	0.118	373: TATGTAAGAGGCTTACACAG	374: AAGATGTTTGGTATGACTGG			
D11S2439	STS	GDB:459728	0.123	375: GATGATTAAGTGTCTGTCG	376: GAGAGAGCTAAGCAGACACTT			Folate receptor2 (FBP2)
D11S1137	STS	GDB:197824	0.196	377: GAGGTGGTGGCCACCTGTA	378: XGAGGGGAGGACACACACTT			Collagen binding protein 2, collagen-2 gene (CBP2)
D11S4351	Gene	GDB:876735	0.141	379: GACGAGTCTGCGCCAGAG	380: TCCCGACGTGCTTCCGACATCC			Reilal outer segment membrane protein 1, ROM1
WI-9219	Gene	GDB:8767179	0.187	381: GAGGAGTGGACAGTGTGA	384: GCMAAGGCTTACATATTG			ZNF126
GTC- ZNF	Gene		0.172	383: TCAAAACACAGCTCTCCA	386: TCCCTGCTCGGGGAAC			
D11S4087	MSAT	GDB:603787	0.158	385: CCTACAGCACCCCCATT	388: GAGACGACAGACTATTGGCC			
D11S4162	MSAT	GDB:611241	0.263	387: GTTCTCCAGAGAGAGACAG	390: CCTGTGTAGAGATGCAGTTGG			
D11S4139	MSAT	GDB:609821	0.151	389: TATAGAGCTTACGCCGTGTGCG	392: GTAGAGCCAGCAATAATGTGAC			
AFM212a63	MSAT	GDB:199282	0.209	391: TCGTAGCCAGCTGCTCTACT	394: CTCGCCCTGTGTCGAGTATT			Serine/threonine kinase
WI-18813	EST		0.252	393: ATCCGTAGACAGAGAGGCCG	396: AGCAGATGCTGCTCTGGCAT			
WI-19549	EST		0.252	395: AACTTCTATTTGCCAAGGGA	398: AAGATGTGAATGGGATGGAGC			
WI-20154	EST		0.25	397: AACAGTTGTCTATCGGTAGCCA	400: CCGTATATCTCCGTTGCTGCC			DRES9
WI-22393	EST	GDB:4583084	0.142	398: GTCCAGGTGGCCTTTATTTT	402: GAATTCGAGGCTGTGCTGTG			Ultra high-sulphur keratin protein (KRN1)
ts1SG45579	EST	GDB:1223732	0.274	401: GCTCTAGTGGGAAAGCCTCAGG	404: CCAATGACTGTGGTGTACCA			
WI-21134	EST		0.273	403: GCTTGTGTCTCAAGGCCAA	406: GTGCTGTGTGTGGGGGAAG			Fas-associated death domain-containing protein, FADD
WI-21696	EST		0.235	405: GCTGCTGTGGATTTCTGTT	408: GGACTGGCCCTTTGGAACCT			
SHGC-7373	STS	GDB:740182	0.225	407: ATTAGAGCTATCCAGAACG	410: AGACGTCGGGAAAAGTGAAGA			
SHGC-36533	STS		0.125	411: ATTGCCAGTGGAAATGCTT	412: TTAATCTTTGTCACTTCCTGATT			
ARHX	Gene		0.242	413: lgcplcdlccgagcaggaagc	414: ggaagc			

【表 3 D】

TABLE 3: HBM STS Table

STS Name	Locus Name	Type	GBS Access	Size (kb)	SEQ ID NO. Forward Primer	SEQ ID NO. Reverse Primer	Gene Name
B235G10-HR		STS			421: CTGGACGTATATGCTGCG	422: AGAGCCGCGAGTACACGAT	
B247F23-HR		STS			423: ATACCTGTGACCTGCGACT	424: CCCTCTGTTTCTGTTT	
B337H24-HL		STS			425: CAGCTTTGAGGAGGAG	426: TAGGACGTAAAGTAGGAG	
B337L5-HL		STS			427: GCTCTGACGTGGGTAAAA	428: ACTCTCCAGAGCTGTGCG	
B382N10-HR		STS			429: CCGCTTCTGAGGCAAGAT	430: GACACACTGGGAGAGAC	
B1211-HR		STS			431: CCGTATGATGCCCATCTG	432: GATCAGCTGCAATGAGG	
B180D17-HR		STS			433: TTGAGTACAGGGGTGAC	434: GGCAGACTGAAGATGA	
B236E5-HR		STS			435: ACCGTCTGCTGCTGCTGG	436: TGGTTTCTGCTGCGGA	
B278E22-HR		STS			437: ATGACGAGCAACCATGT	438: GTACTGGGATACAGGGG	
B312F21-HR		STS			439: GGAAGAGTCTCTTTGGAT	440: TTGCAAGGATTCATGCTT	
B337H24-HR		STS			441: GGAATCTCTCTGAGG	442: ACCTTGCATGTGTGTTT	
B356H6-HR		STS			443: GCACTTTCTCTCTGCTCC	444: TGGTTTCTGCTGCTGG	
B148N18-HL		STS			445: ACAGCTCCAGAGAGGA	446: GCAGTCACTTGAACGAGA	
B172N12-HL		STS			447: AGGCATCAGCTTCTGCTT	448: GGTTCAGAGACGAGCC	
B172N12-HR		STS			449: GTGGTGTCTGCAATACC	450: GGAATCCCTTCTTCTCA	
B215N11-HR		STS			451: GACCATTTGTACGACG	452: GATGGGTGTGAATGAACA	
B236E5-HR		STS			453: CTGAGGTCTGTGCTATGG	454: GCTGTGAGTGTCTTGGCT	
B312B5-HR		STS			455: TACGAAAACGCGAGCTG	456: GCGACCAAGGAAAGATT	
B326G19-HR		STS			457: AAAGAGGAGGATCATGCG	458: TCATTTAGCAGAGGCGAG	
B326G19-HR		STS			459: CTGAGCATCCGATGAGAC	460: GTGCAAAATGAGAGCTT	
B32910-HL		STS			461: TCTACCCCTTACTGCGCG	462: TCGTCAACTGGGAATGA	
B32910-HR		STS			463: TTACAGAGGAGGAGGA	464: ATCTGCGGCGACTCAAG	
B368G19-HL		STS			465: GTGACGCGGCTTATCT	466: TGAGCATTAATTTCTATAGTG	
B368G19-HR		STS			467: GGAAGAGCAATTAATCCA	468: GGTGACAGCAATTTGTCAT	
B36F16-HL		STS			469: AGCAGCGTATTTCTATGG	470: GTACACGAGGAGGAGACA	
B36K11-HR		STS			471: TCGTCTGCAATATGAT	472: GGGGTGTGAGAATGAGAA	
B338D17-HR		STS			473: ATGGGGATTAATACCGG	474: AGTAGCAATTTGGGCTTT	
B268D23-HL		STS			475: CTGAGGAGAGAGGCTGG	476: CCGCTTACAAGGCAATGA	
B268D23-HR		STS			477: AGCATGCTTCTGAGGCTT	478: CCAAGTGTCTGGAAGCG	
B371E15-HR		STS			479: GGTCTCAGAGAGGCTTTA	480: ACATGCCACTCTCTCACTAA	
B312F21-HL		STS			481: AOTTTACAGAGATGGG	482: CACCCGAGAGGATTAAGA	
B338D17-HL		STS			483: TAGGCTCTGCACTCTTGG	484: ACCGAGGAGTCTCTGTC	
B369H19-HL		STS			485: TAAAGGGGTGAGATGAG	486: CTACCGGCTCTCTAGGCT	
B369H19-HR		STS			487: TGGGGCCAGATATCTCT	488: CTGGTGTCTGCTGCTGTT	
B444M11-HR		STS			489: AAGGAGAGGCTCACCAGG	490: CACAAATTCGATTTCCCA	
B269L23-HL		STS			491: TCAATAGGTGATCCACATTT	492: AAGTCCCAAGGGGTC	
B269L23-HR		STS			493: GAGTAGGGGATCTTTT	494: TGTGGAGATTCATTTGGC	
B364H4-HL		STS			495: GTCCCTGGGAAAGTGGAA	496: TCAAGCCCTCTCCCATAA	
B364H4-HR		STS			497: TCTTCCCTGCTACTGCG	498: TGGAGGCTCAGAGTGAATG	
B47333-HR		STS			499: GACAGGTATGTGTTGGG	500: AGCAGGCTGTTTTTGTGA	
B180D17-HL		STS			501: CTTCTTGAATCCGCTGTG	502: CACCGAGATCTCTTAGC	
B200E21-HL		STS			503: GCTGGGAGAGATCAGAA	504: GCTTTCCAGAGAGGACA	
B200E21-HR		STS			505: AGGCTGTCAAGTCACT	506: GAGGATGCTCAGGTGAT	
B141L15-HR		STS			507: TAGGGGATCTTTTCCA	508: GAGCAATTTGAAAGCCA	
B442P6-HR		STS			509: ATGGTCCAGCTGCTGTGT	510: ATAGGACCCCATCTCC	
B188N24-HR		STS			511: ACATTTGCTGTAGCCCA	512: GCAATCCAGAGGATTC	
GTC-ARRB1		Gene		0.087	513: ATGAGTTGGCAGCTGAAG	514: ATGAGGCTTCTGCTCC	Bala-arrestin-1
B509A5-HL		STS			515: GAGGAGAGATCCCAAGCG	516: TCTGTGGGCTACTGAAC	
B369H19-HR		STS			517: CTGAGCTTTTGGCACTGT	518: CTGCTAGGTGACAGAAAT	
B117N18-HL		STS			519: TGTATGATCTGGAGGCTGT	520: ACAGCTGGGTGAGGAAAT	
B141L15-HL		STS			521: GCAGGGGAGCTGATAA	522: TTTTGGTCTGCTACGATG	
B21K22-HL		STS			523: AATATTGTGACAGCTCG	524: TTATATTAAAGTGGCTTGT	
B21K22-HR		STS			525: GTGCAAGCCCAAGAT	526: AGGAAATCCAGAGGAG	
B223E5-HL		STS			527: CCATGTAATTCGATCTTGG	528: TGTGGGTCCAGTCTGCTA	
B278E22-HL		STS			529: AGATTTGGGGAGTCAGG	530: GCGCTCAAGCAATTCG	
B444M11-HL		STS			531: CAAGCCCAAGATGATCA	532: GAATCATCCAACTCAGA	
B543O18-HR		STS			533: AGCCTCCAGGTGATACC	534: GAAGAGCATGCTCAGCAG	
B17N18-HR		STS			535: ATGCTTACAGCATTTCCG	536: TCAATCGGTGATAGGTTA	
		STS			537: GTCGATGGTTTCCAGAA	538: TTTTATGGGAAATTCAGCG	

10

20

30

40

【表 3 E】

TABLE 3: HBM STS Table

STS Name	Locus Name	Type	GDB Access	Size (kb)	SEQ ID NO: Forward Primer	SEQ ID NO: Reverse Primer	Gene Name
B543O19-HL		STS			538: TTGGAAAGAACAAATGT	540: GGCTAGTCTTCTGAGCC	
B442P6-HL		STS			541: GGTAAATGCCGCTGATC	542: GGTCTTACAGCTGAGGA	
B367H4-HR		STS			543: TCAAGCTTGGTCTTCGA	544: GTAGCCGAGAGTGTCT	
B250E21-HR		STS			545: CCGTGGCTGGATAGAT	546: GTTCCGCTGGCTATGT	
B250E24-HL		STS			547: GGCAGCTACTTCTGACCA	548: GGTGCTTCTTACAGGA	
B248C18-HR		STS			549: ACCGAGGCTGGTGTGT	550: AGTGAGTAAATATACGTCCT	
B248C16-HL		STS			551: CATGCATTTGGCTCACC	552: TGTCTTTAGAGCTGTAGC	
B160D8-HR		STS			553: TCAAGCTTCAGAGACGA	554: GGAGTACATCGAGGAGC	
B539L7-HR		STS			555: TGTGCTTTTAAATCCAGA	556: CTCCCTTACTTACTTGAATG	
B47303-HL		STS			557: TCTTCGCGAGGATGT	558: TTATGTCCCTGAGCAG	
AFMa190a49	D154095	STS	GDB:606064	0.193	559: TCCCTGGCTATCTGATATC	560: GTTAGTGGGTCCACG	
ARRB1(2)		STS			561: CGAGAGCGGAGATAGATACA	562: CATCTCCATGCTTTTCAGT	
ARRB1(1)		STS			563: AGTCCGAGAGACGAGACGC	564: GTTGCTATCTCCATGCCCT	
P102F3S		STS			565: GAGCGTGAGAGGTGAGGAG	566: AACCACTCCAGACGACGC	
N172A		STS	GDB:6054145		567: CTGACGACTACCTGTATGACCTG	568: CTAGTACTACTCTACAGGGCCC	
NC6A		STS	GDB:6054146	0.208	569: GAGCAATTCATATCTTAACTG	570: CCAGTCCAGTGGAGGCAATC	
GC111-44A		STS	GDB:6054147	0.23	571: CTCTCTGGCCAGTCTGAGC	572: GGTACTCTTTGAATCCGAGC	
CN187-2A		STS	GDB:6054148	0.239	573: TGAGGATGAATGAGCATAAG	574: TTGTGTGCTCATTTGAGTAGGC	
CC111-524B		STS	GDB:6054150	0.271	575: AGGGGAAGGATGTGCTGG	576: TTGCGCTGAGCGGAGCAGTGT	
P117F3T		STS	GDB:6054151	0.221	577: ATTGAAGTCTGCAAGAGATGCTGAGC	578: AGAAGTGCACATATCTTTTGGGGACAC	
ARRB1(3)		STS			579: TGTATTTGAGACTTTGCTCG	580: CGGTACCATCTCTCTCTTCC	
B21511-HL		STS		0.122	581: TTTTGGCTGATGATGCCCC	582: GGTGACAGAGCAAGACTGCC	
B377G1-HR		STS			583: TTGCTCAAGTCTCTGCGC	584: ACCTTGTCTTGAAGGAGC	
B377G1-HL		STS			585: GTGGCTATTGACAGC	586: GGGCAATTACTGACTTGC	
B29216-HR		STS			587: GTTGCTGAGTGTGACAGG	588: TGGAAATTGTGTCTCTGG	
B10A18-HL		STS			589: CCAATTCACCTGGATGTT	590: ATGGGCTGTGTTTCTCAA	
B527D12-HL		STS			591: CTCCCTATCCCTGGACTT	592: AGTTTGTCTCTAGTGGCC	
B372J11-HR		STS			593: CACACGCTGTGACATCCAT	594: GGATAGTGCACACCA	
B372J11-HL		STS			595: TGGGTGGTACTATTGTCCTCAT	596: AGTCCAGGCCCTTACCAG	
B37E17-HR(GS)		STS			597: GGCCAGTATATCCCTGTGT	598: TTCCAGTGGAGAGACACG	
B37E17-HL(GS)		STS			599: ACAGTGACAGTAGGAGCGG	600: TGCCAGGATGGAGTACCA	
B34F22-HR(GS)		STS			601: CGTGTGGCAGACATATACCG	602: ACACCAAGATGAGGCCAC	
B34F22-HL(GS)		STS			603: TGTGTGTACAGATGCCCA	604: TGACCGAGAGAGCTACCAAG	
B448F22-HR1		STS			605: GCAAGGCTCGAGCTGACTAG	606: GCTGTGAGTTCCTTTTACGC	
B448F22-HL1		STS			607: ACAGTGGGGAACAAAGACAG	608: TACAGGGCACCTCCAGTAG	
B82A4-HR2		STS			609: TCTGTGTAAAGTTCCTCCGC	610: TGTCTCAACCTCCCTCTGC	
B648F22-HL		STS			611: ACATATTTCTTCCCGAGC	612: CAGTCCGAGCCCATGAGAC	
B82L11-HL(GS)		STS			613: CTCTCTGCAATGGAGATC	614: AGACCTGGAGCCAGTCTGTG	
B86J13-HL(GS)		STS			615: GGGAGACGAGCTCACAGAT	616: TGAATTTGGGAAGATGGTGA	
T44A24-HL		STS			617: CAGGCATCTCTATGTGCCA	618: GGGAGGCACAGTCTTTCA	
B82L11-HR(GS)		STS			619: ACTCGTGGCAGTGAATGTG	620: CTTTCTTACGAGTACGCA	
B86J13-HR(GS)		STS			621: GGTGTGCTGAGCTGCTGAT	622: TGGGTCTCTGCTGCTGAGT	
B82L11-HL2(GS)		STS			623: TCACGTAGTTCAGCTTCCG	624: AGACCTGGAGCCAGTCTGTG	
B82L11-HL3(GS)		STS			625: CTCTCTGCAATGGAGATC	626: ATTTCAGGAGCCCTGGAGCC	

【0156】

公に利用可能なゲノム配列、または配列由来のB A Sインサート末端のいずれかから新規のSTSを開発した。クロスマッチ(P.Green, Washington大学)を用いて自動的にベクターおよび反復配列マスキング、および続いてプライマー3 (Rozen, Skaletsky(1996, 1997

10

20

30

40

50

))を用いるプライマーピッキングを実施するスクリプトを使用してプライマーを選択した。プライマー3はwww.genome.wi.mit.edu/genome_software/other/primer3.html <http://www.genome.wi.mit.edu/genome_software/other/primaer3.html>で利用可能である。

【0157】

各々のプライマー対に関するポリメラーゼ連鎖反応 (PCR) 条件を最初に $MgCl_2$ 濃度に関して最適化した。標準的なバッファーは 10 mM のトリス HCl ($pH\ 8.3$)、 50 mM の KCl 、 $MgCl_2$ 、 0.2 mM の各 $dNTP$ 、 $0.2\text{ }\mu\text{M}$ の各プライマー、 $2.7\text{ ng}/\mu\text{l}$ のヒト DNA 、AmpliTaq(Perkin Elmer) 0.25 単位、および $MgCl_2$ 濃度 1.0 mM 、 1.5 mM 、 2.0 mM または 2.4 mM であった。サイクリング条件には 94°C で2分間の最初の変性、続いて 94°C で15秒間、 55°C で25秒間、および 72°C で25秒間の40サイクル、続いて 72°C で3分間の最後の伸長などがある。最適化の最初のラウンドからの結果に依存して、必要な場合さらに条件を最適化した。可変値には、 58°C または 60°C までのアニーリング温度の上昇、42までのサイクル数の増加、および30秒までのアニーリングおよび伸長時間、並びにAmpliTaqGold(Perkin Elmer)の使用などがある。

【0158】

Research Geneticsから購入した全ヒト BAC ライブラリーからの DNA プールのPCR基盤のスクリーニングにより、目的のSTSマーカーを含有する BAC クローン(Kimら、Genomics 32:213-218(1996)、Shizuyaら、Proc.Natl.Acad.Sci.USA 89:8794-8797(1992))を入手した。ヒト DNA の9つのゲノム等価物に相当するライブラリープレート1-596に由来する DNA プールを使用した。最初のスクリーニング方法はスーパープール、すなわち384ウェルライブラリープレートからの全 BAC クローンに由来する DNA の混合物に対する個々のマーカーのPCR反応を含んだ。各々の陽性スーパープールに関して、プレート(8)、列(16)およびコラム(24)のプールをスクリーニングして独特なライブラリーアドレスを同定した。PCR産物を $1\times TBE$ 中 $0.5\text{ }\mu\text{g}/\text{ml}$ 臭化エチジウムを含有する2%アガロースゲル(Sigma)で、150ボルトで45分間電気泳動した。使用した電気泳動ユニットはOwl Scientific ProductsのA3-1システムモデルであった。典型的には、ゲルには50ウェル/段のレーンが10段含まれていた。分子量マーカー(100bpラダー、Life Technologies、Bethesda、メリーランド州)をゲルの双方の末端で負荷した。ゲルの画像をKodak DC40 CCDカメラで捕捉し、Kodak 1Dソフトウェアで処理した。ゲルデータをタブ区切りしたテキストファイルとしてエクスポートした；スクリーニングしたライブラリー、ゲルイメージファイルおよびスクリーニングされたマーカーについての情報をファイルの名前に含めた。カスタマイズされたPerlスクリプトを用いて、データ保存および分析のためのFilemaker(商標)PRO(Claris Corp.)データベースにこれらのデータを自動的にインポートした。不完全なまたは不明瞭なクローンアドレス情報が得られた場合、独特な完全なライブラリーアドレスを回復するために、別の実験を実施した。

【0159】

ライブラリーに由来するクローン BAC 培養物の回収には、 $12.5\text{ }\mu\text{g}/\text{ml}$ クロラムフェニコール(Sigma)を含有するLB寒天(Maniatisら、Molecular Cloning: A Laboratory Manual、Cold Spring Harbor Laboratory、Cold Spring Harbor、ニューヨーク州(1982))上のライブラリーウェルからサンプルを画線することなどがある。2つの個々のコロニーおよび最初の画線の四分円の部分を適当なSTSマーカーでコロニーPCRにより確認のために試験した。 $12.5\text{ }\mu\text{g}/\text{ml}$ クロラムフェニコールおよび15%グリセロールを含有するLBブロス中 -70°C で陽性クローンを保存した。

【0160】

いくつかの異なる型の DNA 調製法を BAC DNA の単離に用いた。以下に列挙した手作業によるアルカリ性溶解ミニプレッププロトコル(Maniatisら、Molecular Cloning: A Laboratory Manual、Cold Spring Harbor Laboratory、Cold Spring Harbor、ニューヨーク州(1982))をたいていの適用、すなわち制限マッピング、CHEFゲル分析、FIS

Hマッピングに用いて成功したが、末端シーケンシングでの再現性は成功しなかった。末端シーケンシングの目的で、特にBAC DNA調製にはAutogenおよびQiagenプロトコルを使用した。

【0161】

50 ml コニカルチューブ中、12.5 $\mu\text{g}/\text{ml}$ クロラムフェニコールを含有する Terrific Broth 15 ml 中、37℃で20時間、300 rpmで振盪しながら細菌を成長させた。Sorvall RT6000D中300 rpm ($\sim 1800\text{g}$)で、4℃で15分間培養物を遠心した。次いで上澄をできるだけ完全に吸引した。この工程で細胞ペレットを2週間まで-20℃で凍結する場合もあった。次いでペレットをボルテックスして細胞をホモジナイズし、凝集を最低限にした。P1溶液(50 mM グルコース、15 mM トリスHCl、pH 8、10 mM EDTA、および100 $\mu\text{g}/\text{ml}$ RNアーゼA) 250 μl を加え、混合物をピペットで上下させて混合した。次いで混合物を2 ml エッペンドルフチューブに移した。次いでP2溶液(0.2 N NaOH、1% SDS)を加え、混合物を穏やかに混合し、室温で5分間インキュベートした。P3溶液(3 N KOAc、pH 5.5) 350 μl を加え、次いで白色沈殿を形成するまで混合物を穏やかに混合した。溶液を氷上で5分間インキュベートし、次いでマイクロヒュージ中4℃で10分間遠心した。上澄を注意深く新しい2 ml エッペンドルフチューブに移し(白色沈殿を避ける)、イソプロパノール0.9 mlを添加し、溶液を混合し、氷上で5分間放置した。サンプルを10分間遠心し、上澄を注意深く除去した。ペレットを70%エタノールで洗浄し、5分間空気乾燥した。ペレットをTE8 (10 mM トリスHCl、pH 8.0、1.0 mM EDTA) 200 μl に再懸濁し、RNアーゼA(Boehringer Mannheim)を100 $\mu\text{g}/\text{ml}$ まで加えた。サンプルを37℃で30分間インキュベートし、次いでC₂H₃O₂Na $\cdot$ 3H₂Oを0.5 Mまで、およびエタノール2容量を添加することにより沈殿させた。サンプルを10分間遠心し、ペレットを70%エタノールで洗浄し、続いて空気乾燥し、TE8 50 μl に溶解した。このDNA調製に関する典型的な収量は細菌培養15 mlあたり3から5 μg であった。10から15 μl をHindIII制限分析に使用した；5 μl をNotI消化およびCHEFゲル電気泳動によるクローンインサートサイジングに使用した。

【0162】

50 ml コニカルチューブ中、12.5 $\mu\text{g}/\text{ml}$ クロラムフェニコールを含有する 2 X LBブロス 15 mlにBACを接種した。4つのチューブを各クローンに関して接種した。培養物を37℃で一晩(~ 16 時間)激しく振盪しながら($> 300\text{ rpm}$)成長させた。BAC DNA単離のための標準的な条件はAutogen740製造者により推奨されるものに従った。培養サンプル3 mlをAutogenチューブに全体で60 mlまたはクローンあたり20チューブに配置した。Autogenプロトコルの一部として、最終的にサンプルを15秒間振盪しながらTE8 100 μl に溶解させた。Autogenプロトコルを終了した後、DNA溶液を各別個のチューブから写し、2 ml エッペンドルフチューブにプールした。多量の細片を含むチューブ(細片をペレット化する工程から持ち越し)を避けた。次いでチューブをTE8 0.5 mlで連続的にすすぎ、この溶液をプールした材料に加えた。DNA溶液を4℃で保存した；凝集は-20℃での凍結時に生じる傾向があった。このDNAを直接制限マッピング、CHEFゲル分析またはFISHマッピングに使用するか、または以下に記載するように末端シーケンシング反応において使用するためにさらに精製した。

【0163】

DNA溶液の容量をTE8で2 mlに調整し、次いでサンプルを穏やかに混合し、65℃で10分間加熱した。次いでDNA溶液を4℃で5分間遠心し、上澄を15 ml コニカルチューブに移した。次いでNaCl濃度を0.75 M (5 M NaCl $\sim$ 0.3 mlをサンプル2 mlに)に調整した。次いで全容量をQiagenカラム平衡バッファー(バッファーQBT)で6 mlに調整した。次いでDNAを含有する上澄をカラムに適用し、重力流により進入させた。カラムをQiagenバッファーQC 10 mlで2回洗浄した。次いで結

合DNAを、65℃を維持したバッファーQFの1mlアリコートで別個に4回溶出した。イソプロパノール1.7容量(〜2.8ml)でDNAを沈殿させた。次いで各サンプルを4個の別個の2.2mlエッペンドルフチューブに移し、室温で2時間かまたは一晩インキュベートした。サンプルをマイクロヒュッジ中4℃で10分間遠心した。上澄を注意深く除去し、70%エタノール1mlを添加した。サンプルを再度遠心し、この段階でDNAペレットがしばしばほぐれるので、上澄を注意深く除去した。サンプルを再度遠心して、マイクロピペットチップで除去された残りの液体を濃縮した。次いでDNAペレットをデシケーター中10分間乾燥させた。滅菌蒸留および脱イオン化H₂O 20μlを各チューブに加え、次いでこれを4℃で一晩放置した。各クローンに関して4つの20μlサンプルをプールし、チューブを別の滅菌蒸留および脱イオン化H₂O 20μlで、最終容量100mlですすいだ。次いでサンプルを65℃で5分間加熱し、次に穏やかに混合した。NotI消化および非切断λDNAとの比較により評価されるように、典型的な収量は培養物60mlあたり2から5μgであった。

10

【0164】

12.5μg/mlクロラムフェニコールを含有するLBブロス 3mlをオートクレーブAutogenチューブに分配した。各クローンに関して1本のチューブを用いた。接種用にはグリセロールストックを-70℃保存から取り出し、ドライアイス上に置いた。グリセロールストックの小部分を元来のチューブから滅菌トウスポックで取り出し、Autogenチューブに移した；トウスポックは廃棄する少なくとも2分前までAutogenチューブ中に放置した。接種後チューブをテープで被覆してきつく密閉されていることを確認した。全サンプルを接種したとき、チューブ単位をAutogenラックホルダーに移し、250rpmで、37℃で16から17時間、回転シェーカーに配置した。成長させた後、製造者により定義されるようなBAC DNA調製に関する標準的な条件を用いてAutogenをプログラムした。プログラムの一部として、サンプルはTE8に溶解せず、DNAペレットを乾燥させたままにした。プログラムを完了したとき、チューブを出力トレイから取り出し、滅菌蒸留および脱イオン化H₂O 30μlをチューブの底に直接添加した。次いでチューブを穏やかに2から5秒間振盪し、次にパラフィルムで被覆し、室温で1から3時間インキュベートした。次いでDNAサンプルをエッペンドルフチューブに移し、シーケンシングに直接用いるか、または後で使用するために4℃で保存した。

20

【0165】

30

VI. 物理マッピングのためのBACクローン特徴付け

制限フラグメントの大きさの分析のために、手作業によるアルカリ性溶解またはAutogenプロトコルのいずれかにより調製したDNAサンプルをHindIIIで消化した。このデータを用いてクローン間の重複の程度を比較した。典型的には1から2μを各反応に使用した。反応混合物には、最終容量25μl中、1xバッファー2(New England Biolabs)、0.1mg/mlウシ血清アルブミン(New England Biolabs)、50μg/mlRNアーゼA(Boehringer Mannheim)、およびHindIII(New England Biolabs)20単位が含まれた。消化物を37℃で4から6時間インキュベートした。CHEFゲル分析によりインサートの大きさを推定するためにBAC DNAをまたNotIでも消化した(以下を参照)。反応条件はNotI 20単位を使用した以外はHindIIIの条件と同一であった。ブROMフェノール・ブルーおよびキシレンシアノールを含有する6xFicoll負荷バッファー6μlを電気泳動の前に添加した。

40

【0166】

0.5μg/ml臭化エチジウムを含有する1xTBE中0.6%アガロース(Seakem、FMC Bioproducts)でHindIII消化を分析した。モデルA4電気泳動ユニット(Owl Scientific)でゲル(20cm x 25cm)を50ボルトで20から24時間電気泳動した。分子量サイズマーカーには未消化λDNA、HindIII消化λDNAおよびHaeIII消化X174 DNAなどがあつた。分子量マーカーをゲルに負荷する前に65℃で2分間加熱した。画像をKodak DC40 CCDカメラで捕捉し、Kodak 1Dソフトウェアで分析した。

50

【0167】

CHEF DRII (BioRad) 電気泳動ユニットで製造者の推奨に従ってNotI消化を分析した。簡単には、1%アガロースゲル (BioRadパルスフィールドグレード) を0.5XTBEで調製し、電気泳動ユニット中14℃で30分間平衡にし、6ボルト/cmで14時間、循環しながら電気泳動した。切り換え時間は10秒から20秒まで傾斜させた。電気泳動の後ゲルを0.5μg/ml 臭化エチジウムで染色した。分子量マーカーには未消化λDNA、HindIII消化λDNA、λラダーPFGラダー、および低値域PFGマーカー (全てNew England Biolabsから) などがあつた。

【0168】

手作業によるアルカリ性溶解またはAutogenプロトコルのいずれかにより調製したBAC DNAを、Bioprime標識キット (BioRad) を用いて、製造者の推奨にわずかな修飾を加えたものに従って、FISH分析のために標識した。DNAおよそ200ngを各反応物50μlに使用した。3μlを2%アガロースゲル上で分析し、標識の程度を決定した。インサイチュ・ハイブリダイゼーションの前にSephadex G50スピンカラムを用いて反応物を精製した。分裂中期のFISHを記載されるように (Maら、Cytogenet. Cell Genet. 74:266-271(1996)) 実施した。

【0169】

VII. BAC末端シーケンシング

BACインサート末端のシーケンシングは前記した2つの方法のいずれかにより調製されたDNAを利用した。DYEnamic Energy TransferプライマーおよびAmershamのDynamic Directサイクルシーケンシングキットをシーケンシング反応に使用した。M13~40正方向シーケンシングプライマーを含む既製シーケンシングミックス (カタログ番号US79730) をT7 BACベクター終端に用いた; 既製シーケンシングミックス (カタログ番号US79530) をSP6 BACベクター終端のためのM13-28逆方向シーケンシングプライマー (カタログ番号US79339) と混合した。シーケンシング反応ミックスには4つの蛍光標識色素プライマーの1つ、4つのジデオキシ終止ミックスの1つ、dNTP、反応バッファー、およびThermosequenaseが含まれた。各BAC DNAサンプルに関して、BAC DNAサンプル 3μlを4PCR ストリップチューブに等分した。次いで4つの色素プライマー/終止ミックスの組み合わせの1つの2μlを4つのチューブの各々に加えた。次いでチューブを密閉し、そしてPCRの前に短時間遠心した。熱循環条件には95℃で1分間の変性、45℃で15秒間のアニーリング、および70℃で1分間の伸長の、全部で35サイクルが含まれる。循環の後、プレートを短時間遠心して全液体をチューブの底に集めた。次いで滅菌蒸留および脱イオン化H₂O 5μlを各チューブに加え、プレートを密閉し、そして再度短時間遠心した。次いで各BACに関する4つのサンプルを一緒にプールした。次いで7.5M NH₄OAc 1.5μlおよび-20℃の100%エタノール 100μlを各チューブに加えることによりDNAを沈殿させた。ピペティングを上下に一度行ってサンプルを混合した。次いでプレートを密閉し、そして氷上で10分間インキュベートした。テーブルトップHaraeus遠心でプレートを4℃、4000rpm (3,290g) で30分間遠心してDNAを回収した。上澄を除去し、そして過剰の液体をペーパータオルで吸い取った。-20℃の70%エタノール 100μlを各チューブに加え、4℃で4000rpm (3,290g) で10分間再遠心することによりペレットを洗浄した。上澄を除去し、そして再度過剰の液体をペーパータオルで吸い取った。残りのわずかな液体は、プレートを上下逆さまにしてペーパータオルの上に置き、そして遠心が800rpmにちょうど達するまで遠心することにより除去した。次いでサンプルを室温で30分間空気乾燥した。チューブにキャップをして、そして電気泳動まで-20℃で乾燥保存した。電気泳動の直前に、DNAをAmersham負荷色素 1.5μlに溶解した。次いでプレートを密閉し、そして2000rpm (825g) で遠心した。次いでプレートをプレートシェーカーで1~2分間ボルテックスした。次いでサンプルを2000rpm (825g) で短時間遠心した。次いでサンプルを65℃で2分間加熱し、そして即座に氷上に置いた。ABI 3

10

20

30

40

50

77 蛍光シーケンサーで、製造者の推奨に従って標準的なゲル電気泳動を実施した。

【0170】

V I I I . H B A B A C D N A のサブクローニングおよびシーケンシング

L R P 5 遺伝子領域の物理地図によりその内部に L R P 5 遺伝子および H B M 遺伝子を含む B A C クローンのセットが提供される。該領域からのいくつかの B A C の D N A シーケンシングが完了した。D N A 配列データは、当業者が L R P 5 遺伝子および H B M 遺伝子を同定するか、または（複数の）遺伝子を同定するためのプローブを調製するか、または（複数の）遺伝子を同定するための D N A 配列多型性を同定するために用いることができるデータを含む独特な試薬である。

【0171】

B A C D N A の Qiagen 精製 (Qiagen, Inc., 製品文献に記載されているように) または標準的なプラスミド D N A のアルカリ性溶解/塩化セシウム調製の変法である手作業による精製の 2 つのプロトコルのいずれか 1 つに従って、B A C D N A を単離した (例えば、Ausubel ら、Current Protocols in Molecular Biology, John Wiley & Sons (1997))。簡単には、手作業によるプロトコルでは、細胞をペレット化し、G T E (50 mM グルコース、25 mM トリス H C l (p H 8)、10 mM E D T A) および リゾチーム (50 mg / ml 溶液)、続いて N a O H / S D S (1 % S D S / 0.2 N N a O H) そして次に 3 M K O A c (p H 4.5 から 4.8) の氷冷溶液に再懸濁した。濾過した上澄に R N アーゼ A を加えて、続いてプロテイナーゼ K および 20 % S D S を加えた。次いで D N A を イソプロパノール で沈殿させ、乾燥し、そして T E (10 mM トリス、1 mM E D T A (p H 8.0)) に再懸濁した。B A C D N A を塩化セシウム密度勾配遠心 (Ausubel ら、Current Protocols in Molecular Biology, John Wiley & Sons (1997)) によりさらに精製した。

【0172】

単離した後、H P L C を用いて (Hengen, Trends in Biochem.Sci.22:273-274 (1997)) B A C D N A を 2000 から 3000 b p のインサートの大きさまで水力学的に剪断した。剪断後、D N A を濃縮し、そして標準的な 1 % アガロースゲル上で分離した。おおよその大きさに相当する単一の分画をゲルから切除し、そして電気溶出により精製した (Sambrook ら、Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring, ニューヨーク州 (1989))。

【0173】

次いで精製された D N A フラグメントを T 4 D N A ポリメラーゼを用いて平滑末端化した。次いで平滑末端 D N A を独特な B s t X I リンカーアダプター (配列番号 627-628) (100 から 1000 倍過剰のモル濃度の 5' G T C T T C A C C A C G G G G および 5' G T G G T G A A G A C) にライゲートした。これらのリンカーは B s t X I 切断 p M P X ベクター (発明者が構築) に相補的であったが、オーバーハングは自己相補的ではなかった。従って、リンカーはコンカテマー化されず、また切断ベクターは容易に自身に再ライゲートしない。リンカー適応インサートを 1 % アガロースゲル上で組み込まれていないインサートから分離し、GeneClean (BI0101, Inc.) を用いて精製した。次いでリンカー適応インサートを修飾 pBlueScript ベクターにライゲートして「ショットガン」サブクローンライブラリーを構築した。ベクターはアダプターダイマーがクローン化されるときに読み枠内になるクローニング部位で読み枠外 l a c Z 遺伝子を含み、それによりこれらを青色により避けることが可能になる。

【0174】

続く全ての工程は A B I 377 自動 D N A シーケンシング法によるシーケンシングに基づいた。プロトコルの主要な変更のみを強調して提示する。簡単には、次いでライブラリーを D H 5 α コンピテント細胞に形質転換した (Life Technologies, Bethesda, メリランド州、D H 5 α 形質転換プロトコル)。アンピシリンおよび I P T G / X g a l を含む抗生物質プレートに播種することによりこれを評価した。プレートを 37 °C で一晩インキュベートした。次いで成功した形質転換体をクローンの播種およびシーケンシ

10

20

30

40

50

ングのための採取に用いた。培養物を 37℃で一晩成長させた。シリカビーズ DNA 調製法 (Ngら、Nucl. Acids. Res. 24:5045-5047(1996)) を用いて DNA を精製した。この様式でクローンあたり DNA 25 µg が得られた。

【0175】

次いで ABI 色素ターミネーター化学を用いてこれらの精製した DNA サンプルをシーケンシングした。ABI 377 機械上で ABI 色素ターミネーター配列解読を行い、データを直接 UNIX 機械に移し、続いてゲルのレーントラッキングを行った。全ての読みを PHRAP (P. Green、Abstracts of DOE Human Genome Program Contractor-Grantee Workshop V、1996年1月、157頁)を用いてデフォルトパラメーターおよびクオリティスコアでアセンブルを行った。最初のアセンブリは6倍被覆で行い、平均8から15コンティグが得られた。最初のアセンブリの後、失くした相手（一本鎖の読みのみを提供したクローンからの配列）を同定し、ABI 技術でシーケンシングしてさらなる重複コンティグの同定を可能にした。ウォーキングのためのプライマーを Genome Therapeutics プログラム Pick プライマーを用いてクローンの末端近くで選択してギャップの閉鎖を促進した。これらのウォークを選択したクローンおよびプライマーを用いてシーケンシングした。データを PHRAP で配列コンティグに再アセンブルした。

10

【0176】

IX. コンピューター法による遺伝子同定

BAC 配列のコンティグへのアセンブリの後、コンティグをコンピューター解析に供し、コード化領域および公知の遺伝子に類似する DNA 配列を担持する領域を同定した。このプロトコルには以下の工程が含まれた。

20

【0177】

1. コンティグのデギャップ：配列コンティグはしばしば、個々の ABI 配列の読みが挿入されているかまたは欠失している位置を表す記号（ピリオドの記号で印が付けられている）を含んでいる。コンティグの自動コンピューター解析の前に、ピリオドを除去した。元来のデータは後で参照するために保持した。

【0178】

2. クロスマッチプログラムを用いて配列内の BAC ベクター配列を「マスキング」した (Phil Green、<<http://chimera.biotech.washington.edu/UWGC>>)。前記で詳記したショットガンライブラリー構築によりショットガンライブラリーにいくつかの BAC ベクターが残るので、このプログラムを用いて BAC コンティグの配列を BAC ベクターと比較し、次の工程の前にいずれかのベクター配列をマスキングした。マスキングされた配列を配列ファイルで「X」と印を付け、続く分析の間不活性のままであった。

30

【0179】

3. BAC 配列を夾雑する大腸菌配列を、BAC コンティグを全大腸菌 DNA 配列と比較することによりマスキングした。

【0180】

4. ヒトゲノムに一般的であることが解っている反復エレメントを、クロスマッチを用いてマスキングした。このクロスマッチの実行において、BAC 配列をヒト反復エレメントのデータベースと比較した (Jerzy Jerka, Genetic Information Research Institute, Palo Alto、カリフォルニア州)。マスキングした反復を X で印を付け、そして続く分析の間不活性のままであった。

40

【0181】

5. MZEF コンピュータープログラムを用いて配列内のエクソンの位置を予測した (Zhang、Proc. Natl. Acad. Sci. 94:565-568(1997))。

【0182】

6. Blastn2 アルゴリズム (Altschulら、Nucl. Acids. Res. 25:3389-3402(1997)) を用いて配列を公に利用可能なユニジェンデータベース (National Center for Biotechnology Information, National Library of Medicine、38A、8N905、8600 Rockville Pike、Bethesda、メリーランド州 20894; www.ncbi.nlm.nih.gov) と比較した。こ

50

の検索のためのパラメーターは、 $E = 0.05$ 、 $V = 50$ 、 $B = 50$ （ここで E は予測される可能性スコアカットフであり、 V は結果の報告時にリターンしたデータベースエントリーの数であり、 B は結果の報告時にリターンした配列アラインメントの数である (Altschulら、J.Mol.Biol.215:403-410(1990))）であった。

【0183】

7. 6個の読み枠全てに関して配列をタンパク質に翻訳し、タンパク質配列を Genpept Swissport PIR(National Center for Biotechnology Information, National Library of Medicine、38A、8N905、8600 Rockville Pike、Bethesda、メリーランド州 20894; www.ncbi.nlm.nih.gov) からコンパイルされた非重複性タンパク質データベースと比較した。この検索のためのパラメーターは、 $E = 0.05$ 、 $V = 50$ 、 $B = 50$ であり、ここで E 、 V 、および B は前記で定義したとおりである。 10

【0184】

8. BAC DNA配列を blastn2(Altschulら、Nucl.Acids.Res.25:3389-3402(1997)) を用いて、直接選択実験（以下で記載する）から誘導した cDNA クローンのデータベースと比較した。この検索のためのパラメーターは、 $E = 0.05$ 、 $V = 250$ 、 $B = 250$ であり、ここで E 、 V 、および B は前記で定義したとおりである。

【0185】

9. BAC配列を blastn2(Altschulら、Nucl.Acids.Res.25:3389-3402(1997)) を用いて、染色体 11q12-13 上の HBM 領域からのその他の全ての BAC の配列と比較した。この検索のためのパラメーターは、 $E = 0.05$ 、 $V = 50$ 、 $B = 50$ であり、ここで E 、 V 、および B は前記で定義したとおりである。 20

【0186】

10. BAC配列を blastn2(Altschulら、Nucl.Acids.Res.25:3389-3402(1997)) を用いて、染色体 11q12-13 上の HBM 領域からの BAC の末端から誘導した配列と比較した。この検索のためのパラメーターは、 $E = 0.05$ 、 $V = 50$ 、 $B = 50$ であり、ここで E 、 V 、および B は前記で定義したとおりである。

【0187】

11. BAC配列を blastn2(Altschulら、Nucl.Acids.Res.25:3389-3402(1997)) を用いて、Genbankデータベース(National Center for Biotechnology Information, National Library of Medicine、38A、8N905、8600 Rockville Pike、Bethesda、メリーランド州 20894; www.ncbi.nlm.nih.gov) と比較した。この検索のためのパラメーターは、 $E = 0.05$ 、 $V = 50$ 、 $B = 50$ であり、ここで E 、 V 、および B は前記で定義したとおりである。 30

【0188】

12. BAC配列を blastn2(Altschulら、(1997)) を用いて、Genbankデータベース(National Center for Biotechnology Information, National Library of Medicine、38A、8N905、8600 Rockville Pike、Bethesda、メリーランド州 20894; www.ncbi.nlm.nih.gov) の STS 部門と比較した。この検索のためのパラメーターは、 $E = 0.05$ 、 $V = 50$ 、 $B = 50$ であり、ここで E 、 V 、および B は前記で定義したとおりである。

【0189】

13. BAC配列を blastn2(Altschulら、Nucl.Acids.Res.25:3389-3402(1997)) を用いて、発現遺伝子配列断片 (EST) Genbankデータベース(National Center for Biotechnology Information, National Library of Medicine、38A、8N905、8600 Rockville Pike、Bethesda、メリーランド州 20894; www.ncbi.nlm.nih.gov) と比較した。この検索のためのパラメーターは、 $E = 0.05$ 、 $V = 250$ 、 $B = 250$ であり、ここで E 、 V 、および B は前記で定義したとおりである。 40

【0190】

X. 直接 cDNA 選択による遺伝子同定

1次リンカー結合 cDNA プールを骨髄、頭蓋冠、大腿骨、腎臓、骨格筋、精巣および脳全体から調製した。ポリ (A) + RNA を頭蓋冠および大腿骨組織から調製し (Chomczy 50

nskiら、Anal.Biochem.162:156-159(1987);D'Alessioら、Focus、9:1-4(1987))、mRNAの残りをClontech(Palo Alto、カリフォルニア州)から購入した。同一組織からオリゴ(dT)およびランダムプライミングされたcDNAプールを作製するために、一方の反応でmRNA 2.5 µgをオリゴ(dT)プライマーと混合し、他方の内で下線を付している反応でmRNA 2.5 µgをランダムヘキサマーと混合し、双方を製造者の推奨(Life Technologies、Bethesda、メリーランド州)に従って第1および第2鎖cDNAに変換した。対形成したリン酸化cDNAリンカー(以下の配列参照)を1:1の比率(各々10 µg)で混合することにより一緒にアニーリングし、65℃で5分間インキュベートし、室温まで冷却した。

【0191】

10

対形成したリンカーオリゴ1/2

OLIGO 1: 5' CTGAGCGGAATTCGTGAGACC3' (SEQ ID NO:12)

OLIGO 2: 5' TTGGTCTCACGTATTCCGCTCGA3' (SEQ ID NO:13)

対形成したリンカーオリゴ3/4

OLIGO 3: 5' CTCGAGAATTCTGGATCCTC3' (SEQ ID NO:14)

OLIGO 4: 5' TTGAGGATCCAGAATTCTCGAG3' (SEQ ID NO:15)

対形成したリンカーオリゴ5/6

OLIGO 5: 5' TGTATGCGAATTCGCTGCGCG3' (SEQ ID NO:16)

OLIGO 6: 5' TTCGCGCAGCGAATTCGCATACA3' (SEQ ID NO:17)

対形成したリンカーオリゴ7/8

20

OLIGO 7: 5' GTCCACTGAATTCTCAGTGAG3' (SEQ ID NO:18)

OLIGO 8: 5' TTGTCAC TGAGAATTCAGTGGAC3' (SEQ ID NO:19)

対形成したリンカーオリゴ11/12

OLIGO 11: 5' GAATCCGAATTCCTGGTCAGC3' (SEQ ID NO:20)

OLIGO 12: 5' TTGCTGACCAGGAATTCGGATTC3' (SEQ ID NO:21)

【0192】

製造者(Life Technologies、Bethesda、メリーランド州)の指示書に従って、全てのオリゴ(dT)およびランダムプライミングしたcDNAプール(以下を参照)にリンカーをライゲートした。

【0193】

30

オリゴ(dT)および骨髓から調製したランダムプライミングしたcDNAプールにオリゴ1/2をライゲートした。オリゴ(dT)および頭蓋冠から調製したランダムプライミングしたcDNAプールにオリゴ3/4をライゲートした。オリゴ(dT)ならびに脳および骨格筋から調製したランダムプライミングしたcDNAプールにオリゴ5/6をライゲートした。オリゴ(dT)および腎臓から調製したランダムプライミングしたcDNAプールにオリゴ7/8をライゲートした。オリゴ(dT)および大腿骨から調製したランダムプライミングしたcDNAプールにオリゴ11/12をライゲートした。

【0194】

各々1:1、1:10および1:100希釈のライゲーション反応物1 µlを用いてPCR増幅によりcDNAプールを長さの分布に関して評価した。Perkin Elmer 9600でPCR反応を実施し、反応物25 µl容量はDNA 1 µl、10 mM トリスHCl (pH 8.3)、50 mM KCl、1.5 mM MgCl₂、0.001%ゼラチン、200 mM 各dNTP、10 µM プライマーおよびTaq DNAポリメラーゼ(Perkin Elmer)1単位を含有し、以下の条件下で増幅された。94℃で30秒間、60℃で30秒間および72℃で2分間の30サイクル。増幅されたcDNAプールの長さの分布を1%アガロースゲル上で電気泳動により評価した。ランダムプライミングおよびオリゴ(dT)プライミングされたcDNAプールを最もよく提示するPCR反応を拡大して、各cDNAプール~2-3 µgを生成した。直接選択反応のための出発cDNAは、オリゴ(dT)プライミングされたcDNA 0.5 µgと混合したランダムプライミングされたcDNA 0.5 µgを含んでなった。

40

50

【0195】

直接cDNA選択手順で用いられた54BACからのDNAをNucleobond AXカラムを用いて製造者(The Nest Group, Inc.)により記載されるように単離した。

【0196】

BACを等モル量でプールし、製造者(Boehringer Mannheim)の指示書に従ってニック翻訳により単離されたゲノムDNA 1 µgをビオチン16-UTPで標識した。当業者に実施できる方法によりビオチンの組み込みをモニター観察した(Del MastroおよびLavett, Methods in Molecular Biology(1996), Humana Press Inc., ニュージャージー州)。

【0197】

当業者に実施できる方法を用いて直接cDNA選択を実施した(Del MastroおよびLavett, Methods in Molecular Biology(1996), Humana Press Inc., ニュージャージー州)。簡単には、2つの別個の反応でcDNAプールを多重化した。一方の反応では骨髄、頭蓋冠、脳および精巣からのcDNAプールを混合し、他方では骨格筋、腎臓および大腿骨からのcDNAプールを混合した。cDNAプールでは反復、酵母配列およびプラスミドの抑制をコット20まで実施した。ビオチン化BAC DNA 100 ngを抑制されたcDNAと混合し、コット200まで溶液中でハイブリダイズさせた。ビオチン化DNAおよび同族cDNAをストレプトアビジンコーティングした常磁性ビーズ上で捕捉した。ビーズを洗浄し、1次選択されたcDNAを溶出した。これらのcDNAをPCR増幅し、直接選択の第2ラウンドを実施した。直接選択の第2ラウンドの生成物を2次選択された材料と称する。以前に11q12-13 (Evans, Genomics 18:473-477(1993))にマッピングされることが示されたガラニンcDNAクローンを用いて選択の2つのラウンドの間の富化をモニター観察した。

【0198】

骨髄、頭蓋冠、大腿骨、腎臓、骨格筋、精巣および全脳からの2次選択材料を以下に示すオリゴ1、3、5、7および11の修飾プライマーを用いてPCR増幅し、製造者の推奨に従ってUDGベクターpAMP10 (Life Technologies, Bethesda, メリーランド州)にクローン化した。

【0199】

修飾されたプライマー配列：

Oligo1-CUA: 5' CUACUACUACUACTGAGCGGAATTCGTGAGACC3' (SEQ ID NO:22)

Oligo3-CUA: 5' CUACUACUACUACTCGAGAATTCTGGATCCTC3' (SEQ ID NO:23)

Oligo5-CUA: 5' CUACUACUACUATGTATGCGAATTCGCTGCGCG3' (SEQ ID NO:24)

Oligo7-CUA: 5' CUACUACUACUAGTCCACTGAATTCTCAGTGAG3' (SEQ ID NO:25)

Oligo11-CUA: 5' CUACUACUACUAGAATCCGAATTCCTGGTCAGC3' (SEQ ID NO:26)

【0200】

各組織供給源からのクローン化された2次選択された材料を、MAX Efficiency DH5aコンピテント細胞(Life Technologies, Bethesda, メリーランド州)に、製造者に推奨されるように形質転換した。各々の形質転換された供給源から384個のコロニーを採取し、96ウェルマイクロタイタープレートに整列させた。

【0201】

全ての2次選択されたcDNAクローンをM13色素プライマー・ターミネーター・サイクル・シーケンシングキット(Applied Biosystems)を用いてシーケンシングし、ABI 377自動蛍光シーケンサー(Applied Biosystems)によりデータを収集した。

【0202】

全配列をBLASTN、BLASTXおよびFASTAプログラムを用いて分析した(Altschulら、J.Mol.Biol.215:403-410(1990)、Altschulら、Nucl.Acids.Res.25:3389-3402(1997))。cDNA配列をヒト反復、ミトコンドリアDNA、リボソームRNA、大腸菌DNAに由来する配列を含有するデータベースと比較し、クロスマッチプログラムを用いてデータセットからバックグラウンドクローンを除去した。既知遺伝子(Genbank)およびHBM領域に由来するBAC配列に対するBLASTN2プログラムを用いて比較の別の

10

20

30

40

50

ラウンドをも実施した。結果およびさらなる分析のためにデータベースに蓄えられたデータに従って、これらの配列に対して>90%相同であるcDNAをファイリングした。同定されたが、HBM領域に由来するBAC配列に対して有意な類似性を有さなかったか、またはクロスマッチにより排除されたcDNA配列を、HBM領域からのBACを含有するナイロン膜にハイブリダイズさせて、これらが標的にハイブリダイズするかどうかを確認した。

【0203】

ハイブリダイゼーション分析を用いてcDNAクローンを、それを選択するBAC標的にマッピングした。HBM領域から同定されたBACを整列させ、96ウェルマイクロタイタープレートで成長させた。25 µg/mlカナマイシンを含有するLB寒天培地を96ウェルマイクロタイタープレートのリッドに注いだ。一度寒天が固化すると、プレカットHybond N+ナイロン膜(Amersham)を寒天の上部に置き、手持ち型の96ウェルレプリカプレーター(V&P Scientific, Inc.)を用いてBACを膜上に2通りずつスタンプした。プレートを37℃で一晩インキュベートした。製造者の推奨に従って膜を処理した。

10

【0204】

ハイブリダイゼーションによりマッピングする必要があるcDNAをそのクローンを増幅するのに適切なプライマー(オリゴ1、3、5、7および11)を用いてPCR増幅した。このPCR増幅のために、プライマーを修飾して、リンカー処理されたジゴキシゲニン分子をオリゴヌクレオチドの5'で含有するようにした。cDNAプール(前記)の調製にて記載したのと同じ条件下でPCR増幅を実施した。PCR反応物の5 µlを負荷することにより1%アガロースゲル上での電気泳動によりPCR産物を品質および量に関して評価した。ハイブリダイゼーション溶液(5X SSPE、0.5X Blot to、2.5% SDSおよび1mM EDTA (pH 8.0)) 10mlが入った50mlコニカルチューブ中でスタンプされたBACを含有するナイロン膜を個々にプレハイブリダイズした。50mlコニカルチューブをロティサリー・オープン(Robbins Scientific)中に65℃で2時間置いた。各cDNAプローブ25 ngを変性し、そしてナイロン膜およびハイブリダイゼーション溶液が入った個々の50mlコニカルチューブに加えた。65℃で一晩ハイブリダイゼーションを実施した。以下の溶液の各々を用いて65℃で20分間、フィルターを洗浄した。3x SSPE、0.1% SDS; 1x SSPE、0.1% SDSおよび; 0.1x SSPE、0.1% SDS。

20

30

【0205】

50mlコニカルチューブから膜を取り出し、皿に置いた。各膜の間にアセテートシートを置き、互いに貼り付くのを防いだ。製造者(Boehringer Mannheim)の推奨に従って膜を抗DIG-APおよびCDP-Starと共にインキュベートした。膜をサランラップで包み、Kodak Bio-Max X線フィルムに1時間暴露した。

【0206】

XI. cDNAクローニングおよび発現分析

直接cDNA選択およびゲノムDNAシーケンシングにより同定された遺伝子の発現を、公に利用可能なデータベースと比較して特徴付けするために、一連の実験を実施してHBM領域の遺伝子をさらに特徴付けした。最初にポリメラーゼ連鎖反応(PCR)において使用するために、cDNA、EST、またはゲノムDNAの部分をDNA分子のプール(cDNAライブラリー)またはRNA集団(RT-PCRおよびRACE)から増幅できるようにオリゴヌクレオチドプライマーを設計した。ゲノムDNAを含有する反応物においてPCRプライマーを使用して、ゲノム(BAC)配列に基づいて予測される大きさの産物が作製されることを確認した。次いで多くのcDNAライブラリーを特異的cDNAまたはESTの存在に関して試験した。特定のcDNAライブラリーにおける転写ユニットのフラグメントの存在により、同一転写ユニットの別の部分が同様に存在する可能性が高いことが示される。

40

【0207】

50

新規の遺伝子を特徴付けするのに必要とされるデータの重要な部分は、プロセシングされた転写物またはメッセンジャーRNA (mRNA) のヌクレオチドの長さである。当業者はmRNAの長さを主にノーザンブロットハイブリダイゼーション(Sambrookら、Molecular Cloning: A Laboratory Manual、Cold Spring Harbor Laboratory、Cold Spring Harbor、ニューヨーク州(1989))により決定する。重要な領域のシーケンシングされたBACに有意な配列類似性を展示するEST群および直接選択されたcDNAクローンを便宜上およそ30キロ塩基のユニットにグループ分けした。各30キロ塩基のユニット内に、1つまたはそれ以上の独立した転写ユニットを含む1から50個までのESTおよび直接選択されたcDNAクローニングがあった。市販により入手可能な試薬(Multiple Tissue Northern blot; Clontech, Palo Alto、カリフォルニア州)を用いて製造者に推奨される条件下で、1つまたはそれ以上のESTまたは直接選択されたcDNAをハイブリダイゼーションプローブとして用いて種々の組織のmRNAの長さを決定した。 10

【0208】

大腿骨および頭蓋冠骨組織から志向的にクローン化されたcDNAライブラリーを当業者に周知の方法(例えばSoares、in Automated DNA Sequencing and Analysis、Adams、FieldsおよびVenter編、Academic Press、ニューヨーク州、110-114頁、(1994))により構築した。最初に骨をハンマーで断片に破碎し、小片を液体窒素中で凍結し、組織パルベライザー(Spectrum Laboratory Products)で粉末まで小さくした。ポリトロンホモジナイザー(Brinkman Instruments)を用いて標準的な酸グアニジニウム・チオシアネート・フェノール・クロロホルム抽出バッファー(例えばChomczynskiおよびSacchi、Anal. Biochem. 162:156-159(1987))で粉末化された骨をホモジナイズすることにより、粉末化された骨からRNAを抽出した。加えて、ヒト脳および肺全RNAをClontechから購入した。ダイナビーズdTを用いて、製造者の推奨に従って(Dynal, Inc.)ポリA RNAを全RNAから単離した。配列: 5'-AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTTTTT-3' (SEQ ID NO:27)のオリゴヌクレオチドプライマーを用いて第1鎖cDNA合成を開始した。このプライマーをcDNAの3'末端でNotI制限部位(下線部)を導入した。「ワンチューブ」cDNA合成キットを用いて、製造者(Life Technologies、Bethesda、メリーランド州)に記載されるように、第1および第2鎖合成を実施した。二本鎖DNAをT4ポリヌクレオチドキナーゼで処理して、分子の末端が平滑であり(Soares、in Automated DNA Sequencing and Analysis、Adams、FieldsおよびVenter編、Academic Press、ニューヨーク州、110-114頁(1994))、次に平滑末端cDNAがBio gelカラム(Huynhら、DNA Cloning、1巻、Glover編、IRL Press、Oxford、49-78頁、(1985))により、またはsize-sep 400セファロースカラム(Pharmacia、カタログ番号27-5105-01)を用いて大きさで選択されることを確実にした。400塩基対またはそれ以上の長さのcDNAのみを次の行程で使用した。次いで当分野で周知の方法により(Soares、1994)、EcoRIアダプター(配列: 5'-OH-AATTCGGCACGAG-OH-3' (配列番号28)および5'-p-CTCGTGC CG-OH-3' (配列番号29))を二本鎖cDNAにライゲートした。次いでNotIで消化することによりcDNAの3'末端からEcoRIアダプターを除去した(Soares、1994)。次いでcDNAをプラスミドベクターpBluescript(登録商標)IISK⁺にライゲートし(Stratagene、La Jolla、カリフォルニア州)、そして当業者に周知のエレクトロポレーション法によりライゲートされた材料を大腸菌DH10BまたはDH12Sに形質転換した(Soares、1994)。37℃で一晩成長させた後、Mega-prepキット(Qiagen、Chatsworth、カリフォルニア)に関して指示されるように処理することによりプレートを擦り取った後、大腸菌コロニーからDNAを回収した。1次形質転換体の全数の一部を計数し、そして平均のインサートの大きさおよびcDNAインサートを含まないプラスミドのパーセンテージを決定することによりcDNAライブラリーの品質を推定した。さらなるcDNAライブラリー(ヒト全脳、心臓、腎臓、白血球、および胎児脳)をLife Technologies、Bethesda、メリーランド州から購入した。 30 40

【0209】

オリゴ(dT)およびランダムヘキサマー(N₆)の双方でプライミングされたcDNA 50

Aライブラリーを用いて、HBM領域：ヒト骨、ヒト脳、ヒト腎臓およびヒト骨格筋（骨格筋（dT）および腎臓（dT）cDNAライブラリー以外は、全てのcDNAライブラリーが発明者により作製された）内に転写されたcDNAクローンを単離するために使用した。各々のcDNAライブラリーの4つの10×10アレイを以下のように調製した。1次形質転換体を用いてcDNAライブラリーを 2.5×10^6 に力価を調節した。適当な容量の凍結ストックを用いてLB／アンピシリン（100mg/ml）2lを接種した。この接種した液体培養を各々4mlの400本のチューブに等分した。各チューブにはおよそ5000cfuが入っていた。チューブを30℃で一晩穏やかに振盪しながらインキュベートした。培養物をOD0.7から0.9まで成長させた。培養物の各々に関して等分した培養物100μlおよび80%グリセロール300μlにより凍結ストックを調製した。ストックをドライアイス／エタノール浴中で凍結させて−70℃で貯蔵した。残りの培養物はQiagen(Chatsworth、カリフォルニア州)スピン・ミニプレップ・キットを用いて製造者の指示に従って調製されたDNAであった。400培養物からのDNAをプールして80カラムを作り、そしてプールを整列させた。PCRにより目的のHBM cDNAのクローンを含有するようにcDNAライブラリーを決定した。推定されるエクソンを増幅するためにマーカーを設計した。一度標準的なPCR最適化を実施し、そして特異的cDNAライブラリーを目的のcDNAクローンを含有するように決定すると、マーカーを使用して整列させたライブラリーをスクリーニングした。同一マーカーを用いる第2のPCRにより、cDNAクローンの存在を示す陽性アドレスを確認した。

【0210】

一度HBM領域に由来する目的の特異的転写物に相当するcDNAクローンを含有する可能性があるようにcDNAライブラリーを同定すると、ESTまたは直接選択された目的のcDNAに同等なcDNAインサートを含むクロンまたは複数のクロンを単離するように操作した。標準的な「コロニースクリーニング」法の変法(Sambrookら、Molecular Cloning:A Laboratory Manual、Cold Spring Harbor Laboratory、Cold Spring Harbor、ニューヨーク(1989))によりこれを達成した。具体的には、20個の150mm LB＋アンピシリン寒天プレートにcDNAライブラリーの20000コロニー形成ユニット（cfu）を広げ、そしてコロニーを37℃で一晩成長させた。コロニーをナイロンフィルター(AmerchamのHybondまたは等価物)に移し、そして本質的には記載されているように(Sambrookら、Molecular Cloning:A Laboratory Manual、Cold Spring Harbor Laboratory、Cold Spring Harbor、ニューヨーク(1989))2つのフィルターを一緒に通過させることにより2通りずつ調製した。次いで「マスター」プレートをさらに6から8時間インキュベートしてコロニーを再び成長させた。次いでフィルターを連続して変性溶液（0.5N NaOH、1.5M NaCl）で2分間、中和溶液（0.5M トリスCl、pH8.0、1.5M NaCl）で2分間（2回）処理することにより細菌コロニーに由来するDNAをナイロンフィルターに貼り付けた。2×SSC／0.1%SDSの溶液で1分間洗浄し、その上ティッシュペーパーでふき取ることにより細菌コロニーをフィルターから除去した。フィルターを空気乾燥し、そして真空化80℃で1から2時間焼いた。

【0211】

ランダムヘキサマー標識(FinebergおよびVogelstein、Anal.Biochem.132:6-13(1983))によるか、または反応に遺伝子特異的プライマーを含めランダムヘキサマーを含めないことにより（小型フラグメント用）cDNAハイブリダイゼーションプローブを調製した。特異活性を算出し、そして $> 5 \times 10^8$ cpm／ 10^8 μg cDNAであった。次いでコロニー膜を10mM トリスCl（pH8.0）、1M NaCl、1mM EDTA 0.1%SDSを用いて55℃で30分間予備洗浄した。予備洗浄の後、フィルターを6×SSC＞2ml／フィルター、50%脱イオン化ホルムアミド、2%SDS、5×デンハート溶液、および100mg/ml変性サケ精子DNAで、42℃で30分間プレハイブリダイゼーションした。次いでフィルターを変性 α^{32} P-dCTP標識cDNAプローブを含むハイブリダイゼーション溶液（6×SSC、2%SDS、5×デンハート溶液、および100mg/ml変性サケ精子DNA）に移し、そして42℃で16か

ら18時間インキュベートした。

【0212】

16から18時間のインキュベーションの後、2X SSC、2% SDS中、常に振盪しながら、室温で20分間フィルターを洗浄し、続いて各々を65℃、15分間で2回洗浄した。2回目の洗浄は0.5X SSC、0.5% SDS中65℃で15分間実施した。次いでフィルターをプラスチックラップで包み、そしてX線写真フィルムに数時間から一晩暴露した。フィルムを現像した後、プレート上の個々のコロニーをオートラジオグラフィプレート上で整列させて、アンピシリン含有LBブrossの溶液1mlに採取した。37℃で1から2時間振盪した後、溶液のアリコートで2次スクリーニング用に150mmプレートに載せた。2次スクリーニングは、個々のコロニーが採取用に明確に同定されるように、~250個のコロニーを含有するプレートで実施したこと以外は、1次スクリーニング（前記）と同一であった。

【0213】

放射線標識プローブでスクリーニングしたコロニーがcDNAクローンを生じた後、制限エンドヌクレアーゼ切断、PCRおよび直接シーケンシングによりクローンを特徴付けして、元来のプローブと単離されたクローンの間の配列同一性を確認した。全長cDNAを得るために、同定されたクローンの末端からの新規の配列を用いてライブラリーを再度プローブした。ノーザンブロット分析によりcDNAクローンの長さが全長であることが推定されるまで、この方法を繰り返した。

【0214】

全長クローンを単離するための別の方法としてRT-PCRを用いた。cDNAを合成し、そして「Superscript One Step RT-PCR」キット(Life Technologies, Gaithersburg, メリーランド州)を用いて増幅した。手順にはRNA 1.5 µgを以下のものに添加することが含まれた。MgSO₄ およびdNTPを伴う専用バッファーミックスである提供された反応物ミックス25 µl、センスプライマー(10 µM) 1 µl および抗センスプライマー(10 µM) 1 µl、逆転写酵素1 µl、並びに提供されたTaq DNAポリメラーゼミックス、ならびに加圧滅菌水で反応物ミックス全量を50 µlにする。次いで反応物をサーモサイクラー内に置いて、予測される生成物のキロ塩基あたり50℃で15から30分、次いで94℃で15秒間、55から60℃で30秒間および68から72℃で1分間の1サイクル、そして最終的に72℃で5から10分間の1サイクルであった。サンプルをアガロースゲル上で分析した。生成物をゲルから切除し、そしてゲル(GeneClean, Bio 101)から精製した。精製された生成物をpCTNP(General Contractor DNA Cloning System, 5Prime-3Prime, Inc.)にクローン化し、そしてシーケンシングしてクローンが目的の遺伝子に対して特異的であったことを確認した。

【0215】

候補遺伝子の5'および3'末端をクローニングする方法として、Marathon cDNA増幅キット(Clontech, Palo Alto, カリフォルニア州)を用いて、製造者の指示に従ってcDNA末端迅速増幅(RACE)を実施した。第1鎖合成を実施することにより全RNAからcDNAプールを調製し、ここで全RNAサンプルのサンプルを修飾オリゴ(dT)プライマーと混合し、70℃まで加熱し、氷上で冷却し、そして続いて以下のものを添加した。5x第1鎖バッファー、10mM dNTPミックス、およびAMV逆転写酵素(20 U/µl)。チューブを42℃で1時間インキュベートし、そして次に反応物チューブを氷上に置いた。第2鎖合成に関しては、以下の成分：5x第2鎖バッファー、10mM dNTPミックス、滅菌水、20x第2鎖酵素カクテルを直接反応チューブに加え、そして反応チューブを16℃で1.5時間インキュベートした。T4 DNAポリメラーゼを反応チューブに加え、そして16℃で45分間インキュベートした。EDTA/グリコーゲンミックスを添加して第2鎖合成を停止させた。サンプルをフェノール/クロロホルム抽出および酢酸アンモニウム沈殿に供した。大きさの分布に関してアガロースゲル上で分析することによりcDNAプールを品質に関して検査した。次いでMarathon cDNAアダプター(Clontech)をcDNA末端にライゲートした。特異的アダプターは5

’ または 3 ’ 末端の増幅を可能にするプライミング部位を含有し、選択した遺伝子特異的プライマー (G S P) の配向に依存した。二本鎖 c D N A のアリコートを下記の試薬に添加した。1 0 μ M Marathon c D N A アダプター、5 x D N A ライゲーションバッファー、T 4 D N A リガーゼ。反応物を 1 6 ° C で一晩インキュベートした。反応物を加熱して不活性化し、反応を停止させた。以下のものを希釈した二本鎖 c D N A プールに添加することにより P C R を実施した。1 0 x c D N A P C R 反応バッファー、1 0 μ M d N T P ミックス、1 0 μ M G S P、1 0 μ M A P 1 プライマー (キット)、5 0 x Advantage c D N A ポリメラーゼミックス。熱循環条件は 9 4 ° C 3 0 秒間、5 サイクルの 9 4 ° C 5 秒間、7 2 ° C 4 分間、5 サイクルの 9 4 ° C 5 秒間、7 0 ° C 4 分間、2 3 サイクルの 9 4 ° C 5 秒間、6 8 ° C 4 分間。P C R の最初のラウンドの後、G S P を用いて実施してアダプターの末端を伸長してアダプタープライマー結合部位を作製し、目的の特異的 c D N A の指数関数的増幅を観察した。通常第 2 のネステッド P C R を実施して特異的 c D N A を確認する。R A C E 産物をアガロースゲル上で分析し、そして次に切除してゲルから精製した (GeneClean、BIO 101)。次に R A C E 産物を p C T N R (General Contracto r DNA Cloning System、5 ’ -3 ’ , Inc.) にクローン化し、そして該クローンを確認するために決定した D N A 配列は目的の遺伝子に特異的であった。

【0216】

X I I . 変異分析

前記の手順を用いて比較遺伝子を同定し、各遺伝子からのエクソンを変異検出分析に供した。比較 D N A シークエンシングを用いて染色体 1 1 q 1 2 - 1 3 から H B M 候補遺伝子における多型性を同定した。候補遺伝子の D N A 配列を患者のリンパ芽球腫細胞株から増幅した。

【0217】

発明者は原因である多型性を検索するための候補領域から増幅した P C R 産物の直接 D N A シークエンシングの分析に基づく方法を開発した。手順は、多型性の分離を見出すための H B M ファミリーの異なるサブセットおよび多型性の頻度を評価するための集団パネルを用いる 3 つの段階から成った。ファミリーリソースは罹患した全ての個体が同一の原因多型性を共有するという仮定に至る 1 人の創始者から導かれる。

【0218】

最初、発端者、娘および彼女の母親、父親および兄弟から成る H B M ファミリーのサブセットにおいて候補領域をスクリーニングした。同時に単一染色体参照配列を生成し、比較に用いた。母親および娘がこの核ファミリーで H B M 多型性を担持し、多型性の伝達をモニター観察する能力を提供した。最終結果は、2 個の H B M 染色体および 6 個の非 H B M 染色体がスクリーニングされたということである。これにより多くの頻出する対立遺伝子の切除が可能になった。罹患した個体に独占的に存在する対立遺伝子のみが分析の次のレベルまで通過した。

【0219】

次いでこの元来のファミリーにおいて H B M 表現型で独占的に分離された多型性を、2 つの別の核ファミリーから成る H B M 家系の拡張された部分において再試験した。これらのファミリーは 5 人の H B M および 3 人の非罹患個体から成る。この群の H B M 個体は 2 つの重要なクロスオーバー個体を含み、重要な領域のセントロメアおよびテロメア境界を提供する。これらの個体と罹患したその親の間の多型性の遺伝をトラッキングすることにより、重要な領域をさらに絞り込むことが可能になった。この群ではスクリーニングして H B M 全染色体が 7 個に、そして非 H B M 全染色体が 1 7 個になった。

【0220】

所定の多型性が拡張された群において H B M 表現型で独占的に分離を続けた場合、次いで集団パネルを試験した。8 4 人のこのパネルは正常な骨ミネラル密度を有していることが解っている 4 2 人の個体、および関連性はないが類型的でない骨ミネラル密度であることが解っている 4 2 人の個体から構成された。この目的で、正常な骨ミネラル密度は B M D Z スコア 0 の 2 標準偏差内である。第 2 の群は広く用いられている個体の C E P H パ

ネルからであった。この集団において稀に見出されたいずれかの分離多型性を、続いて全 HBM 家系およびさらに大きな集団において試験した。

【0221】

ポリメラーゼ連鎖反応 (PCR) を用いて HBM ファミリーの DNA および単一染色体対照からシーケンシング鋳型を作成した。11q12-13 上の HBM 領域内の遺伝子の酵素的増幅を、各エクソンおよび各遺伝子の推定 5' 制御エレメントをフランキングするオリゴヌクレオチドを用いる PCR で達成した。プライマーを選択して、各エクソンおよびスプライシングのいずれかの側の各イントロン内の 15 またはそれ以上の塩基対を増幅した。全ての PCR プライマーをキメラとして作製し、色素プライマーシーケンシングを促進した。合成中に M13-21F (5'-GTA A CGA CGG CCA GT-3') (配列番号 30) および -28REV (5'-AAC AGC TAT G AC CAT G-3') (配列番号 31) プライマー結合部位を各順行性および逆行性 PCR プライマーの 5' 末端に各々構築した。AmpliTaq 2U、500nM プライマーおよび 125μM dNTP を含む PCR 50μl 中ゲノム DNA 150ng を使用した。バッファーおよび循環条件は各プライマーセットに特異的であった。TaqS tart 抗体 (Clontech) を出発 PR に使用してプライマーダイマー形成を最低限にした。10% 生成物をアガロースゲル上で試験した。シーケンシング前に適当なサンプルを脱イオン水で 1:25 に希釈した。

【0222】

標準的な Energy Transfer プライマー (Amersham) プロトコルに従って各 PCR 産物をシーケンシングした。全反応を 96 ウェルトレイで行った。4 つの別個の反応、A、C、G および T に関して 1 つずつを各鋳型に関して実施した。各反応はシーケンシング反応ミックス 2μl、および希釈した鋳型 3μl を含んだ。次いでプレートをホイルテープでヒートシールし、サーマル・サイクラー内に置き、製造者の推奨に従って循環させた。循環後、4 つの反応物をプールした。プールした生成物 3μl を新たな 96 ウェルプレートに移し、製造者の負荷色素 1μl を各ウェルに加えた。96 ウェル全てのピペッティング手順を Hydra 96 ピペッティング・ステーション (Robbins Scientific、米国) で行った。プールした材料 1μl を ABI 377 DNA シーケンサーで 2.4kV 実行で 10 時間作動する 48 レーンゲルに直接負荷した。

【0223】

Polyphred (Washington 大学) を用いて Consed (Washington 大学) で表示するために配列セットをアセンブルした。全ての関係するファミリーのメンバーおよび特定の標的領域の対照を提示する群で配列をアセンブルした。これを 3 段階の各々に関して別個に行った。順行性および逆行性の読みを単一染色体鋳型および着色した参照配列からの読みと一緒に各個体に関して包括した。Polyphred は紫色のフラッグで可能性のある多型性部位を示した。2 つのリーダーは別個に各アセンブリを表示し、紫色のフラッグを付した部位の正当性を評価した。

【0224】

成熟 mRNA に存在する 23 エクソン全ておよび 1 次転写体のいくつかのその他の部分を、2 人の HBM 罹患および 2 人の非罹患個体の核ファミリーのヘテロ接合性に関して評価した。以下の表に示すように 25 個の SNP を同定した。

【0225】

【表 4】

表 4 : LRP5 遺伝子および周辺の単一塩基多型

エクソン名	位 置	塩基変化
b200e21-h_コンティグ 1_1.nt	69169 (309G)	C/A
b200e21-h_コンティグ 4_12.nt	27402 (309G)	A/G
b200e21-h_コンティグ 4_13.nt	27841 (309G)	T/C
b200e21-h_コンティグ 4_16.nt	35600 (309G)	A/G
b200e21-h_コンティグ 4_21.nt	45619 (309G)	G/A
b200e21-h_コンティグ 4_22.nt-a	46018 (309G)	T/G
b200e21-h_コンティグ 4_22.nt-b	46093 (309G)	T/G
b200e21-h_コンティグ 4_22.nt-c	46190 (309G)	A/G
b200e21-h_コンティグ 4_24.nt-a	50993 (309G)	T/C
b200e21-h_コンティグ 4_24.nt-b	51124 (309G)	C/T
b200e21-h_コンティグ 4_25.nt	55461 (309G)	C/T
b200e21-h_コンティグ 4_33.nt-a	63645 (309G)	C/A
b200e21-h_コンティグ 4_33.nt-b	63646 (309G)	A/C
b200e21-h_コンティグ 4_61.nt	24809 (309G)	T/G
b200e21-h_コンティグ 4_62.nt	27837 (309G)	T/C
b200e21-h_コンティグ 4_63.nt-a	31485 (309G)	C/T
b200e21-h_コンティグ 4_63.nt-b	31683 (309G)	A/G
b200e21-h_コンティグ 4_9.nt	24808 (309G)	T/G
b527d12-h_コンティグ 030g_1.nt-a	31340 (308G)	T/C
b527d12-h_コンティグ 030g_1.nt-b	32538 (308G)	A/G
b527d12-h_コンティグ 080C_2.nt	13224 (308G)	A/G
b527d12-h_コンティグ 087C_1.nt	21119 (308G)	C/A
b527d12-h_コンティグ 087C_4.nt	30497 (308G)	G/A
b527d12-h_コンティグ 088C_4.nt	24811 (309G)	A/C
b527d12-h_コンティグ 089C_1HP.nt	68280 (309G)	G/A

【0226】

表 4 で示した多型性に加えて、2 つのさらなる多型性が配列番号 2 にも存在し得る。これらは配列番号 2 の 2002 位置における変化である。グアニンまたはアデニンのいずれかがこの位置に現れ得る。この多型性はサイレントであり、そしてアミノ酸配列におけるなんらかの変化は随伴されない。第 2 の変化はシトシン (C) のチミン (T) への変化に相当する配列番号 2 の 4059 においてである。この多型性の結果、バリン (V) からアラニン (A) への対応するアミノ酸変化に至る。別の多型性は候補遺伝子エクソンおよび隣接するイントロン配列において見出された。表 4 に列挙された多型性のいずれか 1 つもしくはは組み合わせ、または前記で論じた 2 つもまた、配列番号 2 に存在した場合に骨量には小さな影響しかないであろう。

【0227】

本発明は前記で同定した点変異を有する配列番号 1 の核酸配列を有する核酸配列を包含する。

【0228】

好ましくは、本発明は配列番号 2 の核酸を包含する。とりわけ、LRP5 のコード化配列 (HBM 遺伝子) の 582 位置での G から T へ変化する塩基対置換は、全ての HBM 個体のヘテロ接合体として同定され、非罹患個体では見出されなかった (すなわち b527

10

20

30

40

50

d 1 2 - h コンティグ 0 8 7 C 1 . n t) 。図 5 は B 5 2 7 D 1 2 におけるコンティグの順序を示す。HBM 遺伝子に関する転写の方向は左から右である。B 5 2 7 D 1 2 のコンティグ 3 0 8 G の配列は HBM 遺伝子に対するコード化領域の逆相補体である。従って、C から A への塩基変化置換として表 4 に示すコンティグ 3 0 8 G の相対的多型性は、HBM 遺伝子の G から T への置換に対する相補体である。この変異はグリシン 1 7 1 のバリンとの置換を引き起こす (G 1 7 1 V) 。

【 0 2 2 9 】

HBM 多型性は異なる群の個体の DNA 配列を試験することにより確認された。HBM 家系の全メンバー (3 8 人の個体) のうち、罹患した (すなわち骨量が上昇している) 個体のみ (N = 1 8) からのヘテロ接合体で HBM 多型性が観察された。非罹患親族 (N = 2 0) (BMDZ < 2 . 0) では、HBM 多型性は全く観察されなかった。この多型性が HBM 家系の外側の個体でも観察されたかどうかを決定するために、2 9 7 人の表現型個体を HBM 遺伝子の部位で特徴付けした。HBM 多型性の部位でヘテロ接合体であったものはいなかった。非表現型対照群では、6 4 人の個体で 5 8 2 位置でのヘテロ接合体が観察されたものはいなかった。これらを考慮して、このデータにより高骨量表現型を表示する家系において観察された多型性が LRP5 遺伝子の位置 5 8 2 での G → T 多型性に強く相関することが判明した。さらに、この結果を以下で記載する ASO の結果と合わせて、HBM 多型性が HBM 表現型と遺伝的に分離され、HBM 多型性および表現型の双方が一般集団では稀であることが確立された。

【 0 2 3 0 】

X I I I . 対立遺伝子特異的オリゴヌクレオチド (ASO : Allele specific oligonucleotide) 分析

HBM 多型性を含有するアンプリコンを目的のエクソンに特異的なプライマーを用いて PCR 増幅した。個体の適切な集団を以下のように 9 6 ウェルマイクロタイタープレートで PCR 増幅した。1 X Promega PCR バッファー (カタログ番号 M 1 8 8 3 、 1 . 5 m M M g C l ₂ 含有) 、 1 0 0 m M d N T P 、 2 0 0 n M PCR プライマー (配列番号 6 2 9 - 6 3 0) (1 8 6 3 F : C C A A G T T C T G A G A A G T C C および 1 8 6 4 R : A A T A C C T G A A A C C A T A C C T G) 、 Ampli taq 1 単位、およびゲノム DNA 2 0 n g を含有する PCR 反応物 (2 0 μ l) を調製し、以下の PCR 条件下で増幅した。9 4 ° C で 1 分間、 (9 4 ° C 、 3 0 秒間 ; 5 8 ° C 、 3 0 秒間 ; 7 2 ° C 、 1 分間 X 3 5 サイクル) 、 7 2 ° C 、 5 ' 分間、4 ° C を維持。次いで負荷色素を加え、生成物 1 0 μ l を 1 μ g / m l 臭化エチジウムを含有する 1 . 5 % アガロースゲル上で 1 0 0 から 1 5 0 V で 5 から 1 0 分間電気泳動した。変性溶液 (1 . 5 M N a C l 、 0 . 5 N N a O H) 中ゲルを 2 0 分間処理し、水で短時間すすいだ。次いでゲルを 1 M トリス H C l 、 p H 7 . 5 、 1 . 5 M N a C l 中で 2 0 分間中和し、水ですすいだ。ゲルを 1 0 X S S C 中に 2 0 分間浸漬し、1 0 X S S C 中ナイロントランスファー膜 (Hybond N + - A m e r s h a m) に一晩ブロッティングした。フィルターを 6 X S S C で 1 0 分間すすぎ、UV 架橋した。

【 0 2 3 1 】

ほぼ中央で多型性を有する対立遺伝子特異的オリゴヌクレオチド (ASO) を設計した。オリゴヌクレオチドは 5 ' 末端でリン酸不含であり、Gibco BRL から購入した。オリゴヌクレオチドの配列 (配列番号 6 3 1 ~ 6 3 2) は、

2326 ZMAX.ASO.g: AGACTGGGGTGAGACGC

2327 ZMAX.ASO.t: CAGACTGGGGTTGAGACGCC

である。

【 0 2 3 2 】

多型性ヌクレオチドには下線を付している。オリゴを標識するために、1 μ g / μ l ASO オリゴ (2 3 2 6 . Z M A X . A S O . g または 2 3 2 7 . Z M A X . A S O . t) 1 . 5 μ l 、 d d H ₂ O 1 1 μ l 、 1 0 X キナーゼ順行性バッファー 2 μ l 、 γ ³² P - A T P (6 0 0 0 C i / ミリモル) 5 μ l 、および T 4 ポリヌクレオチドキナーゼ

(10 U/μl) 1 μl を混合し、反応物を 37℃ で 30 から 60 分間インキュベートした。次いで反応物を 95℃ で 2 分間放置し、H₂O 30 ml を加えた。G25 マイクロスピナラム (Pharmacia) を用いてプローブを精製した。

【0233】

5xSSPE 10 ml、5X デンハート溶液、2% SDS、および 100 μg/ml 変性ソニケートサケ精子 DNA 中 40℃ で 2 時間ブロットをプレハイブリダイズした。次いでキナーゼ処理したオリゴの全反応物ミックスを新鮮なハイブリダイゼーションバッファー (5xSSPE、5x デンハート溶液、2% SDS) 10 ml に添加し、40℃ で少なくとも 4 時間から一晩ハイブリダイズした。

【0234】

全ての洗浄を 5xSSPE、0.1% SDS で行った。最初の洗浄は 45℃ で 15 分間であった；次いで溶液を交換し、フィルターを 50℃ で 15 分間洗浄した。次いでフィルターを 2 つの増感スクリーニングを有する Kodak バイオマックス・フィルムに -70℃ で 15 分から 1 時間暴露した。必要な場合、フィルターを 55℃ で 15 分間洗浄し、再度フィルムに暴露した。沸騰 0.1xSSPE、0.1% SDS 中 10 分間少なくとも 3 回洗浄することによりフィルターをストリップ状にした。

【0235】

ASO の対立遺伝子特異的アッセイを最もよくとらえた 2 つのフィルムを Adobe Photoshop にスキャンすることによりデジタル画像に変換した。これらの画像を Graphic Converter で互いに積層し、スコア化して FileMaker Pro4.0 に保存した (図参照)。

【0236】

民族的に多様な集団における HBM 対立遺伝子頻度を決定するために、種々の民族群からの 672 人のランダムな個体を対立遺伝子特異的オリゴヌクレオチド (ASO) 法により分類した。この集団には CEPH 祖父母 (本来の白色人種) 96 人、白色人種 192 人、アフリカ系アメリカ人 192 人、ヒスパニック系 96 人、およびアジア系 96 人の個体が含まれた。これらの個体のいずれかの HBM 多型性の存在に関して証拠は得られなかった。全体では、911 人の個体全員が直接シーケンシングまたは ASO ハイブリダイゼーションのいずれかにより分類された。全て HBM 多型性の部位でヘテロ接合性 GG であった (図 14)。この情報は、HBM 対立遺伝子が種々民族集団において希少性であることを説明している。

【0237】

従って本発明は HBM 対立遺伝子を有する個体を同定する迅速な方法を提供する。この方法を骨粗鬆症またはその他の骨傷害に対する罹患性に関して個体を診断およびスクリーニングする分野で用いることができる。該アッセイを用いて HBM 対立遺伝子または本明細書で記載した別の多型性を有する別の個体を同定することもできる。

【0238】

XIV. LRP5 の細胞局在

非同位体インサイチュ・ハイブリダイゼーションによるラット脛骨における遺伝子発現 Pathology Associates International (PAI)、Frederick、メリーランド州によりインサイチュ・ハイブリダイゼーションを行った。この研究は、骨成長およびリモデリングの分野で特に強調されているラットの骨における LRP5 遺伝子を発現する特定の細胞型を決定するために実施した。この研究で用いた LRP5 プローブをヒト (HuZmax1) およびマウス (MsZmax1) 双方の cDNA から作製し、これは 87% の配列同一性を共有する。ヒトおよびマウス LRP5 とラット LRP5 の相同性は公知である。

【0239】

例えば非同位体インサイチュ・ハイブリダイゼーションによる遺伝子発現を以下のとおり実施したが、その他の方法は当業者に公知である。二酸化炭素窒息によって安楽死させた生後 6 から 8 週間の雌の Sprague Dawley ラット 2 匹から脛骨を収集した。死後直ぐに遠位末端を除去し、近接した脛骨を OCT 包埋培地中液体窒素で急速凍結した。組織を -80℃ のフリーザ中で貯蔵した。

10

20

30

40

50

【0240】

cDNAからのPCR産物を増幅するためのプローブ以下のように調製した。cDNAクローンからPCR産物を増幅するためのプライマーを、ヒトLRP5 (Genbankアクセッション番号AB017498) およびマウスLRP5 (Genbankアクセッション番号AF064984) の双方の公開された配列を用いて選択した。LDLレセプターファミリーの別の遺伝子との交差反応を最低限にするために、PCR産物をタンパク質コード化領域の細胞内部分から誘導した。cDNAクローンを鋳型として用いて反応容量50 µlでPCRを実施した。1. 5mM MgCl₂、Amplitaq 1単位、200 µM dNTPおよび2 µM 各プライマーを含有した。PCR循環条件は94℃で1分間、続いて94℃で30秒間、55度で30秒間、72℃で30秒間の35サイクル；続いて72℃で5分間の伸長であった。次いで1.5%アガロース酢酸トリスゲル上で反応物を流した。DNAをアガロースから溶出し、エタノール沈殿し、10mM トリス、pH 8.0に再懸濁した。ゲル精製PCR産物をマウスおよびヒトcDNAの双方に関して調製し、インサイチュ・ハイブリダイゼーションのためにPathology Associates Internationalに供給した。

【0241】

ヒトおよびマウスPCRプライマーおよび産物の配列は以下のとおりであった。

ヒトLRP5センスプライマー (配列番号633) (HBM253)

CCCGTGTGCTCCGCCGCCAGTTC

ヒトLRP5アンチセンスプライマー (配列番号634) (HBM465)

GGCTCACGGAGCTCATCATGGACTT

ヒトLRP5 PCR産物 (配列番号635)

CCCGTGTGCTCCGCCGCCAGTTCCTGCGCGCGGGTTCAGTGTGTGGACCTGCGCTGCGACGGCGAGGCAGACTGTGCTCAGGACCGCTCAGACGAGGTGGACTGTGACGCCATCTGCCTGCCCAACCAGTTCCGGTGTGCGAGCGGCCAGTGTGTCCTCATCAAAACAGCAGTGCAGTCTCTCCCGACTGTATCGACGGCTCCGACGAGCTCATGTGTGAAATCACCAAGCCGCCCTCAGACGACAGCCCGGCCACAGCAGTGCCATCGGGCCCGTCATTGGCATCATCCTCTCTCTTTCGTCATGGGTGGTGTCTATTTTGTGTGCCAGCGGTGGTGTGCCAGCGCTATCGGGGGCCAACGGGCCCTTCCCGCAGAGTATGTCAGCGGACCCCGCAGCTGCCCCTCAATTTTCATAGCCCCGGCGGTTCCAGCATGGCCCTTCACAGGCATCGCATGCGGAAAGTCCATGATGAGCTCCGTGAGCC

マウスLRP5センスプライマー (配列番号636) (HBM655)

AGCGAGGCCACCATCCACAGG

マウスLRP5アンチセンスプライマー (配列番号637) (HBM656)

TCGCTGGTCGGCATAATCAAT

マウスLRP5 PCR産物 (配列番号638)

AGCAGAGCCACCATCCACAGGATCTCCCTGGAGACTAACAACAACGATGTGGCTATCCCACTCACGGGTGTCAAAGAGGCCTCTGCACTGGACTTTGATGTGTCCAACAATCACATCTACTGGACTGATGTTAGCCTCAAGACGATCAGCCGAGCCTTCA TGAATGGGAGCTCAGTGGAGCACGTGATTGAGTTTGGCCTCGACTACCCTGAAGGAATGGCTGTGGACTGGATGGGCAAG AACCTCTATTGGGCGGACACAGGGACCAACAGGATTGAGGTGGCCCGGCTGGATGGGCAGTTCGGGCAGGTGCTTGTGTG GAGAGACCTTGACAACCCAGGTCTCTGGCTCTGGATCCTACTAAAGGCTACATCTACTGGACTGAGTGGGTGGCAAGC CAAGGATTGTGCGGGCCTTCATGGATGGGACCAATTGTATGACACTGGTAGACAAGGTGGGCCGGGCCAACGACCTCACC ATTGATTATGCCGACCAGCGA

【0242】

以下のようにリボプローブを合成した。再増幅産物の上流にT3プロモーターか、または下流にT7プロモーターのいずれかを組み込むように設計したキメラプライマーでPCR産物を再増幅した。得られたPCR産物を鋳型として用いてインビトロ転写 (IVT) によりジゴキシゲニン標識リボプローブを合成した。ジゴキシゲニン-11-UTP (Boehringer-Mannheim) の存在下MAXIscript IVTキット (Ambion) を用いて、製造者に従って、各々T7およびT3ポリメラーゼを用いてアンチセンスおよびセンスリボプローブを合成した。次いでDNAをDnase-1で消化し、組み込まれなかったジゴキシゲニンを超遠心により除去した。変性ポリアクリルアミドゲルを通す電気泳動によりリ

ボプローブの完全性を評価した。分子の大きさを100から1000塩基対(bp)RNAラダー(Ambion)の電気泳動可動性で比較した。ブロット免疫化学によりプローブの収量および標識を評価した。リボプローブを5μlアリコートで、-80℃で貯蔵した。

【0243】

インサイチュ・ハイブリダイゼーションを以下のように実施した。Jung CM3000 クリオスタット(Leica)で凍結ラット骨を5μm切片に切断し、付着性スライド(Instrumedics)にマウントした。mRNA分解を避けるために全スライドを調製するまで切片をクリオスタット中-20℃で保持し、その後4%パラホルムアルデヒド中15分間の後固定した。後固定の後、1ng/ml アンチセンスまたはセンスのいずれかのリボプローブと共にPathology Associates International(PAI)にカスタマイズされたハイブリダイゼーションバッファー中58℃でおよそ40時間切片をインキュベートした。ハイブリダイゼーションの後、スライドを一連のハイブリダイゼーション後ストリンジェンシー洗浄に供し、非特異的プローブ結合を低減させた。アルカリ性ホスファターゼ結合抗ジゴキシゲン抗体(FABフラグメント)を用いる免疫化学によりハイブリダイゼーションを可視化した。アルカリ性ホスファターゼ沈殿基質であるニトロブルー・塩化テトラゾリウム/リン酸ブロモクロロインドリル(Boehringer-Mannheim)、を色素原として用いて染色の程度に依存してハイブリダイズした細胞を紫色からほぼ黒色まで染色した。組織切片を核ファーストレッドで対比染色した。アッセイ対照はプローブの省略、プローブおよび抗ジゴキシゲン抗体の省略を含めた。

10

20

【0244】

アンチセンスプローブとのハイブリダイゼーションを実証するために、mRNAに関する陽性ハイブリダイゼーションシグナルを示す紫色から黒色の細胞質および/または核周囲染色を可視化することにより特異的な細胞型を評価した。各細胞型を同型切片と比較し、これを各々のセンスプローブでハイブリダイズした。アンチセンスプローブで染色が観察され、センスプローブで染色されないかまたは弱いバックグラウンドであった場合、結果は陽性であると考えられた。

【0245】

各々の研究プローブに関するハイブリダイゼーションシグナルの細胞局在を表5にまとめた。LRP5に関するハイブリダイゼーションが骨幹端内の骨内膜および骨梁骨などのリモデリングに関与した骨の部分で主として検出された。骨膜および骨端の選択された骨裏打ち細胞におけるハイブリダイゼーションもまた観察された。成長プレート内の軟骨細胞において、とりわけ増殖性軟骨細胞においても陽性シグナルが認められた。インサイチュ・ハイブリダイゼーション結果の代表的な顕微鏡写真に関しては図10、11および12を参照されたい。

30

【0246】

【表 5】

表 5 : ラット脛骨におけるLRP5インサイチュ・ハイブリダイゼーションの要約

プローブ	部 位	ISH シグナル
HuZmax1	骨 端	
	骨芽細胞	+
	破骨細胞	—
	成長板	
	休止軟骨細胞	—
	増殖軟骨細胞	+
	肥厚軟骨細胞	—
	骨幹端	
	骨芽細胞	+
	破骨細胞	+
	骨 幹	—
	骨内膜	
	骨芽細胞	+
	破骨細胞	+
	骨 膜	—
MsZmax1	骨 端	
	骨芽細胞	+
	破骨細胞	—
	成長板	
	休止軟骨細胞	—
	増殖軟骨細胞	+
	肥厚軟骨細胞	+
	骨幹端	
	骨芽細胞	+
	破骨細胞	+
	骨 幹	—
	骨内膜	
	骨芽細胞	+
	破骨細胞	+
	骨 膜	+

レジェンド:「+」＝ハイブリダイゼーションシグナルが検出された

「—」＝ハイブリダイゼーションシグナルが検出されない

「ISH」－インサイチュウ・ハイブリダイゼーション

【0247】

これらの研究により骨リモデリングおよび骨形成に関与する細胞におけるLRP5の位置的な発現が確認される。増殖帯ならびに近接する骨幹端の骨芽細胞および破骨細胞におけるLRP5発現により、LRP5遺伝子が骨成長およびミネラル化の過程に関与していることが示唆される。骨芽細胞および破骨細胞の活性および分化は、骨が形成されるような発達中に、および成長中に、ならびに骨が持続的にリモデリングしているような成年期に密接に協調される。骨内部構造の形成および骨リモデリングは活性化された破骨細胞による骨再吸収と、続く骨芽細胞による新規材料の沈着とを合わせた結果である。LRP5はLDLレセプター遺伝子に関連し、従って骨リモデリングの過程における機械感覚および続くシグナル発生に関与するレセプターであり得る。従って、この遺伝子の発現レベルの変化は骨のリモデリングの速度およびミネラル化の程度に影響する。

【0248】

X V. アンチセンス

アンチセンスオリゴヌクレオチドは標的RNAに対して相補的な塩基配列を含有する短い合成核酸である。生存細胞におけるRNAのアンチセンスオリゴヌクレオチドとのハイブリダイゼーションはRNA機能に干渉し、最終的にはタンパク質発現を遮断する。従って、部分配列が解っているいずれかの遺伝子をアンチセンスオリゴヌクレオチドにより標的化させることができる。

【0249】

アンチセンス技術は広く用いられる研究手段になりつつあり、ゲノムシーケンシングの試みにより同定された治療用標的の確証および解明に果たす役割の重要性が増してきている。

10

【0250】

アンチセンス技術は、標的遺伝子をコードするmRNAに相補的なオリゴヌクレオチドを利用することにより遺伝子発現を阻止するために開発された。アンチセンスオリゴヌクレオチドの阻止効果に関していくつか可能なメカニズムがある。中でもRNAアーゼHによるmRNAの分解はタンパク質機能阻止の主要なメカニズムであると考えられている。この技術は元来標的遺伝子の機能を解明するために用いられたが、注意深くそして適切に設計された場合治療適用をも有し得る。

【0251】

アンチセンスオリゴヌクレオチドは、例えば約5、10、15、20、25、30、35、40、45または50ヌクレオチドの長さでよい。本発明のアンチセンス核酸を、当分野で公知の手順を用いてキメラ合成および酵素的ライゲーション反応を用いて構築することができる。例えば、アンチセンス核酸（例えばアンチセンスオリゴヌクレオチド）を天然発生ヌクレオチドまたは分子の生物学的安定性を増強するか、またはアンチセンスおよびセンス核酸間で形成された二本鎖の物理学的安定性を増強するように設計した種々に修飾したヌクレオチドを用いて、アンチセンス核酸（例えばアンチセンスオリゴヌクレオチド）を化学的に合成することができ、例えばホスホロチオネート誘導体およびアクリジン置換ヌクレオチドを用いることができる。アンチセンス核酸を作製するために用いることができる修飾されたヌクレオチドの実例としては、5-フルオロウラシル、5-ブロモウラシル、5-クロロウラシル、5-ヨードウラシル、ヒポキサンチン、キサンチン、4-アセチルシトシン、5-（カルボキシヒドロキシメチル）ウラシル、5-カルボキシヒドロキシメチルアミノメチル-2-チオウリジン、5-カルボキシメチルアミノメチルウラシル、ジヒドロウラシル、 β -D-ガラクトシルクエオシン、イノシン、N6-イソペンテニルアデニン、1-メチルグアニン、1-メチルリノシン、2,2-ジメチルグアニン、2-メチルアデニン、2-メチルグアニン、3-メチルシトシン、5-メチルシトシン、N6-アデニン、7-メチルグアニン、5-メチルアミノメチルウラシル、5-メトキシアミノメチル-2-チオウラシル、 β -D-マンノシルクエオシン、5'-メトキシカルボキシメチルウラシル、5-メトキシウラシル、2-メチルチオ-N6-イソペンテニルアデニン、ウラシル-5-オキシ酢酸(v)、ウイブトキソシン、プセウドウラシル、クエオシン、2-チオシトシン、5-メチル-2-チオウラシル、2-チオウラシル、4-チオウラシル、5-メチルウラシル、ウラシル-5-オキシ酢酸メチルエステル、ウラシル-5-オキシ酢酸(v)、t-メチル-2-チオウラシル、3-（3-アミノ-3-N-2-カルボキシプロピル）ウラシル、（acp3）w、および2,6-ジアミノプリンなどがある。

20

30

40

【0252】

加えて、モルフォリノオリゴヌクレオチドを使用することができる。モルフォリノはリボース部分をモルフォリノ基に修飾したオリゴマーである。この技術は米国特許第5,185,444号に網羅されており、そしてSummertonおよびWeller、Antisense Nucleic Acid Drug Dev.7(3):187-195(1997年6月)に記載されている。HBMもしくはLRP5タンパク質または、LRP5および／またはHBMと相互作用してそれにより、例えば転写および／ま

50

たは翻訳を阻止することによりタンパク質の発現を阻止するタンパク質をコードする細胞性mRNAおよび／またはゲノムDNAにハイブリダイズするかまたは結合するように、本発明のアンチセンス核酸分子を典型的には対象に投与するか、またはインサイチュで生成する。ハイブリダイゼーションは安定した二本鎖を形成するために慣用されるヌクレオチド相補性による、例えばDNA二本鎖に結合するアンチセンス核酸分子の場合、二重らせんの主要な溝での特異的相互作用を介するものである。本発明のアンチセンス核酸分子投与経路の実例としては組織部位での直接注射などがある。また別に、アンチセンス核酸分子を修飾して選択した細胞を標的化し、次に全身的に投与することができる。例えば、全身投与用には、アンチセンス分子を、レセプターまたは選択された細胞表面で発現される抗原に特異的に結合するように、例えばアンチセンス核酸分子を細胞表面レセプターまたは抗原に結合するペプチドまたは抗体に連結させることにより修飾することができる。アンチセンス核酸分子を本明細書に記載するベクターを用いて細胞に分配することもできる。

【0253】

さらに別の実施形態では、本発明のアンチセンス核酸分子は α -アノマー核酸分子である。 μ -アノマー核酸分子は相補性RNAと特異的二本鎖ハイブリッドを形成し、通常の γ -ユニットとは反対に、鎖は互いに並行に進む(Gaultierら、Nucleic Acids.Res.15:6625-6641(1987))。アンチセンス核酸分子はまた2'- α -メチルリボヌクレオチド(Inoueら、Nucleic Acids.Res.15:6131-6148(1987))またはキメラRNA-DNAアナログ(Inoueら、FEBS Lett.215:327-330(1987))を含むこともできる。さらに別の実施形態では、本発明のアンチセンス核酸分子はリボザイムである。リボザイムは、それに対する相補性領域を有する一本鎖核酸を切断できるリボヌクレアーゼ活性を有する触媒性RNA分子、例えばmRNAである。従って、リボザイム(例えばハンマーヘッドリボザイム(HaselhoffおよびGerlach、Nature 334:585-591(1988))に記載))を用いてLRP5またはHBM mRNA転写物を切断し、それによりLRP5またはHBM mRNAの翻訳を阻止することができる。LRP5またはHBMコード化核酸に特異性を有するリボザイムを、本明細書に開示するLRP5またはHBM cDNA(すなわち、配列番号1または3)のヌクレオチド配列に基づいて設計することができる。例えば、活性部位のヌクレオチド配列がHBMまたはLRP5コード化mRNAにおいて切断されるヌクレオチド配列に相補的であるTetrahymena L-19 IVS RNAの誘導体を構築することができる。例えば、双方共に参照により本明細書に組み込むCechら、米国特許第4,987,071号およびCechら、米国特許第5,116,742号を参照されたい。また別に、LRP5またはHBM mRNAを用いてRNA分子のプールから特異的リボヌクレアーゼ活性を有する触媒性RNAを選択することができる。例えばBartel,D.およびSzostak,J.W.、Science 261:1411-1418(1993)を参照されたい。また別に、LRP5またはHBM遺伝子の制御領域(例えばLRP5またはHBM遺伝子プロモーターおよび／またはエンハンサー)に相補的であるヌクレオチド配列を標的化して、LRP5またはHBM遺伝子の標的細胞における転写を阻止する三重らせんを形成することによりLRP5またはHBM遺伝子発現を阻止することができる。一般的にはHelene,C.、Anticancer Drug Des.6(6):569-584(1991);Helene,C.ら、Ann.N.Y.Acad.Sci.660:27-36(1992);およびMaher,L.J.、Bioassays 14(12):807-815(1992)を参照されたい。RNA干渉(RNAi)を用いてLRP5またはHBM遺伝子発現を阻止することもできる。これは転写後遺伝子サイレンシング(PTGS)のための技術であり、ここで標的遺伝子活性を同族二本鎖RNA(dsRNA)で特異的に消滅させる。RNAiは多くの態様で植物のPTGSに類似し、そしてトリパノソーマ、ヒドラ、プラナリア、線虫およびショウジョウバエ(Drosophila melanogaster)などの多くの無脊椎動物において検出されている。これは転移エレメント可動性および抗ウイルス状態形成の変調に関与し得る。哺乳動物系のRNAiはPCT出願W000/63364に開示されており、これは参照によりその全てを本明細書に組み込む。基本的には、標的(LRP5またはHBM)に相同的であるdsRNAを細胞に導入し、そして遺伝子活性の配列特異的低減を観察する。小型および干渉性の双方のRNA(siRNA)および短いヘアピンRNA(shRNA)が企図される

。例えば Yu ら、PNAS 99:6047-6052(2002);Paddison ら、Genes and Development 16:948-958(2002);Brummelkamp ら、Science 296:550-553(2002);Tuschl, Nature Biotechnology 20:446-448(2002);およびその参照文献を参照されたい。

【0254】

実例として、アンチセンスオリゴヌクレオチドの調製を以下のように実施できる。骨芽細胞様ネズミ細胞株、MC3T3のアンチセンス技術を用いて研究を実施した。これらの細胞を誘発して骨分化配列と共に発達させることができる。初期増殖期は分化マーカーの最低限の発現およびコラーゲン性細胞外マトリックスの初期合成により特徴付けられる。分化マーカーの導入後、コラーゲンマトリックス合成が必要とされる。一度コラーゲンマトリックス合成が開始されると、骨芽細胞マーカー遺伝子が明確な一過性配列で活性化される。アルカリ性ホスファターゼが初期に誘起されるが、一方骨シアロタンパク質 (bone sialoprotein) およびオステオカルシンが分化過程の後記で出現する。遺伝子発現のこの一過性配列は成熟およびミネラル化過程をモニター観察するのに有用である。マトリックスミネラル化は、成熟化が開始した後数日までは始まらず、細胞層—培養プレートの界面近くのマトリックス内に深在するコラーゲン原繊維上またはその内部でのミネラルの沈着に関与する。培養骨芽細胞により形成されたコラーゲン原繊維随伴ミネラルはインビボで網状骨に見出されるものに類似し、従って研究用試薬として汎用される。

10

【0255】

製造者の説明書に従って（米国特許第5,849,902号）、分化の最初の週にMC3T3細胞をアンチセンスオリゴヌクレオチドでトランスフェクトした。

20

【0256】

LRP5 (Zmax1) のために設計したオリゴヌクレオチドを以下に提示する（配列番号639～641）、

10875: AGUACAGCUUCUUGCCAACCCAGUC

10876: UCCUCCAGGUCGAUGGUCAGCCCAU

10877: GUCUGAGUCCGAGUUCAAAUCAGG

【0257】

図13はMC3T3細胞におけるLRP5のアンチセンス阻止の結果を示す。前記で示した3つのオリゴヌクレオチドをMC3T3にトランスフェクトし、標準的な手順に従ってRNAを単離した。ノーザン分析によりLRP5転写物の著明に低下して安定な状態のレベルが明確に示されるが、一方対照遺伝子GAPDHは未変化のままであった。このように、前記したプライマーを用いるアンチセンス技術により骨生物学におけるLRP5発現の役割の研究が可能になる。

30

【0258】

XVI. 酵母2ハイブリッド

骨密度を変調するためにLRP5が関与するシグナル発生経路を同定するために、酵母2ハイブリッドタンパク質相互作用技術を利用した。この技術によりテストタンパク質が酵母転写系に結合することにより互いに相互作用するタンパク質の同定が促進される (FieldsおよびSong、Nature 340:245-246(1989);FieldsおよびSongによる米国特許第5,283,173号;Johnston、Microbiol.Rev.51:458-476(1987);Keegan ら、Science 231:699-704(1986);Durfee ら、Genes Dev.7:555-569(1993);Chien ら、Proc.Natl.Acad.Sci.USA 88:9578-9582(1991);Fields ら、Trends in Genetics 10:286-292(1994);ならびにGyuris ら、Cell 75:791-803(1993))。第1に、相互作用するタンパク質を探し求めるタンパク質である「ベイト」タンパク質を酵母転写因子のDNA結合ドメインに融合させる。第2に、同一の酵母転写因子の転写活性化ドメインに融合するcDNAを含有するcDNAライブラリーを構築し；これをプレイライブラリーと称する。ベイト構築物およびプレイライブラリーを酵母細胞に形質転換し、次に接合させて二倍体細胞を生成する。ベイトがcDNAライブラリーからの特異的プレイと相互作用する場合、活性化ドメインはこの相互作用を介してプロモーターの近傍にもたらされる。次いで選択マーカー遺伝子により転写が誘導され、選択培地上での成長が相互作用タンパク質の存在を示す。

40

50

【0259】

本明細書で論じる酵母2ハイブリッド実験で用いるアミノ酸配列は細胞質ドメイン全体および膜貫通ドメインの一部から成り、これを以下に示す（アミノからカルボキシへの配向）（配列番号765）：

R V V C Q R Y A G A N G P F P H E Y V S G T P H V P L N F I A P G G S Q H
G P F T G I A C G K S M M S S V S L M G G R G G V P L Y D R N H V T G A
S S S S S S S T K A T L Y P P I L N P P P S P A T D P S L Y N M D M F Y S
S N I P A T V R P Y R P Y I I R G M A P P T T P C S T D V C D S D Y S A
S R W K A S K Y Y L D L N S D S D P Y P P P P T P H S Q Y L S A E D S C
P P S P A T E R S Y F H L F P P P P S P C T D S S

10

【0260】

推定される膜貫通ドメインの最後の6個のアミノ酸を太字で示す。推定SH3ドメインには下線を付している。LRP5またはHBM対立遺伝子によりコードされるタンパク質のいずれかの50個またはそれ以上のアミノ酸の別のアミノ酸配列をベイトして用いることもできる。ベイトとして用いられるポリペプチドの上限の大きさは完全な膜貫通ドメインの存在によってのみ限定され（図4参照）、これによりベイトは酵母2ハイブリッド系における無機能性が付与される。これらのさらなるベイトタンパク質を用いて、限局性付着シグナル発生経路か、またはこれらのHBMまたはLRP5タンパク質が作用し得る別の経路において、HBMまたはLRP5によりコードされるタンパク質と相互作用するさらなるタンパク質を同定することができる。一度同定すると、限局性付着シグナル発生経路か、またはこれらのHBMが作用し得る別の経路においてタンパク質を制御する作用物質を同定する方法を、HBMおよびLRP5タンパク質に関して本明細書に記載するように実施することができる。

20

【0261】

細胞質LRP5シグナル発生経路を同定するために、LRP5の細胞質ドメインをベイトベクターにサブクローニングした。第1のベクターはpDBLeuであり、これを用いて脳をスクリーニングし、HeLaプレイcDNAライブラリーをベクターpPC86(Clontech)にクローニングした。第2のベイトベクターはpDBtrpであり、これを用いてベクターpOP46のTE85骨肉種細胞株に由来するcDNAプレイライブラリーをスクリーニングした。広く利用可能な別の適当なベクターはp86(Gibco, iest(商標)システム)である。当業者に公知の標準的な技術をFieldsおよびSong、Nature 340:245-246(1989);FieldsおよびSongによる米国特許第5,283,173号;Johnston、Microbiol.Rev. 51:458-476(1987);Keeganら、Science 231:699-704(1986);Durfeeら、Genes Dev.7:555-569(1993);Chienら、Proc.Natl.Acad.Sci.USA 88:9578-9582(1991);Fieldsら、Trends in Genetics 10:286-292(1994);ならびにGyurisら、Cell 75:791-803(1993))に記載されるように用いることができる。ベイト構築物およびプレイcDNAライブラリーを標準的な手段を用いて酵母に形質転換した。

30

【0262】

タンパク質相互作用スクリーニングを実施するために、ベイト酵母株の一晩培養物を2%グルコースを含むSD選択培地（pDBLeu、SD-Leu培地、pDBtrp、SD-trp培地）20ml中で成長させた。培養物を30℃で一晩激しく振盪した。培養物を完全培地（2%グルコースを含むYEPD）で1:10に希釈し、次に培養物を振盪しながら30℃で2時間インキュベートした。

40

【0263】

凍結したプレイライブラリーを解凍し、酵母細胞を2%グルコースを含むYEPD培地150ml中30℃で2時間、それらを成長させることにより再活性化した。フィルターユニットを70%エタノールで滅菌し、滅菌水で洗浄してエタノールを除去した。ベイトおよびプレイ双方の培養物の細胞密度を600nmでのODを決定することにより測定した。OD600の1ml=4つの各酵母株の細胞数に相当する適当な容量の酵母細胞、ベイトおよびプレイ（ライブラリー）を50mlFalconチューブに置いた。次いで混

50

合物を滅菌フィルターユニットを通して濾過した。次いでフィルターを予め加温したYEPD寒天プレートに細胞側を上にして移し、フィルターの下部の気泡を全て除去した。次いでプレートを30℃で6時間インキュベートした。1つのフィルターを50ml Falconチューブに移し、2%グルコースを含むSD 10mlを加えた；細胞をボルテックスしながら10秒間再懸濁した。

【0264】

1次二倍体細胞(SD-Leu、-Trpプレート上で成長)の数対SD-TrpおよびSD-Leuプレート上でのみ成長するコロニー形成ユニットの数を決定した。異なる希釈物をプレートし、30℃で2日間インキュベートした。次いでコロニー形成ユニットの数を計数した。二倍体コロニーの数(SD-Leu-Trpプレート上のコロニー)によりプレイ構築物のライブラリー全体がベイトを発現する酵母と接合したかどうかの算出が可能になる。この情報はスクリーニングの品質を判定するために重要である。

10

【0265】

A. 間接的選択

次いで5つのフィルターメーティングからの懸濁細胞をプールし、細胞を遠心により50ml Falconチューブに沈殿させた。次いで細胞を2%Glcを含むSD培地16mlに再懸濁した。滅菌ガラスビーズを伴う8スクエアプレートの各々(SD-Leu、-Trp)にこの細胞懸濁液2mlをプレートし、30℃で18から20時間インキュベートすることにより二倍体を選択した。

【0266】

次いで細胞をスクエアプレートからかき取り、細胞を遠心し、1本の50ml Falconチューブに合わせた。次いで細胞ペレットを2%グルコースを含むSD培地25mlに再懸濁した。次いでNeugebauerチャンバーで適当な希釈物(通常1:100から1:1000)を計数することにより細胞数を決定した。およそ 5×10^7 個の二倍体細胞を選択培地にプレートした。無関係のプレイベクターと組んだベイト株の成長について観察することが、どの選択プレートをライブラリースクリーニングに関して選択しなければならないかの決定の助けとなる。一般に、全てのスクリーニングをSD-Leu、-Trp、-His；SD-Leu、-Trp、His、5mM 3ATを伴うスクエアプレートの各々にプレートし、SD-Leu、-Trp、-His、-Adeが推奨される。

20

【0267】

次いで酵母細胞を滅菌ガラスビーズと共に均質に広げ、30℃で4日間インキュベートした。SD-Leu、-Trpプレートにかき取った細胞懸濁液の異なる希釈物をプレートすることにより酵母細胞を形成するコロニーの数を決定した。通常 10^{-3} および 10^{-4} 希釈物100μlのプレートによりプレートあたり100から1000個のコロニーが得られた。

30

【0268】

B. 直接選択

接合した酵母細胞を伴う5つのフィルターを各々別個の50ml Falconチューブに移し、細胞を2%Glcを含むSD培地10mlに各々再懸濁し、続いて10秒間ボルテックスした。再懸濁した細胞を合わせ、Beckman遠心を用いて3000rpmで遠心した。上澄を捨て、2%Glcを含むSD培地6mlに細胞を再懸濁した。懸濁液2mlを各選択スクエアプレートに広げ、30℃で4から5日間インキュベートした。

40

【0269】

C. 単一のコロニーの単離

単離されたコロニーからの酵母細胞を滅菌トウスポックで採取し、96ウェルプレートの個々のウェルに移した。細胞をSD-Leu、-Trp、-His培地50μlに再懸濁し、30℃で1日間インキュベートした。次いで酵母細胞を96ウェル様式のSD-Leu、-Trp、-Hisプレートにスタンプし、30℃で2日間インキュベートした。酵母細胞をナイロンフィルター被覆YEPDプレートにもスタンプし、30℃で1日間インキュベートした。ナイロンフィルター上の細胞をβ-Galレポーター活性の分析に

50

用いた。

【0270】

滅菌トウスポックでSD-Leu、-Trp、-Hisプレートから酵母コロニーをかき取り、必要な場合、 β -Gal活性に従って、必要な場合再設定し、次に20%グリセロールに再懸濁した。これは-80℃で貯蔵するためのマスタープレートとして提供された。

【0271】

DNA調製に関しては、グリセロールストックからの酵母細胞をSD-Trpプレートにスタンプし、30℃で2日間インキュベートした。インキュベーションの2日後、酵母コロニーはコロニーPCRおよびシーケンシングに備えられていた。標準的なコロニーPCR条件を用いて、相互作用スクリーニングから回収したプレイからのインサートを増幅した。標準的なシーケンシング反応およびABI377(Perkin Elmer)蛍光シーケンシング機械を用いてシーケンシングを行った。

【0272】

D. バイト/プレイ相互作用の確認

目的のプレイのグリセロールストックを解凍し、グルコース-Trpを含むSDの一晩培養物10mlに接種した。一晩成長させた後、BIO 101 RPM 酵母プラスミド単離キットを培養物10mlと共に用いてプラスミドDNA調製を実施した。培養物を遠心し、1.5mlマイクロ遠心チューブに移した。次いで酵母溶解マトリックスをペレットに加え、続いてアルカリ性溶解溶液250 μ lを加えた。次いでサンプルを5分間ボルテックスした。中和溶液250 μ lを加え、サンプルを短時間混合した。マイクロ遠心でサンプルを室温で2分間遠心した。上澄をスピンフィルターに移して細片および溶解マトリックスを回避した。ガラスミルク・スピンバッファ250 μ lを加え、チューブを反転させて混合した。サンプルを1分間遠心し、キャッチ・チューブの液体を捨てた。洗浄溶液500 μ lを加え、サンプルを1分間遠心し、洗浄溶液を捨てた。洗浄工程を1回繰り返し、続いて1分間乾燥遠心してスピンフィルターの外に残りの液体を追いやった。フィルターを新たなキャッチ・チューブに移し、滅菌H₂O 100 μ lを加え；次いでサンプルを短時間ボルテックスして再懸濁し、30秒間遠心してキャッチ・チューブの底部にDNAを回収した。

【0273】

次いで標準的な手順を用いてDNA5 μ lをDH10B Electromax 細胞に形質転換し、グリセロールストックを調製した。Qiagen QIAprep スピン・ミニプレップ・キットを用いてミニプレップDNAを調製した。最後にDNAをQiagen EBバッファ30 μ lで溶出した。次いで標準的な手順を用いて、プラスミドDNAサンプル1 μ lを用いて酵母細胞を形質転換した。SD-Trp培地上で2日間成長させた後、コロニーを採取し、新鮮培地にパッチした。同様に、バイトコロニーをSD-Leu培地にパッチした。双方を30℃で一晩成長させた。

【0274】

接合のために、バイトおよびプレイパッチからの細胞と一緒にYAPD培地上に広げ、30℃で12時間インキュベートした。次いでこのプレートをSD寒天Leu-Trpプレートにレプリカプレートし、30℃で2日間成長させた。相互作用の力を試験するために、これらのプレートをSD寒天Leu-Trp-His、5mM 3ATおよび10mM 3ATを含むSD寒天Leu-Trp-His、SD寒天Leu-Trp-His-Ade、およびSD寒天Leu-Trp-Ura培地上にレプリカプレートし、30℃で2日間成長させた。

【0275】

E. Galacton Star β -ガラクトシダーゼ活性アッセイ

選択プレート上での画線および陽性インタラクターのレプリカプレートの後、SD培地200 μ lが入った96ウェル皿にコロニーを置き、ウェル1および96をブランクにしておいた。最初の96ウェル皿からの10 μ lを、SD培地100 μ lが入った別の平底

10

20

30

40

50

96 ウェル皿に置いた。対照は陰性対照および非常に弱い陽性対照から成った。OD₆₀₀ (96 ウェル分光光度計を利用して1の値は 1×10^7 セルに相当する)で細胞密度を測定した。通常ODは0.03から0.10の間であった。特に照度計のためのマイクロプレートを用いて、反応混合物50 μ lを各ウェルにピペティングした。次いで培養物50 μ lを加え、ピペティングを上下に2回行って混合した。反応物を室温で30分間インキュベートして、続いて照度計を用いて相対光単位を測定した。

【0276】

表6は試験した3つのプレライブラリーからの酵母2ハイブリッドスクリーニングで同定された遺伝子を列挙する。2つの遺伝子、zyxinおよびaxinは3つ全てのスクリーニングのLRP5の細胞性ドメインと相互作用することが見出された。3つの遺伝子、 α -アクチニン、TCBおよびS1-5は3つのスクリーニングの2つと相互作用した。

10

【0277】

細胞-細胞および細胞-マトリックス接触(限局的接触/付着板)の部位で見出された種々のタンパク質がLRP5の細胞質ドメインと相互作用することが示された。これには α -アクチニン、Trio、Pinch様タンパク質およびZyxinなどが含まれる。PINCHはインテグリン結合キナーゼ、インテグリンの初期シグナラーおよび成長因子シグナル発生経路と相互作用することが解っているLIMドメイン含有タンパク質である。酵母2ハイブリッドスクリーニングにおいて密接に関連する遺伝子を見出すことにより細胞外マトリックスシグナルからインテグリンシグナル発生に連結される新規の経路の可能性が出現する。Trioもまた限局性付着に局在することが解っており、細胞-マトリックス相互作用および細胞移動に関与する細胞骨格再編成を協調させるのに重要な役割を果たすと考えられている。Zyxinはもう一つのLIMドメイン含有タンパク質であり、これもまた付着板に局在し、そしてインテグリンシグナル発生を介して誘引が伝達されたときに、細胞骨格の再編成に関与すると考えられている。Zyxinもまた α -アクチニンと相互作用し、これはLRP5と相互作用することが同定された。別の同定されたLIMドメイン含有タンパク質にはマウスアジュバのヒト相同体、LIMD1および新規のLIMD1様タンパク質などがある。

20

【0278】

アクシンもまた2ハイブリッド実験から同定した。このタンパク質はWntシグナル発生経路の阻止に関与し、腫瘍抑制剤APCと相互作用する。ここでは前記した限局性付着シグナル発生との関連がある。2つの経路の1つの一般的な工程はグリコーゲンシンターゼキナーゼ3の阻止に関与し、これが今度は β -カテニン/Lef-1およびAP-1転写因子の活性化に至る。アクシン/APCはこれおよびインテグリン結合キナーゼに関与する。Wnt経路は胚形成中の細胞運命を決定するのに役割を有している。不適切に活性化された場合、Wnt経路もまた癌を導き得る。Wnt経路もまた細胞骨格再編成において役割を有しているようである。アフリカツメガエル胚アッセイでは、HBMおよびWnt5aタンパク質の組み合わせが、前記対照およびWnt5a単独スコアを超えて穏やかであったLRP5およびWnt5aの組み合わせよりも非常に大きな程度でWnt経路を刺激した。HBMおよびLRP5細胞外ドメイン(ECD)はWnt5aの不在下でWntシグナル発生の穏やかな刺激を引き起こし、これはHBM ECDの存在下でWnt5aの存在によりわずかに増加した。限局性付着シグナル発生におけるLRP5関与を表すモデルを図15に示す。

30

40

【0279】

このデータを、インテグリンシグナル発生経路が機械的ストレスおよび付着に対する細胞応答における役割を有していることを示唆する別の研究と合わせた。これにより骨生物学におけるLRP5の作用のメカニズムに関する魅力的なモデルが提供される。機械的ストレスの直接的な検出か、または機械的感覚に関連する細胞外マトリックスの分子の結合のいずれかにLRP5が関与することが可能である。続く経路によるシグナル発生は骨細胞の細胞形態学、細胞付着、遊走、増殖、分化、およびアポトーシスに及ぼす影響のため

50

に骨リモデリングに関与し得る。

【0280】

【表6】

表6： 酵母2ハイブリッドの結果

遺伝子記号	遺伝子	ジェンバンク 受け入れ番号	ヌクレオチド 配列番号：	アミノ酸 配列番号：
ACTN1	アルファ・アクチニン	NM_001102	642	
AES	アミノ末端エンハンサー	NM_001130.	643	
AIP4	アトロフィン-1 相互作用タンパク質	AF038564.1	644	
Novel	Ajuba		645	
AXIN	Wnt シグナル発生	AF009674.1	646	
CDC23	細胞分割サイクル 23、酵母	NM_004661.	647	
HSM800944	TRIO に類似	AL117435.1	648	
HSM800936		AL117427.1	649	
Novel	LIM ドメインに類似		650	
DEEPEST	紡錘体コイルドコイル	NM_006461.	651	
ECM1	細胞外マトリックスプロテイン 1	U65932.1	652	
EF1A	伸長因子 1-アルファ	X16869.1	653	
FN	フィブロネクチン	X02761.1	654	
HOXB13	ホメオドメインタンパク質	U81599.1	655	
Novel	Glu-Lys 富化タンパク質		656	
LIMD1	LIM ドメイン含有 1	NM_014240.	567	
Novel	PINCH 様		568	
RANBPM	中心体タンパク質	NM_005493.	659	
S1-5	細胞外タンパク質	U03877.1	660	
TCB	遺伝子コード化サイトゾル	M26252.1	661	
TID	腫瘍状成虫盤	NM_005147.	662	
ZYX	Zyxin	NM_003461.	663	
TRIO	GTP アーゼ	U42390.1	664	
HUMPITPB	ホスファチジルイノシトール転移	D30037.1	665	
ACTN1	アルファ・アクチニン	NP_001093.1		666
AES	アミノ末端エンハンサー	NP_001121.2		667
AIP4	アトロフィン-1 相互作用タンパク質	AAC04845.1		668
Novel	Ajuba			669
AXIN	Wnt シグナル発生	AAC51624.1		670
CDC23	細胞分割サイクル 23、酵母	NP_004652.1		671
Novel	TRIO に類似 CAB55923.1			672
Novel	LIM ドメインに類似			673
DEEPEST	紡錘体コイルドコイル	NP_006452.1		674
ECM1	細胞外マトリックスプロテイン 1	AAB05933.1		675
EF1A	伸長因子 1-アルファ	CAA34756.1		676
FN	フィブロネクチン	CAA26536.1		677
Novel	Glu-Lys 富化タンパク質			678
HOXB13	ホメオドメインタンパク質 B13	AAB39863.1		679
LIMD1	LIM ドメイン含有 1	NP_055055.1		680
Novel	PINCH 様			681
RANBPM	中心体タンパク質	NP_005484.1		682
S1-5	細胞外タンパク質	AAA65590.1		683
TCB	サイトゾル甲状腺ホルモン	AAA36672.1		684
TID	腫瘍状成虫盤	NP_005138.1		685
ZYX	Zyxin	NP_003452.1		686
TRIO	GTP アーゼ	AAC34245.1		687
PTDINSTP	ホスファチジルイノシトール転移	P48739		688

【0281】

図15に表すモデルおよび表6に示す結果に鑑みて、本発明により意図される別の態様は、限局性付着シグナル発生を制御することにより骨密度および骨量障害制御することである。該制御はDNA、mRNA転写物または酵母2ハイブリッド系により同定されるような限局性付着シグナル発生経路に関与するいずれかのメンバーによりコードされるタンパク質を制御することにより制御を発生させることができる。

【0282】

HBM酵母2ハイブリッド系により同定される新規核酸およびタンパク質もまた企図される。これには、非限定例としては配列番号645 (A j u b a)、配列番号651 (プロテイン1を含有するL I Mドメインをコードする遺伝子に類似の遺伝子)、配列番号656 (G l u - L y s 富化タンパク質)、配列番号658 (P I N C H様遺伝子)、配列番号669 (A j u b aタンパク質)、配列番号672 (T R I Oに類似するタンパク質)、配列番号673、配列番号678 (G l u - L y s 富化タンパク質) および配列番号681 (P I N C H様タンパク質) などがある。

【0283】

X V I I . L R P 5 / L R P 6 およびHBM機能

10

最近の研究ではL R P 5がW n t シグナルトランスダクション経路に関与することが示されている。Gongら、は最近骨発達におけるL R P 5の役割をさらに支持する結果を公表している (Gongら、Cell 107:513-523(2001))。G o n g および共同研究者による研究では常染色体劣性障害骨粗鬆症・偽性神経腫症候群 (O P P G) を引き起こすL R P 5の変異について記載している。彼らはO P P GがL R P 5機能喪失により引き起こされと結論づけ、L R P 6をW n t 経路における重複性のレセプターとして意味づけている。最近、L R P 5機能喪失は低骨量表現型に至ることが示されている (Katoら、J.Cell.Biol.157:303-314(2002))。

【0284】

W n t 経路は肢初期発生学的発達において重要である。Nusse、Nature 411:255-256(2001);およびMaoら、Nature 411:321-325(2001)。W n t タンパク質は膜貫通タンパク質F r i z z l e d (F z) と相互作用する分泌タンパク質である。L R P タンパク質、例えばL R P 5 およびL R P 6 はF z との複合体でW n t シグナルを変調すると考えられている (Tamaiら、Nature 407:530-535(2000))。W n t 経路はD i s h e v e l e d タンパク質 (D s h) を介して細胞内で作用し、今度はリン酸化β-カテニンからのグリコーゲンシンターゼキナーゼ3 (G S K 3) を阻止する。リン酸化β-カテニンはユビキチン化の後急速に分解される。しかしながら安定化β-カテニンは核に蓄積および移動して、そこでT細胞因子 (T C F) 転写アクチベーター複合体のコファクターとして作用する。

20

【0285】

タンパク質d i c k k o p f - 1 (D k k - 1) は、初期発達の頭部形成に必要なW n t 経路のアンタゴニストである (Glinkaら、Nature 391:357-362(1998))。D k k - 1 およびW n t 経路におけるその機能は例えばKrupnikら、Gene 238:301-313(1999);Fediら、J. Biol.Chem.274:19465-19472(1999);Dkk-1およびWnt経路に関してはWuら、Curr.Biol.10:1611-1614(2000)、Shinyaら、Mech.Dev.98:3-17(2000)、Mukhopadhyayら、Dev.Cell 1:423-434(2001)およびPCT特許出願番号W000/52047並びに各々に引用されている参考文献に記載されている。Dkk-1はDshの上流で作用することが解っているが、D k k - 1 による阻止のメカニズムの特性はちょうど解明され始めたばかりである。D k k - 1 はマウス胚性肢芽において発現され、その破壊は肢形態形成異常、とりわけ発生欠損に至る (Gotewoldら、Mech.Dev.89:151-153(1999);およびMukhopadhyayら、Dev.Cell 1:423-434(2001))。

30

【0286】

D k k - 1 とL R P 5 との間の相互作用は、2001年5月17日出願の米国特許出願第60/291,311号;2002年2月1日出願の第60/353,058号、および2002年3月4日出願の第60/361,293号で本発明者により開示された実験でL R P 5 のリガンド結合ドメインと相互作用するタンパク質のための、酵母2ハイブリッド (Y 2 H) スクリーニングにより見出された。2ハイブリッドスクリーニングは当分野で一般的な手順であり、これは例えばGietzら、Mol.Cell.Biochem.172:67-79(1997);Young、Biol.Reprod.58:302-311(1998);BrentおよびFinley、Ann.Rev.Genet.31:663-704(1997);並びにLuおよびHannon編、Yeast Hybrid Technologies、Eaton Publishing、Natick、マサチューセッツ州(2000)に記載されている。さらに最近では、別の研究によりD k k - 1 がL R P の結合パートナーであり、W n t 経路をL R P との直接結合を介して変調することが確認されている (R.Nusse、Nature 411:255-256

40

50

(2001);A.Baficoら、Nat.Cell Biol.3:683-686(2001);M.Semenov、Curr.Biol.11:951-961(2001);B.Mao、Nature 411:321-325(2001)、Zorn、Curr.Biol.11:R592-595(2001)ならびにL.Liら、J.Biol.Chem.277:5977-5981(2002))。

【0287】

Maoおよび共同研究者(2001)はDkk-1をLRP6のリガンドとして同定し、Dkkタンパク質がLRP6のWntコレセプター機能を阻止するという点で、Dkk-1およびLRP6が拮抗的に相互作用することを示唆している。そのグループは同時免疫沈澱を用いて、Dkk-1/LRP6相互作用が直接であることを証明した。Dkk-2もまた直接LRP6に結合することが見出された。しかしながら、Maoらは、いずれかのDkkタンパク質とLRP5の間では相互作用が観察されず、LDLR、VLDLR、ApoeERまたはLRPとの相互作用も見出されなかったことを報告している。加えて、Maoらは市販のTCF-シフェラーゼレポーター遺伝子アッセイ(TOPFLASH)を用いて、LRP6がWntシグナル発生を阻止のDkk-1の効果を決定できることを実証した。類似の結論がアフリカツメガエル胚の類似の研究から引き出された。LRP6機能的ドメインの欠失分析により、EGF反復(β プロペラ)3および4がDkk-1結合に必要であり、LRP6のリガンド結合ドメインはDkk-1結合に影響しないことが示された。Maoらの知見は本発明者により得られたデータとは対照的であり、これはLRP5のリガンド結合ドメインが酵母におけるDkk-1結合に必要であり、十分であることを示している。古典的な生化学的リガンド・レセプター研究を用いて、MaoらはDkk-1/LRP6に関して $K_d = 0.34 \text{ nM}$ 、およびDkk-2/LRP6に関して $K_d = 0.73 \text{ nM}$ を決定した。

【0288】

Semenovら(2001)はMaoのグループの結果を証明し、Dkk-1がWntまたFrizzledに直接結合しないが、むしろLRP6と相互作用することを同時免疫沈澱により確認した。そのスクッチャード分析によりDkk-1/LRP6に関して $K_d = 0.5 \text{ nM}$ が見出された。SemenovらはまたDkk-1がLRP5/Frizzled8複合体を排除し得ることを実証し、これはDkk-1がLRP5との相互作用を介してWntシグナル発生を抑制することもできることを意味している。システイン220がアラニンに変化しているDkk-1変異体はLRP6結合を排除し、Wntシグナル発生を抑制できない。アフリカツメガエル胚の研究によりその結果が確認され、Dkk-1/LRP6の機能的な結果：Wntシグナル発生の抑制が示された。彼らのアフリカツメガエル研究により、LRP6/Dkk-1が、Wnt平面内細胞極性形成経路と相対して、標準的な β -カテニン媒介によるWnt経路に特異的であり得ることもまた示唆された。

【0289】

Baficoら(2001)は¹²⁵I標識Dkk-1分子を用いて、 $K_d = 0.39 \text{ nM}$ であるその唯一の膜レセプターとしてLRP6を同定した。また、Dkk-1/LRP6相互作用の機能的結果は、Dkk-1が極度に低濃度(30 pM)で加えられた場合でさえも、標準的なWntシグナル発生の抑制であった。

【0290】

最初に2001年5月17日出願の米国特許出願第60/291,311号;2002年2月1日出願の第60/353,058号および2002年3月4日出願の第60/361,293号にて本発明者により開示されたように、Dkk-1はLRP5媒介Wntシグナル発生を抑制できるが、HBM媒介Wntシグナル発生の抑制よりも効果が低い。LRP5におけるHBM変異が機能または活性化変異を獲得することであるので、この観察はとりわけ興味深い。すなわち、標準的な経路を介するWntシグナル発生はHBM対LRP5で増強される。理論に縛られるのは望まないが、この相互作用によりHBMとさらに一般的なLRP5/LRP6の間の発達シグナル発生の差異の説明が提供される。

【0291】

別のWntまたはDkkファミリーメンバーのさらなる調査により、Wnt経路による作用の微妙な差異が示され、特定の細胞または組織において発現されるLRP/Dkk/

Wnt / Frizzled レポートリーの機能として達成される Wnt シグナル発生における複雑性および変動性が実証される。これはヒトおよび HBM トランスジェニック動物における HBM 表現型の見かけの骨特異性を証明できる。これらの微妙な変化は可能性のある治療および／または薬物介入の開発において考慮されなければならない。LRP5 シグナル発生を単純にオンまたはオフするのは望ましくない。

【0292】

HBM で観察される LRP5 の Dkk 阻止の有効性の低下が、Dkk の LRP5 / LRP6 / HBM への結合を増強または防御することにより媒介される必要はないであろう。1 つ以上のメカニズムが関与しているであろう。実際に発明者は Dkk-1 が LRP5、LRP6 および HBM に結合するのを観察している。さらに、Dkk ファミリーの別のメンバーが差別的に LRP5 / LRP6 / HBM 活性に影響することが観察されている。むしろ、より好ましい研究法は、できるだけ HBM の微妙な効果に近い再生により、望ましい保護利益に至る薬物または治療を開発することである。HBM の動物モデルへの影響を比較することによる、可能性のある薬物および治療を絞り込み、試験する能力は本発明の主要な特徴である。

【0293】

本発明のトランスジェニック動物および方法を利用して、HBM 表現型の動物モデルの文脈で、可能性のある治療用化合物および方法をスクリーニングすることができる。従って、本発明の動物モデルおよび方法は未来の薬物および治療の開発において非常に貴重な手段を提示する。本発明は骨粗鬆症の有効なモデルを開発し、骨量および脂質変調の理解を深め、そして骨量および脂質代謝を変調するための重要な研究手段を提供する。

【0294】

LRP5 によりコードされるタンパク質は構造的に低密度リポタンパク質レセプター (LDL レセプター) に関連する。Goldstein ら、Ann.Rev.Cell Biology 1:1-39(1985); Brown ら、Science 232:34-47(1986) 参照。LDL レセプターは、コレステロールを含む脂質-タンパク質凝集体である低密度リポタンパク質の取り込みに寄与する。LDL レセプターの欠損を有する個体はコレステロール除去が不十分であり、アテローム性動脈硬化症を進展させる傾向がある。加えて、欠損 LDL レセプターを有する細胞はコレステロールの生成の増加を示し、これは部分的にはコレステロール合成酵素のフィードバック制御の変化により、部分的にはこれらの酵素の遺伝子の転写増加による。いくつかの細胞型ではコレステロールはステロイドホルモンの形成のための前駆体である。従って、LDL レセプターは間接的なシグナルトランスダクションタンパク質としても機能し、遺伝子発現を制御できる。

【0295】

グリシン 171 アミノ酸は、このアミノ酸もまた LRP5 のマウス相同体で見出されるので、LRP5 の機能に重要である可能性がある。密接に関連する LRP6 (Genbank アクセス番号 JE0272) タンパク質もまた対応する位置でグリシンを含有する (Brown ら、Biochemical and Biophysical Research Comm. 248:879-888(1988))。天然の選択により重要な位置でのアミノ酸変化を有する変異が発生から防御されるので、タンパク質の構造または機能において重要であるアミノ酸は種間で観察される傾向がある。

【0296】

加えて、LRP5 の細胞外ドメインは 5 つの YWTD モチーフ、それに続く EGF モチーフから成る 4 つの反復を含有する。この反復もまた LDL レセプターおよびその他の LDL レセプター関連タンパク質において見出されるので、この 5 YWTD + EGF 反復は明確な折り畳みタンパク質ドメインを形成する可能性がある。最初の 3 つの 5 YWTD + EGF 反復はその構造において非常に類似しているが、一方 4 番目の反復は高度に分岐している。グリシン 171 は LRP5 の 1 番目の 5 YWTD + EGF 反復の中央 YWTD モチーフに生じる。LRP5 の別の 2 つの類似の 5 YWTD + EGF 反復もまた対応する位置でグリシンを含有し、LDL レセプタータンパク質における 5 YWTD + EGF 反復でも同様である。しかしながら、アミノ酸の 17.6 % のみが LRP5 の最初の 3 つの 5 Y

W T D + E G F 反復および L D L レセプターの 1 つの反復の中で同一である。これらの観察により、グリシン 171 がこの反復の機能に必須であり、グリシン 171 での変異が L R P 5 の機能的変化を引き起こすことが示される。c D N A およびペプチド配列を図 6 A ~ 6 E に示す。ヌクレオチド位置 582 での重要な塩基は太字で示し、下線を付している。

【0297】

ノーザンブロット分析 (図 7 A - B) により、L R P 5 がヒト骨組織および多くのその他の組織において発現されることが示される。多重組織ノーザンブロット (Clontech、Pal o Alto、カリフォルニア州) を L R P 5 に由来するエクソンでプロービングした。図 7 A に示すように、5.5 kb の L R P 5 転写物は心臓、腎臓、肺、肝臓および脾臓において高度に発現され、骨格筋および脳においては低レベルで発現される。図 7 B に示す第 2 のノーザンブロットにより、5.5 kb の転写物の大きさが確認され、L R P 5 が骨、骨髄、頭蓋冠およびヒト骨芽細胞株において発現されることが示された。

【0298】

これらを考慮に入れて、酵母 2 ハイブリッド結果と合わせたこれらの結果により、L R P 5 遺伝子における H B M 多型性が H B M 表現型に寄与し、L R P 5 遺伝子が骨発達において重要であることが示される。加えて、L R P 5 の変異が骨ミネラル化および発達を変化させ得るので、L R P 5 に結合する分子は骨発達を有用に変化させ得る可能性がある。このような分子には、例えば小型分子、タンパク質、R N A アプタマー、ペプチドアプタマー等がある。

【0299】

X V I I I . 核酸、ベクター、形質転換体および宿主細胞の調製

本発明の大量の核酸を適当な宿主細胞における複製により生成することができる。望ましいフラグメントをコードする天然または合成核酸フラグメントを、原核または真核細胞に導入でき、そこで複製できる組換え核酸構築物、通常 D N A 構築物に組み込む。通常核酸構築物は単核宿主、例えば酵母または細菌における複製に適しているが、培養哺乳動物または植物またはその他の真核細胞株への (ゲノム内での組み込みを伴う、および伴わない) 導入を企図することもできる。本発明の方法により生成される核酸の精製については、例えば Sambrook ら、Molecular Cloning: A Laboratory Manual、第 2 版 (Cold Spring Harbor Laboratory、Cold Spring Harbor、ニューヨーク州 (1989)) または Ausubel ら、Current Protocols in Molecular Biology、J. Wiley & Sons、ニューヨーク州 (1992) に記載されている。

【0300】

本発明の核酸を化学的合成により、例えば Beaucage ら、Tetra. Letts. 22: 1859-1862 (1981) により記載されるホスホラミデート法または Matteucci ら、J. Am. Chem. Soc. 103: 3185 (1981) によるトリエステル法により生成することもでき、市販されている自動化されたオリゴヌクレオチド合成器で実施することができる。相補鎖を合成し、適当な条件下で鎖と一緒にアニーリングするか、または D N A ポリメラーゼを用いて適当なプライマー配列と共に相補的な鎖を加えるいずれかにより化学合成の一本鎖生成物から二本鎖フラグメントを得ることができる。

【0301】

原核または真核細胞宿主に導入するために調製された核酸構築物は、望ましいタンパク質をコードする企図される核酸フラグメントを含む、宿主により認識される複製系を含んでなっており、好ましくはタンパク質コード化セグメントに機能的に連結した転写および翻訳開始制御配列をも含む。発現ベクターには例えば複製起点または自己複製配列 (A R S) および発現調節配列、プロモーター、エンハンサーおよび必要なプロセッシング情報部位、例えばリボソーム結合部位、R N A スプライシング部位、ポリアデニル化部位、転写ターミネーター配列、ならびに m R N A 安定化配列などがある。適当な場合、元来の H B M または L R P 5 タンパク質に由来しても、またはその他のレセプターに由来しても、または同一物もしくは関連種の分泌タンパク質に由来しても、分泌シグナルが含まれ、これ

によりタンパク質が細胞膜を通過するおよび／もしくはそこに留まり、従ってその機能的トポロジーを達成するか、または細胞から分泌されることが可能になる。このようなベクターを当分野で公知の標準的な組換え技術、例えばSambrookら、Molecular Cloning: A Laboratory Manual、第2版(Cold Spring Harbor Laboratory、Cold Spring Harbor、ニューヨーク州(1989))またはAusubelら、Current Protocols in Molecular Biology、John Wiley & Sons、ニューヨーク州(1992)の手段により調製することができる。

【0302】

適当なプロモーターおよびその他の必要なベクター配列を、宿主において機能的であるように選択し、適当な場合天然にLRP5またはHBM遺伝子に随伴されるものを含むことができる。作動可能な細胞株および発現ベクターの組み合わせの実例はSambrookら、Molecular Cloning: A Laboratory Manual、第2版(Cold Spring Harbor Laboratory、Cold Spring Harbor、ニューヨーク州(1989))またはAusubelら、Current Protocols in Molecular Biology、J. Wiley & Sons、ニューヨーク州(1992)に記載されている。多くの有用なベクターが当分野で公知であり、Stratagene、New England Bio Labs、Promega Biotech等のような製造供給元から入手することができる。プロモーター、例えばtrp、lac、およびファージプロモーター、tRNAプロモーターおよび解糖系酵素プロモーターを原核細胞宿主において使用することができる。有用な酵母プロモーターにはメタロチオネイン、3-ホスホグリセレートキナーゼまたはその他の解糖系酵素、例えばエノラーゼまたはグリセルアルデヒド-3-ホスフェートデヒドロゲナーゼ、マルトースおよびガラクトース利用に寄与する酵素等のプロモーター領域などがある。酵母発現に使用するのに適したベクターおよびプロモーターは欧州特許第73,675A号に記載されている。適当な外来の哺乳動物プロモーターにはSV40に由来する初期および後記プロモーター(Fiersら、Nature 273:113(1978))またはネズミモロニー白血病ウイルス、マウス腫瘍ウイルス、トリ肉腫ウイルス、アデノウイルスII、ウシパピローマウイルスまたはポリオーマ由来のプロモーターがある。加えて、構築物を増幅可能な遺伝子(例えば、DHFR)に結合させて、遺伝子の複数のコピーを作成することができる。適当なエンハンサーおよびその他の発現調節配列に関しては、Enhancers and Eukaryotic Gene Expression, Cold Spring Harbor Press、Cold Spring Harbor、ニューヨーク州(1983)をも参照されたい。

【0303】

このような発現ベクターは自己複製できるが、当分野で周知の方法によりこれらを宿主細胞のゲノムに挿入することにより複製することができる。

【0304】

発現およびクローニングベクターは選択マーカー、すなわちベクターで形質転換された宿主細胞の生存または成長に必要なタンパク質をコードする遺伝子を含む可能性がある。この遺伝子の存在によりインサートを発現するこれらの宿主細胞のみの成長が確実になる。典型的な選択遺伝子は、a) 抗生物質またはその他の毒性物質、例えばアンピシリン、ネオマイシン、メソトレキセート等に対する抵抗性を付与する；b) 栄養要求欠損を補完する；またはc) 複合培地から利用できない重要な栄養を供給するタンパク質をコードする、例えば杆菌のD-アラニンラセマーゼをコードする遺伝子である。適切な選択マーカーの選択は宿主細胞に依存し、異なる宿主のための適当なマーカーは当分野で周知である。

【0305】

目的の核酸を含むベクターをインビトロで転写することができ、得られたRNAを周知の方法、例えば注射(Kuboら、FEBS Letts. 241:119(1988)参照)により宿主細胞に導入するか、またはエレクトロポレーション；塩化カルシウム、塩化ルビジウム、リン酸カルシウム、DEAEデキストランまたはその他の物質を用いるトランスフェクション；微粒子銃；リポフェクション；感染(この場合ベクターは感染性物質、例えばレトロウイルスゲノムである)；ならびにその他の方法などの、細胞宿主の型に依存して変化する当分野で周知の方法によりベクターを直接宿主細胞に導入することができる。一般的にはSambrookら(1989)およびAusubelら(1992)を参照されたい。前記したものなどの当分野で公知の

方法による核酸の宿主細胞への導入を本明細書では「形質転換」と称する。前記した核酸を導入した細胞はこのような細胞の子孫を含むことをも意味する。

【0306】

本発明の大量の核酸およびタンパク質を、ベクターまたは原核または真核宿主細胞に適合するその他の発現ベヒクル中のLRP5またはHBM核酸またはその一部を発現させることにより調製することができる。最も一般的に用いられる原核細胞宿主は大腸菌の株であるが、しかしながら、別の原核細胞、例えば枯草菌またはシュードモナスを用いることもできる。

【0307】

哺乳動物またはその他の真核宿主細胞、例えば酵母、繊維状菌類、植物、昆虫または両生類もしくは鳥類の種の宿主細胞が本発明のタンパク質の生成に有用である。培養での哺乳動物細胞の増殖はそれ自体周知である。JakobyおよびPastan(編)Cell Culture.Methods in Enzymology、58巻、Academic Press,Inc.、Harcourt Brace Javanovich、ニューヨーク州(1979)。一般に用いられる哺乳動物宿主細胞株の実例はVEROおよびHeLa細胞、チャイニーズハムスター卵巣(CHO)細胞、ならびにWI38、BHK、およびCOS細胞株であるが、例えば高発現、望ましいグリコシル化パターンまたはその他の特徴を提供するために、その他の細胞株が適当であり得ることは当業者には理解されよう。

【0308】

ベクター構築物の様式に依存してマーカーを使用することにより、クローンを選択する。マーカーは同一または異なるDNA分子、好ましくは同一のDNA分子にあってよい。原核細胞宿主では、例えばアンピシリン、テトラサイクリンまたはその他の抗生物質に対する抵抗性により形質転換体を選択できる。温度感受性に基づく特定の産物の生成が適当なマーカーを提供することもできる。

【0309】

本発明の核酸で形質転換された原核または真核細胞は本発明の核酸およびタンパク質の生成のみならず、例えばLRP5またはHBMタンパク質の特性の研究にも有用である。

【0310】

当業者に理解されるように、アンチセンス核酸配列はLRP5またはHBMの発現を防御または現象させるのに有用である。例えば、LRP5もしくはHBM遺伝子またはLRP5もしくはHBM領域に由来する別の配列の全てまたは一部を含有する核酸ベクターをアンチセンス配向のプロモーターの調節下に置き、細胞に導入することができる。このような細胞内のアンチセンス構築物の発現はLRP5またはHBM転写および／または翻訳および／または複製と干渉する。siRNAまたはshRNAなどのRNA干渉方法学もまた企図される。

【0311】

本明細書に開示するLRP5およびHBM遺伝子配列に基づくプローブおよびプライマーを用いて別の種における相同性LRP5およびHBM遺伝子配列およびタンパク質を同定する。これらのLRP5およびHBM遺伝子配列およびタンパク質を、それが単離された種に関する本明細書で記載した診断／予後、治療および薬物スクリーニング法に使用する。

【0312】

XIX. タンパク質発現および精製

本発明のHBMタンパク質の発現および精製を本質的に以下の概要のように実施することができる。HBM遺伝子、大腸菌における組換えタンパク質のクローニングおよび発現のために遺伝子発現系、例えばpETシステム(Novagen)に由来する膜および分泌タンパク質のクローニング、発現および精製を選択した。またペプチドタグ、His-TagをコードするDNA配列を目的のDNA配列の3'末端に融合し、組換えタンパク質生成物の精製を促進した。いずれかの5'終止シグナル配列の変化を避けるために、融合には3'末端を選択した。

【0313】

例えばHBMをクローニングするために配列番号1、3および5～12に示した核酸から選択した核酸をポリメラーゼ連鎖反応（PCR）により調製した。HBMヌクレオチド配列の5'および3'末端に特異的な合成オリゴヌクレオチドプライマーを設計し、Life Technologies(Gaithersburg、メリーランド州)から購入した。5'末端でNcoIクローニング部位を含むように全ての順行性プライマー（配列の5'末端に特異的）を設計した。NcoI部位内でコードされるメチオニン残基で、続いてバリン残基でタンパク質翻訳を可能にするようにこれらのプライマーを設計し、HBM DNA配列によりタンパク質がコードされる。全ての逆行性プライマー（配列の3'末端に特異的）は5'末端でEcoRI部位を含み、それによりHBM配列のpET-28bの読み枠へのクローニングが可能になる。pET-28bベクターは、6個のヒスチジン残基を含む（C末端で）さらなる20個のカルボキシル末端アミノ酸をコードする配列を提供し、これはヒスチジンアフィニティタグを含んでなった。

10

【0314】

HBM遺伝子からのゲノムDNA調製物をPCR増幅のための鋳型DNAの供給源として使用した(Ausubelら、Current Protocols in Molecular Biology、John Wiley & Sons(1994))。HBMヌクレオチド配列を含有するDNA配列を増幅するために、最終容量100μlで2mM MgCl₂、定義されたHBMに相補的であり、それをフランキングする1μM 合成オリゴヌクレオチドプライマー（順行性および逆行性プライマー）、0.2mM 各デオキシヌクレオチド三リン酸、dATP、dGTP、dCTP、dTTPおよび熱安定性DNAポリメラーゼ 2.5単位(Amplitaq、Roche Molecular Systems, Inc.、Branchburg、ニュージャージー州)が入った反応バイアルにゲノムDNA（50ng）を導入した。

20

【0315】

熱循環反応の完了時に増幅したDNAの各サンプルをQiaquick Spin PCR精製キット(Qiagen、Gaithersburg、メリーランド州)を用いて精製した。増幅したDNAサンプル全てを制限エンドヌクレアーゼ、例えばNcoIおよびEcoRI(New England BioLabs、Beverly、マサチューセッツ州)を用いる消化に供した(Ausubelら、Current Protocols in Molecular Biology、John Wiley & Sons, Inc.(1994))。次いでDNAサンプルを1.0% NuSieve(FMC BioProducts、Rockland、メイン州)アガロースゲルでの電気泳動に供した。臭化エチジウムおよび長波UV照射に暴露することによりDNAを可視化した。アガロースゲルから単離した切片に含まれるDNAをBio 101 Gene Cleanキットプロトコル(Bio 101、Vista、カリフォルニア州)を用いて精製した。

30

【0316】

制限エンドヌクレアーゼ、例えばNcoIおよびEcoRI(New England BioLabs、Beverly、マサチューセッツ州)を用いる消化によりpET-28bベクターをクローニング用に調製した(Ausubelら、Current Protocols in Molecular Biology、John Wiley & Sons, Inc.(1994))。挿入した遺伝子の5'末端に融合できるヒスチジンアフィニティタグをコードするpET-28bベクターを適当な制限エンドヌクレアーゼを用いる消化により調製した。

【0317】

消化に続いて、DNAインサートを予め消化したpET-28b発現ベクターにクローン化した(Ausubelら、Current Protocols in Molecular Biology、John Wiley & Sons, Inc.(1994))。次いで以下に記載するように、ライゲーション反応の生成物を用いて大腸菌のBL21株を形質転換した(Ausubelら、Current Protocols in Molecular Biology、John Wiley & Sons, Inc.(1994))。

40

【0318】

標準的な方法に従ってコンピテント細菌大腸菌株BL21または大腸菌株BL21(DE3)をクローン化されたHBM配列組換えpET発現プラスミドで形質転換した(Ausubelら、Current Protocols in Molecular Biology、John Wiley & Sons, Inc.(1994))。簡単には、ライゲーション反応物1μlをエレクトロコンピテント細胞50μlと混合し、

50

高電圧に供し、その後サンプルをSOC培地（0.5%酵母抽出物、2.0%トリプトン、10mM NaCl、2.5mM KCl、10mM MgCl₂、10mM MgSO₄ および20mM グルコース）0.45ml中37℃で1時間振盪しながらインキュベートした。次いでサンプルを25μg/ml 硫酸カナマイシン含有LB寒天プレートに広げて、一晚成長させた。次いでBL21の形質転換されたコロニーを採取し、以下に記載するように分析してクローン化されたインサートを評価した。

【0319】

元来のPCR増幅クローニング反応で用いたHBM配列に特異的な同一の順行性および逆行性プライマーを用いて、組換えpET-28b HBMヌクレオチド配列で形質転換した個々のBL21クローンをクローン化されたインサートのPCR増幅により分析した。発現ベクター中のHBM配列の組み込みは増幅の成功により証明される(Ausubelら、Current Protocols in Molecular Biology、John Wiley & Sons, Inc.(1994))。 10

【0320】

適切にクローン化されたHBMヌクレオチド配列を担持する組換えpET-28bベクターの個々のクローンを採取し、LBブロス5ml+25μg/ml 硫酸カナマイシン中で一晚インキュベートした。翌日プラスミドDNAを単離し、Qiagenプラスミド精製プロトコル(Qiagen Inc.、Chatsworth、カリフォルニア州)を用いて精製した。

【0321】

pETベクターをクローニングまたはプラスミド調製の目的でいずれかの大腸菌K-12株、例えばEMS174、HB101、JM109、DH5等で増殖させることができる。発現のための宿主にはT7 RNAポリメラーゼの遺伝子の染色体コピーを含有する大腸菌株などがある。これらの宿主は、lacI遺伝子、lacUV5プロモーターおよびT7 RNAポリメラーゼの遺伝子を担持するλ誘導体であるバクテリオファージDE3のリソゲンであった。イソプロピル-β-D-チオガラクトシド(IPTG)の添加によりT7 RNAポリメラーゼを誘導し、T7 RNAポリメラーゼが機能的T7プロモーター、例えばpET-28bを含有し、目的の遺伝子を担持するいずれかの標的プラスミドを転写する。株には例えばBL21(DE3)などがある(Studierら、Meth.Enzymol.185:60-89(1990))。 20

【0322】

組換えHBM配列を発現するために、プラスミドDNA50ngを前記したように単離して、前記したようにコンピテントBL21(DE3)細菌を形質転換した(pET発現キットの一部としてNovagenにより提供される)。HBM組換え構築物に関して記載したように、lacZ遺伝子(β-ガラクトシダーゼ)をpET-Systemに発現させる。形質転換された細胞をSOC培地中1時間培養し、次に培養物を25μg/ml 硫酸カナマイシン含有LBプレートにプレートした。翌日に細菌性コロニーをプールし、硫酸カナマイシン(25μg/ml)含有LB培地にて600nMで光学密度0.5から1.0 O.D.単位まで成長させ、その時点で1mM IPTGを培養物に3時間加えてHBM組換えDNA構築物の遺伝子発現を誘導した。 30

【0323】

IPTGでの遺伝子発現のインキュベーションの後、Sorvall RC-3B遠心を用いて3500xgで4℃で15分間の遠心により細菌を収集した。ペレットをcold mM トリスHCl、pH8.0、0.1M NaClおよび0.1mM EDTA(STEバッファー)50mlに再懸濁した。次に細胞を2000xgで4℃で20分間遠心した。湿ったペレットの重量を測定し、タンパク質精製の準備まで-80℃で凍結した。 40

【0324】

チャイニーズハムスター卵巣(CHO)発現系

また別に、HBMおよびLRP5を真核細胞において発現させることができる。真核細胞、例えば哺乳動物由来の細胞株はシスチン富化ドメイン、例えばEGFおよびLDLRモジュールを含有する適切に折り畳まれたタンパク質を発現することがさらに可能である。

【0325】

構築物の開発

I g G - F c への H B M および L R P 5 細胞外ドメイン融合 (E C D) を調製した。これらの I g G - F c ドメインへの E C D 融合により L R P 5 / H B M レセプターの内因性膜貫通および細胞質部分が除去され、分泌融合タンパク質が生成されるはずである。エンテロキナーゼ認識部位により L R P 5 / H B M E C D から F c 領域が分離され、F c ドメインの存在しない精製された L R P 5 または H B M E C D タンパク質を得ることができる。この構築に用いられたベクターは、アデノ主要後記プロモーターの大部分が t e t オペレーターの 6 個の反復により置換された (Gossen ら、Proc.Natl.Acad.Sci.USA 89:5547-5551(1992)) p E D の誘導体である p H T o p (Kaufman ら、Nuc.Acids.Res.19:4485-490 (1991)) であった。このベクターはデヒドロ葉酸リダクターゼ (d h f r) 遺伝子を含有し、細胞株 C H O / A 2 (以下の記載を参照) に導入されたとき、非常に効率的に機能する。高発現のクローンを、高濃度メソトレキセート (M T X) 中で生存する細胞を単離することにより選択することができる。

【0326】

C H O 発現ベクター p H T O P - F c を S a l I および N o t I で消化した。アクリルアミドゲル切片からの電気溶出により、介在配列を精製して残りのベクターから除去した。S a l I は p H T O P - F c において 5 ' を内在性のミツバチ・メリチンシグナル配列まで切断し、N o t I はちょうど 5 ' でコード化配列 I g G 1 - F c まで切断する。得られた S a l I - N o t I p H T O P - F c ベクターはシグナル配列が除去されており、N o t I クローニング部位は I g G - F c への 5 ' 融合の作製を施しやすい。p C M V S P O R T 6 における L R P 5 全長および p C M V S P O R T 6 における H B M 全長を別個に X m a 1 (シグナル配列をコードする O R F の領域内で切断する) および B a m H I (O R F 内で切断する) で消化して L R P 5 および H B M の 2 2 8 6 b p 5 ' フラグメントを作製した。H B M から L R P 5 を区別する変異体はこのフラグメントに在る。別に、L R P 5 D N A を B a m H I および S a c I で消化して L R P 5 と H B M 遺伝子の双方に共通する 1 8 0 0 b p 3 ' フラグメントを単離した。一緒にするとこれらの 2 個のフラグメントは H B M および L R P 5 細胞外ドメインのコード化配列、少しのシグナル配列の最初の 6 個のアミノ酸のコード化配列ならびに細胞外ドメインの末端の前の 1 8 個の終末アミノ酸を構成し、本発明者は K y t e - D o o l i t t l e プロットから配列 (配列番号 6 9 8) 「S P A H S S」の末端まで推定した。

【0327】

合成二本鎖を設計して、イニシエーターメチオニンおよび K o z a k 配列を含む元来の X m a 1 部位の 5 ' の L R P 5 / H B M シグナル配列コード化配列を再生した。S a l I (5 ') および X m a 1 (3 ') 付着末端を含有するようにこの二本鎖を設計して末端を適応させて前記した遺伝子フラグメントを p H T O P - F c ベクターに適応させた。以下に提示するような 2 つの部分的相補性オリゴヌクレオチド (配列番号 6 9 9 ~ 7 0 0) からこの合成二本鎖を構築した。

5 ' -TCGACCACCATGGAGGCAGCGCCGC-3 ' ,

3 ' -GGTGGTACCTCCGTCGCGCGGGCC-5 ' ,

【0328】

第 2 の合成二本鎖を設計して元来の S a c I 部位から配列「. S P A H S S」のセリンに続く細胞外ドメインの推定される末端までを再生し、また I g G - F c の下流へのフレーム内融合を可能にするクローニング部位をコードした。S a c I (5 ') および N o t I (3 ') 付着末端を含有するようにこの二本鎖を設計して前記した遺伝子フラグメントを p H T O P - F c ベクターに適応させた。以下に提示するような 2 つの部分的相補性オリゴヌクレオチド (配列番号 7 0 1 - 7 0 2) からこの合成二本鎖を構築した。

5 ' -CATGTGTGAAATCACCAAGCCGCCCTCAGACGACAGCCCGGCCACAGCAGTGGC-3 ' ,

3 ' -TCGAGTACACACTTTAGTGGTTCGGCGGGAGTCTGCTGTGGGCCGGGTGTCGTCACCGCCGG-5 ' ,

【0329】

10

20

30

40

50

フラグメント、合成二本鎖、およびベクターと一緒に単一の反応でライゲートした。LRP5およびHBMに関して別個の反応を実施した。ライゲーション混合物を用いてエレクトロコンピテント大腸菌DH10B細胞を形質転換し、一般的なSacI-BamHI 3'フラグメントをプローブとして用いて、得られたコロニーを放射活性コロニーハイブリダイゼーションによりスクリーニングした。挿入したLRP5またはHBMと共にプラスミドを含有するコロニーを同定し、プラスミドを複数の候補から単離し、その配列をDNAシーケンシングにより証明した。次いで証明された構築物をCHO細胞へのトランスフェクションに使用した。

【0330】

CHO安定細胞株の確立

10

TetレプレッサーとヘルペスウイルスVP16転写ドメインの間に転写アクチベーター(tTA)、融合タンパク質を安定して組み込むことにより(Gossenら)、CHO/A2細胞株をCHO-DUKX-B11(UrlaubおよびChasin, Proc.Natl.Acad.Sci.USA 77:4216-4220(1980))から誘導する。pHTopHBM-1.FcをCHO/A2細胞に、およびpHTopZmax1.FcをCHO/A2細胞にトランスフェクトすることにより(リポフェクションを用いて)、細胞外HBM-1.FcおよびLRP5.Fcを発現するCHO細胞株を確立した。細胞を0.02μMメソトレキセート中で培養することによりクローンを選択した。後でクローンを最終濃度0.5μMメソトレキセートまで段階的に増幅した。

【0331】

20

CHO安定細胞株のスクリーニング

複数のクローンを種々の技術によりスクリーニングした。ウェスタンブロットアッセイにより(マウス)抗ヒトIgG.Fc西洋ワサビペルオキシダーゼ(HRP)抗体を用いてクローンをスクリーニングした。また同一クローンを³⁵S-Met/Cysで6時間パルス、または15分間パルスで代謝的に標識し、続いて放射標識Met/Cys不含の培地中で1時間、4時間、または24時間追跡した。条件培地、および細胞抽出物から得たタンパク質に関して免疫沈澱を実施した。次いで精製を実施し、続いて当分野で公知のN末端シーケンシングを用いてタンパク質をシーケンシングした。

【0332】

融合タンパク質精製

30

条件培地またはCHO安定細胞の細胞抽出物からLRP5-IgGまたはHBM-IgG融合タンパク質を精製することができる。プロテインAに対するアフィニティ結合(例えばプロテインAコーティングビーズまたはカラムを使用する)により融合タンパク質を単離する。次いでエンテロキナーゼ消化により続いてZmax/HBM1-ECDTタンパク質からIgG-FCドメインを切断することができる。

【0333】

細胞株およびタンパク質の可能な使用

リガンドハンティング、抗体作製、結晶構造の決定および競合結合アッセイにおいて使用するための精製されたタンパク質を作製するために安定した細胞株を使用することができる。

40

【0334】

当業者に公知の種々の方法学を用いて単離されたタンパク質を精製することができる(Coliganら、Current Protocols in Protein Science、John Wiley & Sons(1995))。例えば凍結細胞を解凍し、バッファーに再懸濁し、小容量マイクロフルイダイザー(M-110Sモデル、Microfluidics International Corp.、Newton、マサチューセッツ州)を数回通過させることにより破壊した。得られたホモジネートを遠心して透明な上澄(粗抽出物)を得て、続いて濾過して粗抽出物をカラムで分画化する。OD_{280nm}での吸光度により分画をモニター観察し、ピーク分画をSDS-PAGEにより分析できる。

【0335】

精製したタンパク質の濃度を分光光度計でアミノ酸含量から算出した吸光計数を用いて

50

定量する(Perkins, Eur. J. Biochem. 157:169-180(1986))。またタンパク質濃度を標準としてウシ血清アルブミンを用いてBfadfordの方法(Anal. Biochem. 72:284-254(1976)およびLowryら、J. Biol. Chem. 193:265-275(1951))により測定する。

【0336】

種々濃度のSDS-ポリアクリルアミドゲルをBioRad(Hercules、カリフォルニア州)から購入し、クーマシー・ブルーで染色した。分子量マーカーはウサギ骨格筋ミオシン(200 kDa)、大腸菌 β -ガラクトシダーゼ(116 kDa)、ウサギ筋肉ホスホリラーゼB(97.4 kDa)、ウシ血清アルブミン(66.2 kDa)、卵白アルブミン(45 kDa)、ウシ炭酸脱水酵素(31 kDa)、大豆トリプシンインヒビター(21.5 kDa)、卵白リゾチーム(14.4 kDa)およびウシアプロチニン(6.5 kDa)を含むことができる。 10

【0337】

一度望ましいタンパク質の十分量が得られると、種々の目的に用いることができる。典型的な使用は結合に特異的な抗体の生成である。これらの抗体はポリクローナルまたはモノクローナルのいずれかでよく、当分野で周知のインビトロまたはインビボ技術により生成することができる。記載するように同定され、単離されたいずれかのペプチドのエピトープに対するモノクローナル抗体をネズミハイブリドーマから調製することができる(Kohler、Nature 256:495(1975))。要約すると、マウスにHBMタンパク質数マイクログラムを2週間にわたって接種する。次いでマウスを屠殺する。抗体を生成する細胞をマウス脾臓から取り出す。次いで脾臓細胞をポリエチレングリコールを用いてマウスミエローマ細胞と融合させる。融合に成功した細胞をマイクロタイタープレートで希釈し、培養物の成長を続けさせる。ウェルあたりの抗体の量をイムノアッセイ、例えばELISA(Engvall、Meth. Enzymol. 70:419(1980))により測定する。抗体産生クローンを広げ、さらに増殖させてHBM抗体を生成させる。その他の適当な技術にはリンパ球の抗原性ポリペプチドへのインビトロ暴露、または別に、ファージもしくは類似のベクター中の抗体のライブラリーの選択などがある。Huseら、Science 246:1275-1281(1989)参照。抗体生成に関するさらなる情報についてはDavisら、Basic Methods in Molecular Biology, Elsevier、ニューヨーク州セクション21-2(1989)参照。 20

【0338】

L RP 5およびL RP 6 ポリクローナル抗体 30

ヒトL RP 5(配列番号3)およびL RP 6(Genbankアクセッション番号AF074264)の双方に対してポリクローナル抗体を作成した。L RP 5アミノ酸配列からのペプチドを5つの目標に基づく免疫原として選択した。1) L RP 5とL RP 6アミノ酸配列の間の差異を最大にする(アミノ酸同一性71%)。図27参照。配列比較のために、典型的には1つの配列が参照配列の機能を果たし、被験配列をそれと比較する。配列比較アルゴリズムを用いる場合、被験および参照配列をコンピューターにインプットし、必要な場合続くコーディネートを指定し、配列アルゴリズムプログラムパラメーターを指定する。次いで配列比較アルゴリズムが(複数の)被験配列に関して、指定されたプログラムパラメーターに基づいて、参照配列に相対する配列同一性パーセントを算出する。2) 配列アラインメントおよび類似性検索を実施することにより別の公知の遺伝子との交差反応の可能性を最低にする。比較のための配列の最適アラインメントを例えばSmithおよびWaterman、Adv. Appl. Math. 2:482(1981)のローカル相同性アルゴリズムにより、NeedlemanおよびWunsch、J. Mol. Biol. 48:443(1970)の相同性アラインメントアルゴリズムにより、PearsonおよびLipman、Proc. Natl. Acad. Sci. USA 85:2444(1988)の類似の方法のための検索により、これらのアルゴリズムおよびWisconsin遺伝学ソフトウェアパッケージ、Genetics Computer Group、585 Science Dr. Madison、ウィスコンシン州に含まれるプログラムにおけるその他のアルゴリズムにより、または肉眼検査(一般的にはAusubelら、Current Protocols in Molecular Biology, John Wiley & Sons(1997))により行うことができる。配列同一性および配列類似性パーセントを決定するのに適したアルゴリズムの別の例はBLASTアルゴリズムであり、これはAltschulら、J. Mol. Biol. 215:403-410(1990) 40 50

に記載されている。B L A S T分析を実施するためのソフトウェアはNational Center for Biotechnology Informationを通して公に入手可能である。3)Wisconsin Geneticsソフトウェアパッケージ、Genetics Computer Group、575 Science Dr.Madison、ウィスコンシン州に含まれるプログラムを用いて、PeptideStructureタンパク質分析により決定されるできるだけ最高の抗原性指標を有するペプチドを入手する。4)ファミリー遺伝子および遺伝子の細胞外および細胞質領域に相対する位置内の高度に相同性のドメイン（例えばEGF様ドメインおよびLDLレセプター反復）に相対するペプチドの位置付け。5)ならびに、ヒトLRP5特異的抗体、ヒトアミノ酸配列（配列番号3）をマウスLRP5配列（Genbankアクセッション番号AF064984）と比較し、前記の基準に基づいてペプチドを選択し、加えて2つの種間の配列類似性を最低にした（図26参照）。 10

【0339】

前記の同一の基準を用いて、LRP6特異的ペプチドをポリクローナル抗体産生に関して選択した。表は選択したアミノ酸配列、ヒト配列とマウス配列間のペプチド内でのアミノ酸相違を列挙する。ニュージーランド白色ウサギにおけるペプチド合成および続くポリクローナル抗体産生のために全てのペプチド配列（アミノ酸12～18の範囲）をSigma/Genosys(St.Louis、ミズーリー州)に提供した。各免疫ウサギの血清に由来するIgG分画をプロテインGセファロース(Amersham)を用いて単離した。これらのペプチドを用いるポリクローナル抗体作製を別の種、例えばニワトリにおいて同様に行うことができる。ヒト（参照）およびネズミ／げっ歯類配列間の類似性の程度が高い場合、これはしばしば有利である。 20

【0340】

【表7】

アミノ酸	アミノ酸	H/M*差異	コメント	配列番号：
171-187	VETPRIERAGMDGSTRK	5	HBM 含有 多型性	703
264-278	NKRTGGKRKEILSAL	3	細胞外	704
290-301	ERQPFHTRCEE	2	EGF-Iに隣接 細胞外	705
532-546	VDGTKRRTLLEDKLP	5	細胞外	706
901-915	DGLNDCMHNNGQCGQ	2	EGF-III 内 細胞外	707
1010-1021	PFVLTSLSQGQN	6	細胞外、ヒト 特異的	708
1415-1429	YAGANGPFPHEYVSG	3	細胞質内	709
1452-1464	ACGKSMMSSVSLM	5	細胞質内、ヒト 特異的	710
1556-1573	RWKASKYYLDLNSDSDPY	1	細胞質内	711
888-902	SGWNECASSNGHCSH		LRP6 特異的	712
1308-1321	NGDANCQDKSDEKN		LRP6 特異的	713

*ヒトとマウス配列間のH/M差異

【0341】

ファージディスプレイにより発展された一本鎖Fv分子

LRP5配列（配列番号3）からペプチドを選択してファージディスプレイにより一本鎖Fv分子をスクリーニングした。合成およびそれに続くscFv分子のためのファージディスプレイスクリーニングのために全部で17個のペプチドをLRP5配列から選択した。Cambridge、英国のCambridge Antibody Technology(CaT)で全てのペプチド合成およ 30

10

20

30

40

50

びファージディスプレイ研究を実施した。前記した基準に基づいてペプチドを選択した。

【0342】

【表8】

タンパク質ドメイン	LRP5 残基	LRP6 残基	同一性%
スパーサー1 (+G171V)	161-181	148-168	76%
スパーサー1 (-G171V)	161-181	148-168	76%
EGF 1	301-321	288-308	76%
スパーサー2	401-421	388-408	52%
EGF 2	611-631	598-618	62%
スパーサー3	781-801	768-788	62%
EGF 3	921-941	908-928	10%
スパーサー4	1000-1021	988-1008	26%
EGF 4	1229-1249	1219-1239	76%
LDLR 1	1261-1282	1252-1272	81%
LDLR 2	1300-1320	1290-1310	57%
LDLR 3	1338-1358	1328-1348	48%
サイトゾル1	1418-1438	1405-1425	14%
サイトゾル1	1516-1536	1503-1525	52%
サイトゾル1	1535-1555	1524-1544	81%
サイトゾル1	1595-1615	1592-1613	82%
スパーサー2 (交差反応)	421-441	408-428	100%

10

20

30

40

50

【0343】

これらの領域の多くは（例えば401～421、421～441、781～801、および1229～1249）はマウスLRP5と100%同一性を共有することに注目されたい（図26参照）。従って、マウスおよびヒトの双方の形態のタンパク質に対してこれらを使用することができる。LRP5およびLRP6の双方を認識する抗体の作製を促進するためにペプチド421-441を含めた（図27参照）。HBM変異部位（LRP5残基161-181）にわたる2個のペプチドを合成し、一方はLRP5配列を有し、他方はHBM配列を含有した。

【0344】

一度scFv分子を単離すると、これを試薬として免疫化学において用いて種々のヒトの正常および疾患組織におけるLRP5タンパク質発現を検出した。LRP5の細胞外ドメインのペプチド1000-1021（IKRAKDDGTQPFVLTSLSQGQN）（配列番号714）に対する3つのファージクローンの詳細なscFV抗体免疫組織化学分析により心筋、腎臓、肺および肝臓での陽性染色が示された。前立腺癌における発現もまた検出した。これらの結果はmRNA組織分布プロファイルおよび公開されたLRP5 mRNA局在性(Kimら、J.Biochem.124:1072-1076(1998))の報告と合致する。得られたファージクローンはプールから生じ、これをシーケンシングして分子のFv領域における変異体の可能性を同定した。一度同定すると、次いで無傷および機能的免疫グロブリンγ（IgG）分子の最終的な組み立てのためにCOS細胞に同時トランスフェクトするため、適当なscFVを可変性重鎖および可変性軽鎖DNA構築物にサブクローン化することができる。次いで当分野で公知なように、さらに特異性および反応性に関して細胞により発現されるIgGを特徴付けすることができる。

【0345】

モノクローナル抗体の作製

例えばBalb/cマウスの完全な細胞およびアデノウイルス免疫によりモノクローナル抗体をLRP5およびHBMおよびその変異体に発展させることができる。例えばBa

l b / c マウスの脾臓から樹状細胞を単離することができ、細胞をインビトロで成長因子 I L - 4 および G M - C S F の存在下で広げた。次いで樹状細胞を H B M または L R P 5 アデノウイルス粒子で感染させることができる。次いで B a l b / c マウスへの静脈内注射の前に 2 4 時間細胞を培養する。3 から 4 週毎に 2 から 3 回、3 か月にわたって樹状細胞 (1×10^6 セル / マウス) を注射した。

【0346】

また別に、例えば p c D N A 3 . 1 発現ベクター中の精製された H B M および L R P 5 D N A をコロイド金粒子にコーティングすることができる。次いで遺伝子銃技術を用いてこの粒子を望ましいマウスに皮下注射することができる。マウスあたり c D N A およそ $5 \mu\text{g}$ を注射することができる。2 週毎に 4 から 6 回、およそ 3 か月にわたって注射を行う。 10

【0347】

別の意見は、H B M および L R P 5 およびその各々のアデノウイルスを過剰発現する細胞 (動物のいずれかの種であるが、好ましくは B a l b / c マウス株または非特異的タンパク質への限定的な抗原応答に関連する、使用されているマウス株と同一の種) を 2 から 3 週毎に必要なに応じて約 1 . 5 から 3 か月間マウスに注射する。マウスからの出血を E L I S A (精製されたタンパク質またはタンパク質融合体を使用する) 、細胞基盤の E L I S A 、免疫組織化学染色およびウェスタンブロッティングにより元来のおよび変性タンパク質との反応性に関して試験できる。動物に由来する血清サンプルを F A C S (蛍光活性化細胞分類) により L R P 5 または H B M アデノウイルスで感染した細胞を用いてスクリーニングすることができる。次いでマウスからの最強の反応性を有する脾臓細胞 (抗体産生細胞) をミエローマと融合させてハイブリドーマ細胞を作製することができる。次いでハイブリドーマからの条件培地を続けるクローニングのための陽性細胞コロニーに関してスクリーニングする。次いで腹水産生のためにこれらのクローン化された細胞をマウスの腹腔内に注射することができる。 20

【0348】

ポリクローナル抗体の適用

L R P 5 および L R P 6 を志向するポリクローナル抗体を発展させてこれらのタンパク質の機能を決定し、種々組織、細胞またはいずれかの生物学的サンプルにおける発現パターンおよびレベルを分析した。L R P 5 、H B M 、L R P 6 および関連する変異体に対するポリクローナル抗体の使用には、非限定例としては、骨横断面マウントの分析、組織分布、罹患 / 非罹患ファミリーメンバーの骨バイオプシサンプルからのタンパク質の発現の評価 (例えば骨細胞消化、骨髓間質細胞培養の外植) 、一過性または安定性トランスフェクト細胞におけるタンパク質発現レベルの評価、組織、血清または体液におけるタンパク質濃度を評価すること、リガンドハンティングおよび機能的アッセイ開発のためのこれらのタンパク質の全長またはフラグメントの精製、これらのタンパク質と相互作用するリガンドまたはタンパク質の同定、ならびに L R P 6 、L R P 5 および H B M 、および関連する変異体のシグナル発現経路の解明などがある。 30

【0349】

例えば L R P 5 を p c D N A 3 . 1 (Invitrogen、Carlsbad、カリフォルニア州) にクローン化した。これを用いて ^{35}S 標識インビトロ翻訳 (Promega、Madison、ウィスコンシン州) L R P 5 を作製した。ペプチド免疫原 : R W K A S K Y Y L D L N S D S D P Y (配列番号 7 1 1) を志向する抗体 ($10 \mu\text{g} / \text{ml}$) 3 1 0 9 および 3 1 1 0 を、 $10 \mu\text{g} / \text{ml}$ 特異的ペプチド (すなわち R W K A S K Y Y L D L N S D S D P Y) (配列番号 7 1 1) もしくは非特異的ペプチド (すなわち S G W N E C A S S N G H C S H) (配列番号 7 1 2) のいずれかの存在下かまたはペプチド不含でインビトロ翻訳産物 $20 \mu\text{l}$ と組み合わせ、そして 4°C で 1 . 5 時間インキュベートした。次いでプロテイン A セファロースをサンプル (予め網状赤血球ライゼートで約 1 . 5 時間遮断した) に加え、そしてサンプルを 4°C で 1 時間振盪した。プロテイン A セファロースを P B S 0 . 5 ml で 3 回洗浄した。続いて結合タンパク質を 4 ~ 1 2 % グラジエント N u P A G E ゲル (Invitroge 40 50

n)上で製造者の指示書に従って分離した。ゲルを80℃で30分間乾燥させ、そして次にKodak X-OMAT-ARフィルムに24から48時間暴露した。特異的プローブが³⁵S標識LRP5免疫沈澱タンパク質に関していずれかの抗体と有意に競合するのが観察された。競合は非特異的タンパク質では観察されなかった。

【0350】

これらの抗体を免疫組織化学に用いることもできる。例えば、HBMトランスジェニックおよび野生型マウスを、CO₂麻酔を用いて屠殺した。マウス頭蓋冠を無傷で取り出し、軟部組織を穏やかに切開した。さらに処理および分析するために、骨を10%リン酸塩緩衝ホルマリン中24時間固定した。固定の後、頭蓋冠をTBD-2脱灰試薬(Sandon, Pittsburgh、ペンシルバニア州)で約7から8時間脱灰し、次に等級別アルコールで脱水した。次いで頭蓋冠を頭頂骨の中央部分を通してラムダ状縫合および冠状縫合まで矢状縫合に対して垂直に二等分し、パラフィン包埋した。4から6個の、5μm厚の典型的な切片を切断した。

【0351】

例えばウサギポリクローナル抗体、LRP5/HBM(すなわち抗体3109および3110)はHBMトランスジェニックおよび野生型の双方のマウス頭蓋冠でLRP5を認識する。タンパク質発現の存在量およびパターンを評価するために抗LRP5または抗HBM抗体を用いてLRP5またはHBMタンパク質を検出することができる。抗体を検出可能な物質に結合させる(すなわち物理的連結)ことにより検出を促進することができる。検出可能な物質の実例としては、種々酵素、補欠分子族、蛍光材料、発光材料、生物発光材料、および放射活性材料などがある。適当な酵素の実例には西洋ワサビペルオキシダーゼ、アルカリ性ホスファターゼ、P-ガラクトシダーゼ、またはアセチルコリンエステラーゼなどがあり;適当な補欠分子族複合体の実例としてはストレプトアビジン/ビオチンおよびアビジン/ビオチンなどがあり;適当な蛍光材料の実例にはウンベリフェロン、フルオレセイン、フルオレセインイソチオシアネート、ローダミン、ジクロロトリアジニルアミンフルオレセイン、塩化ダンシルまたはフィコエリスリンなどがあり;発光材料の実例としてはルミノールなどがあり;生物発光材料の実例としてはルシフェラーゼ、ルシフェリンおよびアクオリエンなどがあり;適当な放射活性材料には¹²⁵I、¹³¹I、³⁵S、および³Hなどがある。また別に、1次LRP5ポリクローナル抗体の存在を検出する2次抗体を用いることができる。実例は全てのウサギ免疫グロブリンを認識する抗体である。この2次抗体を前記したのと同じの様式で結合させて検出を促進することができる。対照はアビジンペルオキシダーゼを伴うサンプルを含むが、しかし抗体は含まなかった。間質細胞および間葉細胞の強い陽性染色が縫合部分で観察された。野生型マウスと比較してHBMの頭蓋冠の骨膜内で前骨芽細胞および骨芽細胞が、いくつかの骨細胞が抗体3109および3110と共に観察された。HBMトランスジェニックマウスの組織頭蓋冠切片の高度な拡大図により骨細胞および縫合部分内の細胞の顕著な細胞膜染色が示された。

【0352】

XX. 使用方法: 遺伝子治療

近年、遺伝的および後天的疾患の遺伝子治療の分野で著明な技術的進歩が為されている。(Kayら、Proc.Natl.Acad.Sci.USA 94:12744-12746(1997))遺伝子治療を、治療目的のための慎重なDNA移動として特徴付けることができる。遺伝子移動方法の改良により疾患の多様な型の処置のための遺伝子治療プロトコルの開発が可能になっている。遺伝子治療はまた新規治療用遺伝子の同定、ウイルス性および非ウイルスの双方の遺伝子分配システムにおける改良、遺伝子制御のよりよい理解、ならびに細胞単離および移植における改良における最近の進歩を巧みに利用している。

【0353】

先行する実験ではHBM遺伝子が骨量上昇を付与する優性変異として同定されている。この変異が優性であるという事実はHBMタンパク質の発現が骨量上昇を引き起こすことを示している。HBM遺伝子を担持する、従ってHBMタンパク質を発現する老齢の個体

は骨粗鬆症を患わない。これらの個体はHBMタンパク質で処置されている個体と同等である。これらの観察はHBMタンパク質での治療的処置は骨粗鬆症を防御するという強力な実験的表示である。HBM遺伝子の骨量上昇活性を「HBM機能」と称する。

【0354】

従って、本発明によれば、間葉幹細胞にHBM機能を供給する方法もまた提供される。(Onyiaら、J.Bone Miner.Res. 13:20-30(1998);Koら、Cancer Res.54:4614-4619(1996))。このような機能を供給することにより骨粗鬆症に対する保護が提供される。HBM遺伝子または該遺伝子の部分を、遺伝子が染色体外に留まるようにベクターの細胞に導入することができる。このような状況では、遺伝子は染色体外位置から細胞により発現される。

【0355】

組換えおよび染色体外維持の双方のための遺伝子を導入するためのベクターは当分野で公知であり、いずれかの適当なベクターを用いることができる。DNAを細胞に導入するための方法、例えばエレクトロポレーション、リン酸カルシウム共沈殿、およびウイルス性形質導入が当分野で公知であり、方法の選択は当業者の能力内である(Robbins編、Gene Therapy Protocols, Human Press、ニュージャージー州(1997))。HBM遺伝子で形質転換された細胞をモデル系として用いて骨粗鬆症および骨成長を促進する薬物処置を研究することができる。

【0356】

一般的に前記で論じたように、HBM遺伝子またはフラグメントを、適用可能な場合、遺伝治療法を用いて間葉幹細胞におけるこのような遺伝子の発現産物の量を増加させることができる。野生型遺伝子が正常に発現されるこれらの細胞においてさえ、所定のHBMタンパク質、またはそのフラグメントの発現レベルを上昇させることも有用であり得る。例えばFriedman、Therapy for Genetic Disease,Friedman編、Oxford University Press、105-121頁(1991)に記載されるように一般に認められている方法に従って遺伝子治療を実施する。

【0357】

発現調節エレメントに連結され、間葉幹細胞内で複製可能なHBM遺伝子のコピーを含むウイルスまたはプラスミドベクターを調製する。適当なベクターは公知であり、例えば米国特許第5,252,479号およびW093/07282(その開示は全て参照により本明細書に組み込む)に記載されている。次いでベクターを骨髄に局所的、または全身的(別の部位に位置するいずれの間葉幹細胞にも到達させるために、すなわち血中で)のいずれかで親に注射する。トランスフェクトされた遺伝子が標的細胞の各々のゲノムに永久的に組み込まれない場合、処置を定期的に繰り返さなければならない。

【0358】

当分野で公知の遺伝子移入系は本発明の遺伝子治療法の実行において有用であり得る。これにはウイルス性および非ウイルス性移動方法が含まれる。ポリオーマ、すなわちSV40(Madzakら、J.Gen.Virol.73:1533-1536(1992))、アデノウイルス(Berkner、Curr.Top.Microbiol.Immunol.158:39-61(1992);Berknerら、Bio Techniques 6:616-629(1988);Gorzigliaら、J.Virol.66:4407-4412(1992);Quantinら、Proc.Natl.Acad.Sci.USA 89:2581-2584(1992)、Rosenfeldら、Cell 68:143-155(1992);Wilkinsonら、Nucl.Acids Res.20:2233-2239(1992);Stratford-Perricaudetら、Hum.Gene Ther.1:241-256(1990))、ワクシニアウイルス(Mackettら、Biotechnology 24:495-499(1992))、アデノ随伴ウイルス(Muzyczka、Curr.Top.Microbiol.Immunol.158:91-123(1992);Ohiら、Gene 89:279-282(1990))、HSVおよびEBVなどのヘルペスウイルス(Margolskee、Curr.Top.Microbiol.Immunol.158:67-90(1992);Johnsonら、J.Virol.66:2952-2965(1992);Finkら、Hum.Gene Ther.3:11-19(1992);Breakfieldら、Mol.Neurobiol.1:337-371(1987);Fresseら、Biochem.Pharmacol.40:2189-2199(1990))、ならびに鳥類のレトロウイルス(Brandyopadhyayら、Mol.Cell Biol.4:749-754(1984);Petropoulosら、J.Virol.66:3391-3397(1992))、ネズミ(Miller、Curr.Top.Microbiol.Immunol.158:1-24(1992);Millerら、Mol.Cell Biol.5:431-437(1985);Sorgeら、Mol.Cell Biol.4:1730-1737(1984);Mannら、J.Virol.54:401-407(1985))および

10

20

30

40

50

ヒト起源 (Pageら、J.Virol.64:5370-5276(1990); Buchschalcherら、J.Virol.66:2731-2739(1992);)などの多くのウイルスが遺伝子移動ベクターとして用いられている。たいいていのヒト遺伝子治療プロトコルは無能化されたネズミレトロウイルスに基づいている。

【0359】

当分野で公知の非ウイルス性遺伝子移動法には化学的技術、例えばリン酸カルシウム共沈殿 (Gahamら、Virology 52:456-467(1973); Pellicerら、Science 209:1414-1422(1980))、機械的技術、例えばマイクロインジェクション (Andersonら、Proc.Natl.Acad.Sci.USA 77:5399-5403(1980); Gordonら、Proc.Natl.Acad.Sci.USA 77:7380-7384(1980); Brinsterら、Cell 27:223-231(1981); Constantiniら、Nature 294:92-94(1981))、リポソームを介する膜融合媒介移動 (Felgnerら、Proc.Natl.Acad.Sci.USA 84:7413-7417(1987); Wangら、Biochemistry 28:9508-9514(1989); Kanedaら、J.Biol.Chem.264:12126-12129(1989); Stewartら、Hum.Gene Ther.3:267-275(1992); Nabelら、Science 249:1285-1288(1990); Limら、Circulation 83:2007-2011(1992))および直接DNA取り込みおよびレセプター媒介DNA移動 (Wolffら、Science 247:1465-1468(1990); Wuら、BioTechniques 11:474-485(1991); Zenkeら、Proc.Natl.Acad.Sci.USA 87:3655-3659(1990); Wuら、J.Biol.Chem.264:16985-16987(1989); Wolffら、BioTechniques 11:474-485(1991); Wagnerら、(1990); Wagnerら、Proc.Natl.Acad.Sci.USA 88:4255-4259(1991); Cottenら、Proc.Natl.Acad.Sci.USA 87:4033-4037(1990); Curielら、Proc.Natl.Acad.Sci.USA 88:8850-8854(1991); Curielら、Hum.Gene Ther.3:147-154(1991))などがある。ウイルス媒介遺伝子移動を、周囲細胞ではなく間葉幹細胞への直接インビボベクターと組み合わせることができる (Romanoら、In Vivo 12(1):59-67(1998); Gonezら、Hum.Mol.Genetics 7(12):1913-1919(1998))。また別に、レトロウイルスベクタープロデューサー細胞株を骨髄に注射することができる (Culverら、Science 256:1550-1552(1992))。次いでプロデューサー細胞の注射によりベクター粒子の継続的な供給源が提供される。この技術は手術不能な脳腫瘍のヒトにおいての使用が認可されている。

【0360】

生物学的および物理学的遺伝子移動方法を組み合わせた研究法では、いずれかの大きさのプラスミドDNAをアデノウイルスヘキソタンパク質に特異的なポリリジン結合抗体と組み合わせ、得られた複合体をアデノウイルスベクターと結合させる。次いで3分子複合体を用いて細胞を感染させる。アデノウイルスベクターにより、結合したDNAが損傷を受ける前に、エンドソームの効率的な結合、内部移行および分解が可能になる。

【0361】

リポソーム/DNA複合体が直接インビボ遺伝子移動を媒介できることが示されている。標準的なリポソーム調製では遺伝子移動方法は非特異的であるが、インビボ取り込みおよび発現が、例えば直接インサイチュ投与に続いて腫瘍沈着物に局在することが報告されている (Nabel、Hum.Gene Ther.3:399-410(1992))。

【0362】

XXI. 使用方法：研究手段としておよび医薬品開発のための形質転換宿主およびトランスジェニック動物

HBM、LRP5またはLRP6遺伝子を担持する細胞および動物はモデル系として用いられ、治療用物質としての可能性を有する物質を研究および試験するための貴重な研究手段である (Onyiaら、J.Bone Miner.Res.13:20-30(1998); Broderら、Bone 21:225-235(1997))。この目的のための細胞は典型的には培養された間葉幹細胞である。これを体細胞系または生殖系HBM遺伝子と共に個体から単離することができる。また別に、前記したように細胞株を操作してHBM遺伝子を担持するようにできる。被験物質を細胞に適用した後、細胞の形質転換された表現型を決定する。形質転換された細胞のいずれかの形質、例えば培養物中の骨マトリックスの形成 (Broderら、Bone 21:223-235(1997))、機械的特性 (Kizerら、Proc.Natl.Acad.Sci.USA 94:1013-1018(1997))、マーカー遺伝子の発現および推定される治療用物質の適用に対する応答を評価することができる。

【0363】

トランスジェニック修飾動物および細胞株を用いて治療用物質を試験することができる。トランスジェニック修飾には、例えばLRP5遺伝子の挿入、ならびにHBM遺伝子および破壊された相同遺伝子の挿入などがある。また別に、慣用される技術、例えばCapecci、Science 244:1288(1989);Valancuisら、Mol.Cell Biol.11:1402(1991);Hastyら、Nature 350:243(1991);Shinkaiら、Cell 68:855(1992);Mombaertsら、Cell 68:869(1992);Philpottら、Science 256:1448(1992);Snouwaertら、Science 257:1083(1992);Donehowerら、Nature 356:215(1992)に記載される技術を用いて挿入された動物の(複数の)LRP5遺伝子および/または(複数の)HGM遺伝子を別の遺伝的変化の挿入または欠失により破壊することができる。被験物質を動物に投与した後、骨の成長を評価しなければならない。被験物質が骨の成長を変調する(例えば増強する)場合、すると被験物質が治療物質候補である。これらの動物モデルが、可能性のある治療用生成物の非常に重要なベヒクルを提供する。

10

【0364】

本発明はまた、内因性遺伝子の発現が活性化されており、さらに増幅され得る動物および細胞株をも提供し、これはLRP5またはHBMタンパク質をコードする内因性DNAのインビトロ操作およびトランスフェクションを必要としない。これらの方法は例えばPCT出願W094/12650および米国特許第5,968,502号に記載されており、これは双方共にその全てを参照により本明細書に組み込む。加えて、本発明は、内因性活性化または過剰発現がインサイチュ相同組換え、相同組換え、または非正統的組換え法により達成される方法を含む。これらの方法は例えばPCT出願W099/15659およびW000/49162に記載されており、その双方ともにその全体に参照により本明細書に組み込む。

20

【0365】

トランスジェニックおよびジーンターゲティング動物の作製

本発明はヒトHBM表現型を反復する遺伝的に修飾された動物を提供する。採用した研究法は、ゲノムに組み込まれたヒトGからTのヌクレオチド置換を有するトランスジェニックおよびジーンターゲティングの双方の動物の作製に関与し、動物はヒトLRP5(Zmax1)を発現するかまたは骨量を変化させる表現型を生み出す変異体を発現する。本発明の製造に続いて、Katoら(Journal of Cell Biology 157:303-314(2002))は最近低骨量表現型を実証するLRP5ノックアウトマウスの作製について記載している。Katoおよび共同研究者により記載された結果はLRP5がピーク骨量の決定因子であるという仮説に合致し、本発明の利用性の態様を実証する。

30

【0366】

HBM多型性を過剰発現するトランスジェニックマウス

CMVbアクチンまたはI型コラーゲンプロモーターのいずれかを利用してヒトHBM cDNAの発現を引き出すプラスミド構築物を調製した。最も一般的に用いられる哺乳動物発現のためのプロモーターはサイトメガロウイルス(CMV)、ラウス肉腫ウイルス(RSV)、シミアンウイルス40(SV40)、およびEF-1a(ヒト伸長因子1aサブユニット)に由来する。CNVはヒトサイトメガロウイルス極初期プロモーターから誘導される。CMVはたいいていの細胞株においてRSVまたはSV40よりも強力なプロモーターである。RSVプロモーターは鳥類ウイルスから誘導され、鳥類細胞株において強力なプロモーターである傾向がある。SV40プロモーターは大きなT抗原、例えばCOS-1を含む細胞株において良好に発現される。これらの細胞株ではプラスミドはより高いコピー数で複製される。EF-1aは組換えタンパク質発現のためにより広範に用いられ始めている。EF-1aはヒト伸長因子1aサブユニットからのプロモーターであり、遺伝子は哺乳動物細胞株において高度に発現され保存される。

40

【0367】

キメラCMVbアクチンプロモーターは骨などトランスジェニックマウスにおける偏在性遺伝子発現を生み出すことが知られているより強力なプロモーターである。このプロモーターを選択して、報告されている内因性マウスLRP5遺伝子の広範な発現と合致する様式でHBMの発現を誘導した。HBM表現型は骨において観察されるが、HBM遺伝子

50

は別の組織において直接的または間接的効果を有し得る。従って、別の強力な偏在性プロモーターを当業者に公知であるように利用することができる。

【0368】

I型コラーゲンプロモーターは組織限定遺伝子発現を提供し、この場合発現は本来骨に限定されている。骨においてHBMの発現に至り得る別の骨特異的プロモーターが利用可能である。例えば、骨芽細胞の増殖に関連するプロモーターにはヒストン、I型コラーゲン、 $TGF\beta 1$ 、 $MSX 2$ 、 $c fos / c Jun$ および $Cbfa1$ を用いることができる。アルカリ性ホスファターゼ、MGP、 $Cbfa1$ 、 Fra / Jun および $Dlx 5$ などの骨マトリックス成熟化に随伴されるプロモーターを用いることもできる。オステオカルシン、オステオポンチン、骨シアロプタンパク質などの骨ミネラル化作用に随伴されるプロモーター、およびコラゲナーゼを用いることもできる。当業者に公知であるように、プロモーター選択を例えば組織発現、制御可能な変調の程度等により決定する。例えば、I型コラーゲンプロモーターを選択してHBMが、骨におけるLRP5発現の内因性パターンに類似する一時的、空間的および骨細胞特異的パターンで骨において発現されることを確認した。

【0369】

野生型LRP5遺伝子を過剰発現するトランスジェニックマウス

LRP5の発現を誘導するCMVbアクチンおよびI型コラーゲンプロモーターを用いてプラスミド構築物を調製した。これらの動物をHBMトランスジェニックマウスの対照動物モデルとして提供できる。さらなる対照には非トランスジェニック同腹仔および野生型同一の遺伝的背景の動物などがある。これらの動物の調製方法はHBM多型性を過剰発現するマウスに関して論じられているものに類似している。

【0370】

HBM多型性を発現するジーンターゲティングマウス

HBMノックイン(KI)対立遺伝子およびLRP5ノックアウト(KO)対立遺伝子を含む動物を作製するために用いることができるジーンターゲティング構築物を調製した。ジーンターゲティング構築物はエクソン3にHBM多型性を含有し、転写停止配列に連結され、 $lox P$ 部位が側面に位置するネオマイシン選択カセットを含んだ。マウスLRP5におけるHBM多型性はマウスLRP5相同体(Genbankアクセッション番号AF064984)の位置170でのグリシンのバリンへのアミノ酸配列の変化に至る。相同組換えを用いて構築物をマウスゲノムに安定して導入する。転写停止配列が機能して修飾されたLRP5対立遺伝子の転写を完全に遮断する場合、次いで機能的LRP5ノックアウト対立遺伝子を作製する。これはLRP5遺伝子に関するホモ接合性ノックアウト動物の作製を促進する。

【0371】

ノックイン対立遺伝子を作製する場合、Creレコンビナーゼを用いて、修飾されたエクソン3および $lox P$ 部位の1つのコピーをそのままにして、ネオマイシン選択カセットを切除することができる。Creを一細胞受精胚に導入してHBMの偏在性発現を促進するか、またはトランスジェニックマウスと動物の交雑により骨特異的HBM発現を得ることができる。HBMノックイン対立遺伝子に関してホモ接合性動物を作製することができる。また別に、当分野で公知であるように、ホモ接合性動物からの核が準備された卵母細胞(例えば脱核されている)に移される核移動技術により動物を作製することができる。例えばCampbellら、Nature 380:64-68(1996)参照。ノックアウトマウスを作製するさらなる方法には、ATG開始コドンが欠失または変異されている相同組換えベクターを操作すること、ベクターにフレームシフト変異を操作すること、プロモーター領域の重要な部分の欠失を操作すること、および/または遺伝子の重要な領域を欠失するようにベクターを操作することなどがある。

【0372】

材料および方法

LRP5プラスミドZMAXGI 3ASの構築

全長LRP5 cDNA構築物をLife TechnologiesからのpCMVSPORT6.0ベクター(Gatewayクローンシステムの一部)のXbaI-NotI部位に操作して、ZMAXGI 3ASを作製した。5'側をHindIIIまたはXbaIのいずれかで、それと共に3'側でEcoRVまたはEcoRIのいずれかで消化することによりベクターからインサート(5,278bp)を放出することができる。

【0373】

4つの独立した粒子クローンからLRP5構築物を作製した。これらのクローンをLRP5特異的プライミングしたcDNAライブラリーから単離した。部分的LRP5 cDNAクローンは内部探索シーケンシングクローンセットに、オリゴdTプライミングHela細胞cDNAライブラリーから単離されたL236B P0049E08として存在した。このクローンを5'プライミング末端でランケートした。全長cDNAを作製するのに必要なより多くの5'プライム含有フラグメントを単離するために、Clontechヒト肝臓ポリA mRNA(カタログ番号6510~1、ロット番号9060032)およびcDNA合成およびプラスミドクローニングキット(カタログ番号18248~013)のLife Technologies SuperScript(登録商標)プラスミドシステムからLRP5遺伝子特異的cDNAライブラリーを作製した。このライブラリーをL401と称した。製造者のプロトコルを以下の修飾を伴って実施した。1)第1および第2鎖合成反応の双方におけるDEPC処理水を α -³²P-dCTPと置き換えた。2)およそ1kb間隔でLRP5遺伝子に特異的であるように選択されたオリゴヌクレオチドを用いて逆転写をプライミングした。公のデータベースに対してプログラムBLASTを用いてこれらの配列を調査してLRP5遺伝子特異性を確認した。3)2つの別個の逆転写反応を実施した。最初の反応(A)は以下の、より多くのLRP5遺伝子の3'領域、(配列番号715~718)47114: 5'-CGTACGTAAAGCGGCCGCTTGGCAATACAGATGTGGGA-3',47116: 5'-CGTACGTAAAGCGGCCGAGTAGCTCCTCTCGGTGGC-3',47118: 5'-CGTACGTAAAGCGGCCGCGCTCATCATGGACTTTCCG-3',および47120: 5'-CGTACGTAAAGCGGCCGCGCACTGCTGTTTGATGAGG-3'にアニーリングするオリゴヌクレオチドでプライミングした。第2の反応(B)は前記した4つのオリゴヌクレオチド、および(配列番号719~721)47108: 5'-CGTACGTAAAGCGGCCGCGAGTGTGGAAGAAAGGCTGC-3',47110: 5'-CGTACGTAAAGCGGCCGCGAGTAGAGCTTCCCCTCCTGC-3',および47112: 5'-CGTACGTAAAGCGGCCGCGTCCATCACGAAGTCCAGGT-3'を使用した。全てのオリゴヌクレオチドはNotIリンカー配列を含有し、0.02μg/μlの濃度で使用した。4)双方の逆転写反応物に由来するSalI適合cDNAを1%アガロース、0.1μg/ml臭化エチジウム、1x TAEゲル上の電気泳動によりサイズ分画化した。0.6から8.0kbの臭化エチジウム染色cDNAをゲルから切除した。製造者のプロトコルを用いてエレクトロエリクション(ISC0 Little Blue Tank Electroelutor(登録商標))によりアガロースゲルからcDNAを精製した。次いで反応AおよびBに由来する精製cDNAを一緒にプールした。5)サイズ分画化SalI適合cDNAをNotI-SalI消化pBluescript(登録商標)(Stratagene、La Jolla、カリフォルニア州)にライゲートした。

【0374】

ライゲートされたライブラリーcDNA(3μl)を用いてエレクトロコンピテント大腸菌(ElectroMAX(登録商標)DH10B細胞およびプロトコル、Life Technologiesカタログ番号18290-015、BioRad E.coliパルサー、電圧1.8KV、3-5ミリ秒パルス)を形質転換した。形質転換された細胞をLBアンピシリン(100μg/ml)寒天プレートにプレートし、37℃で一晩インキュベートした。およそ10⁶コロニー形成単位(cfu)を50,000cfu/150mmプレートの密度でプレートした。LB培地でプレートから細胞を洗い流し、遠心により収集した。QIAGEN プラスミドGigaキットおよびプロトコル(カタログ番号12191)を用いて最終濃度2.05μg/μlでプラスミドRNAを細胞から精製した。

【0375】

ライブラリースクリーニングに使用するための2つのプローブを、鋳型としてライブラリーL401 100ngを用いるポリメラーゼ連鎖反応(PCR)により作製した。標

準的なPCR技術を用いた。反応混合物は各オリゴプライマー 10ピコモル；0.2 mM 各dATP、dTTP、dCTPおよびdGTP (PE Applied Biosystemsカタログ番号N808-0260)；Expand (商標) High Fidelity Taq DNAポリメラーゼ 1.5単位、および1x PCR反応バッファー(Roche Molecular Biochemicals、カタログ番号731-641；10mM Tris-HCl、1.5mM MgCl₂、50mM KCl、pH 8.3)を含有した。混合物を99℃で1分間インキュベートし、続いて96℃で15秒間、50℃で30秒間、72℃で1分間の30サイクル、72℃で7分間の最終インキュベーションを行った(MJResearch DNA Engine(登録商標)Tetrat PCT-225)。LRP5遺伝子のNotI-SalIフラグメントを増幅するオリゴ(配列番号722-723) 107335：5'-CAGCGGCCTGGAGGATGC-3'および107338：5'-CGGTCCAGTAGAGGTTTCG-3'を用いて第1鎖を作製した。LRP5遺伝子のSacI-KpnIフラグメントを増幅するオリゴ(配列番号724-725) 107341：5'-CATCAGCCGCGCCTTCATG-3'および107342：5'-CCTGCATGTTGGTGAAGTAC-3'を用いて第2鎖を作製した。Qiaquick (登録商標) キットおよび製造者のプロトコル (Qiagenカタログ28106)を用いて双方のPCR産物を精製し、次いで製造者のプロトコルに従ってベクターpCRII-TOPO (登録商標) (Invitrogenカタログ番号K4600)にサブクローン化した。精製されたプラスミドDNA (標準的な分子生物学技術を使用)の制限消化および続くDNA配列分析(ABI Prism BigDye Terminator Cycle(登録商標)シーケンシング、カタログ番号4303154、ABI377機器)により陽性サブクローンを同定した。各々の配列確認されたpCRII-TOPO (登録商標) クローンのEcoRI制限消化(New England Biolabs、カタログ番号R0101L)によりプローブDNAを単離した。1%アガロース、0.1 μg/ml 臭化エチジウム、1x TAEゲル上のゲル電気泳動により制限フラグメントをサイズ分画化した。インサートDNAをゲルから切除し、製造者のプロトコルに従ってClontech NucleoSpin(登録商標)核酸精製キット(カタログ番号K3051-2)を用いて精製した。Prime-It II(登録商標)ランダムプライマー標識キットおよびプロトコル(Stratagene、カタログ番号300385)を用いて、精製されたDNAフラグメント(25~50 ng)をRedivue (登録商標) (a³²P)-dCTP(Amersham Pharmacia、カタログ番号AA0005)で標識した。組み込まれなかったdCTPをAmershamのNICK (登録商標) カラムおよびプロトコル(カタログ番号17-0855-02)で除去した。

【0376】

ライブラリーL401のスクリーニングの第2ラウンドを開始し、LRP5遺伝子のフラグメントを単離した。最初に43個の150 cm LB-100 μg/mlアンピシリン寒天プレートをL401に由来する1次形質転換体で、プレートあたり約3,000~4,000コロニーの密度でプレートした。標準的な分子生物学プロトコルを用いて、前記したように500,000から1,000,000 cpm/mlハイブリダイゼーションバッファーで³²P標識プローブ(NotI-SalIフラグメント)を用いてこのライブラリーをスクリーニングした。この1次スクリーニングから、LRP5プローブに対する陽性ハイブリダイゼーションに基づいて、13個の単一コロニーが同定された。QIAprep Spin ミニプレップキットおよびプロトコル(Qiagen Inc.、カタログ番号27106)を用いて調製したプライマーDNAを前記のように制限消化および配列分析により分析した。このスクリーニングからcDNAクローン#44を単離し、配列が部分的LRP5クローンを含有することを確認した。

【0377】

第2のライブラリースクリーニングでは、104個の150 cm LB-100 μg/mlアンピシリン寒天プレートをL401に由来する1次形質転換体で、プレートあたり約3000~4000コロニーの密度でプレートした。³²P標識プローブ(SacI-KpnI)を用いて、正確に前記したようにこのライブラリーをスクリーニングした。この1次スクリーニングで、LRP5プローブに対する陽性ハイブリダイゼーションに基づいて、48個の単一コロニーが同定された。これらのコロニーは単一のコロニー単離体ではないので、48個の1次単離体の各々をプレートあたりおよそ500コロニーの密度で

プレートする2次スクリーニングを開始した。次いでこれらのコロニーを、プローブとして標識 Sac I - Kpn I フラグメントを用いて1次と正確に同様にスクリーニングした。48個の1次クローンのうち34個がLRP5プローブに対して陽性ハイブリダイゼーションになり、単一コロニーから単離された。プラスミドDNAを調製し、前記したように分析した。cDNA単離体#712をこのスクリーニングから単離し、配列が部分的LRP5クローンを含むことを確認した。

【0378】

全例でいずれかのLRP5遺伝子単離体の配列を参照配列（すなわちHBM家系の罹患メンバーに由来する野生型LRP5対立遺伝子の配列）と比較した。文献でDNA多型性がこの遺伝子に関して報告されていたので、この分析は重要であった。この参照配列によりGenbankアクセッション番号AF077820のポリペプチドをコードすることが予測される。

10

【0379】

Zmax IGI 3ASを調製するために用いた4個の別個の部分クローンは以下のとおりである。

【0380】

1) 塩基1-1366: LRP5 cDNA構築物、GTC. Zmax113からXba I - Sal I フラグメントが得られた。GTC. Zmax113は、LRP5の全ORFを含む5057bpインサートを含む。クローンを平滑末端化し、そしてpSTBlue-1のEcoRV部位にクローン化した。pBluescript (商標) II 誘導体の骨ランダムプライミングcDNAライブラリーのスクリーニングに由来する5'クローンを、pBluescript (商標) IIの骨dTプライミングcDNAライブラリーからのPCR産物に由来する3'クローンと融合させることによりこのクローンを作製した。DNAzymeポリメラーゼを用いてLRP5特異的順行性プライマー（配列番号726）5'-GCCCGAAGACCTCTACTGGACCGAC-3'および逆行性プライマー（配列番号727）5'-GCCCAACCCCAATCACAGTTCCACATT-3'を用いてPCRを実施した。得られた3.7kb PCR産物をPCR-XL-TOPOにクローン化した。全長クローン作製するために、5'および3'プラスミドをGibco/BRLのDM1 (dam-) に形質転換した。5'プラスミドをXba Iで消化し、そして3'プラスミドをHind IIIで消化した。消化されたプラスミドをT4ポリメラーゼで埋めて平滑末端を作製し、そしてBcl Iで切断した。1.7kb 5'フラグメントおよび3.5kb 3'フラグメントをゲル精製し、一緒にライゲートし、そしてpSTBlue1のEcoRV部位にクローン化した。これにより塩基100で始まるコード化配列を伴う短い5'UTRが提供される。さらに、これは5'多重クローニング部位でいくつかのさらなる制限部位を担持する。このフラグメントは位置558でGenbankアクセッション番号AF077820配列に相対する、結果的にA (AF077820) からGに変化するDNA多型性をも含有し; この変異はアミノ酸の差異には至らない (Pro)。

20

30

【0381】

2) 塩基1367-2403: 前記の市販のヒト肝臓RNAから作製されたLRP5遺伝子プライミングcDNAライブラリーからこのクローンを入手した。このフラグメントは単離体#44から得られたDNAのSal I - BamHI片である。配列はGenbankアクセッション番号AF077820と同一である。

40

【0382】

3) 塩基2404-4013: 前記したのと同じライブラリーに由来する単離体#71-2からこのBamHI - BssHIIフラグメントを入手した。位置3456にはG (AF077820) からAに至るDNA多型性がある。このヌクレオチドの差異はコード化アミノ酸 (Val) では変化しない。

【0383】

4) 塩基4014-5278: このBssHII - NotIフラグメントは内部クロー

50

ン L 2 3 6 B P 0 0 4 9 E 0 8 に由来する。これをオリゴ d T プライミング H e L a 細胞 c D N A ライブラリーから入手した。停止コドンは塩基 4 9 4 7 に出現する。クローンは、1 2 0 b p ポリ A テイル続いて N o t I 部位を含む 3 ' U T R 配列の 3 3 1 b p を含有する。このサブクローニング工程で用いる 3 ' N o t I 部位は、ライブラリー構築の間にポリ A の末端で導入された、付加されたリンカーの結果である。D N A 多型性は塩基 4 5 1 5 に存在し、アミノ酸レベルではサイレントである (L e u) G (A F 0 7 7 8 2 0) から C への変化に至る。

【0384】

L R P 5 遺伝子の 5 ' セクションを作製するために、X b a I - S a l I フラグメントおよび S a l I - B a m H I フラグメントを pBluescript(登録商標)(Stratagene)の X b a I - B a m H I 部位にライゲートした。遺伝子の 3 ' セクションは L R P 5 単離体 # 7 1 2 からの 1. 6 1 k b B a m H I - B s s H I I フラグメントを L 2 3 6 B P 0 0 4 9 E 0 8 からの 1. 2 6 k b B s s H I I - N o t I フラグメントへのライゲーションにより入手した。これらの 2 つのフラグメントを p B l u e s c r i p t (登録商標)の B a m H I - N o t I 部位にライゲートした。X b a I - B a m H I 5 ' セクションおよび B a m H I - N o t I 3 ' セクションのライゲーションにより、全長 L R P 5 c D N A をベクター p C M V S P O R T 6. 0 (Life Technologies)の X b a I - N o t I 部位に操作した。得られたプラスミド、Z m a x G I 3 A S は X b a I 部位からの 5 2 7 8 b p の、ベクター多重クローニング部位の N o t I 部位へのインサートを含む。このクローンはベクターに存在する C M V プロモーターに関してアンチセンス配向である。L R P 5 コード化配列は 1 0 0 塩基で始まり、4 9 4 7 塩基で終わり、1 2 0 b p ポリ A テイルを含む 3 - プライミング U T R 配列の 3 3 1 b p が続く。この全長 c D N A は、予測されるアミノ酸配列を変化しない、参照配列 (G e n b a n k アクセッション番号 A F 0 7 7 8 2 0) からの 3 つの D N A 多型性を含有する。これらの多型性は位置 5 5 8 ではプロリン残基を維持する A から G に至り；位置 3 4 5 6 ではバリン残基を維持する G から A に至り；位置 4 5 1 5 ではロイシン残基を維持する G から C に至る。

【0385】

Z M A X G I 3 A S (図 2 5) の配列はまた、位置 1 3 3 0 でアラニンのバリンへのアミノ酸変化に至る、C (配列番号 1 および G e n b a n k アクセッション番号 A B 0 1 7 4 9 8) の T への変化に至る配列番号 1 の塩基 4 0 8 8 に相対する D N A 多型性をも含有する。これは M B M 家系の罹患メンバーに由来する野生型対立遺伝子で決定された配列および G e n b a n k アクセッション番号 A F 0 7 7 8 2 0 の公開された配列に合致する。Z M A X G I 3 A S はまた配列番号 1 に相対する 5 ' 末端で 2 9 個のさらなる塩基、ならびに付加的な G、ポリ A トラクトの 1 2 0 塩基、および N o t I 部位から成る 3 ' 末端を有する。

【0386】

H B M 変異 G 1 7 1 V の作製

P C R を用いて、アミノ酸 1 7 1 でグリシンからバリンへの予測されるアミノ酸変化に至る H B M 変異を全長ヒト L R P 5 c D N A (プラスミド Z M A X G I 3 A S) に導入して、位置 6 1 1 の G を T に変化させた。L R P 5 遺伝子の内因性 N o t I 部位をフランキングするオリゴ (配列番号 7 2 8) 1 0 7 3 3 5 : (5 ' - C A G C G G C C T G G A G G A T G C - 3 ') および (配列番号 7 2 9) 4 9 5 1 3 : (5 ' - C G G G T A C A T G T A C T G G A C A G C T G A T T A G C - 3 ') を用いて H B M 変異の導入を行った。この方法により P C R 産物の 3 ' 末端で新規 P v u I I 部位が作製される。該産物の 5 ' 末端に S c a I 部位を導入するオリゴを用いて第 2 の P C R 反応を完了し、3 ' プライミング末端で L R P 5 の内因性 S a l I 部位を含有する。Q i a q u i c k (登録商標) 手順 (Q i a g e n I n c.) を用いて P C R 産物を精製し；前記したベクター p C R I I - T O P O (I n v i t r o g e n) にサブクローン化した。全て前記のとおり、プラスミド D N A を単一の細菌性コロニーから精製し、制限消化および続く配列分析により分析した。配列確認された p C R I I - T O P O クローンを各々 N o t I - P v u I

I および S c a I - S a l I で制限消化した。得られた D N A フラグメントを前記のとおりサイズ分画化し、精製した。次いでこれらの 2 つのフラグメントを N o t I - S a l I 消化により調製したベクター、p B l u e s c r i p t (登録商標) にサブクローン化した。P v u I I および S c a I 産物の双方は二本鎖核酸を消化するのに用いた場合、平滑末端を生じる。従って、得られたライゲートされたフラグメントは P v u I I または S c a I 部位のいずれかを再生することができず、新たに導入された H B M 変異を除いてコンセンサス L R P 5 配列のみを含有する。全長 L R P 5 遺伝子に変異を導入するために、この得られたプラスミドを M s c I および S a l I で消化したが、一方 L R P 5 プラスミド G T C . Z m a x 1 1 3 の X b a I - M s c I 消化により L R P 5 の 5' 領域が得られた。これらの 2 つのフラグメントと一緒に X b a I - S a l I 消化した p B l u e s c r i p t (登録商標) にライゲートし、結局類似の 1. 3 6 6 k b X b a I - S a l I フラグメントを作製した。違いはこの構築物が前記した H B M 変異を含有するということだけである。次いで L R P 5 遺伝子に関して正確に前記のとおり、この新たに作製された H B M 変異を含む X b a I - S a l I フラグメントの置換を伴う全長 H B M c D N A を p C M V S P O R T 6. 0 に集めた。D N A 配列分析により全 c D N A インサートを確認し、H B M 変異の導入を確認した。

【0387】

得られたプラスミド、H B M G I 2 A S (図 2 4) はベクターの多重クローニング部位の X b a I 部位から N o t I 部位への 5, 2 7 8 b p のインサートを含有する。このクローンはベクターにおける C M V プロモーターに関してアンチセンス配向である。H B M コード化配列は 1 0 0 塩基で始まり、4 9 4 7 塩基で終わり、1 2 0 b p ポリ A テイルを含む 3-プライミング U T R 配列の 3 3 1 b p が続く。この全長 c D N A は、アミノ酸配列を変化しない、参照配列からの 3 つの D N A 多型性を含有する。これらの多型性は位置 5 5 8 ではプロリン残基を維持する A から G への変化に至り；位置 3 4 5 6 ではバリン残基を維持する G から A への変化に至り；位置 4 5 1 5 ではロイシン残基を維持する G から C への変化に至る。加えて、H B M 変異は位置 6 1 1 (L R P 5 の G から H B M の T) に存在し、これにより H B M 家系の罹患メンバーに見出されるように、アミノ酸位置 1 7 1 でグリシンからバリンへの予測されるアミノ酸変化に至る。このインサート配列を用いて H B M 過剰発現トランスジェニックマウスに用いられる構築物を作製した。

【0388】

導入遺伝子調製

本明細書で提供される事例はどのようにトランスジェニック動物を調製できるかの説明である。当分野で公知のようにさらなるトランスジェニック動物を調製できる。例えば Glenn Monastersley ら編、Strategies in Transgenic Animal Studies (Amer. Soc. Microbiology 1995) およびそこに引用される参考文献を参照されたい。

【0389】

C M V β アクチンプロモーター-H B M c D N A (H B M M C B A)

C M V β アクチン-H B M 構築物、p C X - E G F P を調製するために、キメラ C M V β アクチンプロモーターを含有するプラスミドを 4 7 7 8 b p E c o R I フラグメントとして精製した。続いて、H B M G I 2 A S から H B M c D N A を 4 9 9 4 b p X b a I / D r a I フラグメントとして切除し、D N A ポリメラーゼのクレノウフラグメントと共に処理し、E c o R I リンカーにライゲートし、E c o R I で消化した。次いでこのフラグメントを p C X - E G F P の E c o R I 部位に挿入した。S p e I / H i n d I I 7 2 6 5 b p C M V β アクチン-H B M フラグメントを精製したマウス胚にマイクロインジェクションを行った。

【0390】

I 型コラーゲンプロモーター-H B M c D N A (H B M M T I C)

最初に p B S (S K-) (Stratagene) ポリリンカーを別のポリリンカー (すなわち K p n I - S p e I - H i n d I I I - B g l I I - N d e I - S a l I - S m a I - E c o R I - P s t I - B a m H I - X b a I - S c a I - N c o I - C l a I - N o t I -

S a c I I - S a c I を含む) と置換することにより、ラット I 型コラーゲンプロモーター-HBM 構築物を作製し、これを BS (SK-) A/D と称する。p c D N A I (Invitrogen, Inc.) からの S V 4 0 スプライシングおよびポリ (A) _n X b a I - N c o I 領域 (750bp) を直接 BS (SK-) A/D にクローン化した。次に、4994bp E c o R I HBM c D N A フラグメント (前記) を E c o R I 部位にクローン化した。3640bp X b a I である I 型コラーゲンプロモーターフラグメントを BS (SK-) (Stratagene) の X b a I 部位にサブクローン化した。次いでプロモーターフラグメントを BS (SK-) から S a c I I で切除し、T4DNAポリメラーゼで平滑末端化し、S p e I で消化し、HBM BS (SK-) A/D 構築物にライゲートし、N d e I で消化し、T4DNAポリメラーゼで平滑し、S p e I で消化した。S p e I / C l a I 9435bp I 型コラーゲン-HBMフラグメントを精製して、マウス胚にマイクロインジェクションを行った。 10

【0391】

CMVβアクチンプロモーター-LRP5 c D N A (Zmax1WT CBA)

前記からCMVβアクチンプロモーター-HBM c D N A 構築物を用いて最終プラスミドを作製した。以下の3つのフラグメントを一緒にライゲートした。1) HBMM C B A に由来する6.34kb X b a I - K p n I バックボーンフラグメント; 2) b アクチンプロモーターの3'末端およびHBM c D N A の5'末端を含有するHBMM C B A に由来する0.64kb X b a I - S a p I フラグメント; ならびに3) 野生型配列を含有するLRP5 c D N A に由来する2.8kb S a p I - K p n I フラグメント 20。7.2kb S p e I - H i n d I I I CMVβアクチン-LRP5フラグメントを精製して、マウス胚にマイクロインジェクションを行った。

【0392】

I 型コラーゲンプロモーター-LRP5 c D N A (Zmax1WTTIC)

前記からのI型コラーゲン-HBM c D N A 構築物を用いて最終プラスミドを作製した。HBMM T I C プラスミドをH i n d I I I で直線化し、S a l I で切断して8.52kb H i n d I I I - S a l I フラグメントを生じるか、またはS a p I で切断して2.98kb S a p I - H i n d I I I フラグメントを生じた。次いで野生型配列を含有するLRP5 c D N A に由来する2.8kb S a p I - S a l I フラグメントを前記の2つのフラグメントにライゲートして最終プラスミドを作製した。9.4kb S p e I - C l a I I 型コラーゲン-LRP5フラグメントを精製して、マウス胚にマイクロインジェクションを行った。 30

【0393】

インビトロ導入遺伝子発現の確認

HBMM C B A HBMM T I C、Xmax1WT CBAおよびZmaxWTTICに関するプラスミド構築物を一過性にヒト骨芽(HOB)細胞にトランスフェクトし、機能性に関する試験としてmRNA発現を測定した。

【0394】

一過性トランスフェクション

HOB-02-02細胞はHOB-02-C1細胞に由来するクローン性、老化後、細胞株である(Bodineら、J.Bone Miner.Res.11:806-819(1996))。親細胞株と同様に、HOB-02-02細胞は温度感受性SV40ラージT抗原変異体、tsA209を発現する。結果的に、これらの細胞は許容温度34℃で増殖するが、非許容温度37℃以上で分割を停止する。親細胞株と同様に、5%CO₂/95%加湿空気インキュベーター(Forma Scientific, Marietta, オハイオ州)中34℃で成長培地(10%熱不活性化ウシ胎仔血清、1%ペニシリン・ストレプトマイシンおよび2mM Glutamax-1を含有するDMEM/F-12)でHOB-02-02細胞を培養する。 40

【0395】

一過性トランスフェクションに関して、HOB-02-02細胞を成長培地に400,000セル/ウェルで6ウェルプレートに播種し、34℃で一晩インキュベートした。L 50

i p o f e c t A M I N E 2 0 0 0 トランスフェクション試薬を用いて、製造者の指示書に従って(Life Technologies、Rockville、メリーランド州)、0.3 mg / ウェル C M V β アクチン-H B M 発現プラスミド、I 型コラーゲン-H B M 発現プラスミド、または対応する空のベクターのいずれかで細胞をトランスフェクトした。34℃で24時間インキュベートした後、培地を交換し、細胞をさらに39℃で24時間インキュベートした。この最後のインキュベーションの終わりに、細胞をハンクス緩衝食塩水ですすいだ。次いでT R I z o l (登録商標)を用いて、製造者の指示書(GibcoBRL、Grand Island、ニューヨーク州)に従って、全細胞R N A を単離した。前記したようにR N A をR N アーゼ不含D N アーゼと処理して夾雑するD N A を除去した(Bodineら、J.Cell.Biochem.65:368-387(1997))。

10

【0396】

m R N A 発現のためのT a q M a n (登録商標)アッセイ

ヒトおよびマウスL R P 5 c D N A 配列に基づいて、T a q M a n (登録商標)プライマーおよびプローブを選択した。図26に説明するように、ヒトおよびマウスL R P 5 (Z m a x 1) 配列のアラインメントの分析により、選択した配列を遺伝子特異的であるように設計した。

【0397】

以下のプライマーおよびプローブセットを用いて、製造者(PE Applied Biosystems、Foster City、カリフォルニア州)に記載されるように、ヒト細胞から単離されたR N A のT a q M a n (登録商標) 定量的逆転写酵素ポリメラーゼ連鎖反応(R T - P C R) 分析を実施した。

20

ヒトZ m a x 1 - 1 / H B M : (配列番号730~732)

順行性プライマー: 5' - G T C A G C C T G G A G G A G T T C T C A - 3'

逆行性プライマー: 5' - T C A C C C T T G G C A A T A C A G A T G T - 3'

プローブ: 6 - F A M - 5' - C C C A C C C A T G T G C C C G T G A C A - 3'

【0398】

P E A p p l i e d B i o s y s t e m s のヒトG A P D H コントロールキットでの多重プロトコルを用いて、実験的プライマー/プローブセットの結果をヒトG A P D H レベルに正規化した。以下のプライマーおよびプローブセットを用いて、製造者(PE Applied Biosystems、Foster City、カリフォルニア州)に記載されるように、ネズミ細胞および組織から単離されたR N A の種特異的T a q M a n (登録商標) 定量的R T - P C R 分析を実施した。

30

ヒトZ m a x 1 - 1 / H B M - 1 : (配列番号733~735)

順行性プライマー: 5' - C G T G A T T G C C G A C G A T C T C - 3'

逆行性プライマー: 5' - T T C C G G C C G C T A G T C T T G T - 3'

プローブ: 6 - F A M - 5' - C G C A C C C G T T C G G T C T G A C G C A G T A C - 3'

マウスZ m a x 1 - 1 / H B M - 1 : (配列番号736~738)

順行性プライマー: 5' - C T T T C C C C A C G A G T A T G T T G G T - 3'

逆行性プライマー: 5' - A A G G G A C C G T G C T G T G A G C - 3'

プローブ: 6 - F A M - 5' - A G C C C C T C A T G T G C C T C T C A A C T T C A T A G - 3'

40

【0399】

P E A p p l i e d B i o s y s t e m s の18SリボソームR N A コントロールキットでの多重プロトコルを用いて、実験的プライマー/プローブセットの結果を18SリボソームR N A レベルに正規化した。これらの結果の要旨を図17~20に提示する。

【0400】

トランスジェニックマウスの作製

D N A マイクロインジェクション

最初にEpicentre TechnologiesのG E L アーゼプロトコルに従って、マイクロインジェ

50

クシヨンのための導入遺伝子フラグメントを1%アガロースゲル上で精製した。次いでフラグメントを塩化セシウム密度勾配でさらに精製し、5 mM トリス (pH 7.4) および0.1 mM EDTAに対して広範に透析した。

【0401】

標準的な手段に従って、直線化DNAをマウス胚にマイクロインジェクションした。DNAを1次的に受精C57BL/6Tマウス胚の雄前核に注射した。注射された胚 (n = 20-35) を交配後0.5日の偽妊娠Swiss Webster胚受体の卵管 (片側的に) 移した。10から14日日齢に子孫の尾のバイオプシを行い、そして遺伝子型に分類した。

【0402】

ジーンターゲティングトランスジェニックマウスの作製

10

ジーンターゲティングベクターおよびプローブ

胚幹 (ES) 細胞のLRP5遺伝子を修飾するために、2つのジーンターゲティングベクターを構築した。図16で説明する、Zmax1-KI/KO A & Bと称する2つの構築物を2つの変異の型、LRP5遺伝子のノックアウト (KO)、およびマウスモデルを作製するためのCreリコンビナーゼ依存性のヌクレオチド置換のノックイン (KI) (すなわちHBM家系のマウスLRP5におけるグリシン170のバリンアミノ酸への置換) を生じるように設計した。

【0403】

初めに、マウスLRP5遺伝子のエクソン5つを含むマウスゲノムDNA BACクローン473P5 (Genbankアクセッション番号AZ095413) のゲノムDNAを用いて骨ジーンターゲティングベクターを構築した。Research Genetics (Huntsville、アラバマ州) によりエクソン3に関するポリメラーゼ連鎖反応 (PCR) スクリーニングを用いてマウス129SvJゲノムBACライブラリーからこのクローンを単離した。順行性プライマーの配列 (5' - GAGCGGGCAGGGATGGATGG - 3') (配列番号739) および逆行性プライマーの配列 (5' - AGGTGGGACGGGTGGATGAAGC - 3') (配列番号740) を用いてPCRによりエクソン3を増幅し；以下の熱循環条件を用いた。95℃で0.5分間、55℃で1分間および72℃で1分間の30サイクル。このクローンのマウスLRP5遺伝子との同一性を、鋳型としてBACクローンDNAを用いてエクソン3をシーケンシングすることにより確認した。pGEM-T-easy T/Aクローニングキットを用いてPCR産物をクローン化した。

20

30

【0404】

LRP5ノックイン/ノックアウトベクター

プローブとしてサブクローン化されたエクソン1、エクソン2およびエクソン3を用いて、エクソン5を通してエクソン1に広がる領域をシーケンシングすることにより、サザンブロット分析によりゲノムBACクローン473P5の組成を特徴付けした。LRP5 (Zmax1) KI/KO標的化のために2つの異なる構築物を調製した。これらの構築物 (AおよびB) は相同性のフランキングアームにおいてのみ異なる。構築物 (A) は相同性の6.5 kb BstEI-XbaI 5' アームおよび相同性の1 kb XbaI-EcoRI 3' アームを含有し；一方構築物 (B) は相同性の1 kb 5' アームおよび相同性の6.0 kb 3' アームを含有する。相同性の短いアームおよび長いアームを、ネオマイシン抵抗性遺伝子 (MC1-Neo、Stratagene) および合成転写中止配列 (syntetic transcriptional pause sequence) (Promega) を含有するLoxPフランキングカセットにライゲートすることにより構築物を調製した。

40

【0405】

双方のZmax1 KI/KO標的化ベクター (AおよびB) を、エクソン3でG170Vアミノ酸置換をコードするGからTへのヌクレオチド置換するように修飾した。鋳型として相同性の短いアームの野生型配列を用いて、重複PCR変異誘発によりこれらの修飾を導入した。加えて、Zmax1 KI/KO (B) 標的化ベクターの1 kbの短いアームを、5' 終端PmeI制限認識部位を含むように修飾した。配列 (5' - AAGCTTGTTTTAAACTGGGCATGGTGGCACCATGGTTGTAAAT - 3') (

50

配列番号741)の順行性プライマーおよび配列(5'-GGGCTTCCACCCAGTCAGTCCAGTACATGTACCT-3') (配列番号742)の逆行性(変異原性)プライマーを用いて5'重複フラグメントを作製した。30サイクルで利用した熱循環条件は95℃で0.5分間、55℃で1分間および72℃で1分間である。配列(5'-CTGACTGGGTGGAGGCAACCCCGGATCGAGC-3') (配列番号743)の順行性プライマーおよび配列(5'-GAATTCATCGGTACCTGTGCGGCGCTTCATTG-3') (配列番号744)の逆行性(変異原性)プライマーを用いて3'フラグメントを作製した。30サイクルで利用した熱循環条件は95℃で0.5分間、55℃で1分間および72℃で1分間である。最終重複PCRは鋳型としてPCR反応の5'フラグメントおよび3'フラグメント各々1mlを利用し、各々5'および3'フラグメントの順行性および逆行性プライマーならびに同一の熱循環パラメーターを用いて増幅を実施した。pGEM-T-easy T/Aクローニングキットを用いて最終PCR産物をクローン化した。変異誘発されたエクソン3を600bp BsmBI-XbaIフラグメントとしてZmax1 KI/KO(B)から切除し、Zmax1 KI/KO(A)に移した。

10

20

30

40

50

【0406】

ES細胞クローンを標的化するZmax1 KI/KO(A)遺伝子のスクリーニングし、特徴付けするためのプローブを、BACクローン473P5の制限フラグメントのサブクローニングにより調製した。5'外側プローブは400bp Nde-BstEIIフラグメントであり、および3'外側プローブは500bp EcoRI-BstXIフラグメントである。

【0407】

相同性の標的化ベクター領域をフランキングする、および直ぐ隣接するゲノムフラグメントのPCRクローニングによりZmax1 KI/KO(B)の外側プローブを調製する。Zmax1 KI/KO(B)に用いられる5'外側プローブは配列(5'-TGAGATGTCTCTGTGTGGC-3') (配列番号745)の順行性プライマーおよび配列(5'-TCCCTTCCCTTCCCTACAGTTG-3') (配列番号746)の逆行性プライマーを用いて作製された498bpフラグメントである。これらのプライマーを用いて30サイクルで利用した熱循環条件は、95℃で0.5分間、55℃で1分間および72℃で1分間である。3'外側プローブは配列(5'-CCTAAGGATCTCTCTGTGTGTGG-3') (配列番号747)の順行性プライマーおよび配列(5'-CTGCAGCAGGTCAGTAGCCCTGC-3') (配列番号748)の逆行性プライマーを用いて作製された600bpフラグメントである。これらのプローブを用いて30サイクルで利用した熱循環条件は、95℃で0.5分間、55℃で1分間および72℃で1分間である。双方のプローブはゲノムサザン分析においてLRP5遺伝子に特異的である。pGEM-T-easy T/Aクローニングキットを用いてPCR産物をクローン化した。

【0408】

LRP5 mRNA構造および転写レベルのリボヌクレアーゼ保護分析のためのプローブをエクソン3からエクソン4を含有するcDNAフラグメントをPCRクローニングすることにより調製した。PCR反応は鋳型として完全cDNA、配列(5'-TGAGATGTCTCTGTGTGGC-3') (配列番号749)の順行性プライマー、配列(5'-TCCCTTCCCTTCCCTACAGTTG-3') (配列番号750)の逆行性プライマーおよび以下の30サイクルのための熱循環条件を使用した。95℃で0.5分間、55℃で1分間および72℃で1分間。pGEM-T-easy T/Aクローニングキットを用いてPCR産物をクローン化した。

【0409】

ES細胞におけるジーンターゲティング

ジーンターゲティングのために胚幹(ES)細胞に直線状標的化ベクター50mgと共にエレクトロポレーションを行い、エレクトロポレーション開始日から7から10日後に

200 mg/ml G418で選択した。G418抵抗性クローンを採取し、広げ、凍結保存した。LRP5の野生型および標的化対立遺伝子を各々4 kbおよび5 kbフラグメントとして検出する、5'外側プローブを用いるEcoRIゲノムサザン制限断片長分析により抵抗性クローニングを相同組換えに関してスクリーニングした。ジーンターゲットイングES細胞クローンを解凍し、広げ、LRP5の野生型および標的化対立遺伝子を各々9 kbおよび8 kbフラグメントとして検出する、3'外側プローブを用いるScaIゲノム制限断片長分析により特徴付けした。ジーンターゲットイングクローンをLRP5エクソン3の配列分析により特徴づけし、GからTへの置換が相同組換えに含まれたことを確認した。

【0410】

10

胚盤胞注射によるジーンターゲットイングマウスの作製

キメラマウスを作製するために、ジーンターゲットイングES細胞クローンを解凍し、広げ、交配後(p.c.)3.5宿主C57BL/6胚盤胞の胞胚腔に9-14 ES細胞を注射した。次いで注射された胚盤胞(12-17)を2.5 p.c.偽妊娠Swiss Webster胚受体の子宮に片側的に移し、出産まで発達させた。キメラ雄は129SvEv雌に戻し交配し、ネオマイシン抵抗性遺伝子に特異的なプライマーでPCR遺伝子型解析することにより標的化対立遺伝子の伝達に関して試験する。

【0411】

Creレコンビナーゼを介するネオマイシン抵抗性カセットのインビトロ欠失

LRP5 KI/KOマウスからLRP5 KIマウスを作製するために、Cre発現プラスミド(2 mg/ml)の、LRP5 KI/KO前融合接合体の雄前核へのマイクロインジェクションにより、ネオマイシン抵抗性(KO)カセットを欠失させた。KOカセットの欠失をカセット挿入部位のPCR分析により確認した。

20

【0412】

トランスジェニックマウスの遺伝子型解析

50 mM トリスHCl (pH 7.2)、50 mM EDTA (pH 8.0)、0.5 % SDSおよび0.8 mg/ml プロテイナーゼKを含有するバッファー500 μlによりマウス尾断片からゲノムDNAを単離した。サンプルを振盪しながら55℃で一晩インキュベートする。10 μl アリコートをして99℃で5分間熱不活化し、1:20で水で希釈した。PCRに関しては、希釈したDNA 1 μlを以下の条件下で増幅した。変性: 96℃で4.5分間; 45サイクル: 96℃で30秒間; 63℃で1分間; 72℃で1分間; 伸長: 72℃で5分間; 4℃を維持。

30

【0413】

以下のプライマーセットを遺伝子型解析に使用する。

HBMMCBA:

5'プライマー: 296塩基対フラグメント

順行性: 5'-GCT TCT GGC GTG TGA CCG GCG-3' (配列番号751)

逆行性: 5'-GCC GCACAG CGC CAG CAG CAG C-3' (配列番号752)

40

3'プライマー: 400塩基対フラグメント

順行性: 5'-CAC CCA CGC CCC ACA GCC AGT A-3' (配列番号753)

逆行性: 5'-ATT TGC CCT CCC ATA TGT CCT TCC-3' (配列番号754)

HBMMTIC:

5'プライマー: 382塩基対フラグメント

順行性: 5'-TTC CTC CCA GCC CTC CTC CAT CAG-3' (配列番号755)

逆行性: 5'-GCC GCA CAG CGC CAG CAG CAG C-3' (

50

配列番号 756)

3' プライマー: 524塩基対フラグメント

順行性: 5' - G A A T G G C G C C C C G A C G A C - 3' (配列番号 757)

逆行性: 5' - G C T C C C A T T C A T C A G T T C C A T A G G - 3' (配列番号 758)

【0414】

サザン分析による遺伝子型の確認

マウスゲノムDNAをEcoRIで消化し、LRP5 cDNAに由来する1.0kb SalI/BamHI制限フラグメントでプロービングした。プローブは導入遺伝子陽性動物の5kbフラグメントにハイブリダイズした。 10

【0415】

表現型解析

インビボおよびエクソビボの双方のアッセイを用いてトランスジェニックマウスの表現型を評価する。C57BL/6および129SvEvと称する野生型マウスの2つの株を研究してトランスジェニックおよびジーンターゲティングマウスにおける表現型評価のための対照データを提供する。加えて非トランスジェニック同腹仔動物を対照として使用する。

【0416】

インビボ分析

pDXA: 野生型およびトランスジェニックマウスを麻酔し、重量測定し、LUNAR小動物PIXImus装置を用いて骨格の全身X線スキャンを作製した。マウスが離乳したときに(すなわち3週齢で)スキャンを始め、2週間隔で繰り返す。野生型動物を3、5、7、9、11、13、15、17、19、21、23、27、および29週でスキャンする。トランスジェニック動物のスキャンは17週までの期間実施する。種々身体領域のBMD(骨ミネラル密度)、BMC(骨ミネラル含量)、TTM(全組織質量)、および脂肪%に関してスキャンを分析する。 20

【0417】

Faxitron X線像: 麻酔動物のpDXAスキャンの後、骨の大きさの測定を可能にするFaxitron装置を使用して別のX線撮影を行った。 30

【0418】

カルセイン標識: 動物に15mg/kgカルセインを2回連続で腹腔内に投与する。1回目の投与は安楽死の9日前に行い、2回目は骨形成速度の測定を可能にする安楽死の2日前に行った。

【0419】

エクソビボ分析

RNA単離: TRIzol(登録商標)を用いて脛骨およびその他の組織から全RNAを単離し、mRNA発現を決定した。

【0420】

pQCT: 右遠位骨幹端の全および骨梁密度ならびに骨幹中央の皮質密度を決定するために右大腿骨の軟部組織を取り除き、70%エタノールに保存した。 40

【0421】

MicroCT: 右大腿骨を用いて遠位骨幹端の骨梁インデックスを決定した。

【0422】

組織学: 右大腿骨を用いて遠位骨幹端の骨面積並びに静的および動的パラメーターを決定した。

【0423】

曲げ強さ: 左大腿骨の軟部組織を取り除き、骨幹中央の曲げ強さの3点解析の前に-20℃で保存した。

【0424】

椎骨の圧縮強度：T 1 0 L 6－7から脊椎全体を取り出した。軟部組織を脊椎に残し、分析まで脊椎を－2 0℃で保存した。L 5椎骨で圧縮強度を測定した。

【0 4 2 5】

血清：動物を安楽死させ、血液から血清を調製して全コレステロール、トリグリセリド、オステオカルシンおよびその他の生化学的代理マーカーを想定する。

【0 4 2 6】

組織学的分析：事例には免疫細胞化学、例えば骨原性マーカーのインサイチュ・ハイブリダイゼーションおよびアポトーシスを被る細胞のT U N E L染色などがある。

【0 4 2 7】

結果

トランスジェニックプラスミド構築物からの発現の確認

非常に低レベルの内因性生成物LRP 5発現を有するHOB-0 2-0 2細胞にHBM (HBMMCB AおよびHBMMT I C) および野生型 (Zmax 1WTCB AおよびZmax 1WTT I C) プラスミド構築物を一過性でトランスフェクトした。トランスフェクションの2日後、RNAを単離し、TaqMan (登録商標) 定量RT-PCRを実施して細胞におけるLRP 5/HBMのmRNAレベルを決定した。プラスミドDNAの夾雑を調節するために、前RT工程伴うかまたは伴わず、RT工程は行わずにPCRを実施し、非常に低レベルのLRP 5/HBM mRNAを検出した。しかしながら、RT工程では、CMV-β-ガラクトシダーゼ対照でトランスフェクトした細胞と比較して、CMV-β-アクチンプロモーター構築物でトランスフェクトした細胞においてHBMおよびLRP 5 mRNAの1 0 0 0倍増加が観察された。I型コラーゲンプロモーター構築物はHBMおよびLRP 5 mRNAのおよそ1 0 倍の増加を示し、これはCMV βアクチンプロモーターと比較して、このプロモーターの特性が弱いことと合致する。図1 7参照。

【0 4 2 8】

HBM/LRP 5発現に関して種特異的なTaqman (登録商標) 試薬

LRP 5/HBMに関する種特異的Taqman (登録商標) プライマーおよびプローブセットを開発した。HOB細胞およびマウスMC-3 T 3-E 1骨芽細胞を用いる一連の実験では、マウスバックグラウンドのLRP 5/HBM mRNAを測定し、逆も同様である。これらの試薬は種特異的発現の検出および定量に有用である。図1 8にて実証するように、プライマーセットはマウスおよびヒト細胞株において種特異的である。さらに、図1 9は、マウスRNAのバックグラウンドにおけるヒトLRP 5 RNAの定量測定を実証する。これらのTaqman (登録商標) セットを用いて、マウストランスジェニック株で発現されているヒトまたはマウスHBMまたはLRP 5 (またはその他のHBM用変異体) メッセージのレベルを決定することができる。

【0 4 2 9】

種特異的Taqman (登録商標) 試薬はトランスジェニックマウス組織における内因性LRP 5 mRNAレベルおよびヒトLRP 5/HBM mRNAレベルの双方を特徴付けるための新規手段である。これらの手段には従来の方法、例えばノーザンハイブリダイゼーションおよび標準RT-PCRに勝るいくつかの利点がある。これらの利点のいくつかは以下の通りである。(1) 特異性、小領域 (< 1 0 0 b p) のみが増幅されるので、プライマーおよびプローブは交差反応性がないかまたは最低であると予測される領域をシーケンシングするために選択される；(2) 速度、手順はそれほど労働集約的ではないので；(3) 正確性、真の定量性であるので；および(4) 感受性、極少量の出発材料 (すなわちRNA) を要求し、およびシグナル対雑音比が高いので。骨から多量のRNAを得ることは困難であるので、骨におけるmRNAレベルを分析するためにこれらの利点はとりわけ重要である。従って、HBMおよびLRP 5発現のTaqman (登録商標) 分析のために開発されたプライマーセットは本発明の重要な実施形態である。当業者は、本明細書に記載されるプライマーが好ましい実施形態であり；得られた核酸が実質的に同一の機能を実施し続ける限り、修飾、例えばプライマーまたは塩基置換の伸長またはトラ

10

20

30

40

50

ンケーションが本発明に包含されることを認識している。

【0430】

トランスジェニックマウスにおけるHBM発現

CMVbアクチン(HBMMCB A)構築物に関して8匹のトランスジェニック創始動物を作製し、繁殖プログラムを開始して株を確立した。図20に示すTaqman(登録商標)分析により決定したmRNAの発現により4つの株における骨発現の変動性レベルが示された。脛骨では、以下の範囲の発現レベル(HOB-03-C5細胞の内因性LRP5に相対して)が示された。株18(x10から11倍);株2(x7から10倍);株13(x1から2倍)および株28(x1倍)。プロモーターの公知の活性に基づいて推定される別の組織においても発現が検出された。株2および13に関しては、HBM発現の最高レベルが心臓において見出された。Taqman(登録商標)遺伝子型解析アッセイは潜在的ホモ接合性動物に関してスクリーニングする。

【0431】

I型コラーゲン(HBMMTIC)構築物に関して6匹のトランスジェニック創始動物を作製し、繁殖プログラムを開始して株を確立した。最初に試験した2つの株でmRNAの発現が見出された。株19では、脛骨および大腿骨においてのHOB-03-C5細胞のLRP5よりも発現が各々7から8倍および19から20倍多かった。株35では脛骨および大腿骨において検出された発現は低レベルであった。

【0432】

HBMトランスジェニックマウスのインビボpDXA測定

HBMMCB A構築物

図21(A~C)で説明する、5週および9週時点でのトランスジェニックマウスの分析により、これまで試験した1つの株のBMD値が対照に比較して大きいことが示された。5週では、大腿骨、椎骨および全身におけるHBMMCB A株2(n=11)BMDが各々21%、24%および10%野生型対照よりも大きかった。9週(n=3)ではBMDにおけるこれらの増加が各々19%、32%および12%に達した。17週以上では、株2のパーセント増加が野生型対照に相対して各々10%、11%、および8%であった。

【0433】

HBMMTIC構築物

図21(D~F)で説明する、5週および9週時点でのトランスジェニックマウスの分析により、試験した2つの株のBMD値が対照動物に比較して有意に大きいことが示された。5週では、大腿骨、椎骨および全身におけるHBMMTIC株19(n=5)BMDが各々63%、70%および41%野生型対照よりも大きかった。9週(n=2)ではBMDにおけるこれらの増加が各々52%、64%および37%に達した。17週以上では各々35%、40%、および28%であった。5週では、大腿骨、椎骨および全身におけるHBMMTIC株35(n=1)BMDが各々4%、47%および6%野生型対照よりも大きかった。9週(n=3)ではこれらの増加が各々32%、43%および18%に達した。17週ではBMDにおける割合の増加が各々20%、33%、および19%であった。I型コラーゲンプロモーターによりHBM発現が誘導されている2つのさらなるHBMトランスジェニック株188および189はHBM表現型を有していることが示された。17週以上では株188により大腿骨、椎骨および全骨におけるBMDの23%、35%および22%の増加が実証される。

【0434】

全体では、トランスジェニックマウスのBMD結果が、HBM罹患家系において観察された表現型と同程度の類似性を示している(Johnsonら、Am.J.Hum.Genetics 60:1326-1332(1997))。例えば、罹患した個体で測定された椎骨BMDは非罹患ファミリーのメンバーよりもおよそ34から63%大きい。トランスジェニック動物由来の椎骨BMDに関するデータは9週齢で正常よりも大きい~30から70%の範囲になる。

【0435】

トランスジェニックマウスのエクソビボ分析

非観血的骨画像診断による動物のモニター観察により、選択されたトランスジェニック株において検出された骨密度の増加をさらに調べるために、5および9週齢で直接骨密度測定および組織学的分析に関してこれらの株の動物において剖検を実施した。左大腿骨を単離し、清浄し、XCT Research 末梢型定量用コンピューター断層撮影(pQCT;Stradtec Medizintechnik、Pforzheim、ドイツ)に配置した。大腿骨の遠位末端を置き、この点から2.5mm近接で全体および骨梁測定に関してpQCTスキニングを開始した。皮質測定のためのpQCTスキャンは最初のスキャンから3.5mm近接(すなわち遠位末端から6mm近接)で開始した。pQCTスキャンは0.5mm厚で、ボクセル(すなわち、3次元画素)のサイズは0.07mmであり、切片を通る360の投影から成る。スキャンが完了した後、画像をモニターに表示し、各スキャンの大腿骨全体を含む目的の領域の輪郭を描いた。双方向性アルゴリズムを用いて軟部組織を自動的に除去し、第1の切片における残りの骨の密度(全体密度)を決定した。次いで骨の外側の55%を求心的らせん状に剥がし、第1の切片の残りの骨の密度(骨梁密度)を mg/cm^3 で報告した。第2の切片では、双方向性アルゴリズムを用いて皮質と骨梁骨間の境界を決定し、皮質骨の密度を決定した。

【0436】

株2F1 CMV β アクチン-HBMの5週齢のトランスジェニックマウスの分析により、トランスジェニック雄マウスにおいて全体密度、骨梁密度および皮質密度が非トランスジェニック雄に対して各々20%、37%および4%高いことが示された。後の時点で、トランスジェニックおよび非トランスジェニック動物間の差異がこの株では減少する。しかしながら、株19に由来する5週齢のI型コラーゲン-HBMトランスジェニック雄では、非トランスジェニック同腹仔よりもさらにもっと劇的な骨密度の増加が明白であった。全体密度、骨梁密度および皮質密度が各々53%、104%および5%高かった。株19の動物では、表7で示すように、全体および骨梁骨密度の増加を伴う表現型が9週齢を超えて維持されることが見出された。17週では、全体および骨梁密度は各々46%、202%増加した。株19の全体骨梁パラメーターに及ぼす影響は3つの時点全てで統計的に有意に高かった($p < 0.001$)。I型コラーゲン-HBMトランスジェニック株35の雄から、骨表現型発現の幾分異なったパターンが明らかであった。5週齢の全体密度、骨梁密度および皮質密度がわずかに高かった(各々7%、4%および4%)。しかしながら、9週齢では表7に示すように、これらのパラメーターの明らかなおよび統計的に有意な増加が明白になった。全体および骨梁骨密度は株35では17週齢を通して上昇したままである。

【0437】

I型コラーゲンプロモーターを伴う2つの別のHBMトランスジェニック株が研究され、株19に類似する劇的な高骨密度表現型を示している。株188および189の双方に由来する雄では、全体骨密度は非トランスジェニック動物と比較して40%まで増加した。骨梁骨密度は9週齢で75%増加した。17週では、株188雄の全体および骨梁骨密度は対照動物、例えば非トランスジェニック同腹仔を越えて、42%および161%であった。これらの値はこの週齢で株19に認められる効果に合致する。株188の雌は9週で36%および144%の増加を、全体および骨梁骨密度は各々26%および148%の増加を示している。株189に由来する雌は9週で27%および64%まで、26週齢では15%および84%まで増加する全体および骨梁密度を有した。

【0438】

とりわけ「骨特異的」I型コラーゲン導入遺伝子を伴う、骨細胞分化の段階により影響を受ける表現型の加齢関連発現の異なるパターンの出現は予期されない。株2および株19の双方の動物は5週齢で脛骨サンプルのHBM mRNAに匹敵するレベルを発現し、これらのレベルはこの週齢でみかけの骨表現型を示さない別の株よりも有意に大きい(>7から8倍)。I型コラーゲンプロモーターにより誘導される株19は株2とは異なって、骨以外の組織における導入遺伝子の非常に低い発現を示す。5週齢では株35の動物は

骨では低レベル発現を示し、別の組織では発現を示さない。HBM/LRP5 特異的抗体を用いる頭蓋冠の骨切片の免疫組織化学により5週齢の株2および株19、ならびに9週齢の株35に由来するトランスジェニック動物の骨細胞では、非トランスジェニック同腹仔に対してより強い染色強度が表される。

【0439】

pQCT分析により表される知見を、 μ CT装置(Scanco)を用いて、より高い解像度でさらに試験した。画像化される領域がおよそ4mm近接に伸びる大腿骨の遠位末端を含み、関節軟骨の軸に対して垂直にするように大腿骨を配置した。 μ CT測定を開始するための参照の株を成長プレートとの重複が最低になるように配置し、200スキャン切片(9mm厚)に関して近接に広げる。 μ CT測定を完了した後、顆状突起が十分に溶け込んでいる第1の切片を同定した。皮質は排除するが、最大量の骨梁腔を含むように目的の領域の輪郭を描いた。最初の30個の切片に関して、5切片ごとに目的の領域を抜き取り、溶け込ませた。残りの105個の切片に関して、10から20切片ごとに目的の領域を抜き取った。骨梁腔が規則正しくなるほど、目的の領域を抜き取る必要性のある頻度が少なくなる。目的の各領域を抜き取った後、その先行したものと溶け込ませた。135個全ての切片で目的の領域が確立された後、350の閾値設定を用いて3次元評価を実施した。

10

【0440】

pQCTにより同定された骨密度増加を μ CTにより確認し、骨構築のエレメントを含むように拡張した。株2トランスジェニック動物では、 μ CT骨容量/全容量、結合性密度および骨梁の厚みが各々50%、83%および12%高かった。結合性密度および骨梁の厚みの双方の指標は、密度の増加が構造強度の増加にも随伴されることを示唆している。骨表面/骨容量はトランスジェニック雄で17%低く、これは吸収性表面および窩がほとんどないのかもしれないことを示唆し得る。骨梁骨応答はさらに非脱灰ゴールドナー染色切片の組織学的評価によりさらに確認され、これはトランスジェニック雄の遠位大腿骨骨幹端の骨ミネラル部分が36%大きいことを示した。動的組織形態計測分析により、カルセイン二重標識により決定されるように、骨ミネラル沈着速度の実質的な増加(+100%)が導入遺伝子の骨増加に部分的に寄与し得ることが示された。pQCTにより明らかな株19の骨梁骨に及ぼす動的効果が μ CTにより支持され、ここでトランスジェニック雄において骨容量/全容量、骨梁数、骨梁の厚みおよび結合性密度が各々130%、45%、30%および121%高いことが見出された。これらの効果の全てが $p < 0.01$ で統計的に有意であった。株19の5週齢で認められた骨表現型は9週齢の動物で保持され、ここで、表7で示すように、骨容量/全容量、骨梁数および結合性密度が非トランスジェニック同腹仔よりも有意に高かった。

20

30

【0441】

株35トランスジェニック動物の μ CT分析は別の2つの株とは幾分異なったパターンを示した。株35の5週齢の雌においてpQCTにより示された、ちょうど中程度の密度増加と対照的に、より大きな画像解析を有し、より大きな容積測定サンプルを含む μ CTを用いて統計的に有意な効果($p < 0.01$)が認められた。骨容量/全容量、骨梁の厚みおよび結合性密度は35%、9%および27%高かった。類似の結果が株355週齢雄において認められ、ここで骨容量/全容量および結合性密度は各々37%および45%増加し、これは μ CT分析により明らかであり、ここではpQCTによりわずかな増加が示された。株35トランスジェニック雄とその非トランスジェニック同腹仔との差異は加齢と共に増加するようで、全密度(28%)および骨梁密度(52%)の統計的に有意な増加が9週齢でのpQCTにより明白である。 μ CTの結果はこの株の骨表現型における加齢に関連する多様性を支持し、トランスジェニックと非トランスジェニック動物とのその差異を示している。骨容量/全容量および結合性密度に関して、これらのパラメーターは5週齢で認められるものの倍化以上であり、各々97%および188%になる。トランスジェニック動物で認められる骨容量増加は、HBM家系の成人男性罹患メンバーの骨バイオプサンプルで検出されたこのパラメーターにおける有意な増加と合致する。トランスジェニック株において影響を受けることが見出された別のパラメーターはピーク/成体骨

40

50

量増加に至る変化を反映することができ、これはマウスのこの株では17から20週齢の間で生じる。

【0442】

株19由来の頭蓋冠の免疫組織化学により、骨膜を裏打ちする前骨芽および骨芽細胞株ならびにミネラル化骨に存在する骨細胞における導入遺伝子の強い発現が示された。トランスジェニック動物における骨膜の骨芽細胞は膨らんでそして立方体様の、活発に細胞外マトリックスを分泌する指示的細胞であるようだ。対照的に、非トランスジェニック同腹仔の骨膜細胞は平坦な裏打ち細胞のように思われる。骨芽細胞分化および機能的マーカーであるアルカリ性ホスファターゼに関する染色が増加し、対照と比較してトランスジェニック動物における細胞の活発な分泌状態が確認される。さらなる分析により、トランスジェニックマウス頭蓋冠におけるTUNEL陽性骨細胞、骨芽細胞および間質細胞の数の低減が示され、これはアポトーシスの減少を示唆している。9週齢雄非トランスジェニックマウスからの頭蓋冠では間質細胞 mm^2 あたりアポトーシス性骨芽細胞が 30.9 ± 1.8 ($n=9$)であり、一方HBM導入遺伝子に由来する頭蓋冠では、間質細胞 mm^2 あたりアポトーシス性骨芽細胞が 11.6 ± 2.8 ($n=9$)であった。これらを考慮に入れて、これらの結果はトランスジェニック動物におけるBMD増加は骨芽細胞数および活性の増加が原因となり、これは部分的には機能的寿命の延長が原因となり得る。

10

【0443】

過剰発現トランスジェニック株で認められる骨密度および骨構築変化は潜在的により大きな生物機械的強度を示唆している。5週齢の株19雄の大腿骨の3点曲げ強度を評価することによりこれを直接試験した。大腿骨から軟部組織を取り除き、デジタルノギスを用いて大腿骨の長さを測定した。pQCTを用いて遠位末端から6mmで骨膜および骨内膜外周および皮質の厚みを測定した。次いで大腿骨を、根幹中央の中心が5mm離れて位置する固定支持体から等距離になるように準備台に置いた。Instron5543の負荷装置のクロスバーを根幹中央に置き、1mm/分の速度で骨折を生じるまで力を適用した。力対転移曲線を作成し、Instron Merlinソフトウェアを用いてピーク負荷を決定した。骨膜外周の増加が原因であると思われる力の75%増加 ($p < 0.01$)があり、これは皮質の厚みの増加に至る。従って、HBMトランスジェニック動物において認められるように、骨密度および骨幾何学における変化は生物機械的強度の増加に翻訳される。

20

【0444】

LDL関連レセプタータンパク質のクラス内のHBM/LRP5の会合を鑑みて、変異が脂質プロファイルに影響し得るかどうかを決定することは興味深いことであった。実際に、HBM家系における脂質研究（すなわち8人の罹患および7人の非罹患メンバー）により、罹患メンバーでトリグリセリドおよびVLDLレベルが統計的に低いことが示された。Hitachi 911装置でBoehringer Mannheim（コレステロール用）およびRoshe（トリグリセリド用）試薬を用いてトランスジェニック株に由来する血清サンプルを分析した。via o-キノン・イミン色素（これはコレステロールとの酵素反応の後に形成される）を介して37℃、505nmでの光度測定によりコレステロールを測定した。トリグリセリド測定のための酵素法は、トリグリセリドおよび脂肪酸の加水分解の後のトリグリセリドのグリセロール部分の決定に基づく。酵素反応の最終色素生成物を505nmで測定した。2株トランスジェニック動物の5週齢雄では、血清コレステロールはわずかに低下したが、血清トリグリセリドレベルはトランスジェニック動物において、その非トランスジェニック同腹仔に対して26%まで低下した。9週齢の株2動物の限定的なサンプルでは、トリグリセリドレベルは20%低いままであった。同様に、5週齢では株19のトランスジェニック雄のトリグリセリドレベルは32%低かった。対照的に5週齢では株35のトランスジェニック雄および雌の双方でトリグリセリドレベルは低くなかった。5週齢株35動物では骨容量/全体容量が統計的に大きいという事実は、脂質変化が直接骨格表現型に関連しないことを示唆している。これは9週齢の株35動物のトリグリセリドレベルが極わずか低下した（11%）という事実により支持されるようだが、5週齢よりも実質的に高い骨密度を呈した。これらの株における導入遺伝子の発現のレベルおよび部位が異

30

40

50

なることが原因で、血清脂質レベルを、HBM/LRP5を介して骨表現型に有利に影響する作用物質に関する代理マーカーとして提供できるという可能性を考慮から外すことができない。

【0445】

HBMまたはHBM様遺伝子に基づくこれらのおよびその他のトランスジェニック株は骨恒常性の特性を解明するための貴重なモデルとして提供される。全ての種における骨密度はその通例の負荷条件に適応する。HBM家系ではおよびトランスジェニック株では、骨格のセンサー／エフェクターシステムが、より大きな骨密度に至るより大きな負荷シグナルを認知するようである。骨負荷の増加が骨密度の増加を導くことができ、非負荷または廃用性が骨密度喪失を導くことを示す実験モデルを確立した。これらの実験パラダイムにおけるトランスジェニック動物の骨格の組織学的、生化学的および遺伝的応答を評価することにより、骨恒常性に寄与するセンサー／エフェクターシステムに関する多くの情報が得られる。骨ターンオーバーが変化した、非限定例としてはステロイド欠乏誘起骨減少症および加齢関連骨減少症などの、別の確立されたモデルにおけるトランスジェニック動物の適用により、骨恒常性におけるLRP5の役割およびHBM変異により誘起される有利な変化の特性を理解するうえで手掛かりとなる。

10

【0446】

トランスジェニックマウスにおけるLRP5過剰発現

野生型LRP5の過剰発現の役割を評価するために、およびHBMの効果との対比のために、I型コラーゲンプロモーターにより誘導されたLRP5を発現するトランスジェニックマウスを作製した。統計的に増加した全体および骨梁大腿骨密度がこれらの株の1つ(LRPWWTIC-19)の9週齢で観察される。HBMトランスジェニック株で認められるほど大きくはないが、骨容量および結合性密度の有意な増加を示す μ CT測定により観察が支持される。9週齢の非トランスジェニックマウスに相対するHBMおよびLRP5トランスジェニックマウスに関する骨格パラメーターにおけるパーセンテージ変化の比較を以下の表7に示す。

20

【0447】

【表9】

表7

株 (性別)	全密度	骨梁密度	BV/TV	結合性密度	骨梁#	骨梁の厚み
HBM-19 (M)	60	146	252	348	55	47
HBM-19 (F)	59	222	206	193	56	44
HBM-35 (M)	28	52	97	188	31	11
LPR5-19 (F)	10	41	35	47	15	4.3

30

【0448】

群は導入遺伝子の発現のレベルが骨格表現型パラメーターに随伴される個々の分析に関して表されているので、もう1つのLRP5トランスジェニック株のさらなる評価は有意なpQCT値を示さなかった。野生型レセプターの過剰発現がタンパク質同化性骨表現型を生み出すが、表現型の変化はHBM変異により大きく拡大される。

40

【0449】

HBMジーンターゲティング

LRP5 KI/KOジーンターゲティングベクターを129 SvEv, C57BL/6 ESおよび129ES細胞にエレクトロポレートする。ゲノムDNAの制限断片長分析およびPCR増幅したフラグメントのシーケンシングを用いてジーンターゲティングされたクローンを同定することができる。ジーンターゲティングベクターのノックイン版により、マウスゲノムへの影響を最小にしてHBM変異の内因性LRP5ゲノム遺伝子座への導入が可能になる。それによりさらに天然の環境で、すなわち、例えばトランスジェニックマウスまたはトランスフェクトされた細胞株のような過剰発現モデルではなく、

50

HBMタンパク質の生成が可能になる。ジーンターゲティングベクターのノックアウト版を、1つの機能的LRP5対立遺伝子の喪失に至る可能性のある転写停止配列を含有するように操作した。この変異を有するヘテロ接合性動物を繁殖させ、ヌル対立遺伝子に関する胚ホモ接合性の生成に至る。ジーンターゲティングベクターの異なる設計では、lox

P部位を位置決定して内因性LRP5遺伝子の条件的ノックアウトの生成を促進することができる。Creレコンビナーゼの存在下、LRP5遺伝子の重要な領域をlox P部位の間で欠失させ、従って、結果的に1つの機能的対立遺伝子を喪失する可能性ができる。次いで動物の繁殖を用いてヌル対立遺伝子を伴うホモ接合体を作製する。別のリコンビナーゼ酵素系、例えば同族f rt部位と組み合わせたf lpリコンビナーゼを用いて欠失を作製することができる。胚へのプラスミド注射、およびCreを発現するトランスジェニック動物の使用などの前記した多くの方法でリコンビナーゼを投与することができる。Creの発現を誘導するために用いられるプロモーターを、LRP5遺伝子の広範なまたは組織特異的な欠失に至る様式で選択することができ、従って条件的ノックアウトに至る。別の実施形態では、誘導系、例えばGeneSwitchおよびテトラサイクリンパラダイムを用いてCre酵素そのものを発現させることができる。

10

【0450】

LRP6 ジーンターゲティングされたノックアウトマウス

LRP6 遺伝子の最初のイントロンに挿入された遺伝子トラップベクターを担持するOmnibank (登録商標) 胚性幹 (ES) 細胞を用いてLRP6 ノックアウトマウスを作製した。挿入部位の3'で宿主遺伝子転写物にスプライシングされた5'遺伝子トラップベクター配列から成る融合転写物の逆転写PCR (RT-PCR) により、Omnibank配列タグ (OST) により挿入された位置はLRP6 遺伝子であることが決定された。遺伝子トラップベクターはLRP6 エクソン1を遺伝子トラップのIRES-LacZ-PolyAエレメントの強制スプライシングにより、機能的にマウスLRP6 遺伝子をノックアウトし、LRPの転写を防御する。

20

【0451】

C57BL/6 アルビノ宿主胚盤胞に注射することによりES細胞でキメラマウスを作製し、OST38808と同定し、これを次いで偽妊娠雌に写し、誕生まで発達させた。生殖系列キメラを129SvEvBrd株マウスに戻し交配させ、近交系1298SvEvBrdの遺伝的背景でLRP6のノックアウト対立遺伝子を維持した。LRP6-KO対立遺伝子の生殖系列伝達を遺伝子トラップ特異的配列のPCR増幅により同定した。ヘテロ接合性LRP6-KOの交配対を連続的繁殖に使用した。wtおよびLRP6-KO子孫の遺伝子型をテイルDNA PCRにより決定する。

30

【0452】

雌ヘテロ接合性ノックアウトマウスの9週齢での骨密度の測定により、μCTで測定されたような骨容量、骨梁数、および骨梁の厚さに有意 ($p < 0.05$) な低下が示された。これらの結果は、LRP6もまた骨密度の変調に関与し、治療および薬物の開発の標的であるという仮説に合致する。従って、LRP6トランスジェニック動物およびLRP6の骨変調変異体を発現するトランスジェニック動物は本発明の範囲内であることが企図される。

40

【0453】

トランスジェニック動物および細胞の使用

本発明のトランスジェニック動物および細胞の使用はHBM表現型の代理マーカーを同定するための方法における有用な手段である。本発明により提供される代理マーカーはまた予定される処置の評価およびスクリーニングのための有用な手段でもある。HBM遺伝子を担持する個体は骨量が上昇している。HBM遺伝子は、骨発達に関与する別の分子の活性、レベル、発現パターン、および修飾状態を変化させることによりこの表現型を引き起こす。種々の確立された技術を用いて、その活性、レベル、発現パターン、および修飾状態が、LRP5遺伝子を含有する株とHBM遺伝子を含有する株との間で異なっている分子、好ましくはタンパク質またはmRNAを同定することが可能である。このような株

50

は例えば、細胞不含抽出物、細胞、組織または生存生物、例えばマウスまたはヒトでよい。LRP5の変異形態に関しては、LRP5の完全欠失、タンパク質の細胞外もしくは細胞内部分を欠く変異、またはLRP5遺伝子のいずれか別の変異を用いることができる。アンチセンスLRP5 RNAまたはオリゴヌクレオチドの発現を用いてLRP5タンパク質の生成を阻止することも可能である。HBMの変異に関しては、HBMの完全欠失、HBMタンパク質の細胞外もしくは細胞内部分を欠く変異、またはHBM遺伝子のいずれか別の変異を用いることができる。アンチセンスHBM RNAもしくはオリゴヌクレオチドの発現をまたはRNA干渉方法学を用いてHBMタンパク質の生成を阻止することも可能である。

【0454】

10

LRP5株およびHBM株の比較により同定された分子を、医薬品開発またはヒトもしくは動物の骨疾患の診断における代理マーカーとして使用することができる。また別に、このような分子を骨疾患の処置に用いることができる。Schnaら、Science 270:467-470(1995)参照。

【0455】

例えばマウス相同体遺伝子座にHBM遺伝子を担持するトランスジェニックマウスを構築する。遺伝子型HBM/+のマウスは生存可能で、健康であり、骨量が上昇している。骨量上昇の代理マーカーを同定するために、HBM/+（すなわちヘテロ接合性）およびアイソジェニック+/+（すなわち野生型）マウスを剖検する。双方の組織mRNAを各動物から抽出し、+/+個体で発現されたmRNAに相当する「遺伝子チップ」を構築する。異なる組織に由来するmRNAを各遺伝子型の動物から単離し、逆転写し、蛍光標識し、次に固体支持体に貼り付けた遺伝子フラグメントにハイブリダイズさせる。2つの集団間の蛍光強度の比率は+/+およびHBM/+動物の特異的mRNAの相対的存在度の指標である。また別に、野生型およびトランスジェニック動物からmRNAを単離することができる。これらのサンプルから調製したcDNAはインビトロで転写されて、特注のまたは市販により入手可能な、例えばAffymetrixにより製造されるような遺伝子アレイチップで使用するための標識mRNAが得られる。表現型の機能として発現が変化している遺伝子のセットを種々の通常のコンピューター分析により同定することができる。野生型対照と相対して過剰または過少発現するmRNAをコードする遺伝子は、HBM遺伝子により協調的に制御される遺伝子の候補である。

20

30

【0456】

既に見出されているタンパク質と同一のシグナル発生カスケードの一部である新規タンパク質を同定するための1つの標準的な手順は以下の通りである。細胞を放射活性亜リン酸と反応させ、既に見出されているタンパク質を多かれ少なかれ活性であるように操作する。次いで細胞の別のタンパク質のリン酸化状態をポリアクリルアミドゲル電気泳動およびオートラジオグラフィー、または類似の技術によりモニター観察する。既知のタンパク質の活性レベルを、例えば特異的インヒビター、例えば薬物もしくは抗体を用いて野生型を変異タンパク質と比較すること、単純に既知の細胞外タンパク質を添加するか、もしくはは添加しないこと、または既知タンパク質発現のアンチセンスインヒビターを使用することなどの多くの方法により操作することができる(Tamuraら、Science 280(5369):1614-1617(1998);Meng、EMBO J.17(15):4391-4403(1998);Cooperら、Cell 1:263-273(1982))。

40

【0457】

別の実例では、LRP5のセンスまたはアンチセンスいずれかのcDNAを発現するTE85骨肉腫細胞においてリン酸化のレベルが異なるタンパク質を同定する。TE85細胞は正常では高レベルのLRP5を発現する(Dongら、Biochem.&Biophys.Res.Comm.251:784-790(1998))。センス構築物を含有する細胞はさらに高レベルのLRP5を発現するが、アンチセンス構築物を発現する細胞は低レベルを発現する。³²Pの存在下で細胞を成長させ、収穫し、溶解し、ライゼートをSDSポリアクリルアミドゲルに流し、タンパク質を分離し、ゲルをオートラジオグラフィーに供する(Ausubelら、Current Protocols in Molecular Biology、John Wiley & Sons(1997))。センスとアンチセンス細胞株32の間

50

の強度が異なるバンドはそのリン酸化状態または絶対レベルがLRP5のレベルにตอบสนองして変化するリンタンパク質を表す。^{3 2} P標識の代替として、未標識タンパク質をSDS-PAGEにより分離し、市販により入手可能な抗ホスホチロシン抗体をプローブとして用いて免疫ブロッティングに供することができる(Thomasら、Nature 376(6537):267-271(1995))。アンチセンスRNAの発現の代替として、化学的に修飾されたアンチセンスオリゴヌクレオチドでのトランスフェクションを用いることができる(Woolfら、Nucleic Acids Res.18(17):1763-1769(1990))。

【0458】

多くの骨障害、例えば骨粗鬆症は発症が遅く、そして処置に対する応答が遅い。従って骨発達およびミネラル化の代理マーカーを開発することが有用である。このようなマーカーは骨障害の処置を開発するのに、および骨障害を後に発達させるリスクがある患者を診断するのに有用であり得る。好ましいマーカーの実例は、例えば米国特許第5,455,179号、同第5,641,837号および同第5,652,112号(その開示は参照によりその全てを本明細書に組み込む)に記載されているN-およびC-末端テロペプチドマーカーである。HIV疾患の分野では、CD4計数およびウイルス負荷が疾患の進行の有用な代理マーカーである(Vlahovら、JAMA 279(1):35-40(1998))。骨疾患の分野で類似の代理マーカーが必要とされている。

10

【0459】

代理マーカーは容易に試験され、非特異的影響に対しては相対的に非感受性であるという何かしら特徴的であり得る。例えば、代理マーカーは組織または血清におけるタンパク質またはmRNAのような分子でよい。また別に、代理マーカーは診断徴候、例えば痛みに対する感受性、反射応答等でよい。

20

【0460】

さらに別の例では、HBM遺伝子を担持するヒトの家系を用いて骨量上昇の代理マーカーを同定する。HBM遺伝子を担持する3人の個体から、およびHBM遺伝子を担持しない近親の3人の個体から血液サンプルを採取する。これらの個体の血清中のタンパク質を2次元ゲル系で電気泳動し、ここで一方の次元では大きさによりタンパク質を分離し、もう一方の次元では等電点によりタンパク質を分離する(Epsteinら、Electrophoresis 17(11):1655-1670(1996))。タンパク質に対応するスポットを同定する。いくつかのスポットが異なる量で、または正常なその親族に比較してHBM個体ではわずかに異なる位置で存在することが予期される。これらのスポットは代理マーカーの候補であるタンパク質に相当する。タンパク質の同定はマイクロシーケンシングにより決定され、診断試験手順に使用するための標準的な方法によりタンパク質に対する抗体を生成することができる。HBMタンパク質またはその他の代理マーカー候補に関する診断アッセイには、本発明で記載した抗体およびヒト体液、膜、骨、細胞、組織またはその抽出物中のHBMを検出するためのリポーター分子の使用などがある。抗体を検出可能なシグナルを提供する物質と共に共有結合的または非共有結合的に結合させることにより抗体を標識することができる。多くの科学文献および特許文献では、種々のレポーター分子または、放射性核種、酵素、蛍光、化学ルミネサンスまたは色素原物質などの標識が記載されている(米国特許第3,817,837号;第3,850,752号;第3,939,350号;第3,996,345号;第4,277,437号;第4,275,149号;および第4,366,241号)。トランスジェニックまたは遺伝的に修飾された動物を代理マーカー同定のための方法に提供することもできる。

30

40

【0461】

これらの抗体を用いて、正常な個体および骨障害、例えば骨粗鬆症、骨粗鬆症偽性神経膠腫、エンゲルマン病、リビング病、高ホスファターゼ血症、ファン・ブッヘム病、メロレオストーシス、大理石骨病、濃化異骨症、硬化狭窄症、オステオポイキローシス、先端巨大症、パジェット病、繊維性骨異形成症、管状狭窄症、骨形成不全症、副甲状腺機能低下症、偽性副甲状腺機能低下症、偽性偽性副甲状腺機能低下症、1次および2次性副甲状腺機能亢進症および随伴される症候群、高カルシウム尿症、甲状腺の髄様癌、骨軟化症ならびにその他の疾患を罹患する患者における代理マーカー候補のレベルを測定する。臨床

50

の場で抗体を用いて血清中のタンパク質のレベルを測定する技術は十分に確立されている。特定の疾患または疾患の型を担持する個体における高レベルまたは低レベルで一貫して存在するタンパク質が代理マーカーとして有用である。

【0462】

骨障害の診断において代理マーカーを用いることができる。例えば、高頻度の骨折で医師のもとを訪れる子供を検討する。根底の原因は子供の虐待、不適切な子供の挙動または骨障害であろう。骨障害を迅速に試験するために、前記した抗体を用いて代理マーカータンパク質のレベルを測定する。

【0463】

代理マーカーの修飾状態のレベルを開発中の薬物の、可能性のある有効性の指標として測定することができる。骨発達またはミネラル化における変化を観察するのに長い時間を必要とするかもしれないので、骨障害の処置を行うときに代理マーカーを使用することがとりわけ便利である。例えば、前記したように、「HBM誘導mRNAセット」と称する骨mRNAのセットが+/+マウスに比較してHBM/+マウスにおいて過剰発現されることが見出されている。このセットの発現を代理マーカーとして使用することができる。具体的には、+/+マウスの化合物での処理がHBM誘導mRNAセットの過剰発現に至る場合、次いでその化合物はさらなる開発のための候補として有望であると考えられる。

【0464】

本発明は、種々の薬物スクリーニング技術のいずれかでLRP5またはHBMタンパク質を用いることにより、またはそのフラグメントを結合させることにより化合物をスクリーニングするのにとりわけ有用である。

【0465】

このような試験に用いられるLRP5またはHBMタンパク質またはフラグメントは、溶液中で遊離である、固体支持体に貼り付けられている、または細胞表面上の骨に位置する、のいずれでもよい。薬物スクリーニングの1つの方法はタンパク質またはフラグメントを発現する組換え核酸で安定して形質転換された真核または原核生物宿主細胞を、好ましくは競合結合アッセイにおいて利用する。このような細胞を生存可能なまたは固定されたいずれかの形態で標準的な結合アッセイに使用することができる。例えばLRP5またはHBMタンパク質またはフラグメントと被験物質との間で複合体を形成するために測定するか、またはLRP5またはHBMタンパク質またはフラグメントと既知のリガンドとの間の複合体の形成が被験物質により干渉される程度を試験することができる。

【0466】

従って、本発明はこのような作用物質をLRP5またはHBMタンパク質またはそのフラグメントと接触させること、および当分野で公知の方法により(i)作用物質とLRP5またはHBMタンパク質またはそのフラグメントとの間の複合体の存在、または(ii)LRP5またはHBMタンパク質またはそのフラグメントとリガンドとの間の複合体の存在に関して検定することからなる薬物のスクリーニングの方法を提供する。このような競合結合アッセイでは、LRP5またはHBMタンパク質またはフラグメントは典型的には標識されている。遊離のLRP5またはHBMタンパク質またはフラグメントは、タンパク質：タンパク質複合体に存在するものから分離されており、遊離(すなわち非複合化)の標識の量は各々、被験物質のLRP5もしくはHBMへの結合、またはLRP5もしくはHBM：リガンド結合とのその干渉の測定値である。

【0467】

薬物スクリーニングのための別の技術は、LRP5もしくはHBMタンパク質に対して適当な結合親和性を有する化合物に関する高速大容量スクリーニングを提供し、これは国際公開公報第号84/03564に詳細に記載されている。簡単に述べると、異なる小型ペプチド試験化合物の多くを固体基質、例えばプラスチックピンまたはその他の表面上で合成する。ペプチド試験化合物をLRP5またはHBMタンパク質と反応させ、洗浄する。次いで結合LRP5またはHBMタンパク質を当分野で公知の方法により検出する。前記した薬物スクリーニング技術において使用するために精製されたLRP5またはHBMをプレー

10

20

30

40

50

トに直接コーティングすることができる。しかしながら、タンパク質に対する非中和抗体を用いて抗体を捕捉してLRP5またはHBMタンパク質を固相上に固定することができる。

【0468】

本発明はまたLRP5またはHBMタンパク質を特異的に結合できる中和抗体が被験化合物と、LRP5またはHBMタンパク質またはそのフラグメントへの結合に関して競合する競合薬物スクリーニングアッセイの使用をも企図する。この様式では、抗体を用いてLRP5またはHBMタンパク質の1つまたはそれ以上の抗原決定基を共有するいずれかのペプチドの存在を検出することができる。

【0469】

薬物スクリーニングのための別の技術は、非機能的LRP5またはHBM遺伝子を有する宿主真核細胞株または細胞（例えば前記したもの）の使用に關与する。これらの宿主細胞株または細胞はLRP5またはHBMタンパク質レベルで欠損がある。宿主細胞株または細胞は薬物化合物の存在下で成長する。宿主細胞の成長速度を測定して、化合物がLRP5またはHBM欠損細胞の成長を制御できるかどうかを決定する。

【0470】

合理的な薬物設計の目標は、目的の生物学的に活性なタンパク質またはそれが相互作用する小型分子（例えばアゴニスト、アンタゴニスト、インヒビター）の構造アナログを、例えばそのタンパク質のより活性なまたは安定した形態である薬物、または、例えばインビボでそのタンパク質の機能を増強または干渉する薬物を作るために、生成することである。例えばHodgson、Bio/Technology 9:19-21(1991)参照。1つの研究法では、最初に目的のタンパク質（例えばLRP5またはHBMタンパク質）の、または例えばLRP5またはHBMレセプターもしくはリガンド複合体の3次元構造を、X線結晶学により、コンピューターモデリングにより、または最も典型的には研究法の組み合わせにより決定する。それほど頻繁ではないが、タンパク質の構造に関して有用な情報が、相同タンパク質の構造に基づくモデリングにより得ることができる。合理的な薬物設計の実例は、HIVプロテアーゼインヒビターの開発である(Ericksonら、Science 249:527-533(1990))。加えて、ペプチド（例えばLRP5またはHBMタンパク質）をアラニンスキャンにより分析する(Wells、Methods in Enzymol.202:390-411(1991))。この技術では、アミノ酸残基をAlaと置換し、ペプチドの活性に及ぼすその影響を決定する。ペプチドの各々のアミノ酸残基をこの様式で分析してペプチドの重要な領域を決定する。

【0471】

機能アッセイにより選択される標的特異的抗体を単離し、次にその結晶構造を解析することも可能である。原理的には、この研究法により続く薬物設計が基づき得るファルマコアが得られる。機能的、薬理学的に活性な抗体に対する抗イディオタイプ抗体（抗id）を作製することにより結局タンパク質結晶学を迂回することが可能である。鏡像の鏡像として、抗idの結合部位を元来のレセプターのアナログであると予測される。次いで抗idを用いて化学的または生物学的に生成されたペプチドバンクのバンクからペプチドを同定および単離することができる。次いで選択されたペプチドはファルマコアとして働く。

【0472】

従って、例えば望ましいLRP5またはHBMタンパク質活性または安定性を有するか、またはLRP5またはHBMタンパク質活性のインヒビター、アゴニスト、アンタゴニスト等として作用する薬物を設計できる。クローン化されたLRP5またはHBM配列の利用性のおかげで、LRP5またはHBMタンパク質の十分量を利用してX線結晶学のような分析研究を実施することができる。加えて、本明細書で提供されるLRP5またはHBMタンパク質配列の知識により、X線結晶学の代わりに、またはそれに加えてコンピューターモデリング技術を用いる研究が導かれる。

【0473】

トランスジェニック動物の使用により、同定された薬物候補（「リード」として公知である）をさらに研究することができる。本発明のトランスジェニック動物は薬物リードを

10

20

30

40

50

試験し、さらに精密にするために用いることができる、骨密度変調の動物モデルを作製するのに有用である。前記したトランスジェニック動物は、本明細書で企図されるLRP5およびHBMまたはHBM様トランスジェニック動物の1つの実例を提示するものである。当業者は本発明の修飾において特定の目的に通常的に適用される変動および考慮すべき事項を認識している。トランスジェニック動物モデルの開発の実例は、例えばStrategies in Transgenic Animal Science(1995, MonasterskyおよびRobl編、Washington, DC: American Society for Microbiology)で示され、その参考文献は全て参照によりその全てを本明細書に組み込む。

【0474】

実例としては、トランスジェニック動物の少なくとも2つの群を記載するように作製して、一方の群がHBMを、他方の群がLRP5を発現するようにできる。これらの動物を数日から動物の余命までにわたる期間、候補物質で処置することができる。骨量の変化および／またはHBM表現型の代理マーカーに関して動物をモニター観察する。このような研究で用いられるトランスジェニック動物はヒトHBMタンパク質およびLRP5タンパク質または各々の種もしくはその変異体で定義されるHBMおよびLRP5相同タンパク質を発現することができる。偏在性プロモーターまたは公知の骨特異的プロモーターにより発現を誘導することができる。異なるプロモーターを利用して動物の群を比較することは情報価値がある。

【0475】

本発明による方法のトランスジェニック動物は、HBM、LRP5を発現するか、または動物の元来のプロモーターの調節下ではいずれも発現しない、ノックイン(KI)および／またはノックアウト(KO)動物例えばマウスを含むこともできる。前記したように(ならびに当分野の他の文献、例えば米国特許第6,187,991号および第6,187,992号、およびそこに引用される参考文献(これは参照によりその全てを本明細書に組み込む))ES細胞における相同組換えによりこのような動物を作製することができる。候補薬物で処置したトランスジェニック動物の実験群を非観血的手段により、前記したように代理マーカーのモニター観察により、および／または所定の時点で剖検された動物の骨のエクソビポ分析によりモニター観察することができる。

【0476】

同様に、食事調節(例えばビタミン、ミネラル、タンパク質、脂質等の摂取の変化)、卵巣切除、精製されたHBMまたはLRPタンパク質の全てもしくは一部の直接投与、アンチセンスヌクレオチドの投与、LRP5に対する抗体、または成人の遺伝子治療のような処置の効果を、本発明によるトランスジェニック動物に対する処置の全身投与により調査することができる。このような処置には、エストロゲン、タモキシフェン、ラロキシフェン、(またはその他の選択的エストロゲンモデュレーター、SERM)、ビタミンDアナログ、カルシトニン、カテプシンKインヒビター、スタチン(例えばシムバスタチン、パラバスタチンおよびロバスタチン)、ビスホスフォネート、副甲状腺ホルモン(PTH)、米国特許第6,190,880号および第5,866,364号に記載される骨形態形成タンパク質(BMP)ならびに前記の化合物の組み合わせなどがある。

【0477】

LRP5およびLRP6間の、およびLDLレセプターに対する相同性、ならびにHBM対象において変調されている心臓健康のためのマーカーのさらなる観察を考慮すると、本明細書に開示された新規研究方法を、公知の心臓保護処置から骨変調効果をスクリーニングするための用いることは本発明の実施形態である。それにより本発明は、骨保護および骨質を改善する両方の治療方法を提供する。モデルは薬物を試験し、LRP5およびHBMに関連する脂質変調効果を研究するのに有用である。

【0478】

LRP5およびHBM遺伝子の種々変異の影響を、本発明に従ってトランスジェニック動物のさらなる株を作製することにより調査でき、ここでこれらの動物はこのような変異を含む。骨発達または代理マーカーの直接測定と比較により、本発明の実施形態は遺伝子

10

20

30

40

50

治療試薬をスクリーニングするのに有用な研究手段、候補薬物治療、および骨発達変調の分子メカニズムの解明を提供する。当業者はこれらの目標を達成するための本発明の方法の使用の仕方を知っている。

【0479】

本発明は予期される遺伝子治療を試験するための方法および有用な研究手段を提供する。トランスジェニックノックアウトマウスは予期される遺伝子治療を試験するのに有用である。例として、内因性LRP5またはHBMを発現しないトランスジェニックノックアウト動物例えばマウスを前記のように作製する。予期される遺伝子治療、例えばCMV β アクチンプロモーターにより誘導される組換え複製欠損アデノウイルスコード化ヒトHBMタンパク質を投与する。骨密度のパラメーターおよび／または代理マーカーを、治療を行っている間モニター観察する(Ishibashiら、J.Clin.Invest.92:883-893(1993))。TaqMan (登録商標) プライマーセット、例えば前記したものを用いてトランスジェニックHBMの発現を測定することができる。当業者は代替方法、例えばノーザンブロット法を知っている。生殖系列トランスジェニック動物、ノックアウト(ヌル対立遺伝子)バックグラウンド、および野生型内因性LRP5バックグラウンド動物の群内および群間の双方の処置および未処置動物の比較により、遺伝子治療の種々の様式の相対的有効性を評価するための補足的対照が提供される。

10

【0480】

本明細書で企図されるトランスジェニック動物モデルの使用には、非限定例としては、(1) 骨細胞(例えば骨芽細胞および破骨細胞)機能および数を研究するためのトランスジェニック動物の頭蓋冠に由来する骨細胞培養を作製するための供給源；(2) トランスジェニック動物の卵巣切除(ovx)によりエストロゲン喪失の影響を研究するためのモデル；(3) 骨に及ぼす機械的負荷およびその他のストレス／強度試験を試験するためのモデル；(4) 骨異常を展示する、別の遺伝的修飾がなされたまたは天然発生変異動物に繁殖させる繁殖モデル；(Chipmanら、PNAS 90:1701-1705(1993);Phillipsら、Bone 27:219-226(2000);Kajkenovaら、J.Bone Min.Res.12:1772-1779(1997);Jilkaら、J.Clin.Invest.97:1732-1740(1996);Takahashiら、Bone and Mineral 24:245-255(1994))；(5) 重量負荷すなわち重力の影響を試験するための骨不適切／廃用(misuse)モデル；(6) 骨代謝を変調できるかまたは変調することが解っている試薬を同定およびスクリーニングするためのモデル(例えばPTH、エストロゲン、ビタミンDアナログ、ビスホスフォネート、スタチン、レプチン、BMP、アポE、SERMS)；(7) 骨折修復を改善するための予期される処置を調査するためのモデルなどがある。トランスジェニック動物を骨疾患、脂質疾患、Wntシグナル発生等の別の遺伝的(または遺伝的に修飾された)マウスモデルと交雑繁殖させることができる。これらの実例としては骨形成不全症(oi)マウス、特発性骨折(sfx)マウス、アポE異常を伴う動物、TCF-LacZまたはその他のレポーター遺伝子(GFP、ルシフェラーゼ、CAT)でWntシグナル発生をモニター観察するトランスジェニック動物等がある。

20

30

【0481】

非限定的な例としては、pDEXA、pQCTおよびマイクロCTによる骨密度測定；組織学；分子マーカー分析、アポトーシス、細胞増殖、細胞サイクル、ミネラル化、血清生化学、転写プロファイル等のような方法を用いてトランスジェニック動物モデルを分析することができる。

40

【0482】

XXII. 使用方法：鳥類および哺乳類動物牧畜

LRP5 DNAおよびLRP5タンパク質および／またはHBM DNAおよびHBMタンパク質を脊椎動物および好ましくはヒト治療薬に、および家畜繁殖用を含めて鳥類および哺乳類脊椎動物薬に用いることができる。例えばニワトリ、雄ドリ、雌ドリ、シチメンチョウ、ダチョウ、アヒル、キジおよびウズラなどの鳥類を高骨量の遺伝子および経路の同定から利益を得ることができる。文献に引用されている多くの例(例えばMcCoyら、Res.Vet.Sci.60(2):185-186(1996))では、牧畜条件を原因とする脆弱な骨により、ケージ

50

病、骨粗鬆症および高死亡率が引き起こされる。鳥類の骨粗鬆症またはその他の骨障害を処置するためのさらなる治療薬は牧畜産業、例えば食肉および卵の生産の福利および経済状態にかなり有利な影響を及ぼし得る。

【0483】

XXII. 使用方法：骨発達に影響する遺伝的変化の検出のためのLRP5特異的オリゴヌクレオチドを用いる診断アッセイ

骨発達の変化または疾患がLRP5遺伝子またはHBM遺伝子の変化に関与すると推定される場合、特定のオリゴヌクレオチドを構築し、それを用いて骨発達に影響する骨組織またはその他の組織におけるLRP5 mRNAまたはHBM mRNAのレベルを各々評価することができる。

10

【0484】

例えば骨密度に影響するHBM遺伝子をヒトが有しているかどうかを試験するために、ポリメラーゼ連鎖反応を使用することができる。標準的な方法により2つのオリゴヌクレオチドを合成するか、または特注のオリゴヌクレオチドの販売者から入手する。Oligo 4.0プライマー・ピッキングプログラム(Wojchick Rychlik, 1992)またはいずれかの適当な代替を用いて標準的な基準により長さおよび塩基組成を決定する。オリゴヌクレオチドの1つを、用いるPCR条件下でHBM DNAにのみハイブリダイズするように設計する。その他のオリゴヌクレオチドをLRP5ゲノムDNAのセグメントにハイブリダイズするように設計して、これらのオリゴヌクレオチドプライマーを用いるDNA増幅が便宜的に同定されたDNAフラグメントを生成するようにする。例えば、以下の条件を用いる場合、プライマー対CCAAGTTCTGAGAAAGTCC（配列番号32）およびAATACCTGAACCATACCTG（配列番号33）はDNAサンプルから530塩基対のDNAフラグメントを増幅する。工程1では95℃で120秒間；工程2では95℃で30秒間；工程3では58℃で30秒間；工程4では72℃で120秒間；ここで工程2から4を35回繰り返す。卵胞、全血、または頬側口腔から組織サンプルを入手できる。

20

【0485】

前記の手順で作製したフラグメントを標準的な技術によりシーケンシングする。HBM遺伝子に関する個々のヘテロ接合性はグリシン171のコドンの2番目の位置で等量のGおよびTを示す。正常なすなわちホモ接合性野生型の個体はこの位置でGのみを示す。類似の通常的な手順を用いて本発明による別の多型性および変異体のためのアッセイを開発することができる。

30

【0486】

PCRに加えて別の増幅技術、例えばライゲーション媒介PCRまたはQ-βレプリカーゼに関与する技術(Cahillら、Clin.Chem.37(9):1482-1485(1991))を代替として用いることができる。例えばオリゴヌクレオチドAGCTGCTCGTAGCTGTCTCTCCCTGGATCACGGGTACATGTA CTGGACAGACTGGGT（配列番号34）およびTGAGACGCCCCGATTGAGCGGGCAGGGATAGCTTATTCCTGT GCCGATTACGGC（配列番号35）を変性ヒトDNAサンプルとハイブリダイズさせ、DNAリガーゼと処置し、次にプライマーオリゴヌクレオチドAGCTGCTCTCGTAGCTGTCTCTCTCCCTGGTA（配列番号36）およびGCCGTAAATGCGGCACAGGGAATAAGCT（配列番号37）を用いるPCR増幅に供することができる。最初の2つのオリゴヌクレオチドでは、外側の27塩基がプライマー結合部位に相当するランダム配列であり、内側の30塩基がLRP5遺伝子の配列に相当する。最初のオリゴヌクレオチドの末端のTはHBM遺伝子に相当する。最初の2つのオリゴヌクレオチドは、HBM遺伝子を担持するヒトDNAにハイブリダイズしたときにのみライゲートし、これは増幅可能な114塩基DNAフラグメントの形成に至る。

40

【0487】

アガロースゲル電気泳動、定量用ハイブリダイゼーション、または分子生物学の当業者に公知の核酸検出のための等価な技術により増幅産物を検出することができる(Sambrookら、Molecular Cloning:A Laboratory Manual、Cold Spring Harbor Laboratory、Cold S

50

pring、ニューヨーク州(1989))。

【0488】

L R P 5 遺伝子またはH B M 遺伝子における別の変化を同一の型の増幅検出手順により、これらの変化を同定するように設計されたオリゴヌクレオチドを使用することにより診断することができる。これらの手順を動物およびヒトにおいて使用して、骨発達に影響するL R P 5 またはH B M における変化を同定することができる。

【0489】

脊椎動物細胞を遺伝的に操作するためのベクターの状況で、L R P 5 またはH B M の c D N A を各々骨特異的プロモーターに融合させることにより骨組織におけるL R P 5 またはH B M の発現を達成することができる。D N A をウイルスキャプシッドにパッケージングすることにより、カチオン性リポソーム、エレクトロポレーションを使用することにより、またはリン酸カルシウムトランスフェクションによりD N A 構築物を細胞に導入する。トランスフェクトされた細胞、好ましくは骨芽細胞を培養物中で研究するか、または骨への直接注射により、もしくは骨芽細胞の静脈内注射により、続いて骨組織への組み込みにより動物の骨組織に導入することができる(Koら、Cancer Research 56(20):4614-4619(1996))。例えば骨芽細胞で特異的に活性であるオステオカルシンプロモーターを用いてL R P 5 遺伝子またはH B M 遺伝子の転写を導くことができる。いくつかのベクターおよびトランスフェクション方法、例えばレトロウイルスベクター、アデノウイルスベクターもしくはカチオン性リポソームを用いるトランスフェクションの後に維持されるベクター、または本明細書に記載したその他の方法およびベクターのいずれかを用いることができる。

10

20

【0490】

機能的L R P 5 タンパク質またはH B M タンパク質のレベル変化は骨ミネラル化のレベルに影響する。機能的L R P 5 タンパク質またはH B M タンパク質のレベルを操作することにより、骨発達に影響し、骨ミネラル化のレベルを上昇または低下させることが可能である。例えば骨粗鬆症の患者において骨ミネラル化を増大させることは有用であろう。また別に、大理石骨病またはパジェット秒の患者において骨ミネラル化を低下させることは有用であろう。L R P 5 レベルまたはH B M レベルの変化を研究手段として用いることもできる。具体的には、そのレベルまたは修飾状態がL R P 5 またはH B M の機能的レベルにおける変化に応答して変化するタンパク質、m R N A およびその他の分子を同定することが可能である。骨障害の病理および病因は公知であり、例えばRubinおよびFarber(編)、Pathology、第2版、S.B.Lippincott Co.、Philadelphia、ペンシルバニア州(1994)に記載されている。

30

【0491】

種々の技術を用いて機能的L R P 5 またはH B M のレベルを変化させることができる。例えばL R P 5 の細胞外部分もしくはその変異体、またはH B M もしくはその変異体の静脈内または骨内注射により処置したヒト、動物または鳥類の体内のL R P 5 活性またはH B M 活性の各々のレベルが変化する。L R P 5 タンパク質またはH B M タンパク質のトランケート版を注射して機能的L R P 5 タンパク質またはH B M タンパク質のレベルを各々変化させることもできる。L R P 5 またはH B M の特定の形態は内因性タンパク質の活性を増強するが、一方別の形態は抑制性である。

40

【0492】

好ましい実施形態では、H B M タンパク質を用いて骨粗鬆症、骨折またはその他の骨障害を処置する。別の好ましい実施形態では、H B M タンパク質の細胞外部分を使用する。このH B M タンパク質は任意でタンパク質を細胞表面に付着させる基の付加により修飾されていてもよい。タンパク質を医薬的に許容される溶液で調製し、注射または許容される薬物動態および分配を達成する別の方法により投与する。

【0493】

この方法の第2の実施形態では、L R P 5 またはH B M レベルは遺伝子治療技術により上昇または低下する。L R P 5 またはH B M レベルを上昇させるために、前記したように

50

、骨芽細胞または別の有用な細胞型を遺伝的に操作して高レベルのLRP5またはHBMを発現される。また別に、LRP5またはHBMレベルを低下させるために、LRP5またはHBM mRNAの翻訳可能なレベルを特異的に低減させるアンチセンス構築物を使用することができる。概して、組織非特異的プロモーター、例えばCMVプロモーターまたは発現ベクターにおいて見出される、市販により入手可能な別のプロモーターを用いることができる(Wuら、Toxicol.Appl.Pharmacol.141(1):330-339(1996))。好ましい実施形態では、LRP5 cDNAまたはそのアンチセンスが骨特異的プロモーター、例えばオステオカルシンまたは別のプロモーターにより転写され、骨組織において特異的発現を達成する。この方法で、LRP5発現DNA構築物またはHBM発現構築物を非骨細胞に導入する場合、これは発現されない。

10

【0494】

この方法の第3の実施形態では、LRP5またはHBMに対する抗体を用いてその機能を阻止する。このような抗体は本明細書で同定されている。

【0495】

この方法の第4の実施形態では、LRP5機能またはHBM機能のアゴニストまたはアンタゴニストである薬物を使用する。このような薬物は本明細書にて記載されており、医薬品開発の当業者に周知の医薬品化学の技術により最適化される。

【0496】

LRP5およびHBMはいくつかのタンパク質、例えばアポEと相互作用する。LRP5またはHBMとアポEまたはその他の結合パートナーとの相互作用を阻止する分子は骨発達およびミネラル化を変化させると予測される。このようなインヒビターは骨粗鬆症、骨大理石病、またはその他の骨ミネラル化の疾患の処置において薬物として有用であろう。このようなインヒビターは低分子量化合物、タンパク質またはその他の型の分子でよい。Kimら、J.Biochem.(Tokyo)、124(6):1072-1076(1998)参照。

20

【0497】

標準的な薬物スクリーニング技術によりLRP5またはHBMと相互作用するタンパク質との相互作用のインヒビターを単離することができる。例えばLRP5タンパク質(もしくはそのフラグメント)またはHBM(もしくはそのフラグメント)を固体支持体例えばマイクロタイターウェルのベースに固定することができる。第2のタンパク質またはタンパク質フラグメント、例えばアポEを誘導体化して例えば蛍光を用いる検出を補助する。LRP5またはHBMのこのタンパク質-タンパク質ドメインを各々特異的に阻止できる候補化合物の存在下、ヨウ素またはビオチンを次いでLRP5またはHBMに添加し、従ってその膜セグメントに随伴される問題を回避する。この型の薬物スクリーニングは医薬品開発の分野の当業者に周知である。

30

【0498】

LRP5およびHBMは骨発達に関与するので、LRP5およびHBMに結合するタンパク質もまた骨発達に関与することが予測される。標準的な方法、例えば同時免疫沈澱、同時分画化、または2ハイブリッドスクリーニング(Ausubelら、Current Protocols in Molecular Biology、John Wiley & Sons (1997))によりこのような結合タンパク質を同定することができる。例えば2ハイブリッド系を用いてLRP5相互作用タンパク質またはHBM相互作用タンパク質を同定するために、LRP5またはHBMの細胞外ドメインをLexAに融合させ、酵母ベクターpEG202(「bait」)に関して発現させ、酵母株EGY48において発現させた。酵母株を適当なベクター中、候補相互作用タンパク質に融合したガラクトース誘導転写活性化配列をコードする「prey」ライブラリーで形質転換する。この方法により相互作用タンパク質を最初に選択し、続いて確認するための技術は分子生物学の当業者に周知である(Ausubelら、Current Protocols in Molecular Biology、John Wiley & Sons (1997))。

40

【0499】

好ましい実施形態では、HBMと相互作用するが、LRP5とは相互作用しないタンパク質を前記の手順の変法を用いて同定する(Xuら、Proc.Natl.Acad.Sci.USA 94(2):12473-

50

12478(1977年11月))。この2ハイブリッド系の変法は2つのベイトを用い、LRP5およびHBMを各々LexAおよびTetRに融合させる。また別にHBMと相互作用するがLRP5とは相互作用しないタンパク質を単離することができる。これらの手順は分子生物学の当業者に周知であり、標準2ハイブリット手順の単純な変法である。

【0500】

LRP5またはHBMと相互作用する物質を単離するまた別の方法として生化学的研究方法を用いる。LRP5タンパク質またはそのフラグメント、例えば細胞外ドメイン、またはHBMタンパク質またはそのフラグメント、例えば細胞外ドメインを化学的にセファロースビーズに結合させる。LRP5またはHBM結合ビーズをカラムに注ぐ。生物学的抽出物、例えば脂質分画、血清タンパク質、骨バイオプシの上澄のタンパク質、または穏やかに溶解した骨芽細胞に由来する細胞内容物をカラムに添加する。非特異的結合化合物を溶出し、カラムを低塩バッファーで数回洗浄し、次に強固に結合した化合物を高塩バッファーで溶出することができる。これらはLRP5またはHBMに結合する候補化合物であり、標準的な試験および対照実験により特異的結合に関して試験することができる。タンパク質を結合させるのに用いるセファロースビーズおよび結合を実施するための方法は市販により入手可能であり(Sigma)、本明細書に記載した手順はタンパク質生化学の分野の当業者に周知である。

10

【0501】

前記の手順の変法として、高塩によりLRP5またはHBMセファロースカラムから溶出されるタンパク質を次いでHBM-LRP5セファロースカラムに添加する。付着せずに流れるタンパク質はLRP5に結合するが、HBMには結合しないタンパク質である。また別に、カラムを使用する順序を逆にすることによりHBMタンパク質に結合するが、LRP5タンパク質には結合しないタンパク質を単離することができる。

20

【0502】

単離された化合物を標準的な方法、例えば2Dゲル電気泳動、クロマトグラフィーおよび質量分析により同定することができる。

【0503】

XXIV. 使用方法：形質転換随伴組換え(TAR)クローニング

LRP5の新規対立遺伝子変異体の同定に必須なものは、個体における遺伝子の双方のコピーの配列を試験する能力である。これを達成するために、2つの「フック」、すなわち有意な類似性の領域をゲノム配列内で同定して、これがクローン化されるべきDNAの部分をフランキングするようにする。最も好ましくは、このフックの1つ目を目的の最初のエクソンに対して5'の配列から誘導し、2つ目を目的の最後のエクソンに対して3'の配列から誘導する。これらの2つの「フック」を細菌/酵母シャトルベクター、例えばLarionovら、Proc.Natl.Acad.Sci.USA 94:7384-7387(1997)に記載されるベクターにクローン化する。別の類似のベクター系を用いることもできる。LRP5遺伝子の全ゲノムコピーを回収するために、2つの「フック」を含有するプラスミドを制限エンドヌクレアーゼで直線化するか、またはPCRのような別の方法により生成する。この直線状DNAフラグメントをヒトゲノムDNAと一緒に酵母細胞に導入する。典型的には、サッカロマイセス・セレビシエを宿主細胞として用いるが、Kouprinaら(Genome Res.8:666-672(1998))は同様にニワトリ宿主細胞の使用を報告している。形質転換の過程の間および後に、組換え事象により内因性宿主細胞が直線状プラスミドを環状に変換し、それにより「フック」に相同なヒトゲノムDNAの領域がプラスミドに挿入される。このプラスミドを回収し、当業者に周知の方法により分析することができる。明らかに、この反応の特異性には、直線状フラグメントに存在する「フック」に類似する配列を認識する、宿主細胞の機構が必要である。しかしながら、ヒトゲノムに共通する重複反復配列を用いてげっ歯類/ヒトハイブリッド細胞株からヒトDNAのフラグメントを回収することについて記載するKouprinaら(Genomics 53(1):21-28(1998年10月))に示されるように、100%配列同一性は必要ではない。

40

【0504】

50

別の例では、Larionovら、Proc.Natl.Acad.Sci.USA 95(8):4469-4474(1998年4月)に記載されるように、「フック」のみが必要である。「ラジアルT A Rクローニング」と称するこの型の実験に関して、組換えを誘導するための配列類似性の別の領域をゲノムの反復配列から誘導する。この方法で、L R P 5 遺伝子コード化領域に隣接するD N Aの領域を回収し、機能に影響し得る変化に関して試験することができる。

【0505】

X X V. 使用方法：ゲノムシーケンシング

H B M 遺伝子またはL R P 5 に連結した多型性遺伝子マーカーの使用は骨粗鬆症またはその他の骨疾患に対する感受性を予測するのに非常に有用である。Kollerら、Amer.J.Bone Min.Res.13:1903-1908(1998)、多型性遺伝子マーカーの使用が連鎖解析に有用であることを実証した。同様に、高骨量遺伝子内の多型性遺伝子マーカーの同定により、骨発達に影響する別の遺伝的損傷と連鎖不均衡にある特異的対立遺伝子変異体の同定が可能になる。B A C からのD N A 配列を使用して、ジヌクレオチドC A n 反復を同定し、以下に示す、この反復を含有するゲノムD N A を増幅する2つの独特なP C R プライマーを設計し、

B 2 0 0 E 2 1 C 1 6 L : G A G A G G C T A T A T C C C T G G G C (配列番号38)

B 2 0 0 E 2 1 C 1 6 R : A C A G C A C G T G T T T A A A G G G G (配列番号39)

遺伝子マッピング研究に使用した。

【0506】

この方法は別の当業者に用いられて成功している(例えばSheffieldら、Genet.4:1837-1844(1995);LeBlanc-Straceskiら、Genomics 19:341-349(1994);Chenら、Genomics 25:1-8(1995))。集団または個人でのこれらの試薬の使用により彼らの骨粗鬆症の危険性が予測される。同様に、例えば前記の表4で示す一塩基多型(S N P)を同様に用いて骨疾患を進展させる危険性またはH B M 遺伝子の場合骨粗鬆症に対する抵抗性を予測することができる。

【0507】

X X V I. 使用方法：組織石灰化の変調

ヒト体内での組織の石灰化は十分文献化されている。Towlerら、J.Biol.Chem.273:30427-30434(1998)はモデル系の発達中の頭蓋の石灰化を制御することが解っているいくつかのタンパク質が石灰化大動脈において発現されることを実証した。骨芽前駆細胞で転写される遺伝子であるM s x 2 の石灰化血管組織における発現は、骨発達において重要である遺伝子その他の組織の石灰化に関与していることを示している。H B M タンパク質、アゴニストまたはアンタゴニストでの処置は、その実証された骨ミネラル密度に及ぼす影響のために石灰化(例えば脈管構造、象牙質および頭蓋内臓の骨)を緩解する可能性がある。組織の石灰化が実証される実験系では、L R P 5 活性の過剰発現または抑制により、L R P 5 遺伝子により直接制御される分子の同定が可能になる。これらの遺伝子は組織石灰化を変調することを目指す治療の潜在性標的である。例えば、動物、例えばL D L R - / -、マウスに高脂肪食を与え、観察して、L R P 5 を含む組織石灰化のマーカーの発現を実証する。次いでL R P 5 もしくはH B M タンパク質に対する抗体、L R P 5 もしくはH B M c D N A に対して志向するアンチセンスオリゴヌクレオチドで、またはL R P 5 もしくはH B M タンパク質に結合することが解っている化合物またはその結合パートナーもしくはリガンドでこれらの動物を処置する。血管組織からR N A またはタンパク質を抽出し、組織で発現される遺伝子の相対発現レベルを当分野で周知の方法により決定する。組織において制御される遺伝子は組織石灰化のモデュレーターとして医薬品開発に可能性のある治療標的である。

【0508】

本発明の、個体に与えるための核酸、タンパク質、ペプチド、アミノ酸、小型分子またはその他の医薬上有用な化合物を組成物の形態で、当分野で周知の医薬上許容される担体、賦形剤または希釈剤と共に投与することができる。個体は哺乳動物または鳥類、好まし

10

20

30

40

50

くはヒト、ラット、マウスまたは鳥類でよい。このような組成物を個体に医薬上有効な量で投与することができる。投与される量は処置される症状および処置される患者に依存して変化する。組成物を単独でまたはその他の処置と組み合わせて投与することができる。

【実施例】

【0509】

[実施例]

本発明を以下の実施例を参考にして記載するが、該実施例は説明のために提供され、本発明をいかなる様式でも限定することを意図するものではない。当分野で周知である標準的な技術か、または具体的に以下で記載した技術を利用した。

【0510】

[実施例1]

医師が発端者をCreighton Osteoporosis Centerに送り、骨密度異常であると思われるものを評価した。彼女は18歳であり、背中の痛みのために2年前に治療に来たが、これは彼女が乗客として乗っていた自動車に追突された自動車事故によるものであった。彼女の損傷は下背部への軟部組織損傷であり、痛みおよび筋肉圧痛の徴候が現れた。放射線写真での骨折または不全脱臼の証拠はなかった。彼女は終日学校に通うことができたが、2年間痛みが続いた。彼女がセンターに来院するまでに痛みはほぼ消散し、彼女は高校生として通常生活に戻った。身体診察により身長66インチ、および体重128ポンドの健全な若い女性であることが示された。全体骨格の放射線写真により骨は皮質が厚く高密度に見えることが示された。骨格の全ての骨が関係していた。最も重要なことに、全ての骨の形状は全く正常であった。脊髄のBMCはL1-4で94.48グラムであり、脊髄BMDはL1-4で1.667 gm/cm²であった。BMDは女性のピーク骨格質量を超えて標準偏差5.62 (SD)であった。これらはHologic 2000を用いるDXAにより測定した。次いで彼女の母親をスキャンし、腰椎BMCは58.05グラムおよびBMDは1.500 gm/cm²であることが見出された。彼女の母親の値はピーク質量を超えて4.12 SDになり、同等のヒトを超えて4.98の値になる。彼女の母親は51歳で、身長65インチ、および体重140ポンドであった。彼女の母親は極めて良好な健康状態であり、筋骨格疾患またはその他の症状の病歴がなかった。彼女の父親は腰椎BMC 75.33グラムおよびBMDは1.118 gm/cm²であった。これらの値は男性のピーク骨質量を超えて0.25 SDになる。彼は良好な健康状態であり、身長72インチ、および体重187ポンドであった。

【0511】

これらの臨床データにより、発端者が彼女の母親から形質を遺伝していることが示唆され、これは非常に高い骨量に至るが、それ以外は正常な骨格であり、母方の家系が目された。米国特許第5,691,153号では、これらのメンバーの22人がDXAによる骨量の測定値を有した。一例では、発端者の母方の祖父が死去したが、しかしながら医学的記録、生前の骨格放射線写真およびDNA遺伝子型解析のためのパラフィン包埋された胆嚢標本は得られた。彼の放射線写真は、大腿骨および脊髄を含む試験可能な全ての骨で明らかに極端な密度が示され、彼は罹患メンバーに含まれた。本発明では、37人の情報価値の在る個体を含むように家系を広げている。元来の試験以来、追加された14人の個体の中で2人の個体が重要なクロスオーバーを宿しているので、この追加により元来の親族関係よりも有意に改善された(Johnsonら、Am.J.Hum.Genet.60:1326-1332(1997))。個体12から14および15への男性から男性への伝達の存在によりX連鎖を除外する。

【0512】

[実施例2]

本発明はHBM遺伝子領域に由来する2つのBACクローンから誘導されたDNA配列について記載し、これらのクローンの集合である以下の表8で明らかである。クローンb200e21-h (ATCC番号980812; 配列番号10~11)は1997年12月30日にAmerican Type Culture Collection(ATCC)、10801 Blvd大学、Manassas、バージニア州20110-2209米国に寄託された。クローンb527d12-h (ATCC番号980720

10

20

30

40

50

；配列番号 5～9）は1998年10月2日にAmerican Type Culture Collection(ATCC)、10801 Blvd大学、Manassas、バージニア州20110-2209米国に寄託された。これらの配列は独特な試薬であり、当業者はこれを用いてLRP5遺伝子に関するDNAプローブ、遺伝子を増幅するためのPCRプライマー、LRP5遺伝子におけるヌクレオチド多型性、またはLRP5遺伝子の制御エレメントを同定することができる。

【0513】

【表10】

表8

コンティグ	ATCC No.	配列番号	長さ(塩基対)
b527d12-h_コンティグ 302G	980720	5	3096
b527d12-h_コンティグ 306G	980720	6	26928
b527d12-h_コンティグ 307G	980720	7	29430
b527d12-h_コンティグ 308G	980720	8	33769
b527d12-h_コンティグ 309G	980720	9	72049
b200e21-h_コンティグ 1	980812	10	8705
b200e21-h_コンティグ 4	980812	11	66933

10

【0514】

20

[実施例3]

HBM過剰発現トランスジェニックおよび非トランスジェニックマウスに由来する頭蓋冠および脛骨移植片培養物の転写プロファイリング

HBM表現型の代理マーカーおよび骨量変調治療のための推定標的および本発明の方法による薬物の同定ならびに転写プロファイリングによるHBMに関連する遺伝子の同定および特徴付けのための本発明のトランスジェニック動物の使用を実証する。

【0515】

骨特異的I型コラーゲンプロモーター下でHBMを発現するトランスジェニックマウス(株19)を含む生仔(12日齢)マウスから頭蓋冠および脛骨を入手する。4匹のトランスジェニックおよび4匹の非トランスジェニックマウスから頭蓋冠をプールし、コラゲナーゼで消化した。消化物を培養中に置いた。頭蓋冠培養物をアスコルビン酸およびβリン酸グリセロール含有または不含で19日間維持した。RNAを19日に単離した。

30

【0516】

骨髄間質細胞を脛骨から洗い流し、個々の動物からの脛骨を次いで2連続コラゲナーゼ消化に供した。コラゲナーゼ消化に続いて骨チップを培養に置き3つの連続する播種が得られた。播種3からの細胞は播種1および2からの細胞よりも非常にゆっくりと成長した。播種1から3(継代1)の全面成長細胞からRNAを単離した。頭蓋冠および脛骨培養物の双方からのRNAをU74Av2転写プロファイリングアレイで分析した。

【0517】

アスコルビン酸およびβリン酸グリセロールでの処理により、遺伝子のセットが全く異なる差次的発現に至った。処置細胞の分化を示す処置の後、アルカリ性ホスファターゼ(AKP)遺伝子発現が増加した。処置細胞から得られたデータは未処置細胞から得られたデータとは質が異なっており、同型培養間で遺伝子発現の変動もあった。

40

【0518】

非トランスジェニックおよびトランスジェニックマウスの転写プロファイルはいくつかの関連遺伝子、例えばS100A1、MMP9、およびMT1の発現における差異を示した。

【0519】

各群(すなわち非トランスジェニックおよびHBMトランスジェニック)4匹のマウスからの脛骨移植片培養物の転写プロファイルにマウス-マウス間の変動がある。この変動

50

は正規化の後にも依然認めることができる。しかしながら変動はなんら有意な程度までデータの解釈に影響しない。変動はこれらの細胞のアルカリ性ホスファターゼ活性の測定においても認められる。

【0520】

結果を以下に要約する。その転写プロファイルにおいて播種1および播種2は類似している（播種3とは異なっている）。これは、これらの細胞の成長特性における差異による可能性がある。播種1はまたトランスジェニックおよび非トランスジェニックプロファイル間でも大きな差異を示すが、恐らく播種1が播種2または3のいずれかよりも相対的により混合された細胞集団であるのが原因である。

【0521】

非トランスジェニックおよびHBMトランスジェニックマウス間の遺伝子発現におけるいくつかの差異はヒトHBM1家系からの罹患および非罹患個体間で認められる差異と合致する。一例として、S100A1（Genbankアクセッション番号AF087687）はトランスジェニック骨芽細胞培養において上方制御される。この遺伝子によりコードされるタンパク質は2EF-ハンド・カルシウム結合モチーフを含有するタンパク質のS100ファミリーのメンバーである。S100タンパク質は広範な細胞の細胞質および／または核に局在し、多くの細胞性過程、例えば細胞サイクルの進化および分化の制御に関与する。マトリックス・メタロプロテイナーゼ9（MMP9）（Genbankアクセッション番号X72794）もまたトランスジェニック骨芽細胞培養において上方制御される。マトリックス・メタロプロテイナーゼ（MMP）ファミリーのタンパク質は正常な生理学的過程、例えば胚発達、生殖、および組織リモデリングにおいて、ならびに疾患過程、例えば関節炎および転移において細胞外マトリックスの分解に関与する。たいていのMMPは不活性なプロタンパク質として分泌され、これは細胞外プロテイナーゼにより切断されたときに活性化する。この遺伝子によりコードされる酵素はI VおよびV型コラーゲンを分解する。

【0522】

HBMトランスジェニック骨芽細胞培養物で下方制御される遺伝子の中で、システイン富化金属結合タンパク質であるメタロチオネイン1（MT1）（Genbankアクセッション番号S62785）が抗酸化剤として重要な働きをすることが示されている。その活性は炎症過程の細胞毒性を媒介する。骨および軟骨の双方でこれは発現される。

【0523】

当業者は本明細書に記載する方法を用いてトランスジェニック骨組織において差次的に発現される別の遺伝子を決定することができる。転写プロファイリングデータから、いくつかの方法でトランスジェニックマウス脛骨のプロファイルがヒトHBM家系の罹患メンバーのプロファイルに類似していることを示すことができる。

【0524】

[実施例4]

トランスジェニックマウスを用いる負荷試験

トランスジェニック修飾、例えばLRP5、LRP6およびHBM発現、過剰発現またはノックアウトの骨発達に及ぼす影響を負荷または非負荷プロトコルを用いて評価することができる。負荷または非負荷に供された骨成長速度、遺伝子発現応答プロファイリング、およびHBM表現型の生化学的パラメーターをこれらの方法によりさらに特徴付けすることができる。本発明のトランスジェニック動物を用いるこれらの方法は、HBM表現型の望ましい特性を再現する処置および薬物の開発における貴重な手段である。

【0525】

4点屈曲装置を用いてトランスジェニックまたは非トランスジェニックマウスの脛骨に機械的負荷をかける。インビボで正確な外力適用に関して装置を校正する。装置はバルサ材から成り、1mm厚の外科用チューブで被覆した4個の丸いパッドを通して力を適用する。上部のパッドは4.5mm離れており、12mm離れている下部のパッド間の中心にある。4点屈曲で、脛骨の外側で圧迫、および内側で張力を伴う2個の内側のパッドの間

10

20

30

40

50

の骨組織に一定の屈曲モーメントが限なく分配される。装置の説明を図28に示す。上部遠位パッドは脛骨－腓骨接合部（TFJ）に対して1mm近位で脚と接触し、下部遠位パッドはTFJに対して2.5mm遠位の内側表面と接触する。最大屈曲領域はTFJに1から6mm近位から、または遠位末端から8.5から13.5mmである。この領域は放射線写真により定義されている。負荷装置は各実験の前に較正されており、負荷は各動物に関して毎日記録される。機械をゼロに合わせ、負荷において何らかのドリフトがある場合は調整する。脚位置および適用される負荷は動物および日の間で一致し、脚位置による下部の変動は10%未満であり、負荷では変動は1.6%未満である(Haginoら、J.Bone Miner.Res.8:347-357(1993))。

【0526】

10

機械的負荷は右下肢に適用され、その間マウスは希イソフルラン麻酔（2%）下にある。信頼性のある脚位置は、プラットフォーム上の標準的な位置にマウスを配し、右脚を籠に置き、膝を負荷装置と並べることにより達成される。イソフルランは短時間作用性であるので、負荷の間の動きを防御できるが、負荷の数秒以内に正常な体重負荷作用が戻る。イソフルランからの適当な回復および脚に損傷を生じていないことに関して作用をモニター観察する。片側性神経切断により脱負荷を達成することができる(Kodamaら、Bone 25:183-190(1999))。加えて、4点負荷の適用の間にインビボで脛骨に株ゲージを適用することができる。

【0527】

蛍光色素標識：組織形態計測に関して研究すべき全マウスは、組織収集の10および3日前に投与される二重カルセイン標識を受ける。長期間負荷研究のマウスは負荷の前に、最終二重カルセイン注射に加えてテトラサイクリンのベースライン注射を受ける。これらの注射を注射に適した希釈で1ml/kgで調製する。25gマウスのための注射容量は0.25mlになる。全注射は皮下注射であり、穏やかな沈静下で与えられる（2%イソフルラン）。

20

【0528】

カルセイン標識(Sigma、St.Louis、ミズーリ州)を6.2mg/kgで注射する。2つのカルセイン標識を異なる2日で、すなわち剖検前3および10日に投与する。この蛍光色素標識を用いて非脱灰組織の表面のミネラル化を同定し、処置の最終週の間の骨形成の速度を定量する。標識は脱灰骨を試験するBrdUまたはインサイチュウ・ハイブリダイゼーション研究には用いられない。

30

【0529】

テトラサイクリン(Pfizer、コネチカット州)を25mg/kgで注射する。単回テトラサイクリン標識を0日に長期間研究（5週以上）の全動物に投与する。この蛍光色素標識は研究の開始時に表面のミネラル化を標識し、実験中の全骨形成の定量を可能にする。

【0530】

BrDU（プロモデオキシウリジン、Boehringer Mannheim）を40mg/kgで注射し、ベヒクルは静菌性である。マウスに6時間間隔で5回注射し、DNA合成を24時間にわたって標識する。最後の注射は組織収集の1時間前である。

【0531】

40

動物に固定液を灌流する場合を除いて、CO吸入により死亡を誘起する。全ての負荷および廃用研究用に右および左脛骨を切除する。右肢は負荷または処置された脚であり、左は処置対照である。右脛骨の負荷領域から、および左脛骨の類似の領域から組織を収集する。マウスでは、平均負荷領域が脛骨－腓骨接合部（TFJ）に対して1から6mm近位であるように決定している。

【0532】

非脱灰皮質骨サンプル：脛骨、大腿骨および脊椎を非脱灰骨切片の標準的な組織形態測定のために収集する。大部分の筋肉を除去し、骨を70%EtOH中に48時間置く。のこぎりで骨を切断し、分析用に以下のサンプルを作製する。a) TFJを含む脛骨骨幹；b) 遠位大腿骨、およびc) ディスクを含まない脊椎体。脛骨骨幹をビラヌエバ色素中に

50

72時間置いて、次いで80%エタノールに戻す。全てのその他の骨を直接脱水に移す。次の14日の間、標本を等級別エタノールおよびアセトン中で脱水し、次いで別個の修飾メタクリル酸メチルで包埋する。包埋された脛骨皮質サンプルをミクロトームのこぎり(1600モデル、Leica、ドイツ)で、70 μ mの横断切片にし、切片は領域から収集する。負荷領域から切片を取り、ラットTFJに対して5から7mm近位で、遠位末端から9から13mmで、0.8mmの切片間距離を有する切片を作製する。各脛骨に由来する2つの切片をマウントし、無作為な数字を付け、分析する。

【0533】

脱灰皮質骨サンプル：脚の軟部組織が固化するまで4%パラホルムアルデヒドで動物組織を灌流する。脛骨を切除し、筋肉を剪刀で取り除き、一方骨は冷4%パラホルムアルデヒドに浸す。骨膜および小型の筋肉層は無傷のまま残す。インサイチュウ・ハイブリダイゼーション研究用に、RNAアーゼ不含有材料で全ての作業を行う。脛骨の負荷領域からの4から5mmの切片をのこぎりで無傷脛骨から切除し、4%パラホルムアルデヒド中4℃で24時間固定する。固定後、骨を7%EDTA(Sigma)中pH6.5で2から3週間脱灰する。次いで骨を1%MgCl中に6時間置き、アルカリ性ホスファターゼ活性を回復させる。骨幹セグメントをJB-4プラス(Polyscience)またはパラフィンに包埋する。横断切片をミクロトーム(Reichert Jung)でタングステン・カーバイドナイフを用いて5 μ mの厚さに切断する。切片をポリ1-リジン(Sigma)またはコーティングスライドにマウントする。

【0534】

図29に示すように、負荷に供したマウスのカルセイン染色により、ヘテロ接合性HBMトランスジェニックマウス(19株)では、非トランスジェニック対照マウスよりも大きな成長が示される。単一株(5Nまたは7N、2Hzで36サイクル)を5日間投与した。15日に収穫された組織でカルセイン標識を5および12日に生じた。

【0535】

本発明を詳細に記載してきたが、多くの変更および修飾を為すことができ、このような変更および修飾を本発明の精神および範囲から逸脱することなく為すことができることは当業者に認識されよう。

【0536】

本明細書に引用した特許、特許出願および公開は参照によりその全てを全ての目的のために本明細書に組み込む。さらに2000年4月5日出願された米国特許出願第09/543,771号および第09/544,398号、1999年1月13日出願の出願第09/229,319号、1998年1月13日出願の米国仮出願第60/071,449号ならびに1998年10月23日出願の米国仮出願第60/105,511号は参照によりその全てを全ての目的のために本明細書に組み込む。

【0537】

加えて、本出願は2001年5月11日出願の出願第60/290,071号;2001年5月17日出願の第60/291,311号;2002年2月1日出願の第60/353,058号、および2002年3月4日出願の第60/361,293号の優先権を主張する;各々の開示は参照によりその全てを全ての目的のために本明細書に組み込む。

【図面の簡単な説明】

【0538】

【図1】図1は遺伝子連鎖研究に用いられた個体の家系を示す。各々の個体の下に、ID番号、脊髄BMDに関するzスコアがあり、対立遺伝子が、染色体11上の重要なマーカーを示す。黒ぬりの記号は「患っている」個体を表す。「N」を含む記号は「患っていない」個体である。37人の個体からのDNAについて遺伝子型を同定した。クエションマークは遺伝子型が同定されなかった未知の遺伝子型または個体を示す。

【図2】図2は11q13.3におけるHBM領域のBAC/STS含量物理地図を表す。STSマーカーは遺伝子、EST、マイクロサテライト、ランダム配列、およびBACエンドシーケンスを長い水平方向の線の上に記す。GDBに存在するマーカーに関しては、同一の命名法を用いている。遺伝子座の名前(D11S####)は利用可能な場合

10

20

30

40

50

、第1次名の後に括弧の中に列挙する。BACエンドシーケンスから誘導したSTSをBAC名と共に列挙し、クロンの左および右末端に関して各々LまたはRと続ける。2つの大きな矢印はHBMの重要な領域を定義する遺伝子マーカーを示す。STSの下の方の水平方向の線は、9倍の適用範囲のBACライブラリーのPCR基盤のスクリーニングにより同定されたBACクロンを示す。白丸はライブラリースクリーニング中、マーカーが対応するBACライブラリーアドレスを増幅しなかったことを示している。クロン名は以下の慣用法を用いる。BACに関してはB、プレート、列および囲みアドレス、続いてHがHBMプロジェクトを示す（すなわちB36F16-H）。

【図3】図3A～3Fはイントロン配列の側面に位置しているLRP5のゲノム構造を示す。翻訳は下線を付したエクソン1の「ATG」により開始される。HBM遺伝子の多型性の部位はエクソン3であり、下線を付した「G」により表され、それによりこのヌクレオチドはHBM遺伝子の「T」である。mRNAの3'非翻訳領域にはエクソン23内で下線を付している（エクソン1、配列番号40；エクソン2、配列番号41；エクソン3、配列番号42；エクソン4、配列番号43；エクソン5、配列番号44；エクソン6、配列番号45；エクソン7、配列番号46；エクソン8、配列番号47；エクソン9、配列番号48；エクソン10、配列番号49；エクソン11、配列番号50；エクソン12、配列番号51；エクソン13、配列番号52；エクソン14、配列番号53；エクソン15、配列番号54；エクソン16、配列番号55；エクソン17、配列番号56；エクソン18、配列番号57；エクソン19、配列番号58；エクソン20、配列番号59；エクソン21、配列番号60；エクソン22、配列番号61およびエクソン23、配列番号62）。

【図4】図4はYWTDスパーサー、細胞外付着部位、LDLおよびカルシウムの結合部位、システイン富化成長因子リピート、膜貫通領域、CK-IIリン酸化部位を伴う理想的なPEST領域および内部移行ドメインを含むLRP5のドメイン構成を示す。図4はまたHBMタンパク質で生じるグリシンのバリンへの変化の部位を示す。シグナルペプチドはアミノ酸1～31に位置し、細胞外ドメインはアミノ酸32～1385に位置し、膜貫通セグメントはアミノ酸1386～1413に位置し、細胞質ドメインはアミノ酸1414～1615に位置する。

【図5】図5はHBM遺伝子に関連するBACコンティグB527D12およびB200E21の略図である。

【図6】図6A～6Jは野生型遺伝子LRP5のヌクレオチド（配列番号1）およびアミノ酸（配列番号3）配列を示す。ヌクレオチド582でのグアニンのチミンへの塩基対置換の位置（配列番号2、4）には下線を付している。この対立遺伝子変異体はHBM遺伝子である。HBM遺伝子は171位置でグリシンのバリンへのアミノ酸置換を有するタンパク質をコードする。5'非翻訳領域（UTR）は塩基1から70の境界であり、3'UTRは塩基4916～5120の境界である。

【図7】図7Aおよび7Bは種々組織におけるLRP5の発現を示すノーザンブロット分析である。

【図8】図8はPCR産物分析である。

【図9】図9はLRP5エクソン3変異の対立遺伝子特異的オリゴヌクレオチド検出である。

【図10】図10は100倍拡大した、センスおよびアンチセンスプローブを用いるインサイチュ・ハイブリダイゼーションによるマウスLRP5の細胞局在化である。

【図11】図11は400倍拡大した、センスおよびアンチセンスプローブを用いるインサイチュ・ハイブリダイゼーションによるマウスLRP5の細胞局在化である。

【図12】図12は400倍拡大した、センスおよびアンチセンスプローブを用いる、骨内膜での骨芽細胞のインサイチュ・ハイブリダイゼーションによるマウスLRP5の細胞局在化である。

【図13】図13はMC-3T3細胞でのLRP5発現のアンチセンス阻止を示す。

【図14】図14は、野生型LRP5対立遺伝子（左パネル、G特異的オリゴ；55℃、

洗浄)と比較したHBM対立遺伝子(右パネル、T特異的オリゴ; 58℃、洗浄)の希少性を説明するLRP5エクソン3対立遺伝子特異的オリゴヌクレオチド(ASO)アッセイを示す。右パネルで現れる陽性スポットは陽性対照である。

【図15】図15は接着斑情報伝達におけるLRP5(Zmax1)の潜在的役割を表すモデルを示す。

【図16】図16は内因性マウスLRP5のノックアウトまたはHBM多型性の条件的ノックインのための2つのLRP5ジーンターゲティングベクターを模式的に表す。B、X、およびRは各々DNA BAC4735P5におけるBamHI、XbaIおよびEcoRI部位を示す。エクソン3、4および5を黒色の長方形で示す。AG→T塩基変化はHBM多型性を作製するためにエクソン3の塩基24で操作される。ネオマイシン抵抗性遺伝子および合成休止配列を含有するLoxPフランキングされるカセットならびにES細胞クローンのスクリーニングおよび特徴付けに用いられるプローブの位置をも示す。 10

【図17】図17はトランスジェニック(すなわちHBMMCBAおよびHBMMTIC)および野生型(すなわちZmaxWT CBAおよびZmaxWT TIC)プラスミド構築物により発現を確認する。これらの構築物をHOB-02-02細胞に一過性にトランスフェクトし、mRNAレベルをTaqMan(登録商標)定量PCRを用いて決定した。HBMMCBAおよびZmaxWT CBAを表の左のコラムで示し(すなわちCMVβアクチン)ならびにHBMMTICおよびZmaxWT TICを表の右のコラムで示す(すなわちI型コラーゲン)。

【図18】図18はヒトおよびマウスTaqMan(登録商標)プライマー/プローブセット間の比較を示す。HOB(HOB-03-C5)およびマウス(MC-3T3-E1)骨芽細胞mRNAをプローブおよびプライマーを用いて分析した。 20

【図19】図19はTaqMan(登録商標)プライマー/プローブセットを用いて、ヒトおよびマウスRNA混合物バックグラウンドで発現されたヒトZmax-1 mRNAの定量を示す。結果は添加したヒトLRP5 mRNA(ng)対測定したヒトLRP5 mRNA(ng)を表す。

【図20】図20はTaqMan(登録商標)により分析されたmRNA発現に基づくトランスジェニックマウスにおけるHBMの発現を表す。

【図21】図21A~Cは脊椎(図21A)、大腿骨(図21B)および全身(図21C)においてHBMMCBA構築物を発現する種々トランスジェニックマウス株の分析を表す。図21D~Fは脊椎(図21D)、大腿骨(図21E)および全身(図21F)においてHBMMTIC構築物を発現する種々トランスジェニックマウス株の分析を表す。 30

図21G~Lは脊椎、大腿骨および全身において17週を通じてHBMMTIC構築物を発現するトランスジェニックマウス株の分析を表す。

【図22】図22はインビボpDXA分析*を用いて、5週でのHBMトランスジェニックマウスにおけるBMD(すなわちHBMMCBAおよびHBMMTIC構築物)の変化を表す。BMD変化を、これもまた5週齢のみである野生型動物と比較して提示する。

【図23】図23はインビボpDXA分析*を用いて、9週でのHBMトランスジェニックマウスにおけるBMD(すなわちHBMMCBAおよびHBMMTIC構築物)の変化を表す。BMD変化を、これもまた9週齢のみである野生型動物と比較して提示する。 40

【図24】図24(A~D)はHBMGI 2ASベクターインサートの配列(配列番号759)を提示する。

【図25】図25(A~D)はZMAXGI 3ASベクターインサートの配列(配列番号760)を提示する。

【図26】図26(A~C)はヒト(配列番号761)およびマウス(配列番号762)LRP5アミノ酸配列のアラインメントを提示する。

【図27】図27(A~C)はヒトLRP5(配列番号763)およびLRP6(配列番号764)アミノ酸配列のアラインメントを提示する。

【図28】図28はマウスの骨成長に及ぼす負荷の影響を試験するための装置を説明する。

【図 2 9】図 2 9 は骨負荷の H B M トランスジェニックおよび非トランスジェニックマウスの骨成長に及ぼす影響の組織形態学的説明を提示する。

【配列表】

SEQUENCE LISTING

<110> Babij, Philip
Yaworsky, Paul
Bex, Frederick J. III
Bodine, Peter Van Nest

<120> Transgenic Animal Model of Bone Mass Modulation

<130> 032796-133

<150> US 60/290,071

<151> 2001-05-11

<150> US 60/291,311

<151> 2001-05-17

<150> US 60/353,058

<151> 2002-02-01

<150> US 60/361,293

<151> 2002-03-04

<160> 765

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 5120

<212> DNA

<213> Homo sapiens

<400> 1

actaaagcgc	cgccgcgcgc	ccatggagcc	cgagtgcgcg	cggcgcgggc	ccgtccggcc	60
gccggacaac	atg gag gca gcg ccg ccc ggg ccg ccg tgg ccg ctg ctg					109
	Met Glu Ala Ala Pro Pro Gly Pro Pro Trp Pro Leu Leu					
	1 5 10					
ctg ctg ctg ctg ctg ctg ctg gcg ctg tgc ggc tgc ccg gcc ccc gcc						157
Leu Leu Leu Leu Leu Leu Leu Ala Leu Cys Gly Cys Pro Ala Pro Ala						
15 20 25						
gog gcc tcg ccg ctc ctg cta ttt gcc aac cgc cgg gac gta cgg ctg						205
Ala Ala Ser Pro Leu Leu Leu Phe Ala Asn Arg Arg Asp Val Arg Leu						
30 35 40 45						
gtg gac gcc ggc gga gtc aag ctg gag tcc acc atc gtg gtc agc ggc						253
Val Asp Ala Gly Gly Val Lys Leu Glu Ser Thr Ile Val Val Ser Gly						
50 55 60						
ctg gag gat gcg gcc gca gtg gac ttc cag ttt tcc aag gga gcc gtg						301
Leu Glu Asp Ala Ala Ala Val Asp Phe Gln Phe Ser Lys Gly Ala Val						
65 70 75						

10

20

30

tac tgg aca gac gtg agc gag gag gcc atc aag cag acc tac ctg aac 349
 Tyr Trp Thr Asp Val Ser Glu Glu Ala Ile Lys Gln Thr Tyr Leu Asn
 80 85 90
 cag acg ggg gcc gcc gtg cag aac gtg gtc atc tcc ggc ctg gtc tct 397
 Gln Thr Gly Ala Ala Val Gln Asn Val Val Ile Ser Gly Leu Val Ser
 95 100 105
 ccc gac ggc ctc gcc tgc gac tgg gtg ggc aag aag ctg tac tgg acg 445
 Pro Asp Gly Leu Ala Cys Asp Trp Val Gly Lys Lys Leu Tyr Trp Thr
 110 115 120 125
 gac tca gag acc aac cgc atc gag gtg gcc aac ctc aat ggc aca tcc 493
 Asp Ser Glu Thr Asn Arg Ile Glu Val Ala Asn Leu Asn Gly Thr Ser
 130 135 140
 cgg aag gtg ctc ttc tgg cag gac ctt gac cag ccg agg gcc atc gcc 541
 Arg Lys Val Leu Phe Trp Gln Asp Leu Asp Gln Pro Arg Ala Ile Ala
 145 150 155
 ttg gac ccc gct cac ggg tac atg tac tgg aca gac tgg ggt gag acg 589
 Leu Asp Pro Ala His Gly Tyr Met Tyr Trp Thr Asp Trp Gly Glu Thr
 160 165 170
 ccc cgg att gag cgg gca ggg atg gat ggc agc acc cgg aag atc att 637
 Pro Arg Ile Glu Arg Ala Gly Met Asp Gly Ser Thr Arg Lys Ile Ile
 175 180 185
 gtg gac tcg gac att tac tgg ccc aat gga ctg acc atc gac ctg gag 685
 Val Asp Ser Asp Ile Tyr Trp Pro Asn Gly Leu Thr Ile Asp Leu Glu
 190 195 200 205
 gag cag aag ctc tac tgg gct gac gcc aag ctc agc ttc atc cac cgt 733
 Glu Gln Lys Leu Tyr Trp Ala Asp Ala Lys Leu Ser Phe Ile His Arg
 210 215 220
 gcc aac ctg gac ggc tcg ttc cgg cag aag gtg gtg gag ggc agc ctg 781
 Ala Asn Leu Asp Gly Ser Phe Arg Gln Lys Val Val Glu Gly Ser Leu
 225 230 235
 acg cac ccc ttc gcc ctg acg ctc tcc ggg gac act ctg tac tgg aca 829
 Thr His Pro Phe Ala Leu Thr Leu Ser Gly Asp Thr Leu Tyr Trp Thr
 240 245 250
 gac tgg cag acc cgc tcc atc cat gcc tgc aac aag cgc act ggg ggg 877
 Asp Trp Gln Thr Arg Ser Ile His Ala Cys Asn Lys Arg Thr Gly Gly
 255 260 265
 aag agg aag gag atc ctg agt gcc ctc tac tca ccc atg gac atc cag 925
 Lys Arg Lys Glu Ile Leu Ser Ala Leu Tyr Ser Pro Met Asp Ile Gln
 270 275 280 285
 gtg ctg agc cag gag cgg cag cct ttc ttc cac act cgc tgt gag gag 973
 Val Leu Ser Gln Glu Arg Gln Pro Phe Phe His Thr Arg Cys Glu Glu
 290 295 300
 gac aat ggc ggc tgc tcc cac ctg tgc ctg ctg tcc cca agc gag cct 1021
 Asp Asn Gly Gly Cys Ser His Leu Cys Leu Leu Ser Pro Ser Glu Pro
 305 310 315
 ttc tac aca tgc gcc tgc ccc acg ggt gtg cag ctg cag gac aac ggc 1069
 Phe Tyr Thr Cys Ala Cys Pro Thr Gly Val Gln Leu Gln Asp Asn Gly
 320 325 330
 agg acg tgt aag gca gga gcc gag gag gtg ctg ctg ctg gcc cgg cgg 1117
 Arg Thr Cys Lys Ala Gly Ala Glu Glu Val Leu Leu Leu Ala Arg Arg

10

20

30

335	340	345	
acg gac cta cgg agg atc tcg ctg gac acg ccg gac ttc acc gac atc			1165
Thr Asp Leu Arg Arg Ile Ser Leu Asp Thr Pro Asp Phe Thr Asp Ile			
350	355	360	365
gtg ctg cag gtg gac gac atc cgg cac gcc att gcc atc gac tac gac			1213
Val Leu Gln Val Asp Asp Ile Arg His Ala Ile Ala Ile Asp Tyr Asp			
370	375	380	
ccg cta gag ggc tat gtc tac tgg aca gat gac gag gtg cgg gcc atc			1261
Pro Leu Glu Gly Tyr Val Tyr Trp Thr Asp Asp Glu Val Arg Ala Ile			
385	390	395	
cgc agg gcg tac ctg gac ggg tct ggg gcg cag acg ctg gtc aac acc			1309
Arg Arg Ala Tyr Leu Asp Gly Ser Gly Ala Gln Thr Leu Val Asn Thr			
400	405	410	
gag atc aac gac ccc gat ggc atc gcg gtc gac tgg gtg gcc cga aac			1357
Glu Ile Asn Asp Pro Asp Gly Ile Ala Val Asp Trp Val Ala Arg Asn			
415	420	425	
ctc tac tgg acc gac acg ggc acg gac cgc atc gag gtg acg cgc ctc			1405
Leu Tyr Trp Thr Asp Thr Gly Thr Asp Arg Ile Glu Val Thr Arg Leu			
430	435	440	445
aac ggc acc tcc cgc aag atc ctg gtg tgc gag gac ctg gac gag ccc			1453
Asn Gly Thr Ser Arg Lys Ile Leu Val Ser Glu Asp Leu Asp Glu Pro			
450	455	460	
cga gcc atc gca ctg cac ccc gtg atg ggc ctc atg tac tgg aca gac			1501
Arg Ala Ile Ala Leu His Pro Val Met Gly Leu Met Tyr Trp Thr Asp			
465	470	475	
tgg gga gag aac cct aaa atc gag tgt gcc aac ttg gat ggg cag gag			1549
Trp Gly Glu Asn Pro Lys Ile Glu Cys Ala Asn Leu Asp Gly Gln Glu			
480	485	490	
cgg cgt gtg ctg gtc aat gcc tcc ctc ggg tgg ccc aac ggc ctg gcc			1597
Arg Arg Val Leu Val Asn Ala Ser Leu Gly Trp Pro Asn Gly Leu Ala			
495	500	505	
ctg gac ctg cag gag ggg aag ctc tac tgg gga gac gcc aag aca gac			1645
Leu Asp Leu Gln Glu Gly Lys Leu Tyr Trp Gly Asp Ala Lys Thr Asp			
510	515	520	525
aag atc gag gtg atc aat gtt gat ggg acg aag agg cgg acc ctc ctg			1693
Lys Ile Glu Val Ile Asn Val Asp Gly Thr Lys Arg Arg Thr Leu Leu			
530	535	540	
gag gac aag ctc ccg cac att ttc ggg ttc acg ctg ctg ggg gac ttc			1741
Glu Asp Lys Leu Pro His Ile Phe Gly Phe Thr Leu Leu Gly Asp Phe			
545	550	555	
atc tac tgg act gac tgg cag cgc cgc agc atc gag cgg gtg cac aag			1789
Ile Tyr Trp Thr Asp Trp Gln Arg Arg Ser Ile Glu Arg Val His Lys			
560	565	570	
gtc aag gcc agc cgg gac gtc atc att gac cag ctg ccc gac ctg atg			1837
Val Lys Ala Ser Arg Asp Val Ile Ile Asp Gln Leu Pro Asp Leu Met			
575	580	585	
ggg ctc aaa gct gtg aat gtg gcc aag gtc gtc gga acc aac ccg tgt			1885
Gly Leu Lys Ala Val Asn Val Ala Lys Val Val Gly Thr Asn Pro Cys			
590	595	600	605
gcg gac agg aac ggg ggg tgc agc cac ctg tgc ttc ttc aca ccc cac			1933

10

20

30

Ala	Asp	Arg	Asn	Gly	Gly	Cys	Ser	His	Leu	Cys	Phe	Phe	Thr	Pro	His		
				610					615					620			
gca	acc	egg	tgt	ggc	tgc	ccc	atc	ggc	ctg	gag	ctg	ctg	agt	gac	atg	1981	
Ala	Thr	Arg	Cys	Gly	Cys	Pro	Ile	Gly	Leu	Glu	Leu	Leu	Ser	Asp	Met		
				625				630						635			
aag	acc	tgc	atc	gtg	cct	gag	gcc	ttc	ttg	gtc	ttc	acc	agc	aga	gcc	2029	
Lys	Thr	Cys	Ile	Val	Pro	Glu	Ala	Phe	Leu	Val	Phe	Thr	Ser	Arg	Ala		
				640				645						650			
gcc	atc	cac	agg	atc	tcc	ctc	gag	acc	aat	aac	aac	gac	gtg	gcc	atc	2077	
Ala	Ile	His	Arg	Ile	Ser	Leu	Glu	Thr	Asn	Asn	Asp	Val	Ala	Ile			
				655				660						665			
ccg	ctc	acg	ggc	gtc	aag	gag	gcc	tca	gcc	ctg	gac	ttt	gat	gtg	tcc	2125	
Pro	Leu	Thr	Gly	Val	Lys	Glu	Ala	Ser	Ala	Leu	Asp	Phe	Asp	Val	Ser		
				670				675						680			
aac	aac	cac	atc	tac	tgg	aca	gac	gtc	agc	ctg	aag	acc	atc	agc	cgc	2173	
Asn	Asn	His	Ile	Tyr	Trp	Thr	Asp	Val	Ser	Leu	Lys	Thr	Ile	Ser	Arg		
				690				695						700			
gcc	ttc	atg	aac	ggg	agc	tgc	gtg	gag	cac	gtg	gtg	gag	ttt	ggc	ctt	2221	
Ala	Phe	Met	Asn	Gly	Ser	Ser	Val	Glu	His	Val	Val	Glu	Phe	Gly	Leu		
				705				710						715			
gac	tac	ccc	gag	ggc	atg	gcc	gtt	gac	tgg	atg	ggc	aag	aac	ctc	tac	2269	
Asp	Tyr	Pro	Glu	Gly	Met	Ala	Val	Asp	Trp	Met	Gly	Lys	Asn	Leu	Tyr		
				720				725						730			
tgg	gcc	gac	act	ggg	acc	aac	aga	atc	gaa	gtg	gag	cgg	ctg	gac	ggg	2317	
Trp	Ala	Asp	Thr	Gly	Thr	Asn	Arg	Ile	Glu	Val	Ala	Arg	Leu	Asp	Gly		
				735				740						745			
cag	ttc	egg	caa	gtc	ctc	gtg	tgg	agg	gac	ttg	gac	aac	ccg	agg	tgc	2365	
Gln	Phe	Arg	Gln	Val	Leu	Val	Trp	Arg	Asp	Leu	Asp	Asn	Pro	Arg	Ser		
				750				755						760			
ctg	gcc	ctg	gat	ccc	acc	aag	ggc	tac	atc	tac	tgg	acc	gag	tgg	ggc	2413	
Leu	Ala	Leu	Asp	Pro	Thr	Lys	Gly	Tyr	Ile	Tyr	Trp	Thr	Glu	Trp	Gly		
				770				775						780			
ggc	aag	ccg	agg	atc	gtg	cgg	gcc	ttc	atg	gac	ggg	acc	aac	tgc	atg	2461	
Gly	Lys	Pro	Arg	Ile	Val	Arg	Ala	Phe	Met	Asp	Gly	Thr	Asn	Cys	Met		
				785				790						795			
acg	ctg	gtg	gac	aag	gtg	ggc	cgg	gcc	aac	gac	ctc	acc	att	gac	tac	2509	
Thr	Leu	Val	Asp	Lys	Val	Gly	Arg	Ala	Asn	Asp	Leu	Thr	Ile	Asp	Tyr		
				800				805						810			
gct	gac	cag	cgc	ctc	tac	tgg	acc	gac	ctg	gac	acc	aac	atg	atc	gag	2557	
Ala	Asp	Gln	Arg	Leu	Tyr	Trp	Thr	Asp	Leu	Asp	Thr	Asn	Met	Ile	Glu		
				815				820						825			
tgc	tcc	aac	atg	ctg	ggt	cag	gag	cgg	gtc	gtg	att	gcc	gac	gat	ctc	2605	
Ser	Ser	Asn	Met	Leu	Gly	Gln	Glu	Arg	Val	Val	Ile	Ala	Asp	Asp	Leu		
				830				835						840			
ccg	cac	ccg	ttc	ggt	ctg	acg	cag	tac	agc	gat	tat	atc	tac	tgg	aca	2653	
Pro	His	Pro	Phe	Gly	Leu	Thr	Gln	Tyr	Ser	Asp	Tyr	Ile	Tyr	Trp	Thr		
				850				855						860			
gac	tgg	aat	ctg	cac	agc	att	gag	cgg	gcc	gac	aag	act	agc	ggc	cgg	2701	
Asp	Trp	Asn	Leu	His	Ser	Ile	Glu	Arg	Ala	Asp	Lys	Thr	Ser	Gly	Arg		
				865				870						875			

10

20

30

aac cgc acc ctc atc cag ggc cac ctg gac ttc gtg atg gac atc ctg 2749
 Asn Arg Thr Leu Ile Gln Gly His Leu Asp Phe Val Met Asp Ile Leu
 880 885 890
 gtg ttc cac tcc tcc cgc cag gat ggc ctc aat gac tgt atg cac aac 2797
 Val Phe His Ser Ser Arg Gln Asp Gly Leu Asn Asp Cys Met His Asn
 895 900 905
 aac ggg cag tgt ggg cag ctg tgc ctt gcc atc ccc ggc ggc cac cgc 2845
 Asn Gly Gln Cys Gly Gln Leu Cys Leu Ala Ile Pro Gly Gly His Arg
 910 915 920 925
 tgc ggc tgc gcc tca cac tac acc ctg gac ccc agc agc cgc aac tgc 2893
 Cys Gly Cys Ala Ser His Tyr Thr Leu Asp Pro Ser Ser Arg Asn Cys
 930 935 940
 agc ccg ccc acc acc ttc ttg ctg ttc agc cag aaa tct gcc atc agt 2941
 Ser Pro Pro Thr Thr Phe Leu Leu Phe Ser Gln Lys Ser Ala Ile Ser
 945 950 955
 cgg atg atc ccg gac gac cag cac agc ccg gat ctc atc ctg ccc ctg 2989
 Arg Met Ile Pro Asp Asp Gln His Ser Pro Asp Leu Ile Leu Pro Leu
 960 965 970
 cat gga ctg agg aac gtc aaa gcc atc gac tat gac cca ctg gac aag 3037
 His Gly Leu Arg Asn Val Lys Ala Ile Asp Tyr Asp Pro Leu Asp Lys
 975 980 985
 ttc atc tac tgg gtg gat ggg cgc cag aac atc aag cga gcc aag gac 3085
 Phe Ile Tyr Trp Val Asp Gly Arg Gln Asn Ile Lys Arg Ala Lys Asp
 990 995 1000 1005
 gac ggg acc cag ccc ttt gtt ttg acc tct ctg agc caa ggc caa aac 3133
 Asp Gly Thr Gln Pro Phe Val Leu Thr Ser Leu Ser Gln Gly Gln Asn
 1010 1015 1020
 cca gac agg cag ccc cac gac ctc agc atc gac atc tac agc cgg aca 3181
 Pro Asp Arg Gln Pro His Asp Leu Ser Ile Asp Ile Tyr Ser Arg Thr
 1025 1030 1035
 ctg ttc tgg acg tgc gag gcc acc aat acc atc aac gtc cac agg ctg 3229
 Leu Phe Trp Thr Cys Glu Ala Thr Asn Thr Ile Asn Val His Arg Leu
 1040 1045 1050
 agc ggg gaa gcc atg ggg gtg gtg ctg cgt ggg gac cgc gac aag ccc 3277
 Ser Gly Glu Ala Met Gly Val Val Leu Arg Gly Asp Arg Asp Lys Pro
 1055 1060 1065
 agg gcc atc gtc gtc aac gcg gag cga ggg tac ctg tac ttc acc aac 3325
 Arg Ala Ile Val Val Asn Ala Glu Arg Gly Tyr Leu Tyr Phe Thr Asn
 1070 1075 1080 1085
 atg cag gac cgg gca gcc aag atc gaa cgc gca gcc ctg gac ggc acc 3373
 Met Gln Asp Arg Ala Ala Lys Ile Glu Arg Ala Ala Leu Asp Gly Thr
 1090 1095 1100
 gag cgc gag gtc ctc ttc acc acc ggc ctc atc cgc cct gtg gcc ctg 3421
 Glu Arg Glu Val Leu Phe Thr Thr Gly Leu Ile Arg Pro Val Ala Leu
 1105 1110 1115
 gtg gtg gac aac aca ctg ggc aag ctg ttc tgg gtg gac gcg gac ctg 3469
 Val Val Asp Asn Thr Leu Gly Lys Leu Phe Trp Val Asp Ala Asp Leu
 1120 1125 1130
 aag cgc att gag agc tgt gac ctg tca ggg gcc aac cgc ctg acc ctg 3517
 Lys Arg Ile Glu Ser Cys Asp Leu Ser Gly Ala Asn Arg Leu Thr Leu

10

20

30

1135	1140	1145	
gag gac gcc aac atc gtg cag oct ctg ggc ctg acc atc ctt ggc aag			3565
Glu Asp Ala Asn Ile Val Gln Pro Leu Gly Leu Thr Ile Leu Gly Lys			
1150	1155	1160	1165
cat ctc tac tgg atc gac cgc cag cag cag atg atc gag cgt gtg gag			3613
His Leu Tyr Trp Ile Asp Arg Gln Gln Gln Met Ile Glu Arg Val Glu			
1170	1175	1180	
aag acc acc ggg gac aag cgg act cgc atc cag ggc cgt gtc gcc cac			3661
Lys Thr Thr Gly Asp Lys Arg Thr Arg Ile Gln Gly Arg Val Ala His			
1185	1190	1195	
ctc act ggc atc cat gca gtg gag gaa gtc agc ctg gag gag ttc tca			3709
Leu Thr Gly Ile His Ala Val Glu Glu Val Ser Leu Glu Glu Phe Ser			
1200	1205	1210	
gcc cac cca tgt gcc cgt gac aat ggt ggc tgc tcc cac atc tgt att			3757
Ala His Pro Cys Ala Arg Asp Asn Gly Gly Cys Ser His Ile Cys Ile			
1215	1220	1225	
gcc aag ggt gat ggg aca cca cgg tgc tca tgc cca gtc cac ctc gtg			3805
Ala Lys Gly Asp Gly Thr Pro Arg Cys Ser Cys Pro Val His Leu Val			
1230	1235	1240	1245
ctc ctg cag aac ctg ctg acc tgt gga gag ccg ccc acc tgc tcc ccg			3853
Leu Leu Gln Asn Leu Leu Thr Cys Gly Glu Pro Pro Thr Cys Ser Pro			
1250	1255	1260	
gac cag ttt gca tgt gcc aca ggg gag atc gac tgt atc ccc ggg gcc			3901
Asp Gln Phe Ala Cys Ala Thr Gly Glu Ile Asp Cys Ile Pro Gly Ala			
1265	1270	1275	
tgg cgc tgt gac ggc ttt ccc gag tgc gat gac cag agc gac gag gag			3949
Trp Arg Cys Asp Gly Phe Pro Glu Cys Asp Asp Gln Ser Asp Glu Glu			
1280	1285	1290	
ggc tgc ccc gtg tgc tcc gcc gcc cag ttc ccc tgc gcg cgg ggt cag			3997
Gly Cys Pro Val Cys Ser Ala Ala Gln Phe Pro Cys Ala Arg Gly Gln			
1295	1300	1305	
tgt gtg gac ctg cgc ctg cgc tgc gac ggc gag gca gac tgt cag gac			4045
Cys Val Asp Leu Arg Cys Asp Gly Glu Ala Asp Cys Gln Asp			
1310	1315	1320	1325
cgc tca gac gag gtg gac tgt gac gcc atc tgc ctg ccc aac cag ttc			4093
Arg Ser Asp Glu Val Asp Cys Asp Ala Ile Cys Leu Pro Asn Gln Phe			
1330	1335	1340	
cgg tgt gcg agc ggc cag tgt gtc ctc atc aaa cag cag tgc gac tcc			4141
Arg Cys Ala Ser Gly Gln Cys Val Leu Ile Lys Gln Gln Cys Asp Ser			
1345	1350	1355	
ttc ccc gac tgt atc gac ggc tcc gac gag ctc atg tgt gaa atc acc			4189
Phe Pro Asp Cys Ile Asp Gly Ser Asp Glu Leu Met Cys Glu Ile Thr			
1360	1365	1370	
aag ccg ccc tca gac gac agc ccg gcc cac agc agt gcc atc ggg ccc			4237
Lys Pro Pro Ser Asp Asp Ser Pro Ala His Ser Ser Ala Ile Gly Pro			
1375	1380	1385	
gtc att ggc atc atc ctc tct ctc ttc gtc atg ggt ggt gtc tat ttt			4285
Val Ile Gly Ile Ile Leu Ser Leu Phe Val Met Gly Gly Val Tyr Phe			
1390	1395	1400	1405
gtg tgc cag cgc gtg gtg tgc cag cgc tat gcg ggg gcc aac ggg ccc			4333

10

20

30

```

Val Cys Gln Arg Val Val Cys Gln Arg Tyr Ala Gly Ala Asn Gly Pro
1410 1415 1420
ttc ccg cac gag tat gtc agc ggg acc ccg cac gtg ccc ctc aat ttc 4381
Phe Pro His Glu Tyr Val Ser Gly Thr Pro His Val Pro Leu Asn Phe
1425 1430 1435
ata gcc ccg ggc ggt tcc cag cat ggc ccc ttc aca ggc atc gca tgc 4429
Ile Ala Pro Gly Gly Ser Gln His Gly Pro Phe Thr Gly Ile Ala Cys
1440 1445 1450
gga aag tcc atg atg agc tcc gtg agc ctg atg ggg ggc cgg ggc ggg 4477
Gly Lys Ser Met Met Ser Ser Val Ser Leu Met Gly Gly Arg Gly Gly
1455 1460 1465
gtg ccc ctc tac gac cgg aac cac gtc aca ggg gcc tcg tcc agc agc 4525
Val Pro Leu Tyr Asp Arg Asn His Val Thr Gly Ala Ser Ser Ser Ser
1470 1475 1480 1485
tcg tcc agc acg aag gcc acg ctg tac ccg ccg atc ctg aac ccg ccg 4573
Ser Ser Ser Thr Lys Ala Thr Leu Tyr Pro Pro Ile Leu Asn Pro Pro
1490 1495 1500
ccc tcc ccg gcc acg gac ccc tcc ctg tac aac atg gac atg ttc tac 4621
Pro Ser Pro Ala Thr Asp Pro Ser Leu Tyr Asn Met Asp Met Phe Tyr
1505 1510 1515
tct tca aac att ccg gcc act gcg aga ccg tac agg ccc tac atc att 4669
Ser Ser Asn Ile Pro Ala Thr Ala Arg Pro Tyr Arg Pro Tyr Ile Ile
1520 1525 1530
cga gga atg gcg ccc ccg acg acg ccc tgc agc acc gac gtg tgt gac 4717
Arg Gly Met Ala Pro Pro Thr Thr Pro Cys Ser Thr Asp Val Cys Asp
1535 1540 1545
agc gac tac agc gcc agc cgc tgg aag gcc agc aag tac tac ctg gat 4765
Ser Asp Tyr Ser Ala Ser Arg Trp Lys Ala Ser Lys Tyr Tyr Leu Asp
1550 1555 1560 1565
ttg aac tcg gac tca gac ccc tat cca ccc cca ccc acg ccc cac agc 4813
Leu Asn Ser Asp Ser Asp Pro Tyr Pro Pro Pro Thr Pro His Ser
1570 1575 1580
cag tac ctg tcg gcg gag gac agc tgc ccg ccc tcg ccc gcc acc gag 4861
Gln Tyr Leu Ser Ala Glu Asp Ser Cys Pro Pro Ser Pro Ala Thr Glu
1585 1590 1595
agg agc tac ttc cat ctc ttc ccg ccc cct ccg tcc ccc tgc acg gac 4909
Arg Ser Tyr Phe His Leu Phe Pro Pro Pro Pro Ser Pro Cys Thr Asp
1600 1605 1610
tca tcc tgacctcggc cgggccactc tggcttctct gtgcccctgt aaatagtttt 4965
Ser Ser
1615
aaatatgaac aaagaaaaaa atatatttta tgatttaaaa aataaatata attgggattt 5025
taaaaacatg agaaatgtga actgtgatgg ggtgggcagg gctgggagaa ctttgtacag 5085
tggagaaata tttataaaact taattttgta aaaca 5120

```

10

20

30

```

<210> 2
<211> 5120
<212> DNA
<213> Homo sapiens

```

<400> 2

actaaagcgc	cgccgcccgcg	ccatggagcc	cgagtgcgcg	cgccgcgggc	ccgtccggcc	60
gccggacaac	atg gag gca gcg ccg ccc ggg ccg ccg tgg ccg ctg ctg	109				
	Met Glu Ala Ala Pro Pro Gly Pro Pro Trp Pro Leu Leu					
1	5	10				
ctg ctg ctg ctg ctg ctg ctg gcg ctg tgc ggc tgc ccg gcc ccc gcc	157					
Leu Leu Leu Leu Leu Leu Leu Ala Leu Cys Gly Cys Pro Ala Pro Ala						
15	20	25				
gcg gcc tcg ccg ctc ctg cta ttt gcc aac cgc cgg gac gta cgg ctg	205					
Ala Ala Ser Pro Leu Leu Phe Ala Asn Arg Arg Asp Val Arg Leu						
30	35	40	45			
gtg gac gcc ggc gga gtc aag ctg gag tcc acc atc gtg gtc agc ggc	253					
Val Asp Ala Gly Gly Val Lys Leu Glu Ser Thr Ile Val Val Ser Gly						
50	55	60				
ctg gag gat gcg gcc gca gtg gac ttc cag ttt tcc aag gga gcc gtg	301					
Leu Glu Asp Ala Ala Val Asp Phe Gln Phe Ser Lys Gly Ala Val						
65	70	75				
tac tgg aca gac gtg agc gag gag gcc atc aag cag acc tac ctg aac	349					
Tyr Trp Thr Asp Val Ser Glu Glu Ala Ile Lys Gln Thr Tyr Leu Asn						
80	85	90				
cag acg ggg gcc gcc gtg cag aac gtg gtc atc tcc ggc ctg gtc tct	397					
Gln Thr Gly Ala Ala Val Gln Asn Val Val Ile Ser Gly Leu Val Ser						
95	100	105				
ccc gac ggc ctc gcc tgc gac tgg gtg ggc aag aag ctg tac tgg acg	445					
Pro Asp Gly Leu Ala Cys Asp Trp Val Gly Lys Lys Leu Tyr Trp Thr						
110	115	120	125			
gac tca gag acc aac cgc atc gag gtg gcc aac ctc aat ggc aca tcc	493					
Asp Ser Glu Thr Asn Arg Ile Glu Val Ala Asn Leu Asn Gly Thr Ser						
130	135	140				
cgg aag gtg ctc ttc tgg cag gac ctt gac cag ccg agg gcc atc gcc	541					
Arg Lys Val Leu Phe Trp Gln Asp Leu Asp Gln Pro Arg Ala Ile Ala						
145	150	155				
ttg gac ccc gct cac ggg tac atg tac tgg aca gac tgg gtt gag acg	589					
Leu Asp Pro Ala His Gly Tyr Met Tyr Trp Thr Asp Trp Val Glu Thr						
160	165	170				
ccc cgg att gag cgg gca ggg atg gat ggc agc acc cgg aag atc att	637					
Pro Arg Ile Glu Arg Ala Gly Met Asp Gly Ser Thr Arg Lys Ile Ile						
175	180	185				
gtg gac tcg gac att tac tgg ccc aat gga ctg acc atc gac ctg gag	685					
Val Asp Ser Asp Ile Tyr Trp Pro Asn Gly Leu Thr Ile Asp Leu Glu						
190	195	200	205			
gag cag aag ctc tac tgg gct gac gcc aag ctc agc ttc atc cac cgt	733					
Glu Gln Lys Leu Tyr Trp Ala Asp Ala Lys Leu Ser Phe Ile His Arg						
210	215	220				
gcc aac ctg gac ggc tcg ttc cgg cag aag gtg gtg gag ggc agc ctg	781					
Ala Asn Leu Asp Gly Ser Phe Arg Gln Lys Val Val Glu Gly Ser Leu						
225	230	235				
acg cac ccc ttc gcc ctg acg ctc tcc ggg gac act ctg tac tgg aca	829					
Thr His Pro Phe Ala Leu Thr Leu Ser Gly Asp Thr Leu Tyr Trp Thr						
240	245	250				

10

20

30

gac tgg cag acc cgc tcc atc cat gcc tgc aac aag cgc act ggg ggg	877
Asp Trp Gln Thr Arg Ser Ile His Ala Cys Asn Lys Arg Thr Gly Gly	
255 260 265	
aag agg aag gag atc ctg agt gcc ctc tac tca ccc atg gac atc cag	925
Lys Arg Lys Glu Ile Leu Ser Ala Leu Tyr Ser Pro Met Asp Ile Gln	
270 275 280 285	
gtg ctg agc cag gag cgg cag cct ttc ttc cac act cgc tgt gag gag	973
Val Leu Ser Gln Glu Arg Gln Pro Phe Phe His Thr Arg Cys Glu Glu	
290 295 300	
gac aat ggc ggc tgc tcc cac ctg tgc ctg ctg tcc cca agc gag cct	1021
Asp Asn Gly Gly Cys Ser His Leu Cys Leu Leu Ser Pro Ser Glu Pro	
305 310 315	
ttc tac aca tgc gcc tgc ccc acg ggt gtg cag ctg cag gac aac ggc	1069
Phe Tyr Thr Cys Ala Cys Pro Thr Gly Val Gln Leu Gln Asp Asn Gly	
320 325 330	
agg acg tgt aag gca gga gcc gag gag gtg ctg ctg ggc cgg cgg	1117
Arg Thr Cys Lys Ala Gly Glu Glu Val Leu Leu Leu Ala Arg Arg	
335 340 345	
acg gac cta cgg agg atc tcg ctg gac acg ccg gac ttc acc gac atc	1165
Thr Asp Leu Arg Arg Ile Ser Leu Asp Thr Pro Asp Phe Thr Asp Ile	
350 355 360 365	
gtg ctg cag gtg gac gac atc cgg cac gcc att gcc atc gac tac gac	1213
Val Leu Gln Val Asp Asp Ile Arg His Ala Ile Ala Ile Asp Tyr Asp	
370 375 380	
ccg cta gag ggc tat gtc tac tgg aca gat gac gag gtg cgg gcc atc	1261
Pro Leu Glu Gly Tyr Val Tyr Trp Thr Asp Asp Glu Val Arg Ala Ile	
385 390 395	
cgc agg gcg tac ctg gac ggg tct ggg gcg cag acg ctg gtc aac acc	1309
Arg Arg Ala Tyr Leu Asp Gly Ser Gly Ala Gln Thr Leu Val Asn Thr	
400 405 410	
gag atc aac gac ccc gat ggc atc gcg gtc gac tgg gtg gcc cga aac	1357
Glu Ile Asn Asp Pro Asp Gly Ile Ala Val Asp Trp Val Ala Arg Asn	
415 420 425	
ctc tac tgg acc gac acg ggc acg gac cgc atc gag gtg acg cgc ctc	1405
Leu Tyr Trp Thr Asp Thr Gly Thr Asp Arg Ile Glu Val Thr Arg Leu	
430 435 440 445	
aac ggc acc tcc cgc aag atc ctg gtg tcg gag gac ctg gac gag ccc	1453
Asn Gly Thr Ser Arg Lys Ile Leu Val Ser Glu Asp Leu Asp Glu Pro	
450 455 460	
cga gcc atc gca ctg cac ccc gtg atg ggc ctc atg tac tgg aca gac	1501
Arg Ala Ile Ala Leu His Pro Val Met Gly Leu Met Tyr Trp Thr Asp	
465 470 475	
tgg gga gag aac cct aaa atc gag tgt gcc aac ttg gat ggg cag gag	1549
Trp Gly Glu Asn Pro Lys Ile Glu Cys Ala Asn Leu Asp Gly Gln Glu	
480 485 490	
cgg cgt gtg ctg gtc aat gcc tcc ctc ggg tgg ccc aac ggc ctg gcc	1597
Arg Arg Val Leu Val Asn Ala Ser Leu Gly Trp Pro Asn Gly Leu Ala	
495 500 505	
ctg gac ctg cag gag ggg aag ctc tac tgg gga gac gcc aag aca gac	1645
Leu Asp Leu Gln Glu Gly Lys Leu Tyr Trp Gly Asp Ala Lys Thr Asp	

10

20

30

510	515	520	525	
aag atc gag gtg atc aat gtt gat ggg acg aag agg cgg acc ctc ctg				1693
Lys Ile Glu Val Ile Asn Val Asp Gly Thr Lys Arg Arg Thr Leu Leu				
	530	535	540	
gag gac aag ctc cgg cac att ttc ggg ttc acg ctg ctg ggg gac ttc				1741
Glu Asp Lys Leu Pro His Ile Phe Gly Phe Thr Leu Leu Gly Asp Phe				
	545	550	555	
atc tac tgg act gac tgg cag cgc cgc agc atc gag cgg gtg cac aag				1789
Ile Tyr Trp Thr Asp Trp Gln Arg Arg Ser Ile Glu Arg Val His Lys				
	560	565	570	
gtc aag gcc agc cgg gac gtc atc att gac cag ctg ccc gac ctg atg				1837
Val Lys Ala Ser Arg Asp Val Ile Ile Asp Gln Leu Pro Asp Leu Met				
	575	580	585	
ggg ctc aaa gct gtg aat gtg gcc aag gtc gtc gga acc aac cgg tgt				1885
Gly Leu Lys Ala Val Asn Val Ala Lys Val Val Gly Thr Asn Pro Cys				
	590	595	600	605
gcg gac agg aac ggg ggg tgc agc cac ctg tgc ttc ttc aca ccc cac				1933
Ala Asp Arg Asn Gly Gly Cys Ser His Leu Cys Phe Phe Thr Pro His				
	610	615	620	
gca acc cgg tgt ggc tgc ccc atc ggc ctg gag ctg ctg agt gac atg				1981
Ala Thr Arg Cys Gly Cys Pro Ile Gly Leu Glu Leu Leu Ser Asp Met				
	625	630	635	
aag acc tgc atc gtg cct gag gcc ttc ttg gtc ttc acc agc aga gcc				2029
Lys Thr Cys Ile Val Pro Glu Ala Phe Leu Val Phe Thr Ser Arg Ala				
	640	645	650	
gcc atc cac agg atc tcc ctc gag acc aat aac aac gac gtg gcc atc				2077
Ala Ile His Arg Ile Ser Leu Glu Thr Asn Asn Asp Val Ala Ile				
	655	660	665	
ccg ctc acg ggc gtc aag gag gcc tca gcc ctg gac ttt gat gtg tcc				2125
Pro Leu Thr Gly Val Lys Glu Ala Ser Ala Leu Asp Phe Asp Val Ser				
	670	675	680	685
aac aac cac atc tac tgg aca gac gtc agc ctg aag acc atc agc cgc				2173
Asn Asn His Ile Tyr Trp Thr Asp Val Ser Leu Lys Thr Ile Ser Arg				
	690	695	700	
gcc ttc atg aac ggg agc tgc gtg gag cac gtg gtg gag ttt ggc ctt				2221
Ala Phe Met Asn Gly Ser Ser Val Glu His Val Val Glu Phe Gly Leu				
	705	710	715	
gac tac ccc gag ggc atg gcc gtt gac tgg atg ggc aag aac ctc tac				2269
Asp Tyr Pro Glu Gly Met Ala Val Asp Trp Met Gly Lys Asn Leu Tyr				
	720	725	730	
tgg gcc gac act ggg acc aac aga atc gaa gtg gcg cgg ctg gac ggg				2317
Trp Ala Asp Thr Gly Thr Asn Arg Ile Glu Val Ala Arg Leu Asp Gly				
	735	740	745	
cag ttc cgg caa gtc ctc gtg tgg agg gac ttg gac aac ccg agg tgc				2365
Gln Phe Arg Gln Val Leu Val Trp Arg Asp Leu Asp Asn Pro Arg Ser				
	750	755	760	765
ctg gcc ctg gat ccc acc aag ggc tac atc tac tgg acc gag tgg ggc				2413
Leu Ala Leu Asp Pro Thr Lys Gly Tyr Ile Tyr Trp Thr Glu Trp Gly				
	770	775	780	
ggc aag ccg agg atc gtg cgg gcc ttc atg gac ggg acc aac tgc atg				2461

10

20

30

Gly	Lys	Pro	Arg	Ile	Val	Arg	Ala	Phe	Met	Asp	Gly	Thr	Asn	Cys	Met	
			785					790					795			
acg	ctg	gtg	gac	aag	gtg	ggc	cgg	gcc	aac	gac	ctc	acc	att	gac	tac	2509
Thr	Leu	Val	Asp	Lys	Val	Gly	Arg	Ala	Asn	Asp	Leu	Thr	Ile	Asp	Tyr	
		800					805					810				
gct	gac	cag	cgc	ctc	tac	tgg	acc	gac	ctg	gac	acc	aac	atg	atc	gag	2557
Ala	Asp	Gln	Arg	Leu	Tyr	Trp	Thr	Asp	Leu	Asp	Thr	Asn	Met	Ile	Glu	
		815				820					825					
tcg	tcc	aac	atg	ctg	ggc	cag	gag	cgg	gtc	gtg	att	gcc	gac	gat	ctc	2605
Ser	Ser	Asn	Met	Leu	Gly	Gln	Glu	Arg	Val	Val	Ile	Ala	Asp	Asp	Leu	
		830				835				840					845	
ccg	cac	ccg	ttc	ggc	ctg	acg	cag	tac	agc	gat	tat	atc	tac	tgg	aca	2653
Pro	His	Pro	Phe	Gly	Leu	Thr	Gln	Tyr	Ser	Asp	Tyr	Ile	Tyr	Trp	Thr	
				850					855					860		
gac	tgg	aat	ctg	cac	agc	att	gag	cgg	gcc	gac	aag	act	agc	ggc	cgg	2701
Asp	Trp	Asn	Leu	His	Ser	Ile	Glu	Arg	Ala	Asp	Lys	Thr	Ser	Gly	Arg	
			865						870					875		
aac	cgc	acc	ctc	atc	cag	ggc	cac	ctg	gac	ttc	gtg	atg	gac	atc	ctg	2749
Asn	Arg	Thr	Leu	Ile	Gln	Gly	His	Leu	Asp	Phe	Val	Met	Asp	Ile	Leu	
		880					885					890				
gtg	ttc	cac	tcc	tcc	cgc	cag	gat	ggc	ctc	aat	gac	tgt	atg	cac	aac	2797
Val	Phe	His	Ser	Ser	Arg	Gln	Asp	Gly	Leu	Asn	Asp	Cys	Met	His	Asn	
		895				900					905					
aac	ggg	cag	tgt	ggg	cag	ctg	tgc	ctt	gcc	atc	ccc	ggc	ggc	cac	cgc	2845
Asn	Gly	Gln	Cys	Gly	Gln	Leu	Cys	Leu	Ala	Ile	Pro	Gly	Gly	His	Arg	
		910			915				920					925		
tgc	ggc	tgc	gcc	tca	cac	tac	acc	ctg	gac	ccc	agc	agc	cgc	aac	tgc	2893
Cys	Gly	Cys	Ala	Ser	His	Tyr	Thr	Leu	Asp	Pro	Ser	Ser	Arg	Asn	Cys	
				930					935					940		
agc	ccg	ccc	acc	acc	ttc	ttg	ctg	ttc	agc	cag	aaa	tct	gcc	atc	agt	2941
Ser	Pro	Pro	Thr	Thr	Phe	Leu	Leu	Phe	Ser	Gln	Lys	Ser	Ala	Ile	Ser	
				945					950					955		
cgg	atg	atc	ccg	gac	gac	cag	cac	agc	ccg	gat	ctc	atc	ctg	ccc	ctg	2989
Arg	Met	Ile	Pro	Asp	Asp	Gln	His	Ser	Pro	Asp	Leu	Ile	Leu	Pro	Leu	
		960					965					970				
cat	gga	ctg	agg	aac	gtc	aaa	gcc	atc	gac	tat	gac	cca	ctg	gac	aag	3037
His	Gly	Leu	Arg	Asn	Val	Lys	Ala	Ile	Asp	Tyr	Asp	Pro	Leu	Asp	Lys	
		975				980					985					
ttc	atc	tac	tgg	gtg	gat	ggg	cgc	cag	aac	atc	aag	cga	gcc	aag	gac	3085
Phe	Ile	Tyr	Trp	Val	Asp	Gly	Arg	Gln	Asn	Ile	Lys	Arg	Ala	Lys	Asp	
		990			995				1000					1005		
gac	ggg	acc	cag	ccc	ttt	gtt	ttg	acc	tct	ctg	agc	caa	ggc	caa	aac	3133
Asp	Gly	Thr	Gln	Pro	Phe	Val	Leu	Thr	Ser	Leu	Ser	Gln	Gly	Gln	Asn	
			1010						1015				1020			
cca	gac	agg	cag	ccc	cac	gac	ctc	agc	atc	gac	atc	tac	agc	cgg	aca	3181
Pro	Asp	Arg	Gln	Pro	His	Asp	Leu	Ser	Ile	Asp	Ile	Tyr	Ser	Arg	Thr	
		1025					1030					1035				
ctg	ttc	tgg	acg	tgc	gag	gcc	acc	aat	acc	atc	aac	gtc	cac	agg	ctg	3229
Leu	Phe	Trp	Trp	Cys	Glu	Ala	Thr	Asn	Thr	Ile	Asn	Val	His	Arg	Leu	
		1040					1045					1050				

10

20

30

agc ggg gaa gcc atg ggg gtg gtg ctg cgt ggg gac cgc gac aag ccc 3277
 Ser Gly Glu Ala Met Gly Val Val Leu Arg Gly Asp Arg Asp Lys Pro
 1055 1060 1065
 agg gcc atc gtc gtc aac gcg gag cga ggg tac ctg tac ttc acc aac 3325
 Arg Ala Ile Val Val Asn Ala Glu Arg Gly Tyr Leu Tyr Phe Thr Asn
 1070 1075 1080 1085
 atg cag gac cgg gca gcc aag atc gaa cgc gca gcc ctg gac ggc acc 3373
 Met Gln Asp Arg Ala Ala Lys Ile Glu Arg Ala Ala Leu Asp Gly Thr
 1090 1095 1100
 gag cgc gag gtc ctc ttc acc acc ggc ctc atc cgc cct gtg gcc ctg 3421
 Glu Arg Glu Val Leu Phe Thr Thr Gly Leu Ile Arg Pro Val Ala Leu
 1105 1110 1115
 gtg gtg gac aac aca ctg ggc aag ctg ttc tgg gtg gac gcg gac ctg 3469
 Val Val Asp Asn Thr Leu Gly Lys Leu Phe Trp Val Asp Ala Asp Leu
 1120 1125 1130
 aag cgc att gag agc tgt gac ctg tca ggg gcc aac cgc ctg acc ctg 3517
 Lys Arg Ile Glu Ser Cys Asp Leu Ser Gly Ala Asn Arg Leu Thr Leu
 1135 1140 1145
 gag gac gcc aac atc gtg cag cct ctg ggc ctg acc atc ctt ggc aag 3565
 Glu Asp Ala Asn Ile Val Gln Pro Leu Gly Leu Thr Ile Leu Gly Lys
 1150 1155 1160 1165
 cat ctc tac tgg atc gac cgc cag cag cag atc gag cgt gtg gag 3613
 His Leu Tyr Trp Ile Asp Arg Gln Gln Met Ile Glu Arg Val Glu
 1170 1175 1180
 aag acc acc ggg gac aag cgg act cgc atc cag ggc cgt gtc gcc cac 3661
 Lys Thr Thr Gly Asp Lys Arg Thr Arg Ile Gln Gly Arg Val Ala His
 1185 1190 1195
 ctc act ggc atc cat gca gtg gag gaa gtc agc ctg gag gag ttc tca 3709
 Leu Thr Gly Ile His Ala Val Glu Glu Val Ser Leu Glu Glu Phe Ser
 1200 1205 1210
 gcc cac cca tgt gcc cgt gac aat ggt ggc tgc tcc cac atc tgt att 3757
 Ala His Pro Cys Ala Arg Asp Asn Gly Gly Cys Ser His Ile Cys Ile
 1215 1220 1225
 gcc aag ggt gat ggg aca cca cgg tgc tca tgc cca gtc cac ctc gtg 3805
 Ala Lys Gly Asp Gly Thr Pro Arg Cys Ser Cys Pro Val His Leu Val
 1230 1235 1240 1245
 ctc ctg cag aac ctg ctg acc tgt gga gag ccg ccc acc tgc tcc ccg 3853
 Leu Leu Gln Asn Leu Leu Thr Cys Gly Glu Pro Pro Thr Cys Ser Pro
 1250 1255 1260
 gac cag ttt gca tgt gcc aca ggg gag atc gac tgt atc ccc ggg gcc 3901
 Asp Gln Phe Ala Cys Ala Thr Gly Glu Ile Asp Cys Ile Pro Gly Ala
 1265 1270 1275
 tgg cgc tgt gac ggc ttt ccc gag tgc gat gac cag agc gac gag gag 3949
 Trp Arg Cys Asp Gly Phe Pro Glu Cys Asp Asp Gln Ser Asp Glu Glu
 1280 1285 1290
 ggc tgc ccc gtg tgc tcc gcc gcc cag ttc ccc tgc gcg cgg ggt cag 3997
 Gly Cys Pro Val Cys Ser Ala Ala Gln Phe Pro Cys Ala Arg Gly Gln
 1295 1300 1305
 tgt gtg gac ctg cgc ctg cgc tgc gac ggc gag gca gac tgt cag gac 4045
 Cys Val Asp Leu Arg Leu Arg Cys Asp Gly Glu Ala Asp Cys Gln Asp

10

20

30

1310	1315	1320	1325	
cgc tca gac gag gtg gac tgt gac gcc atc tgc ctg ccc aac cag ttc				4093
Arg Ser Asp Glu Val Asp Cys Asp Ala Ile Cys Leu Pro Asn Gln Phe				
	1330	1335	1340	
cgg tgt gcg agc ggc cag tgt gtc ctc atc aaa cag cag tgc gac tcc				4141
Arg Cys Ala Ser Gly Gln Cys Val Leu Ile Lys Gln Gln Cys Asp Ser				
	1345	1350	1355	
ttc ccc gac tgt atc gac ggc tcc gac gag ctc atg tgt gaa atc acc				4189
Phe Pro Asp Cys Ile Asp Gly Ser Asp Glu Leu Met Cys Glu Ile Thr				
	1360	1365	1370	
aag ccg ccc tca gac gac agc ccg gcc cac agc agt gcc atc ggg ccc				4237
Lys Pro Pro Ser Asp Asp Ser Pro Ala His Ser Ser Ala Ile Gly Pro				
	1375	1380	1385	
gtc att ggc atc atc ctc tct ctc ttc gtc atg ggt ggt gtc tat ttt				4285
Val Ile Gly Ile Ile Leu Ser Leu Phe Val Met Gly Gly Val Tyr Phe				
	1390	1395	1400	
gtg tgc cag cgc gtg tgc tgc cag cgc tat gcg ggg gcc aac ggg ccc				4333
Val Cys Gln Arg Val Val Cys Gln Arg Tyr Ala Gly Ala Asn Gly Pro				
	1410	1415	1420	
ttc ccg cac gag tat gtc agc ggg acc ccg cac gtg ccc ctc aat ttc				4381
Phe Pro His Glu Tyr Val Ser Gly Thr Pro His Val Pro Leu Asn Phe				
	1425	1430	1435	
ata gcc ccg ggc ggt tcc cag cat ggc ccc ttc aca ggc atc gca tgc				4429
Ile Ala Pro Gly Gly Ser Gln His Gly Pro Phe Thr Gly Ile Ala Cys				
	1440	1445	1450	
gga aag tcc atg atg agc tcc gtg agc ctg atg ggg ggc cgg ggc ggg				4477
Gly Lys Ser Met Met Ser Ser Val Ser Leu Met Gly Gly Arg Gly Gly				
	1455	1460	1465	
gtg ccc ctc tac gac cgg aac cac gtc aca ggg gcc tcg tcc agc agc				4525
Val Pro Leu Tyr Asp Arg Asn His Val Thr Gly Ala Ser Ser Ser Ser				
	1470	1475	1480	
tcg tcc agc acg aag gcc acg ctg tac ccg ccg atc ctg aac ccg ccg				4573
Ser Ser Ser Thr Lys Ala Thr Leu Tyr Pro Pro Ile Leu Asn Pro Pro				
	1490	1495	1500	
ccc tcc ccg gcc acg gac ccc tcc ctg tac aac atg gac atg ttc tac				4621
Pro Ser Pro Ala Thr Asp Pro Ser Leu Tyr Asn Met Asp Met Phe Tyr				
	1505	1510	1515	
tct tca aac att ccg gcc act gcg aga ccg tac agg ccc tac atc att				4669
Ser Ser Asn Ile Pro Ala Thr Ala Arg Pro Tyr Arg Pro Tyr Ile Ile				
	1520	1525	1530	
cga gga atg gcg ccc ccg acg acg ccc tgc agc acc gac gtg tgt gac				4717
Arg Gly Met Ala Pro Pro Thr Thr Pro Cys Ser Thr Asp Val Cys Asp				
	1535	1540	1545	
agc gac tac agc gcc agc cgc tgg aag gcc agc aag tac tac ctg gat				4765
Ser Asp Tyr Ser Ala Ser Arg Trp Lys Ala Ser Lys Tyr Tyr Leu Asp				
	1550	1555	1560	
ttg aac tcg gac tca gac ccc tat cca ccc cca ccc acg ccc cac agc				4813
Leu Asn Ser Asp Ser Asp Pro Tyr Pro Pro Pro Pro Thr Pro His Ser				
	1570	1575	1580	
cag tac ctg tcg gcg gag gac agc tgc ccg ccc tcg ccc gcc acc gag				4861

10

20

30

Gln Tyr Leu Ser Ala Glu Asp Ser Cys Pro Pro Ser Pro Ala Thr Glu
 1585 1590 1595
 agg agc tac ttc cat ctc ttc ccg ccc cct ccg tcc ccc tgc acg gac 4909
 Arg Ser Tyr Phe His Leu Phe Pro Pro Pro Pro Ser Pro Cys Thr Asp
 1600 1605 1610
 tca tcc tgacctcggc cgggccactc tggcttctct gtgccctgt aaatagtttt 4965
 Ser Ser
 1615
 aaatatgaac aaagaaaaaa atatatttta tgatttaaaa aataaatata attgggattt 5025
 taaaaacatg agaaatgtga actgtgatgg ggtgggcagg gctgggagaa cttgtacag 5085
 tggagaaata tttataaact taattttgta aaaca 5120

<210> 3
 <211> 1615
 <212> PRT
 <213> Homo sapiens

10

<400> 3
 Met Glu Ala Ala Pro Gly Pro Pro Trp Pro Leu Leu Leu Leu Leu
 1 5 10 15
 Leu Leu Leu Leu Ala Leu Cys Gly Cys Pro Ala Pro Ala Ala Ser
 20 25 30
 Pro Leu Leu Leu Phe Ala Asn Arg Arg Asp Val Arg Leu Val Asp Ala
 35 40 45
 Gly Gly Val Lys Leu Glu Ser Thr Ile Val Val Ser Gly Leu Glu Asp
 50 55 60
 Ala Ala Ala Val Asp Phe Gln Phe Ser Lys Gly Ala Val Tyr Trp Thr
 65 70 75 80
 Asp Val Ser Glu Glu Ala Ile Lys Gln Thr Tyr Leu Asn Gln Thr Gly
 85 90 95
 Ala Ala Val Gln Asn Val Val Ile Ser Gly Leu Val Ser Pro Asp Gly
 100 105 110
 Leu Ala Cys Asp Trp Val Gly Lys Lys Leu Tyr Trp Thr Asp Ser Glu
 115 120 125
 Thr Asn Arg Ile Glu Val Ala Asn Leu Asn Gly Thr Ser Arg Lys Val
 130 135 140
 Leu Phe Trp Gln Asp Leu Asp Gln Pro Lys Ala Ile Ala Leu Asp Pro
 145 150 155 160
 Ala His Gly Tyr Met Tyr Trp Thr Asp Trp Gly Glu Thr Pro Arg Ile
 165 170 175
 Glu Arg Ala Gly Met Asp Gly Ser Thr Arg Lys Ile Ile Val Asp Ser
 180 185 190
 Asp Ile Tyr Trp Pro Asn Gly Leu Thr Ile Asp Leu Glu Gln Lys
 195 200 205
 Leu Tyr Trp Ala Asp Ala Lys Leu Ser Phe Ile His Arg Ala Asn Leu
 210 215 220
 Asp Gly Ser Phe Arg Gln Lys Val Val Glu Gly Ser Leu Thr His Pro
 225 230 235 240
 Phe Ala Leu Thr Leu Ser Gly Asp Thr Leu Tyr Trp Thr Asp Trp Gln
 245 250 255

20

30

Thr Arg Ser Ile His Ala Cys Asn Lys Arg Thr Gly Gly Lys Arg Lys
 260 265 270
 Glu Ile Leu Ser Ala Leu Tyr Ser Pro Met Asp Ile Gln Val Leu Ser
 275 280 285
 Gln Glu Arg Gln Pro Phe Phe His Thr Arg Cys Glu Glu Asp Asn Gly
 290 295 300
 Gly Trp Ser His Leu Cys Leu Leu Ser Pro Ser Glu Pro Phe Tyr Thr
 305 310 315 320
 Cys Ala Cys Pro Thr Gly Val Gln Met Gln Asp Asn Gly Arg Thr Cys
 325 330 335
 Lys Ala Gly Ala Glu Glu Val Leu Leu Leu Ala Arg Arg Thr Asp Leu
 340 345 350
 Arg Arg Ile Ser Leu Asp Thr Pro Asp Phe Thr Asp Ile Val Leu Gln
 355 360 365
 Val Asp Asp Ile Arg His Ala Ile Ala Ile Asp Tyr Asp Pro Leu Glu
 370 375 380
 Gly Tyr Val Tyr Trp Thr Asp Asp Glu Val Arg Ala Ile Arg Arg Ala
 385 390 395 400
 Tyr Leu Asp Gly Ser Gly Ala Gln Thr Leu Val Asn Thr Glu Ile Asn
 405 410 415
 Asp Pro Asp Gly Ile Ala Val Asp Trp Val Ala Arg Asn Leu Tyr Trp
 420 425 430
 Thr Asp Thr Gly Thr Asp Arg Ile Glu Val Thr Arg Leu Asn Gly Thr
 435 440 445
 Ser Arg Lys Ile Leu Val Ser Glu Asp Leu Asp Glu Pro Arg Ala Ile
 450 455 460
 Ala Leu His Pro Val Met Gly Leu Met Tyr Trp Thr Asp Trp Gly Glu
 465 470 475 480
 Asn Pro Lys Ile Glu Cys Ala Asn Leu Asp Gly Gln Glu Arg Arg Val
 485 490 495
 Leu Val Asn Ala Ser Leu Gly Trp Pro Asn Gly Leu Ala Leu Asp Leu
 500 505 510
 Gln Glu Gly Lys Leu Tyr Trp Gly Asp Ala Lys Thr Asp Lys Ile Glu
 515 520 525
 Val Ile Asn Val Asp Gly Thr Lys Arg Arg Thr Leu Leu Glu Asp Lys
 530 535 540
 Leu Pro His Ile Phe Gly Phe Thr Leu Leu Gly Asp Phe Ile Tyr Trp
 545 550 555 560
 Thr Asp Trp Gln Arg Arg Ser Ile Glu Arg Val His Lys Val Lys Ala
 565 570 575
 Ser Arg Asp Val Ile Ile Asp Gln Leu Pro Asp Leu Met Gly Leu Lys
 580 585 590
 Ala Val Asn Val Ala Lys Val Val Gly Thr Asn Pro Cys Ala Asp Arg
 595 600 605
 Asn Gly Gly Cys Ser His Leu Cys Phe Phe Thr Pro His Ala Thr Arg
 610 615 620
 Cys Gly Cys Pro Ile Gly Leu Glu Leu Leu Ser Asp Met Lys Thr Cys
 625 630 635 640
 Ile Val Pro Glu Ala Phe Leu Val Phe Thr Ser Arg Ala Ala Ile His
 645 650 655

10

20

30

Arg Ile Ser Leu Glu Thr Asn Asn Asn Asp Val Ala Ile Pro Leu Thr
 660 665 670
 Gly Val Lys Glu Ala Ser Ala Leu Asp Phe Asp Val Ser Asn Asn His
 675 680 685
 Ile Tyr Trp Thr Asp Val Ser Leu Lys Asn Ile Ser Arg Ala Phe Met
 690 695 700
 Asn Gly Ser Ser Val Glu His Val Val Glu Phe Gly Leu Asp Tyr Pro
 705 710 715 720
 Glu Gly Met Ala Val Asp Trp Met Gly Lys Asn Leu Tyr Trp Ala Asp
 725 730 735
 Thr Gly Thr Asn Arg Ile Glu Val Ala Arg Leu Asp Gly Gln Phe Arg
 740 745 750
 Gln Val Leu Val Trp Arg Asp Leu Asp Asn Pro Arg Ser Leu Ala Leu
 755 760 765
 Asp Pro Thr Lys Gly Tyr Ile Tyr Trp Thr Glu Trp Gly Gly Lys Pro
 770 775 780
 Arg Ile Val Arg Ala Phe Met Asp Gly Thr Asn Cys Met Thr Leu Val
 785 790 795 800
 Asp Lys Val Gly Arg Ala Asn Asp Leu Thr Ile Asp Tyr Ala Asp Gln
 805 810 815
 Arg Leu Tyr Trp Thr Asp Leu Asp Thr Asn Met Ile Glu Ser Ser Asn
 820 825 830
 Met Leu Gly Gln Glu Arg Val Val Ile Ala Asp Asp Leu Pro His Pro
 835 840 845
 Phe Gly Leu Thr Gln Tyr Ser Asp Tyr Ile Tyr Trp Thr Asp Trp Asn
 850 855 860
 Leu His Ser Ile Glu Arg Ala Asp Lys Thr Ser Gly Arg Asn Arg Thr
 865 870 875 880
 Leu Ile Gln Gly His Leu Asp Phe Val Met Asp Ile Leu Val Phe His
 885 890 895
 Ser Ser Arg Gln Asp Gly Leu Asn Asp Cys Met His Asn Asn Gly Gln
 900 905 910
 Cys Gly Gln Leu Cys Leu Ala Ile Pro Gly Gly His Arg Cys Gly Cys
 915 920 925
 Ala Ser His Tyr Thr Leu Asp Pro Ser Ser Arg Asn Cys Ser Pro Pro
 930 935 940
 Thr Thr Phe Leu Leu Phe Ser Gln Lys Ser Ala Ile Ser Arg Met Ile
 945 950 955 960
 Pro Asp Asp Gln His Ser Pro Asp Leu Ile Leu Pro Leu His Gly Leu
 965 970 975
 Arg Asn Val Lys Ala Ile Asp Tyr Asp Pro Leu Asp Lys Phe Ile Tyr
 980 985 990
 Trp Val Asp Gly Arg Gln Asn Ile Lys Arg Ala Lys Asp Asp Gly Thr
 995 1000 1005
 Gln Pro Phe Val Leu Thr Ser Leu Ser Gln Gly Gln Asn Pro Asp Arg
 1010 1015 1020
 Gln Pro His Asp Leu Ser Ile Asp Ile Tyr Ser Arg Thr Leu Phe Trp
 1025 1030 1035 1040
 Thr Cys Glu Ala Thr Asn Thr Ile Asn Val His Arg Leu Ser Gly Glu
 1045 1050 1055

10

20

30

Ala Met Gly Val Val Leu Arg Gly Asp Arg Asp Lys Pro Arg Ala Ile
 1060 1065 1070
 Val Val Asn Ala Glu Arg Gly Tyr Leu Tyr Phe Thr Asn Met Gln Asp
 1075 1080 1085
 Arg Ala Ala Lys Ile Glu Arg Ala Ala Leu Asp Gly Thr Glu Arg Glu
 1090 1095 1100
 Val Leu Phe Thr Thr Gly Leu Ile Arg Pro Val Ala Leu Val Val Asp
 1105 1110 1115 1120
 Asn Thr Leu Gly Lys Leu Phe Trp Val Asp Ala Asp Leu Lys Arg Ile
 1125 1130 1135
 Glu Ser Cys Asp Leu Ser Gly Ala Asn Arg Leu Thr Leu Glu Asp Ala
 1140 1145 1150
 Asn Ile Val Gln Pro Leu Gly Leu Thr Ile Leu Gly Lys His Leu Tyr
 1155 1160 1165
 Trp Ile Asp Arg Gln Gln Gln Met Ile Glu Arg Val Glu Lys Thr Thr
 1170 1175 1180
 Gly Asp Lys Arg Thr Arg Ile Gln Gly Arg Val Ala His Leu Thr Gly
 1185 1190 1195 1200
 Ile His Ala Val Glu Glu Val Ser Leu Glu Glu Phe Ser Ala His Pro
 1205 1210 1215
 Cys Ala Arg Asp Asn Gly Gly Cys Ser His Ile Cys Ile Ala Lys Gly
 1220 1225 1230
 Asp Gly Thr Pro Arg Cys Ser Cys Pro Val His Leu Val Leu Leu Gln
 1235 1240 1245
 Asn Leu Leu Thr Cys Gly Glu Pro Pro Thr Cys Ser Pro Asp Gln Phe
 1250 1255 1260
 Ala Cys Ala Thr Gly Glu Ile Asp Cys Ile Pro Gly Ala Trp Arg Cys
 1265 1270 1275 1280
 Asp Gly Phe Pro Glu Cys Asp Asp Gln Ser Asp Glu Glu Gly Cys Pro
 1285 1290 1295
 Val Cys Ser Ala Ala Gln Phe Pro Cys Ala Arg Gly Gln Cys Val Asp
 1300 1305 1310
 Leu Arg Leu Arg Cys Asp Gly Glu Ala Asp Cys Gln Asp Arg Ser Asp
 1315 1320 1325
 Glu Val Asp Cys Asp Ala Ile Cys Leu Pro Asn Gln Phe Arg Cys Ala
 1330 1335 1340
 Ser Gly Gln Cys Val Leu Ile Lys Gln Gln Cys Asp Ser Phe Pro Asp
 1345 1350 1355 1360
 Cys Ile Asp Gly Ser Asp Glu Leu Met Cys Glu Ile Thr Lys Pro Pro
 1365 1370 1375
 Ser Asp Asp Ser Pro Ala His Ser Ser Ala Ile Gly Pro Val Ile Gly
 1380 1385 1390
 Ile Ile Leu Ser Leu Phe Val Met Gly Gly Val Tyr Phe Val Cys Gln
 1395 1400 1405
 Arg Val Val Cys Gln Arg Tyr Ala Gly Ala Asn Gly Pro Phe Pro His
 1410 1415 1420
 Glu Tyr Val Ser Gly Thr Pro His Val Pro Leu Asn Phe Ile Ala Pro
 1425 1430 1435 1440
 Gly Gly Ser Gln His Gly Pro Phe Thr Gly Ile Ala Cys Gly Lys Ser
 1445 1450 1455

10

20

30

Met Met Ser Ser Val Ser Leu Met Gly Gly Arg Gly Gly Val Pro Leu
 1460 1465 1470
 Tyr Asp Arg Asn His Val Thr Gly Ala Ser Ser Ser Ser Ser Ser
 1475 1480 1485
 Thr Lys Ala Thr Leu Tyr Pro Pro Ile Leu Asn Pro Pro Pro Ser Pro
 1490 1495 1500
 Ala Thr Asp Pro Ser Leu Tyr Asn Met Asp Met Phe Tyr Ser Ser Asn
 1505 1510 1515 1520
 Ile Pro Ala Thr Ala Arg Pro Tyr Arg Pro Tyr Ile Ile Arg Gly Met
 1525 1530 1535
 Ala Pro Pro Thr Thr Pro Cys Ser Thr Asp Val Cys Asp Ser Asp Tyr
 1540 1545 1550
 Ser Ala Ser Arg Trp Lys Ala Ser Lys Tyr Tyr Leu Asp Leu Asn Ser
 1555 1560 1565
 Asp Ser Asp Pro Tyr Pro Pro Pro Thr Pro His Ser Gln Tyr Leu
 1570 1575 1580
 Ser Ala Glu Asp Ser Cys Pro Pro Ser Pro Ala Thr Glu Arg Ser Tyr
 1585 1590 1595 1600
 Phe His Leu Phe Pro Pro Pro Pro Ser Pro Cys Thr Asp Ser Ser
 1605 1610 1615

<210> 4
 <211> 1615
 <212> PRT
 <213> Homo sapiens

<400> 4
 Met Glu Ala Ala Pro Pro Gly Pro Pro Trp Pro Leu Leu Leu Leu Leu
 1 5 10 15
 Leu Leu Leu Leu Ala Leu Cys Gly Cys Pro Ala Pro Ala Ala Ser
 20 25 30
 Pro Leu Leu Leu Phe Ala Asn Arg Arg Asp Val Arg Leu Val Asp Ala
 35 40 45
 Gly Gly Val Lys Leu Glu Ser Thr Ile Val Val Ser Gly Leu Glu Asp
 50 55 60
 Ala Ala Ala Val Asp Phe Gln Phe Ser Lys Gly Ala Val Tyr Trp Thr
 65 70 75 80
 Asp Val Ser Glu Glu Ala Ile Lys Gln Thr Tyr Leu Asn Gln Thr Gly
 85 90 95
 Ala Ala Val Gln Asn Val Val Ile Ser Gly Leu Val Ser Pro Asp Gly
 100 105 110
 Leu Ala Cys Asp Trp Val Gly Lys Lys Leu Tyr Trp Thr Asp Ser Glu
 115 120 125
 Thr Asn Arg Ile Glu Val Ala Asn Leu Asn Gly Thr Ser Arg Lys Val
 130 135 140
 Leu Phe Trp Gln Asp Leu Asp Gln Pro Lys Ala Ile Ala Leu Asp Pro
 145 150 155 160
 Ala His Gly Tyr Met Tyr Trp Thr Asp Trp Val Glu Thr Pro Arg Ile
 165 170 175

10

20

30

Glu Arg Ala Gly Met Asp Gly Ser Thr Arg Lys Ile Ile Val Asp Ser
 180 185 190
 Asp Ile Tyr Trp Pro Asn Gly Leu Thr Ile Asp Leu Glu Gln Lys
 195 200 205
 Leu Tyr Trp Ala Asp Ala Lys Leu Ser Phe Ile His Arg Ala Asn Leu
 210 215 220
 Asp Gly Ser Phe Arg Gln Lys Val Val Glu Gly Ser Leu Thr His Pro
 225 230 235 240
 Phe Ala Leu Thr Leu Ser Gly Asp Thr Leu Tyr Trp Thr Asp Trp Gln
 245 250 255
 Thr Arg Ser Ile His Ala Cys Asn Lys Arg Thr Gly Gly Lys Arg Lys
 260 265 270
 Glu Ile Leu Ser Ala Leu Tyr Ser Pro Met Asp Ile Gln Val Leu Ser
 275 280 285
 Gln Glu Arg Gln Pro Phe Phe His Thr Arg Cys Glu Glu Asp Asn Gly
 290 295 300
 Gly Trp Ser His Leu Cys Leu Leu Ser Pro Ser Glu Pro Phe Tyr Thr
 305 310 315 320
 Cys Ala Cys Pro Thr Gly Val Gln Met Gln Asp Asn Gly Arg Thr Cys
 325 330 335
 Lys Ala Gly Ala Glu Glu Val Leu Leu Ala Arg Arg Thr Asp Leu
 340 345 350
 Arg Arg Ile Ser Leu Asp Thr Pro Asp Phe Thr Asp Ile Val Leu Gln
 355 360 365
 Val Asp Asp Ile Arg His Ala Ile Ala Ile Asp Tyr Asp Pro Leu Glu
 370 375 380
 Gly Tyr Val Tyr Trp Thr Asp Asp Glu Val Arg Ala Ile Arg Arg Ala
 385 390 395 400
 Tyr Leu Asp Gly Ser Gly Ala Gln Thr Leu Val Asn Thr Glu Ile Asn
 405 410 415
 Asp Pro Asp Gly Ile Ala Val Asp Trp Val Ala Arg Asn Leu Tyr Trp
 420 425 430
 Thr Asp Thr Gly Thr Asp Arg Ile Glu Val Thr Arg Leu Asn Gly Thr
 435 440 445
 Ser Arg Lys Ile Leu Val Ser Glu Asp Leu Asp Glu Pro Arg Ala Ile
 450 455 460
 Ala Leu His Pro Val Met Gly Leu Met Tyr Trp Thr Asp Trp Gly Glu
 465 470 475 480
 Asn Pro Lys Ile Glu Cys Ala Asn Leu Asp Gly Gln Glu Arg Arg Val
 485 490 495
 Leu Val Asn Ala Ser Leu Gly Trp Pro Asn Gly Leu Ala Leu Asp Leu
 500 505 510
 Gln Glu Gly Lys Leu Tyr Trp Gly Asp Ala Lys Thr Asp Lys Ile Glu
 515 520 525
 Val Ile Asn Val Asp Gly Thr Lys Arg Arg Thr Leu Leu Glu Asp Lys
 530 535 540
 Leu Pro His Ile Phe Gly Phe Thr Leu Leu Gly Asp Phe Ile Tyr Trp
 545 550 555 560
 Thr Asp Trp Gln Arg Ser Ile Glu Arg Val His Lys Val Lys Ala
 565 570 575

10

20

30

Ser Arg Asp Val Ile Ile Asp Gln Leu Pro Asp Leu Met Gly Leu Lys
 580 585 590
 Ala Val Asn Val Ala Lys Val Val Gly Thr Asn Pro Cys Ala Asp Arg
 595 600 605
 Asn Gly Gly Cys Ser His Leu Cys Phe Phe Thr Pro His Ala Thr Arg
 610 615 620
 Cys Gly Cys Pro Ile Gly Leu Glu Leu Leu Ser Asp Met Lys Thr Cys
 625 630 635 640
 Ile Val Pro Glu Ala Phe Leu Val Phe Thr Ser Arg Ala Ala Ile His
 645 650 655
 Arg Ile Ser Leu Glu Thr Asn Asn Asn Asp Val Ala Ile Pro Leu Thr
 660 665 670
 Gly Val Lys Glu Ala Ser Ala Leu Asp Phe Asp Val Ser Asn Asn His
 675 680 685
 Ile Tyr Trp Thr Asp Val Ser Leu Lys Asn Ile Ser Arg Ala Phe Met
 690 695 700
 Asn Gly Ser Ser Val Glu His Val Val Glu Phe Gly Leu Asp Tyr Pro
 705 710 715 720
 Glu Gly Met Ala Val Asp Trp Met Gly Lys Asn Leu Tyr Trp Ala Asp
 725 730 735
 Thr Gly Thr Asn Arg Ile Glu Val Ala Arg Leu Asp Gly Gln Phe Arg
 740 745 750
 Gln Val Leu Val Trp Arg Asp Leu Asp Asn Pro Arg Ser Leu Ala Leu
 755 760 765
 Asp Pro Thr Lys Gly Tyr Ile Tyr Trp Thr Glu Trp Gly Gly Lys Pro
 770 775 780
 Arg Ile Val Arg Ala Phe Met Asp Gly Thr Asn Cys Met Thr Leu Val
 785 790 795 800
 Asp Lys Val Gly Arg Ala Asn Asp Leu Thr Ile Asp Tyr Ala Asp Gln
 805 810 815
 Arg Leu Tyr Trp Thr Asp Leu Asp Thr Asn Met Ile Glu Ser Ser Asn
 820 825 830
 Met Leu Gly Gln Glu Arg Val Val Ile Ala Asp Asp Leu Pro His Pro
 835 840 845
 Phe Gly Leu Thr Gln Tyr Ser Asp Tyr Ile Tyr Trp Thr Asp Trp Asn
 850 855 860
 Leu His Ser Ile Glu Arg Ala Asp Lys Thr Ser Gly Arg Asn Arg Thr
 865 870 875 880
 Leu Ile Gln Gly His Leu Asp Phe Val Met Asp Ile Leu Val Phe His
 885 890 895
 Ser Ser Arg Gln Asp Gly Leu Asn Asp Cys Met His Asn Asn Gly Gln
 900 905 910
 Cys Gly Gln Leu Cys Leu Ala Ile Pro Gly Gly His Arg Cys Gly Cys
 915 920 925
 Ala Ser His Tyr Thr Leu Asp Pro Ser Ser Arg Asn Cys Ser Pro Pro
 930 935 940
 Thr Thr Phe Leu Leu Phe Ser Gln Lys Ser Ala Ile Ser Arg Met Ile
 945 950 955 960
 Pro Asp Asp Gln His Ser Pro Asp Leu Ile Leu Pro Leu His Gly Leu
 965 970 975

10

20

30

Arg Asn Val Lys Ala Ile Asp Tyr Asp Pro Leu Asp Lys Phe Ile Tyr
 980 985 990
 Trp Val Asp Gly Arg Gln Asn Ile Lys Arg Ala Lys Asp Asp Gly Thr
 995 1000 1005
 Gln Pro Phe Val Leu Thr Ser Leu Ser Gln Gly Gln Asn Pro Asp Arg
 1010 1015 1020
 Gln Pro His Asp Leu Ser Ile Asp Ile Tyr Ser Arg Thr Leu Phe Trp
 1025 1030 1035 1040
 Thr Cys Glu Ala Thr Asn Thr Ile Asn Val His Arg Leu Ser Gly Glu
 1045 1050 1055
 Ala Met Gly Val Val Leu Arg Gly Asp Arg Asp Lys Pro Arg Ala Ile
 1060 1065 1070
 Val Val Asn Ala Glu Arg Gly Tyr Leu Tyr Phe Thr Asn Met Gln Asp
 1075 1080 1085
 Arg Ala Ala Lys Ile Glu Arg Ala Ala Leu Asp Gly Thr Glu Arg Glu
 1090 1095 1100
 Val Leu Phe Thr Thr Gly Leu Ile Arg Pro Val Ala Leu Val Val Asp
 1105 1110 1115 1120
 Asn Thr Leu Gly Lys Leu Phe Trp Val Asp Ala Asp Leu Lys Arg Ile
 1125 1130 1135
 Glu Ser Cys Asp Leu Ser Gly Ala Asn Arg Leu Thr Leu Glu Asp Ala
 1140 1145 1150
 Asn Ile Val Gln Pro Leu Gly Leu Thr Ile Leu Gly Lys His Leu Tyr
 1155 1160 1165
 Trp Ile Asp Arg Gln Gln Gln Met Ile Glu Arg Val Glu Lys Thr Thr
 1170 1175 1180
 Gly Asp Lys Arg Thr Arg Ile Gln Gly Arg Val Ala His Leu Thr Gly
 1185 1190 1195 1200
 Ile His Ala Val Glu Glu Val Ser Leu Glu Glu Phe Ser Ala His Pro
 1205 1210 1215
 Cys Ala Arg Asp Asn Gly Gly Cys Ser His Ile Cys Ile Ala Lys Gly
 1220 1225 1230
 Asp Gly Thr Pro Arg Cys Ser Cys Pro Val His Leu Val Leu Leu Gln
 1235 1240 1245
 Asn Leu Leu Thr Cys Gly Glu Pro Pro Thr Cys Ser Pro Asp Gln Phe
 1250 1255 1260
 Ala Cys Ala Thr Gly Glu Ile Asp Cys Ile Pro Gly Ala Trp Arg Cys
 1265 1270 1275 1280
 Asp Gly Phe Pro Glu Cys Asp Asp Gln Ser Asp Glu Glu Gly Cys Pro
 1285 1290 1295
 Val Cys Ser Ala Ala Gln Phe Pro Cys Ala Arg Gly Gln Cys Val Asp
 1300 1305 1310
 Leu Arg Leu Arg Cys Asp Gly Glu Ala Asp Cys Gln Asp Arg Ser Asp
 1315 1320 1325
 Glu Val Asp Cys Asp Ala Ile Cys Leu Pro Asn Gln Phe Arg Cys Ala
 1330 1335 1340
 Ser Gly Gln Cys Val Leu Ile Lys Gln Gln Cys Asp Ser Phe Pro Asp
 1345 1350 1355 1360
 Cys Ile Asp Gly Ser Asp Glu Leu Met Cys Glu Ile Thr Lys Pro Pro
 1365 1370 1375

10

20

30

Ser Asp Asp Ser Pro Ala His Ser Ser Ala Ile Gly Pro Val Ile Gly
 1380 1385 1390
 Ile Ile Leu Ser Leu Phe Val Met Gly Gly Val Tyr Phe Val Cys Gln
 1395 1400 1405
 Arg Val Val Cys Gln Arg Tyr Ala Gly Ala Asn Gly Pro Phe Pro His
 1410 1415 1420
 Glu Tyr Val Ser Gly Thr Pro His Val Pro Leu Asn Phe Ile Ala Pro
 1425 1430 1435 1440
 Gly Gly Ser Gln His Gly Pro Phe Thr Gly Ile Ala Cys Gly Lys Ser
 1445 1450 1455
 Met Met Ser Ser Val Ser Leu Met Gly Gly Arg Gly Gly Val Pro Leu
 1460 1465 1470
 Tyr Asp Arg Asn His Val Thr Gly Ala Ser Ser Ser Ser Ser Ser
 1475 1480 1485
 Thr Lys Ala Thr Leu Tyr Pro Pro Ile Leu Asn Pro Pro Pro Ser Pro
 1490 1495 1500
 Ala Thr Asp Pro Ser Leu Tyr Asn Met Asp Met Phe Tyr Ser Ser Asn
 1505 1510 1515 1520
 Ile Pro Ala Thr Ala Arg Pro Tyr Arg Pro Tyr Ile Ile Arg Gly Met
 1525 1530 1535
 Ala Pro Pro Thr Thr Pro Cys Ser Thr Asp Val Cys Asp Ser Asp Tyr
 1540 1545 1550
 Ser Ala Ser Arg Trp Lys Ala Ser Lys Tyr Tyr Leu Asp Leu Asn Ser
 1555 1560 1565
 Asp Ser Asp Pro Tyr Pro Pro Pro Pro Thr Pro His Ser Gln Tyr Leu
 1570 1575 1580
 Ser Ala Glu Asp Ser Cys Pro Pro Ser Pro Ala Thr Glu Arg Ser Tyr
 1585 1590 1595 1600
 Phe His Leu Phe Pro Pro Pro Pro Ser Pro Cys Thr Asp Ser Ser
 1605 1610 1615

10

20

<210> 5
 <211> 3096
 <212> DNA
 <213> Homo sapiens

<400> 5
 catcttctca cacgatctct cgtcttcgac tccttccttt gattgggttt caccatttac 60
 tcagacgacg gtctctcttc gatctttgca cattcttcta tcatctacta ccttcatacc 120
 cagctccgtc ccctaataatt catgcgcgga tggcccatte cgtggtgaaa attcccttct 180
 actctgctaa tctgctgttc tctctccctc ccgtcgggtt ctgctcctgc cacgttctcc 240
 cctctcccca ccaaaggctg ggttttcttt gtcagggttc ctttccctt tggagaagag 300
 ggggctgtat ggccttggtg cgaggccctc cagtgcacag atcccccatc acccagagtt 360
 ccacaggccc tggtagggag gagggggagc agaagaggag gtgccatctt tgcctgctgg 420
 ggaagggcag gggccaccca cacagagctc tcccatttgc tgtggaccct ggggccactg 480
 ccagttcct tccaaaggaa agccagctcc ccaggtggtg ggagagtgat atggcttctc 540
 cttaaaactta ggggaattgag tgtgtggttg cttctaagtg ccttagaagc cgggagcggc 600
 tcttggaag agcctgctg ccacagcggg ccttaccctg gctgtgccca cagatgtccc 660
 tggggcctgc cgtcctgccc cggctctctc ggcctccccc ggtgtgggtt gggaaaagca 720

30

```

cagcaaat  aaaaacacct  ccattctctgg  cctttgaaga  atgcattctga  acagccgaga  780
gtgtaaaccg  tgggtgaatg  tgggtctttcc  agttttgggga  gaagcagggc  agagctgggg  840
cttttgtacc  caggggtttcc  aagagctcct  gcttccctcog  gctgggctgg  ccagggcccc  900
ccgctgggac  ctccagctgt  aataggggaag  gttttactgg  gttgctggcc  actgtggact  960
gcccctaagg  gcaggatatg  ctgcctttac  cogggttccc  ctctgcctg  gaagatacag  1020
cccatgggag  gctgtttgtc  tgtgggatcc  tccagcatca  gagacactgg  ggccagcgtc  1080
tgcttggtga  ggtgcaggcc  tggcaggccc  ggtcccccac  ctgcttgagc  acccacggtg  1140
gtgggggctc  gctgcctccc  gagacaatct  atgtcattgt  tgtccaagga  agctaattta  1200
gagtagaaag  ttccgtgtcc  agtcccactc  tgtgcgtgtg  ttagcagggg  actctcgggc  1260
cggagctggg  tccaccctgg  tagggggact  tcatggggcc  tgggcgacag  cactgtgtat  1320
ttgtgtgtgt  gtgtgtttgt  gtgtgtgtgt  gtctgaggag  gtggaccagt  ttctcaaaag  1380
gctgtgacc  ccaagaacca  aggaatttca  gcctgggtgg  atcacacctt  cactggtgag  1440
tgggacaagc  tgggggcccc  cgccacagga  gcagccaggg  catggggcac  agttggcctc  1500
attcacaana  tgggagtata  agtgatccct  gctctggcgg  ccaggacgat  gagtgggaac  1560
acaccgtgtg  ggggctgcct  ggcctgggtg  tgccgcgggt  gtccctgttg  gtgatggttc  1620
cacctgcttg  tgccaccagt  gcccctcggg  tctcacacac  aactctcttc  ccagcgaagg  1680
cccctcctgc  cctcaggcct  cagtgcctgt  tccgtctcgg  aaggccccag  gagctcctgc  1740
atcctgggag  tgattcctgt  gtgcctgcag  accccctcgc  ggctgccatc  tcatcctttg  1800
gtgcacctgt  tggccagacc  tcttggttag  ggggtgctga  ctcccctgaa  tgtgcggggg  1860
cctgggggca  gggacctggg  ctccctccct  actgagtgga  gggaactcag  tgtcttgagg  1920
ttgggggtgc  tgcaggtcgg  gtggtgcagg  tgaatgcag  acctctcagc  tgggtgtcca  1980
ggcagctgc  ctccccccgc  ccgagggact  tcaccgcag  ccagtcagg  ggtggcgcc  2040
gggtgcctgc  cccgcaggct  gggtaggggt  ggagcctggg  tggccctgcc  tgtgagctgc  2100
atagttgtcg  cctttgaccc  tgagttttct  tggttatctg  tttggacctg  tttggggcag  2160
gcaggggatg  agatctgaag  ataaatgcct  tagctgtgac  catctccttt  tgtgagaggt  2220
caatgtccag  ttccgtgtga  gttataacat  cccatttttt  gatttctttt  tattttttcc  2280
tttttctttt  tgagatggag  tctcgctctg  tcaccagggc  tggagtgaac  tgggggtgac  2340
tcagctcact  gcaacctcca  cttctcgggt  tcaagtgatt  ctctgcctc  agcctcctga  2400
ctagcagggg  ttacaggcgt  gagccaccac  gccagctaa  tttttgtatt  tttagtagag  2460
gcaaggtttc  gtcagtgttg  ccaggctggg  ctcaaacctc  tggccttaag  tgatctgccc  2520
gcctcggcct  cccaaagtgc  tgagatgaca  ggtgtgagcc  accgtgccc  gcccagaact  2580
ctttaattcc  cacctgaaac  ttgccgcctt  aagcaggtcc  ccagtctccc  tcccctagtc  2640
cctgggtccc  ccattctgct  ttctgtctca  atgaatttgc  ctaccgtaag  taactcatat  2700
aaattgaatc  ataaagtatt  tgtcttttta  tatctggctt  atttcaacta  gcataacatt  2760
cttaagtttc  atccatgttg  tagcatgtgt  cagaatctct  ctcttttttt  tttttttttt  2820
tttttttttt  ttttgagac  agagtctcgc  tctgtcatct  agactggagt  tcagtggcac  2880
gatctcgggt  cactgcaaca  totgcctcct  gggtecaagg  aattctcctg  cctcagcctc  2940
cttagcagct  ggaactacag  gcgcgtgcca  ccagtccttg  ctaatttttt  tattttttatg  3000
tggaggcagg  gtttcccat  cttggccagg  ctgggtctga  attcctgggt  ttcaccacgg  3060
gggcccgaag  gaccggggca  aagcgtggag  gggagg

```

<210> 6

<211> 26928

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> (12044), (12489), (26433), (26434), (26435), (26436), (26439), (26441)

<223> Identity of nucleotide sequences at the above locations are unknown.

10

20

30

<400> 6

gaagaccaag	ggcacacagc	gaggcagttt	cagggcgggc	agcctggggc	cccacggggc	60
ggccccggac	acttgttctc	acctgtggag	ggcagagaag	ggaacaggga	gagaagtggc	120
cggctggggg	tggaggtggg	tttgaggttt	tactgtaaac	taaatgtgta	ccctctacct	180
tagttatgaa	ttatgagaca	cgaagactgc	gaaacagaca	cactcctcta	aaagtgcctc	240
taggctgaca	gggagaaagt	cccgccaggc	tcccagacgc	cacctttgag	tccttcaaca	300
agcccggccag	ggcctcttgc	ccaccgggtg	cagctcagcc	actgaaccct	ccaggaaagaa	360
gaagtgtctg	taggagaaga	atctcaccca	ggcacagcct	ggaaggggca	cagaaggggc	420
tccggaacca	gcaagcccaa	gttggaactc	ccagtctgct	actttctaga	acgactgtgc	480
ccttggcggg	tctaagtaga	acctctccgc	gcactctttc	ctcctttgta	aagtggggac	540
agcaatggcc	accttcaggg	ttcagagagg	gcttgagta	cctcacagaa	ctgagtgccc	600
gtgaacgtgt	gtgttccctc	agatttgtga	cagctttgcc	aggctggagt	caggctgaac	660
gcctctgccc	tcattggggt	tattttctag	gaagaccaac	aaaaacaaga	agacggaaaa	720
ttaaaaaac	aaaagcccca	ttgacaggcc	gtgaagaatg	ccatgaaaaa	tgaatggcgt	780
tgtgctgac	tctttgggga	aacgggctta	cggaaagaag	gacacttgag	ctgctaccaa	840
tgagcagccg	tccgggtggg	gggcagttca	ggaagagcag	acatccactg	aggaggcgct	900
ggggcagagg	gcagcctggg	cgctggatcc	ggggggaggaa	ccacatcagg	ccatgagctg	960
gagctgggtg	tagaatgtac	aggagaggcc	agccaggggc	agctcatgtc	agacctcaag	1020
cggggaagat	gaatcgagaa	tgcacccccc	gagcaatggg	aagccagtct	acgatttaag	1080
cagcaaaaat	attttccctt	cttccaccct	gcatecagct	ctaccagcac	agcctggggg	1140
tctattttca	agatagaata	gaccagagct	cccagctctt	cttactcttc	tactactgcc	1200
acctgtccac	cactcatggg	tccccacttg	cagcctcgac	ccccctccac	ctgatctcat	1260
ggcagccagg	gaagctccag	ggctcgtgag	ggctgccatc	tcaggaaaga	agcaaaaagcc	1320
ttcggcacct	gcaggggcctg	ctccaaccac	acttcttctt	tgaaccttca	gcttcccttag	1380
ccactccctt	cccacatctc	accctgctcc	agccacagtg	gtgtctctgt	gggttctcaa	1440
acacaccagg	tgcactcctg	cctcagggcc	tttgtgcttg	ctgttctctg	ctgggactct	1500
tttttttttt	tttttttttg	agacagggtc	tcactctgtg	gcccaggctg	gagtgtagtg	1560
gtgtgatcgt	agctcattgc	aaacctcaaac	tcctgggctc	aagcaatcct	cccacctcag	1620
cctctcaagt	agttagcttt	tgtttgtttg	ttttgagatg	ggatctcact	ctgttgccca	1680
ggctggagtg	cagtggggca	atcttggttc	accacaacct	ctgcctccca	ggctcaagca	1740
attctcctgc	ctcagcctcc	caagtagctg	ggattacagg	catgtgccac	cacgcccagc	1800
ttatttttgt	attttttagta	gagacagggg	ttcaccatgt	tggctctggc	ggtcttgaac	1860
tcttggcctc	agatgatcca	cctgcctcgg	cctcccaaag	tgctgggatg	acaggcatga	1920
gcctgtctct	agtagttagg	actacagaga	ggggccatca	tgccctggtga	tcctcccacc	1980
ttttctgctc	caactctttc	accccactta	gcctcgtggc	tcactctctt	acctcttcag	2040
ctcctcagtc	aggcctgagg	acccctgttg	aaaattgcaa	accacacccc	ccaccaccac	2100
caccactat	tgccagcact	ttctactcca	tttctctgct	ttacttttct	cctttgtact	2160
catcaccacc	tgactcatta	catgtttaag	tatctttctt	ctctccacta	gcattggaagc	2220
tccaggagag	cagagagtgt	agttttatcc	cctgatgtgt	ttcctgtgcc	cgtaccaggg	2280
cctagcacac	agtaggtgct	cagtaaatgt	gtgttggtatg	aacaaatata	gtgaaaggat	2340
ctgatctaca	tttataaaga	aggcactctg	gctgctgagt	ggggatgaga	ctgtcaggag	2400
gaaagaggcc	cctgtggggg	cctggccagc	aggtgggtac	aatggttagca	gccaggagag	2460
agggcctctt	ggactcaagt	ggatggggcc	tgctcagggc	tcgggcccaca	ggaacaaagg	2520
gaaggggggc	caggatggcc	tgtcatagag	gacacattac	aactggccca	aagttcaagt	2580
caggtttcta	aattttggga	gggatacaga	aaaactaaag	actctactgg	acagtcagtt	2640
attgaaatga	ttacatagaa	aatgtaccaa	gaattaaaaa	aaaaaaaaaa	aagcattatg	2700
aaggggccac	cagagactcc	cagagaggaa	agggactatg	ggctggatgc	ggtgactcac	2760
acctataatc	ccagcacttt	gggaggccga	ggaggggtga	tcacgaggtc	aggagttcaa	2820
aaccagccta	ggcaacatgg	taaaaccccc	gtttctacta	aaaatacaaa	aaattagctg	2880
ggcatggcag	catgtgcctg	taatcccagc	tactcgggag	gctgaggcag	gagagttgct	2940

10

20

30

agaaccaggg	aggcagaggt	tgcagtggagc	cgagattggag	ccactatgct	ccagcttggg	3000
cgacagagca	agactccgtc	tctaaaaaaa	agaaaaaaa	ggccagatga	ggtggctcat	3060
gctgtaatc	ccagcacttt	gggaggccga	ggtgggtgga	tcacgaggtc	aggagatcga	3120
gaccatcctg	gctaactcgg	tgaaactcca	tctctactta	aaatacaaaa	aattagccgg	3180
gcgtgggtggc	gggcacccgt	agtcccagct	acttgggagg	ctgaggcagg	agaatggcgt	3240
gaacctggga	ggcggagcct	gcagtggagc	gagattggcg	cactgcactc	catccagcct	3300
gggcgacaga	gttagactcc	gtctcaaaaa	aaaaaaaaa	aaaaaaatta	gctgattagt	3360
tgggcttggg	ggcgggccc	tgtaatccca	actactcggg	aggctgaggc	gggagaatca	3420
cttgaacccg	ggaggcagag	gttgcaatga	gccgatatca	cgccactaca	ctccagcctg	3480
ggcgacagag	caagactcca	tctcaaaaaa	gaaaaaaaaa	aagaaagggg	ctgtgctgtg	3540
gcctgggacc	caaaacacac	tactgcaagg	tcccagggtg	cctgactcca	accggagcct	3600
tgagaacatt	catttgcaaa	gaatgaatta	aaattcagca	ctattttatt	ctgcaggatt	3660
ccagcaccoc	aaggacagtc	attttttagac	ccttcagtaa	cgtaataagt	aaccggaggga	3720
tgtgctgagc	ttccacttcc	ccagacgggt	gcctgtcaca	gctcatcagg	ccaacaaact	3780
ttctcttaggc	ctcaaatttg	gaaatgttca	ctctcagttc	gttccttaga	tgcaagtcca	3840
ttccaatgaa	gtaacagggg	ctcagcacct	gtccaatctc	attgcttccg	gggacagggg	3900
cccatgagga	tgtcgtttca	gcccggtgac	acttgggcaa	agtgcctttt	ggtttccctc	3960
ccaggctgga	acgtgctggc	tctgtgaagt	tacgttgggc	acaagagccc	cccccaacc	4020
ggcaggactg	actgctgtgg	tcagaggcgc	ccctggggct	ttgggagcca	cagaatcttc	4080
ctgagggcag	cgccggagga	ggccccagtg	agagtgccca	ctgccaggct	cattcctcag	4140
gctgccgcag	gcctctcccc	aaaacaggca	atgcttctca	gcaacctgcc	ccaggagcag	4200
ggcagggaag	gcccgcactc	gcctacagtg	ctgggctctg	gagggttggg	ttggtaacag	4260
gccatggttt	ctatgagcca	gctggggtgt	gaaggacaca	ggctggattc	acctctctgg	4320
gcctcagttt	ctgcattcaa	aaagtgggaa	tcattgatatc	tgctctatct	cttatctctc	4380
agtgtgtatg	tgaacctcca	ataagacttt	taaaaatact	ctttctacct	tacttttatt	4440
tttcatttat	tttaagataa	tgtctagctg	tctcaccag	gctggagtgc	agtgggtgtga	4500
ttacggctca	ctacagcctt	aacctcccag	gctcaagtga	tcctcctacc	acagcctccc	4560
aagtagctgg	aactacaggc	atgcaccacc	gcacctggat	aatTTTTTct	tttgagacaa	4620
ggtttcactc	tgttgcccag	gctggagtgc	agtgggtgcac	tcttggctca	ctgcagcctc	4680
aacctccctg	ggcttaggtg	atcctcacac	ttcagttctc	caagtagctg	ggactacagg	4740
tatgtgccag	tacaccagc	taatatTTTT	gaaggatggg	gtttcactat	attgcccagg	4800
ctggtcttga	actccagggt	ttaagcaatc	taccttctct	agcctgccaa	agtgttagga	4860
ttatagggtat	gagccacccc	ccggcctata	atcctaccac	tttaaaaaag	cctgtaattt	4920
tagcacttta	aaaaattttt	ctaaattttt	tatagagatg	ggggacagct	gtgggtctcac	4980
tgtgttgccc	agggtggtct	tgaaactccta	ggatcaagcc	atcctcctgg	cctggcctcc	5040
caaagtgttg	ggattataag	cataagcctt	accttacctt	ttttttttga	gttgacgttt	5100
tgttcttgtt	gctcaggctg	gagtgcattg	gcaagatctt	ggctcactgc	aacctccacc	5160
tcccgggttc	aagcaattct	cctgcctcag	cctcccaggt	agctgggatt	acaggcatgc	5220
gccaccacac	ccagctaatt	ttgtattttt	agtagagatg	gggtttctct	atatacctta	5280
atTTTaaagc	actgcattca	tgtaaattgt	gattaaacatg	gattcaagag	agggagtggag	5340
gatgaatgag	ccaggcagtc	acctcggtcg	tcacctccca	cttctctcct	ccttctgaca	5400
gtcatcgtcc	atccgtttct	gcagctgttt	gtttgactct	cctgatcatt	ttgcttgcca	5460
cataacttgc	ctcctgggaa	agaatgccct	gggcaggccc	acatgagtag	tgaaaaataa	5520
tctgcagtga	aaaaataaac	taagtagtct	ggtccacaga	gcagtcttat	tttttccactg	5580
cagatgaagg	agttgacatt	caggcttcat	tctcattttat	aagtgtttta	aagacacata	5640
cagtggattg	aacagtggcc	ttcaaaaaga	tgtatctaca	tcctaattccc	tgggacctgt	5700
gaatgttaac	caagttagga	aaagggtctt	cccgggtgtc	attaagttag	agatcttgag	5760
atgaggagct	catcgtggat	tatccaggtg	gaccttgcct	ccaaggacaa	atggtcctta	5820
gaaaagaaaa	gcagaggctg	ggcacagtgg	ctcaagcctg	taatcccagc	actttgagag	5880
gccgagggtg	gtggatcacc	taaggtcactg	agttcgagag	cagcctggcc	aacatgatga	5940

10

20

30

aatcccatct	ctactaaaaa	tacaaaaatt	agcaaggcat	ggtggcgggt	gcctataatc	6000
ccagctactc	aggaagctga	ggcaggagaa	tggcttgcac	ctgggaggcg	gaggttgcag	6060
tgagccaaga	tcgcgcact	gcactccagc	ctgagggaga	aaagtgaac	tctgtctcat	6120
aaaagaaaag	aaaagcagac	agagatctga	gacagaagag	gagagtgaag	gaaaaaaggc	6180
catgtgaaga	tgaggcagag	gttggagcca	tgacagccaca	agccaaggaa	tacctggagc	6240
cccagaaagt	gcaagaggta	ggaagaagcc	tcccctagag	cctccagacg	gagcacagcc	6300
ctgccaacac	ctccacctca	gacttctggc	ctccagcact	gtgagataat	caactgctgt	6360
tgttttaagc	caccagattt	gtggtaattt	gttatggcag	ccacaggaaa	ctaatacagt	6420
acctaattct	cacaaaccca	tcttacagaa	aaggaaactg	aagtccagaga	ggtagtggct	6480
tgtgcagtg	gttaggccat	tcttgtatta	ctataaagaa	atacctgagg	ccgggcatgg	6540
tggtcacgc	ctgtaatccc	agcactttgg	gaggccaagg	tgagtggatc	acttgaggtc	6600
aggagttaa	gaccagcctg	gacaacatgg	tgaaacccca	tttctactga	aaatatgaaa	6660
attagccagg	catggtggcg	tgcatctgta	gtcccagcta	ctcaggaggc	tgaggcagga	6720
gaatcacttg	cgcccgagg	gaggaggttg	tagtgagcca	agattgtgcc	actgcactcc	6780
agcctgggag	acaagagaga	aacctgtct	caaaataaat	aaaaacaaa	taaacacctg	6840
agactgggta	gtttataaag	aaaggggtta	actggctccc	ggttctgcag	gctgtacaag	6900
catggtgccg	gcactctgct	ggttgctggg	aaggcttcag	ggagttttac	tcactgtgga	6960
aggcagagcc	agagcaggtg	catcacacag	caaaagcagg	agcgagagag	agagagagca	7020
gggaggtgtg	cacactttta	aatgagcaga	tctcacgaga	actcaccatt	gcaaggacag	7080
caccaagcca	cgagggtgt	gccccatga	cccaaacctc	ccactaggcc	ccaccccca	7140
cattgggaat	tacagttcaa	catgaggttt	gggggggaca	atatccaaac	tatatcattc	7200
caccctgggc	ccccagatc	tcattgtctt	ctcacattgc	aaaatatagt	catgccttcc	7260
cagtagcccc	ccaaagtctt	aactcacc	agcattaact	caaaaatccc	attcccaagt	7320
ccaacgtctc	atctgaagat	gagttccttt	cacctacaag	actgtaaaaa	tgaaaacagt	7380
tatttactgc	tgagatacaa	tgggggcata	ggcattaggt	aaacattcct	gttccaaaag	7440
ggagaaatcg	gtcaaaagaa	aggggctata	ggccccaaag	aagtccaaaa	cccagcagag	7500
caatcattca	atcttaaagc	tccaaaataa	cctccttaaa	ctccatgtcc	catagccagg	7560
gcacactggg	gcaaggggca	ggctcccaag	gcottgggca	gctctattcc	tgcggttttg	7620
cagaattcag	tcccctggc	tgctcttaca	gattggagat	gagggcctgc	ggcttttcca	7680
ggtgcagggt	gcaagctgct	ggtgatctac	cattctgggg	tgtggatgg	ggcgccccc	7740
tcccgagct	ccactaggca	ttgtccag	ggggactcta	tgtggggcct	ccaacccccc	7800
atttccctc	caatgggaag	gctctgccc	tgacagagcc	ttcttctctg	gctcccaggc	7860
tttctcatac	atcctctgac	atctaggtgg	atgggtgcaa	gcttcttcca	ctcttgca	7920
ctgcacacct	acaggcttaa	caccacatgg	aagctgcca	ggtgtatggc	tggaacctcc	7980
tgaagcagca	gcctgagctg	tgactatggc	cctttgagcc	aaggctggag	ctggaacagt	8040
ctagatgcag	gcaggagcca	gtgtcctgag	gctgtgcaga	gcagcagggc	cctgtgcctg	8100
gacaatgaaa	ccattctttc	ctcctcatcc	tctgggctgt	tgatgggagg	gttgtggaag	8160
atctctgaaa	tgcttttgag	gcctttttgc	ctctgaggcc	tatttcttat	tgtctcagtt	8220
attggcagtc	ggctcctttt	tagttatgca	aatcctctag	caagaggtta	ctccactgcc	8280
ggcttgaact	cctctcctga	aaaagctttt	tctttctttg	tcacatggcc	aggetgcaaa	8340
ttttccaaac	ttttatgctc	tgttttacct	ttaaatataa	cttctaactt	taattcattt	8400
atttgctcct	gcatttgagc	atagggaatt	caaagaagct	gggccacatc	ttgaatgctt	8460
tgctgcttca	aaatttatgg	ccacgcttgg	tggctcacac	ctgtaatccc	agcactttgg	8520
gaggcctagg	tgggcagatc	acgagatcag	gagatcgaga	ccatcctgg	caacatgggtg	8580
aaaccatct	ctactaaaaa	tacaaaaaaa	ttagcttgg	gtgggtggcg	agacctgtag	8640
tcccagctac	tgagagagct	gaggcaggag	aattacttga	acctgggagg	cagaggttgc	8700
agtgcagcca	gatcatgcca	ctgcactcca	gcctgggtgac	agaataagat	ttgatctcga	8760
aaggaaaggaa	ggaaggagga	agggaaagaa	tgtcttcccc	ccagatgtcc	tggttcctcc	8820
ctcttatgtt	caaacttcaa	cagatcccta	gggcagtaaa	ataatacagc	caaattattt	8880
gctaaggcat	aacgaaagtg	acctttgctc	cagttcccaa	taagttcttc	atttccatct	8940

10

20

30

gagactcacc	accctggcct	tggcttgtcc	atatcactgt	cagcattttg	gtcaccaatca	9000
tttaaccagc	taatcgggag	gctgaggcaa	gaggatcact	tgaacccagg	aggttgaggc	9060
tgcagtggag	tgtgatcaca	tactgcagct	ccagcttggg	caacagagca	agatcctgtc	9120
tcaataaata	aataaataaa	tacataaata	acttaagttt	atttaaagct	gcattcttgc	9180
caccatggag	aaaggccagg	ccagctcctt	ctctctttct	gcacgtgttc	ctccacctc	9240
agctgcctct	gtccctcaag	gaggaacaga	gggagtagga	aaggccatcc	caggaggccc	9300
agcaccocat	gacctggctc	tggggccttg	tgggtttatg	gattcccagt	gctgagtcac	9360
ccctcacagg	ctcttgtggg	caccttggac	attggtcaga	agcatgtggt	ccccgggaac	9420
acaccttttc	ctgatcatct	gggaagggca	gcttgtgcc	gcgaggccac	ctgttcagcg	9480
ccacggcccg	ccagacagct	gcagccacag	ccttgccctt	gatcagagca	aacaccagac	9540
atgtgtgtca	tgcccccaac	ccatctccag	gggacacatg	tcctttcttg	ccaggccctga	9600
gatgaacaag	agagggacaa	gtcccccaag	ctctctctcc	ttcctgcctc	acccactccg	9660
ctgttagatt	ctcaaggtgg	atgggtgggt	aactagggca	accgaccatc	ctggtttacc	9720
tagaactgag	ggggcatttt	caggaataaa	actgcaaaag	tctggagcaa	acaggagcaa	9780
gttggctact	ctggggctgg	tggagtcagg	tttctctctg	caggccccct	ccccgcagc	9840
atgggtggaa	cccaggacag	gaacacagag	caggccccag	gaccgggctt	gtcacttaca	9900
agtctttttt	tttttttttt	tttttgatg	gagctctgct	ctgtcatcag	ggctggagta	9960
cagtgggtgcc	atcttagctc	actgcaacct	ctgcctctct	ggttcaagtg	atccccctgc	10020
ctcagcctcc	tgagttagctg	ggactacagg	tggcaccacc	acgcccagct	aattttttgt	10080
atcttagta	gagatgagat	ggccaggctg	gtcttgaact	cctgacctca	agtgatctgc	10140
ccgccttggc	ctcccaaggt	gctgggatta	cagggtgtgag	ccactgtgco	tggcccccact	10200
cacaagtctt	aaacctgccc	tcagcacatc	aatgccatct	acaaaaaggt	agagggattt	10260
tccaggcaca	aatagatgaa	agacatagga	tgattgatca	tgtcctgctt	aaacataggt	10320
ctgatgctat	taagaattga	gggtctgggag	cggtgggtca	cgctgtaat	cccagcactt	10380
tgggaggccg	aggcggggcg	atcacgaggt	caggagatcg	agaccatcct	ggctaacacg	10440
gtgaaccccc	atcttacta	aaaatacaaa	aatggccgcg	gcgcggtgac	tcacgcctgt	10500
aatcccagca	ctttggggag	ccaaggcggg	cggatcacga	ggtcaggaga	tcgagaccat	10560
cctggctaac	acagtgaagc	cccgtctcta	ctaaaaaata	caaaaaaat	tagccaggca	10620
tgggtggcgg	cgctgtagt	cccagcaact	tgggaggctg	aggcaggaga	agaatggtgt	10680
gaacctggga	ggtggagctt	ccagttagcc	gagatcacac	cactgcactc	cagcctgggc	10740
gacagagtga	aactccatct	caaaaaaaa	ataataaat	aaataagaat	tgttagtatt	10800
ttgcaggtgt	gacaaatgat	tctgtttctg	tggcagaatg	ttctcaggag	atctcttttg	10860
aactctcatg	gaaagcatca	tgtgtttggc	aacatcacat	ttatttttat	ttatttatta	10920
tttttttagag	acagggtcct	gctctgttgc	ccaggctgga	gtgcagtggc	acaatcacag	10980
ctcactgcag	cctcaacctc	ctgggctcaa	gcaatcctcc	tgcctcagcc	tcccaaagta	11040
gctgggacca	caggcgtgag	ccactgcact	cagcccaatg	taccttcaat	atttacattt	11100
ctggcaagg	tagcaaaacc	ttaacaaatt	ttgaatctag	ataataaat	tatgaggctg	11160
ggtgcagtgg	ccctgacagg	gatggctcac	atctgtaatc	tcaacatttt	gggaggccaa	11220
ggtaggcgga	tcacctgagg	ccaggagtgt	gagaccagcc	tggccaacat	ggtgtaaccc	11280
tgtctctaac	aaaaatacaa	aaaaattagc	cagacgtggt	ggtgcacgto	tgtcatccca	11340
gctactaggg	aggctgaggc	aggagaattg	cttgaacccg	agaggcagag	gttgtgatga	11400
gccgagatcg	cgtcattgca	ctccagcctg	ggcaaaagca	agagcgaaac	tctctctcca	11460
aaaaataaaa	aaaaataaaa	ttaatgaatt	aattaaaata	aaataaaaata	atggatagtc	11520
actgtaaaga	aaaaataaat	gtatatatca	gccaacaagt	gatggaatag	agcaccctat	11580
ctccctggct	ggacagatac	atcccacaac	acctggaagg	cggctccatg	tagaactttc	11640
tggactgctt	gaggtgctgt	gctggagcac	ggtgacagag	gagctggacc	atggacctcc	11700
cccgcccccc	accaaggggc	aggtccccc	gtgggtgggt	tgaggggagg	atccgtatgg	11760
cctctgcggc	ttgggcaggg	aatttggggt	ccaagtactt	ggtgcaaaag	ctggaaagag	11820
ggtttgggtg	ctgaggggcat	atccctggg	ccacatgggg	gcagaagtgg	ggccccctga	11880
agcttggagt	cctgggcagg	ggcatctatt	tgtctgtctg	aggccttcag	tacttgaagc	11940

10

20

30

aaaatggagg	cagaatgtcc	caccttaatg	cccctgattc	ctccaaacca	attccagaga	12000
cagcaagggc	cagaacaggg	atggccctgc	ccagggtcac	gcancagga	agtggccagg	12060
ctgggatctg	aacccaggct	aatccctccc	cttgtccctc	tccaggccct	caacctcgca	12120
tagagccctc	cagctcactc	atcctcgccc	agctccatct	cctcagcttg	taaacccccc	12180
cgggattttc	ctttcttaaa	aaacaaaggc	ttggccaggc	acggtggctc	acgcctgtac	12240
tttgggggtg	gtccccagca	ctttggggag	ccaaggtggg	cggatcatga	ggtcaagaga	12300
ttgagaccat	tctggccagc	atggtgaaac	cctgtattta	ctaaaaaaa	aaaaataac	12360
tgggcatggt	ggctagctac	ttagggagct	gaggcaggag	aatcgcttga	acctgggaga	12420
aagaggtttg	agttagccaa	gacgcggcca	ctccacttta	acctggcaac	agaacaagat	12480
tccgtttcna	aaaacaaaca	aacaaacaaa	taaacaaaaa	aaggcggagc	gcgatggctc	12540
gcgcctgcaa	ccccagcact	ttggggaggt	gaggcgggag	gatcacttga	ggttaggagt	12600
ttgagaccag	cttggccaac	atggtgaaac	cccatttcca	ctaaaagtac	aaaaatcagc	12660
cagggtgtgt	gggtgggtgc	tgtaatccca	gctactcagg	aggctgaggg	aggagaatcg	12720
cttgaaccca	tgacctggag	gctacagtga	gctgagattg	cgccactgta	ctccagcttg	12780
ggcaacaaga	tttgtttctc	taaaaaaaaa	aaaaaaaaaga	ctggcccttc	cccttcagct	12840
cttcctcagg	gtccctgagc	actctacacc	cccgctctaca	ctgagcactc	cacctctgctg	12900
tctacactga	gcactccacc	ctgccatcta	cactgaggac	tccaccccac	tgtctacact	12960
ggctgcctcc	cgccctcacc	tctgtetaag	gccattccccc	gctgcactctg	tcttctagat	13020
tctgcagcct	tcagcaogct	gggcccctcc	tttgtccctc	tgagccacct	ccagcctccc	13080
cctgagctgc	tactcctctc	ccagcagcct	ccacccaagc	ccctccagtc	cccaagctgt	13140
cccttgcatc	cagcactgac	cttccacgtg	ccccttccct	ccagcttcac	agcagggtgg	13200
ggcctccagg	ccctgcccac	tgtgcccac	cacaagttgt	ggtgggagct	ccgaggggag	13260
gcaggggtgt	gcattggactt	gggacgtcca	agtctgggac	cagggggcagc	tgggttggtg	13320
agtgtggggc	gggataaggga	ctttcaggta	gagaggctgt	aggggcaaga	tcgggacggc	13380
ggatgtccct	aaggagggtc	ctgaacctggg	aaatattgtg	cagcttccct	tttgccattc	13440
ctggagctca	gacactggcc	ggctctcacc	ccgcccctcc	tgccaggacac	agctccatcc	13500
cagttagttc	ctagtgtaga	catctccagc	agcaccggatg	ggaaaaggag	tcatcaaagg	13560
tgccagggac	cggaggcttt	ttctggagggt	ggcagaggag	ggtgtgggtc	tcagggtctc	13620
ggctgagggc	aagcgtggga	ggtcttaggt	ctgcaccagc	cccgtaagg	cccctcctgc	13680
tccctgggtg	agtccctagag	ggaaacagcag	cccctaggct	ctagcaggag	tgggttagggg	13740
ctttctctggc	ttctactgt	gccagcagga	tagctggggc	tggcactgag	cccaaagatc	13800
acatgccggg	gcattggcgc	agttaggaac	agacccttgc	caaagctggc	aaagaagacc	13860
ccatgggggtg	cagctgggtga	agctgagagc	tcaatgtttg	ggggagcctg	gcaaaagggg	13920
tcctccctcc	cctctgcagg	ccaggatcgc	aggttttccc	tacatgtttg	taattctcaa	13980
acaatcccat	ggcoactgga	gcaaaagatca	cagtgggcgg	cgccctcggg	agcagtgagc	14040
agggcacgca	gtgcctttga	tgccagagcc	ctcgccccaa	agtcaacaaa	ctctgcagcg	14100
gactttgcac	ccggactttg	ttttcaccat	acaaggaaag	ggacagatca	caggccctct	14160
cgctgccctc	gctgagccgg	aagctgcagc	gtgagctctc	tcaagcccca	tttctaggtt	14220
ccccaggcgc	acccctgagc	ccctactcgc	ctattaagtt	ctcctaatag	cccttcaagg	14280
tcttaatgta	tgtccattag	acagagggga	aaactgaggc	gagggcaagt	gacttgaccg	14340
aggttcctcg	gcgagcaggg	cgtggagctg	agaacctcgt	tattactgct	ccccacacaa	14400
ccctctgggc	gttcttggaa	gaaggctgag	ccccgggggg	gccagagtga	cccaaacacc	14460
atgggcccgc	tcgggtaaca	cgtgcggcca	cgaaggggca	gcagtttccc	gcccggccgg	14520
gctctctccg	gcgctcagta	tccgtccag	gccaagaaga	agaaaactcg	ggaggagggc	14580
ggagggggct	gcgtgggagg	gcgtgggaaga	tggacgtggc	caggggagtg	gcagctgcac	14640
acagtggatg	ctgttaagat	gaagggaag	aacgtgggct	ccgagatcac	tggacacggg	14700
tccacctttc	ttcccgctca	ctgcatggcc	ctgggcgggt	tgttgaaccc	ttggaaacct	14760
gtttttccct	ttttcccttt	tttttgagac	aggtctctgc	tctgtggccc	agactggagt	14820
gccgtggcac	gatcttggct	cactgctgcc	tcccaggctc	aagtatcct	ccagctcag	14880
cctcctcggt	agctgggacc	ccaggtatgt	gtcaccacag	ccggctaatt	tttgtatttt	14940

10

20

30

tttgtagaga	cgggatttgc	ccgtattgco	caggctgggc	tcaaaactcct	gagttcacgc	15000
gatcttctcg	cctcagcctc	ccaaagtgc	gggattactg	gcattgagcca	ccgcacccag	15060
cagagacctc	agttttctaa	cctgtgcccag	caggaataat	gatagctgcc	tagcttggct	15120
gtgctgggaa	ttaagtaaga	tgaccgggta	gcaaataatga	agtattactg	gacacagagg	15180
gccccaggct	gggttagcag	cggtggtcag	ggctgctgct	tctggccctg	agctcgaagg	15240
agggccctca	ttaccacctg	ggtgagtcct	cgtccaagcc	tggcaactgct	gcgtgggaat	15300
aacttctgcc	acccaagttg	gcagattgtg	tgcaaaagta	agtcctgact	ctgtgggggtg	15360
gacttctgagg	cctcttcatc	ggacctgctt	ccggtgactg	cattcgcacc	tctctctgtt	15420
cctgggttaa	cacagcccag	ctttcctcct	gctgagccct	cctggggcct	gctgtcacc	15480
tctgtccgct	gtgcctcgca	gtgccactcc	ctgtaccctg	aatactttgc	cctgcctctc	15540
cacccagctg	agagtcagg	ccctgtgag	gctctgccc	gcccgtctc	cgggtttctg	15600
cctctgctga	gcacttccct	gcattgattg	ttctgagagt	ccccccagcc	tgtgagcttc	15660
tcaggactgg	gacagcttct	caggaccgag	gcttctctgt	ctgcttgcaa	ttttacaggc	15720
gggcacattt	tcccttggcc	aacatcagag	actggacatc	tcagatctg	tgtagccac	15780
tgagcacc	ggcaccocag	caggtagctc	tgtaaccaac	ccattctgta	aagctgaggc	15840
tcagagaggt	gaagcgccct	gcctggggcc	acagcctgog	tcagctgcag	agccaggagc	15900
tgagatctgc	acctgcggct	ctgctcacag	ggctcctgcac	agactgctgc	tggagccacc	15960
tatgtagagt	caagagagtt	catgttaact	ccctctcaca	tccctcagcc	aggggtggggg	16020
ctgacgatag	acactcaggg	atggcctacc	ctccccaaca	acccccgtca	ggtttgccc	16080
atctccttgg	aagaaaagtt	ctgggcagaa	ttccaccgtt	ggcctggcct	acactctcct	16140
tagtggctta	ggaccctcag	cggtggataa	gttgtgggca	gaagagatgc	aatcaggatt	16200
ctcaccact	cacccttgc	cagccccaat	aagctcaata	agctgggctc	ggtctgagga	16260
agtgctcagg	aatgtgcaa	atggcctggg	acagccctgt	gttcttttca	gtaagggtgc	16320
tgaagggtgag	gctgaaagtt	ggagaaacag	aagccagtgc	ttatggtttt	aattaagata	16380
atggaatgta	tgtatgtatg	tatgtatgta	tgtatgtatt	tatgtattta	tcttttagaga	16440
tagagtctca	ctctgttgcc	caggctggaa	tgcggtgaca	caatcatagc	tccctgcagc	16500
ctcagcttcc	tatgcccaca	tgatcctcct	acctcagcct	cctgagtagc	tgggactaca	16560
gacacacgcc	aactatgcct	agctaatttt	tatttctatt	ttttgtggag	actgggttct	16620
cactttgttg	cccaggctgg	tcttgaaccc	ctagcttcaa	gcaatcctcc	tgcctcagcc	16680
tcccaaagtg	gagggattac	aggtgtgagc	caccacacct	ggcctggaat	ttatttgtat	16740
tctgcttata	aaattaatac	attcttattg	cagaaaagtt	tgaaaataaa	agaaaggaca	16800
aagaacaaaa	agcgtatata	atttcacagc	tcagatctca	ctgctattaa	catttttatt	16860
tactttcagg	cttttttctt	tctaggtaca	tatgcagaga	ttattttatt	ttattttatt	16920
tattttat	tttattttat	attttttatt	tcattatttt	attttatttt	attttattat	16980
tttttagagac	agggcctcac	tctgtcacc	aggctggagt	acaatggagt	gatcatagct	17040
cactgcagcc	tcaaacacct	gggctcaagc	aatcccccca	ctcagccttc	tgagtagttg	17100
ggactaaagt	gtgagctctg	ctaatttttt	ttactttttg	tattgacaga	ggtctcacta	17160
tgttgcccag	gctgatctca	aactcctggg	ttcaagcgat	cctccacct	tggactccca	17220
aagtgcctggg	attacaggca	tgagccacca	tgcttggcct	aaaatgccac	tttttgtcat	17280
ttactaaaa	cccatggaca	ctttgacatg	tctgtattct	atgctattga	tctgactgtt	17340
ggcatctaca	tcattatggc	catctatcat	ctatcataat	ccattttaac	attaaaattg	17400
tgtctgtgct	tagatttttc	tggcctgtct	cctatttgta	ttcttccaga	taaattttag	17460
aatcatttta	tcaaatcccc	cttgccagaaa	aagccctatt	ggattttggg	tgaaaaatac	17520
tgaattttta	cattaactta	ggaaagggct	gggcacgggtg	gctcacgcct	gtaatcccta	17580
cacttttoga	ggccaaggca	ggtggatcac	ttgaggttgg	gagtttgaga	ccagcctggc	17640
caacatggtg	aaactcggtc	tttactaaaa	atacaaaaaat	tgccaggcgc	attggctcac	17700
ctgtaatccc	agcactttgg	gaggccgagg	tgggtggatc	acgaggtcag	gagatagaga	17760
ccatcctggc	taacacgggtg	caaccccgtc	tctcctaaaa	atacaaaaaa	ttagccaggc	17820
gtggtggtgg	gcgcctgtgg	tctcagctac	ttaggaggct	gaggcaggag	aatgggtgtga	17880
accagggagg	cggagcttgc	agtgagccaa	gatcgcgcca	ctgcactcca	gcctggggca	17940

10

20

30

cagagtgaga	ctccatctca	aaaaaaaaata	ataataataa	tacaaaaaatt	agccggggggt	18000
cgtggcgtgc	acctataatc	ccagttactt	gggaggctga	ggcaggagaa	tcgcttgaat	18060
ccaggagggtg	gaggttgcaa	tgagcagaga	tcgtgccact	gtactccagc	ctgggtgaca	18120
gagtgacact	ctgtgaaaaa	aaaaaaaaaaa	ttctgaagga	ttgagactct	tagactctta	18180
ggtcttctca	tccaagagca	caatatagct	tttcatgtat	tcaagccttt	ttcaatgcac	18240
caacagaatt	ttacagtttt	tttcatgata	tcctgctatt	tcttataaaa	tgtattccca	18300
gatattctgc	atgttttccg	gttgttttgt	aataaatatt	tttcatttgt	cattatttcc	18360
taattggctg	ttatttgtat	atatgacatc	tgttgaattt	tttgattact	ttgaaaatgg	18420
ccattctttt	gtgttttttt	ttaactttct	attttgagat	aattttgact	tacagaagat	18480
ttgcaaaaat	agtacagaga	gttctctgtt	cccccttatg	ttaaccagct	ttctctttat	18540
gttaacatct	tacataacta	cagaacaatt	gtcaaatcta	agaatcaacc	tgggcacaa	18600
gctattaact	aaactgcaga	agctgttcag	atctcaccag	ttcttctact	gtcccccttt	18660
tctcttccag	tgttcaatcc	ggaatcctac	attatattta	gttgtcattt	ctctttgggtg	18720
tcttccaatc	tgtgacagtt	cctcagtcct	tctttgtctt	tcatgacttt	cattttttta	18780
tacttttgaa	aaatactggc	cggttggttt	gtagaacgcc	ctcagtttgg	gtttgcctga	18840
agttttttgt	gattagatcg	aggtcatgca	ttattggaga	gggtgccacc	gcctcgatgt	18900
gcaagctcaa	tgcatcatat	cagagggttt	gtaatgtcag	tttataccgc	cggagaccct	18960
aacctggagc	atltctgtaa	ggtgctgtct	gccaggatcc	tccactagaa	agttactatt	19020
tttccctttt	taattactga	atgtctgagg	ggaaatactt	tgagactatg	caaatatcct	19080
gtttctcgct	taactcggc	tcactaagtt	tagcattcat	ctatggatct	cgcttatagc	19140
aagtattact	gtggagtctt	aatggtaatt	ttctgtttct	ctcattcctt	caacctttat	19200
taatatgctt	cttctcact	tattcatttt	gtttcagttg	tttataccaa	catggatttg	19260
tggtatattg	ttttattctt	tgggttgcaa	ttgaatccta	tcattatttt	gttagtcagt	19320
tgttccatcc	gaccttggtc	attaggagcc	cttgaaattt	ggctcccatg	cctttttttt	19380
tttttttgag	accgagtctc	actctgtcac	ccaggtttga	gtgcagtggc	atgatcttgg	19440
cttctctcaa	cctccgcctc	ccaggttcaa	gcaattctcc	tgccctcagc	tcctgagtag	19500
ctgggtattat	aggcgctcca	ccaccttgcc	cggctaattt	tttgattttt	tagtagagat	19560
gggggttttat	tatgttggtc	aggctggctc	caaacctctg	acctcaggtg	atctgcccgc	19620
ctgggcctcc	caaagtgtcg	ggactacagg	ctgagccac	cacacctggc	ctctatgccc	19680
attttaacat	gcccgtcttt	tctttttctt	tcctacttcc	tgtagctgta	agaagctcca	19740
ggatacatct	ctgctgccc	agacttagcc	tcaatcagtt	ctcagaaaag	ctctgggtct	19800
ttttatggga	tacttagaaa	actagctctg	tatggcctgg	cgcggtggct	cagcctgta	19860
atcccagtag	tttgggaggc	cgagggtggc	agatcacaga	tcacgaagtc	aggagatcaa	19920
gacctacctg	gctaacatgg	tgaaactctg	tctctactaa	acatacaaaa	aattagtcca	19980
ggcgcggtgg	cgggcgctcg	tagtcccagc	tactcaggag	gctgaggcag	gagaacggca	20040
tgaaccgggg	aggcgagct	tgcagtgagc	cgagatcggc	agccactgca	ctccagcctg	20100
ggccacagag	cgagactccg	tctcaaaaaa	aaaaaaagga	aaaagaaaaa	agaaaactag	20160
ctctgtatgc	tagttttttt	tttaagacag	ggtctctctt	gccccagctg	gagtgtagca	20220
gcacgatcac	agctcactgt	agcctcaacc	ttctgggctc	aagcaatcct	cctgcctcag	20280
tctcctaagt	agctgggtct	acaggcatgc	accaccgtac	gtggcaattt	ttaaaaactg	20340
tttgtagaga	tggagtctcc	ctatgttgcc	tggctctggaa	ctcctggcct	caagtgatcc	20400
tcttgctcgc	gcctcccaaa	gtgctgagat	tacaggcatg	agccactgta	cctggcctgg	20460
ccaaggctcg	tctttttttt	aaagaagttg	ttgtatagtt	gttttttttt	ttattttttt	20520
ttctgagacg	gagtctcgct	ctgtcgccca	ggctggagtg	cagtgggtgcg	atctcggtct	20580
actgcaagct	ccgcctccca	ggttccacgc	attctcctgc	ctcagcctcc	cgagttagctg	20640
ggcctacagg	cgcccgctac	cacgcctggc	taattttttg	catttttagt	agagacgggg	20700
tttccacgtg	ttagccagga	tggctctgat	ctcctgacct	cgtgatccgc	ccgcctcggc	20760
ctcccaaagt	gctgggatta	caggcgtgag	ccaccgcgcc	cggcctgttg	tatagttttt	20820
atctcgagtt	ttctagcgat	ttaatcatat	tggttacaaa	aaaggatgat	tttactacct	20880
cctttccaat	gtttctacat	attttttcat	tttatctaac	tgcattttta	aataaacttt	20940

10

20

30

taatthttaga	atgggtttcat	atttacagaa	aatgtgcaaa	gatatgtacag	agagttcctg	21000
tgtactccac	accoggtttc	cttattatta	tcttaacgtg	atacacaatt	aataaaccag	21060
taacattatt	attcactgaa	gtccacactt	tctttttttt	ttttttctgag	acggagtcta	21120
cttctgtcac	ccaggctgga	gtgcagtggc	gcaatctcgg	ctcactgcaa	cctccacctc	21180
ctgggttcag	gcaattctgt	ggctcagcat	cccaagtagc	tgggaataca	gggtcccgcc	21240
accacgccc	gctaattttt	tgtattttta	gtagagatgg	ggtttcacca	tgtagccag	21300
gatggtcttg	aactcctgac	ctcgtgatct	gcctgcctca	gcctcccaaa	gtgctgggat	21360
tacaggcgtg	agccaccgcg	cccggcgctc	atactttctt	tagatatcct	tcctttttac	21420
ctaactgtct	tcttctggtt	caggatccca	tccagaaagc	aacattaccc	ctcgcctatca	21480
cgtcttcaca	ggctccccct	gacgggaaga	gttcctcaga	ctttccttgt	ttttgttgac	21540
cttgacagtt	ttgaggagga	ctggtatcct	agtctgtttt	gtgctgctat	cacagactag	21600
ctgagaccga	tacatgatac	atgaaaaaaa	atgtattctt	acagttgtgg	aggctgggaa	21660
gttcaagacg	aagttgctgg	ttggttttgt	ctctggtttc	aagatggcgc	cttgctgctg	21720
catcctctgg	agaagaagaa	tgcggtgtcc	tctcactgca	gaagatggaa	gcgctaaaag	21780
gaatgaactc	cctttgccaa	gccattttat	aatgggcatt	aatccacaaa	ggatgaaacc	21840
ctgagaaaca	tcaagcttta	aagcaactgg	tctcaacctt	tttgggtctca	ggagcccttt	21900
ataactctta	aacgttttga	ggatcccaaa	aaaaggcttc	tacaggttcc	atcttttaat	21960
atttaccata	tcaaaaatta	aactgaaaaa	attttaaatt	atttattcat	ttaaaataac	22020
aaggataaac	ccattacatg	ctaacataaa	tcatgtatct	tatgaaaaat	agctatatct	22080
atcaaaacaa	aaattagtga	gaagagtggc	atgtataaatt	ttttttgttt	atttttgttt	22140
tttagatgga	atcttattct	gtcgccagg	ctggagtggc	gtggtgtgat	ctcggtctac	22200
tgcgaagctc	gcctcccagg	ttcacaccat	tctcctgcct	cagcctcctg	agtagctggg	22260
actgcaggtg	cctgccacca	cgcccggtca	attttttgta	tttttagtag	agatggagtt	22320
tcaccgtggt	agccaggatg	gtcttgatct	cctgaccttg	tgatccaccc	gcctcagcct	22380
cccaaagtgc	tgggattaca	ggcttgagcc	actgcgtctg	gcctaaattt	ttgtgaatgt	22440
ctttaatgcc	tgccttctca	tatttggttc	tgcattcaag	ttattgcaaa	atggtgtgtt	22500
ggttgaagtt	tgtaaagaaa	atgtggctc	atacagttgt	gtagttggaa	aggcaagagt	22560
atthttgattc	tctcttcaaa	caactatgga	caacctgctg	ttacaaaacc	agaatgcaaa	22620
aagttgtagt	aaatacaggt	taggtgtagt	gtggaatctg	aaagcatgtg	aatgaacctt	22680
ctgagttttg	taacattaaa	gtccagtgtc	gttaagctac	tgtgatagca	tatagcattg	22740
tcctaatact	ggaattagta	tcagaagtgg	ggtgctactg	ttataaata	aaaagaaata	22800
aataaatcat	gtgatactgg	ctcagaagtc	aggcagtagg	ctgtgtggaa	cctgacatca	22860
cgccatgtaa	tacattggca	accatttgat	ccagctgtct	gtcatgatga	cttggaaggt	22920
caaccacata	cttacagagc	ctgtagacat	aggggaaaat	agtataaaac	agaatactaa	22980
cagtggaact	tggttctctg	cagttgcatt	tagccaaata	ttaaacaaaa	gagatattct	23040
tgggcagcaa	ctggaccatc	ttcaagtaaa	agtgaaggt	aataaacaga	gtccagacat	23100
ttgtgcccac	gcgggttaag	aaaaatccag	ttgctcttag	acaccgtata	tgaaaacaac	23160
gctgaaaaca	agcctttgag	tggtaaaggc	cgattaacac	tcagcgcggt	aacaaagacc	23220
aggtgggcta	acccgaaatg	aatgagaag	cctgtggtga	tgaggaggca	gagaagttaa	23280
atcaagtttg	agcatttctg	ttaggagagt	ttgggctctg	attacttgca	catgcaaacg	23340
aactggaaac	aaacagatca	gatgtctacc	acttctctga	gggaattgca	ttgccaaaga	23400
agtcatgaaa	gcagactcta	tactgattag	gcattaaaac	aaaaacaatc	tttaggcccc	23460
taaacttgca	tgggcaggaa	gtgggctgtc	aaagctgttc	atcctctaag	gtggacctag	23520
ttcctagctc	ccagtataca	cttcagatgt	ggccctggag	gacactggac	atggaggacc	23580
tcccagagga	tgaggctagg	gcttcatttc	tccaatgacc	tcagctgcct	ctatttcccc	23640
ttcttctctc	ggaagtccca	tcactgttat	tattattatt	atcatcattt	ttattttgag	23700
ataaggtctc	gctctgttgc	ccaggctgga	gtgcagtgtc	atgatcatgg	ctcactgcag	23760
ccctcccagg	ctcaagtgtg	cctcctgcct	cagcctcctg	agtagctggg	agtacaggca	23820
catgccacca	tgccttggtc	tttttttttt	cagtagagat	agggtctctc	ctatgttgcc	23880
agggctgatc	tcaacctcct	gggttcaaga	gatcctccta	cctcagctcc	tgagttagctg	23940

10

20

30

ggattcgggt	gcacaccacc	atgccaaacta	atttttaatt	tttttttgta	tggacaggat	24000
gtacagtgtt	agaaatggat	tgcttgcaga	ggcaggagga	tcacttgagc	ccaggagttt	24060
gatcacactg	tgaaccatga	tcgcacccct	gcactccaat	ctggggcaaca	gagtgagacc	24120
ttgtctcaaa	aaaaaaaaaa	aagagagaga	gagagagact	caaagatagg	caaaaaagtg	24180
ggaaagcttt	atagtggaca	aaaagggaacg	ctctaagtct	gccctatttg	catgggtgctg	24240
aagggtgggct	aactagagat	aggggggtact	atgtgggtga	ctatgggtgc	atctttggct	24300
ttccctgggt	gacctaagt	tgggaagcagg	gacaaaaatt	aggggaagctg	ttagttattc	24360
atcacgttct	ggcagtagtg	gactggttgt	gatagaagtt	attgttttgg	ccagggtgcgg	24420
tggctcatgc	ctgtaatcct	agccctttca	gagttcaacg	tgggtggatc	aggaaggagg	24480
gaggatttgg	gaggtcagga	gttagcctgg	ctaacctggc	gaaatcccat	ctctactaaa	24540
aatacaaaaa	ttagctgggc	gtggtggtgc	atgcctataa	tcccagctac	tggggacgct	24600
gaggcaggag	aatcagttga	acctggggag	goggagggttg	cagtgaacca	agatcgtgcc	24660
caatttcac	tcaaaaaaaa	aaaaaaagtt	atcgtttagc	ttcctcgatt	gttactggac	24720
gtagtaatct	ggcttcctgc	aagtctaact	ttcagcagac	tggctacatg	ggctgtgtac	24780
tgtagataag	gcagtaagta	aagcaaaaat	tgatagagca	tcaaggataa	atagaaaatc	24840
cgtaatcaag	cagaagattt	gaacacttca	ctttcagtaa	ctgataaaaa	aagtagacaa	24900
aaaaaatcag	taaggatgta	gaagatttga	acaacgtaat	taacaaactt	gacttgattt	24960
acacgtctag	aacctgcag	aacacacact	ttttcaagca	tactcagaac	atttatataa	25020
agtgaccata	tgggtggacca	taaagcagtt	tcaacaaatc	tcacaggagt	aaaataacag	25080
accgtgtttt	ctgacogtaa	gtacagttaa	cctagaaatt	gaaaacaaaa	agctagaaaa	25140
accccatgta	tctggaaatt	ttaatatata	ctttgaaata	acaaatggat	cagagattaa	25200
ttcaaatagg	aatttagaaa	taccttgaac	tgaaaaataa	tgagaatact	atacccaaaa	25260
actgtggggt	gcagctgaac	agtatataga	cgaaaagtat	actcatatgt	gcatacctta	25320
aggagcgggg	aggattgaaa	gttaatggga	ggcaaaaagca	ggtggatcac	ttgaggttag	25380
gagttcaaga	tcagcctggc	taacagggtg	aaaccccatc	tctactaaaa	atacaaaaaa	25440
ttatccaggc	gtagtggagg	tgaggcaaga	gaatcgttgg	aaacccaggag	gcagagggtg	25500
cagtgagccg	cgattgcgcc	actgcacccc	agcctgggag	acagagcgag	actccatctc	25560
aagaaagaaa	aaaaaaaaag	aaaaggccag	gcgcggtggc	tcattgcctgt	aatcccagca	25620
ttttggggagg	ccgaggtggg	cggtatcaga	ggtcaggaga	tcgagactat	cctggctagc	25680
acggtgaaac	cccgcctcta	ctaaaaatac	aaaaaaatta	gccaggcgtg	gtggcggttg	25740
cctgtagtcc	cagctactca	ggaggctgag	gcaggagaat	gtcatgaacc	caggaggcag	25800
agcttgcagt	gagccgagat	cgccccaactg	tactccagcc	tgggcaacag	agagagactc	25860
tgtctcaaaa	aaaaaaaaaa	gttaatggga	taaacatoca	tctcaagaag	ttagaaagga	25920
atgacaaata	aaccaaaaaa	aaaaaaatca	aaagaagaaa	atcataaggt	caagactata	25980
aagagagtgg	ctgggtgcag	tggctcaggc	ctgtaatctc	agcattttgg	gaagcagagg	26040
tgggcagatc	acttgagccc	aggagtcca	gaccagcctg	agtaacatag	agagacctca	26100
tctttgctga	aaataaaaaat	aaaaaattag	ccaggcatgg	tggtagctgag	gtgggaggat	26160
cacttgagcc	taggaggttg	aggctgcagt	aagccatgat	tgtgccactg	cacttcagcc	26220
tgggtgacag	agtgggaccc	tgtctctaaa	aaactaaaat	aaggctgggc	gcgggtggctc	26280
aaatctgtaa	tcccaccact	ttgggaggcc	aaggctgagg	tcagcagttt	gagaacagct	26340
tggccaacaa	gatgaaacct	catctctact	aaaaatacaa	aaaattagtt	gggtgtggtg	26400
gcattgtcct	gtaatcccag	ctacttagga	ggnnnnctnt	ngattatatt	ttctccttcc	26460
tacgtcgtaa	ttggactgaa	ttcagaatga	tgactctcat	tggagctctt	cctgtctcct	26520
aactacagtg	gcttcggacc	ccactctggt	tttcaattca	cccctctgct	gctcatacga	26580
gtagatactt	ccctccttct	ttctcacttg	ttgctcttcc	tcaacccccc	ccgttggtgt	26640
ccctcctctc	ttatcttttt	ctcgcgacac	ctgcgttctc	ttgcccctct	atcatccctt	26700
tctcgaggcg	gtcctttcct	ttatccagct	taaatacctt	ctcctctgtt	tatttggggg	26760
ttgggttttt	atctctcacc	ctccctctaa	tttctttcct	ctttccgcac	ccatcaagcc	26820
tctcgtgggt	tctcttctct	tactctcggg	tccccccctt	ctcccttctt	ttttttcttc	26880
acccccccaa	gcgccttgcc	ttttttttct	ttgcccttta	ttcccccc		26928

10

20

30

<210> 7
 <211> 29430
 <212> DNA
 <213> Homo sapiens

 <220>
 <221> unsure
 <222> (4336), (4345), (4349), (4392), (4447), (4490)
 <223> Identity of nucleotide sequences at the above locations are unknown.

 <400> 7
 aggggaaggg cgggctccgt agctcacacc tataatccca gcactttccg aggagagagg 60
 atcatctcag gccaggaggt caagaccagc ctgggcaaca cagcaagacc gcactctctac 120
 aaaaacttct tttaaagctt aaaaaaaaaa aaaaaagcaa agaggacagt tcaggagaaa 180
 agcctgtaga ggcagcacac taaggaggag acgcagccca ggcaccagga ggggctggcc 240
 atgggcactc actcctccag caggcgagtg cccagacca gctggccac ccagacacc 300
 aggacacggc ctgaatggct ccgtattcac gtgggtggta ataaacaagc aatacacata 360
 gccataagg acaccttagt aatgttacat cataaacgct gcagatcagg gaaatggtgc 420
 agggtaagt ggggtgggg gctgcagtgc acatgagaag tgggtcggg ggctgcagtc 480
 tacctgagac agagcaggcc ttgctgggaa agaaggagcc ggcaggcctg ggcaaggctc 540
 ctggggggg agcacactgg agcagagtgt gggggtagca tggcgggtgc tggctcctctg 600
 ggcgccctcc caccacgtca tgtgcccatg tgcccaaggt ctctcgtttc acagccccc 660
 gaagctcagg ggtcacagct acacagcccc cagatacctt ggcctgcccc aggtcattcc 720
 atccagtgat ggacctgctg acctctagcc tgacctctgg gcagcgtaat ttgagaagga 780
 ggagaaggga gggcaacaga cctggggcga tgagggatgc acagggtggc agacacctga 840
 ggctgcacct tggagcctca gttctgggtg tgggtggggg atggacagga tgagggtgga 900
 agcagctggg cccggccacc atcacacccc aggacccacc agatcaccat gaaaaaccga 960
 atgtcaactg gcagcccaga gtgcagaaca aacctttcag aaacacgggt gtgactgccc 1020
 catcatgaac ataaaaataa tacgcctctc cccagggat caccctgca ggagtttgtc 1080
 ccaagaaaca ccagaagaa ggaaaacgtc tgagtcacaa tatttctgta ggcttattt 1140
 gtaatagcaa aaaaaaaaaa aaaaaagaa caatctccag cggcaggggt aactagacta 1200
 ttgtctccgt ggaaaggtag caccaattaa ctagtaacaa aatgactgag gtaacaacaa 1260
 aacgttcgac atgtcaacac caaaaaccac acaccagca taaccgtgaa ccatgattc 1320
 tactagaatg aatggcagtt atgagaaagc accagcggag acaaagattg aaaaagttaa 1380
 ggtggcctca ttagggagac aagtctctgg gtaatatatt gtaatactgg taaatatata 1440
 gtttttaata tattttttta ttccaaattc catatatgtt cctatgaagc ttttctgca 1500
 aatatttttt tcaggacagt acatcacaaa ggcaaaaggg ccaggtcagc tctccagctg 1560
 agagtgaaca cttcagagca gacggcagac tccagggtta gcaagcctgg ctgagacctg 1620
 gccatgaca atcactcaac cctctgacc tcaacatcct gtctgtgaaa tggggataat 1680
 tactgcacct ccacatcaca gagtgcgagg cttaaacagg atgcttcata gaaaagcgt 1740
 caagaggtaa cagccgggag ggggtagtgg ttttcattaa ttaaatgttg cttcatcca 1800
 gccctgggac agctccaaca caaagcacac accatccact cagactcagt tgcctggatt 1860
 caaagcccgg cctggcctcc agctgtgaga ttccgggagc gatttcccat ctcccagagc 1920
 ctcaagttcc tcattcatga aacaggaagt gatcattcct tttattttta tttttatttt 1980
 tattttgaga cggagtttca ctctagtgtc ccaggctgga gtatgatggc gcaatctcag 2040
 ctactgcaa cctcggtccc ccagtttcaa gcgattctcc caccctcagtc tctgagtag 2100
 ctgggattac aggcacacgc caccacgccc agctaatttt gtatttttag tagagacggg 2160
 gttttgccat gttggtcagg ctggtctcga actcctgacc tcagggtgac cggccgcctt 2220
 ggcattccaa agtgcctggg ttacaggtgt gagccaccaa gccaggttga caactgcttt 2280
 taaagacacc tctggctgct gtggaaaaca gcctggtagt gcctcaaaaa gttacacata 2340

10

20

30

gaatgatcct	atgaccagta	attccactcc	tacatatata	cccaaaagaa	ctgaacccct	2400
ctactcatgt	atgtacacat	acaggtaac	gcattgttaac	agcagtggtc	acaaagccaa	2460
aacatggaaa	cagctcaaat	gtccataacc	gatgaacgga	taaatgaaac	gtagtctatt	2520
caccacctga	cggaggtgag	agggggccata	aaaaggaatg	atgcataaaa	acgaatatta	2580
tggccaggta	tgggtggtca	cgctgttaat	cccaggactt	tgggaggctg	aggcgggagg	2640
atcacgaggt	aaggagttcg	agaccagcct	ggccaacacg	gtgaaacccc	atctctacta	2700
aaaatacaca	aattagctgg	gcattggtgga	ggggcgctgt	aataccagct	actccggagg	2760
ctgaggcaag	agaatccctt	gaacctggga	aacagagggt	gcagtgagct	gagattgcac	2820
cactgcactc	cagcctgggc	gacagaccaa	aactccgttt	cggaaaaaaa	agaaaaaatt	2880
agccagggtg	ggtggcgggt	gggtccctgt	aatccagct	ctacttggga	tactgaggca	2940
ggagaaccac	ttgaaccggg	gaggtggagg	tagcggtgag	ctgagattgt	gccactgcgc	3000
tccagcctgt	gtgacagaag	gagactctgt	ctctaaaaaa	caaaaaacaa	aaaggccoga	3060
cgcggtgtct	tacacctgta	atgccaacac	tttgggaagc	caaggcaggc	agatcatctg	3120
aggtcaggag	tttgagagca	gcctgggcaa	cacggtgaaa	ccccatctct	actaaaaata	3180
cagaaattag	ccagggtgtg	tggcacatgc	ctgtaatccc	agctactcgg	gaggctgagg	3240
caggagaacc	gcttgaaacc	aggaagcggg	gggtgcagtg	agccgacatt	gcaccattat	3300
actccagcct	gggtgacaga	gtgagattct	gtctcaaaaa	aaaaaaaaaa	aaaaaaaaaa	3360
ctaaacaaaa	gcaaaaaaac	caatgagtaa	tgttgtaaac	tgaacttcat	cccaatggga	3420
atgcagataa	tttgtttaaa	aggcaccatg	cacactgggc	aggctggctt	ccctgggaa	3480
cgtcttcttt	tgcctggatt	cccagttggt	ttaatcgggc	gtagaacact	ttcttcaatc	3540
cggtgattcag	gcacccctgc	tcagcacaaa	ctcagtaac	cccgcactct	gctgtgggtt	3600
cttggcacta	ttaggagaat	gtgaggggtg	gattcagatc	tatctctagt	gggtgcattg	3660
ctgccactcc	caggaaacgc	cacttctggc	aagtcaagtg	cagagaaagg	ccagctcgtg	3720
gcccctcctg	ccttgagtc	caggaccctg	gatcagtcct	accggagca	gaatcaggag	3780
tttgaatacc	caagtgcaca	caatctcatt	ttaaccatg	taagcatatc	caatatttat	3840
atatagaatt	cataacagat	gtctgggctt	ccattccaat	agcctatatt	ttacactggt	3900
tatttacatg	gttacaccaa	acaagactca	attcaaggta	accacatcct	ttgtactat	3960
acccaaataa	gcaacatttt	cagtcctatg	cttatatata	ttcaccagc	attacactag	4020
gcctccaact	gctcatcgga	gcaagctgca	gcctggacac	aagctagaga	ttaatcagtc	4080
aggaatgatc	ctgcgtccag	tgccagcatg	atggaagaga	cagagaaaca	gaagacatca	4140
gggtccocaga	gtcaaggagc	ctgcagggtt	ggtgggcagg	atatacacac	atacacacac	4200
acacgcacac	acaaaaccac	ccaagaagaa	aaggtgggat	gaatgcatgg	acaggtaatg	4260
cctggagcct	ggggatggat	aagctgactg	caggtggccc	aggcaggctt	cctggaggaa	4320
gaagacctgg	ctgtangtgg	ggtangcang	ctttctaaat	ggggaaaatc	tggctgtggg	4380
tggagtgggc	angtttcoga	aaagaagaaa	agctgactat	gggtacacct	ggctgttggg	4440
ggaacangca	ggcttcttgg	aagaagaaaa	tctggctgtg	ggtggatcan	gcaagcttct	4500
tggagaagat	aaacctgact	atgggtggac	caggcaggct	tcctagagga	agaagaccgg	4560
ctgtgggtga	accaggcagg	cttcttagac	agaggaagat	ctggctgogg	ttagagtggg	4620
caggcttcta	agaagaggaa	gggtgactg	tgggttagacc	tggctgtggg	tagactgggc	4680
aggcttctctg	gaggaggaag	agctggagca	ttgaaaaaca	aacatgactt	ggtgaatggt	4740
gagcatgccc	aggcctgac	cccagaggca	attacgcact	caagttactt	aattctactc	4800
acaatgcctc	acaaacaact	tctctgacac	ctaacacagc	tctgggcacc	ttctagcttc	4860
agctcctcaa	agcagttatt	cagctacta	ccctgcacac	ctcctcacac	cccaacccca	4920
gggacaggag	ttctgccaga	tgccaaagct	cctgatgcca	aagcctgggt	ctgcttccgg	4980
gctcctcttg	gtctaaactg	ccaccccgca	tggcatgat	gtgcaaaaac	aaggcttctg	5040
aatctgccct	gatgcctggc	ggagcgagtc	cctccgatit	cgtctccttc	agaaacacct	5100
gggtgcccct	ggtcctgtta	tacccccaac	acattctaca	gtcagctccg	caagttccac	5160
aaagatcaac	gctggcggtt	ttatggcatt	ttatttacag	tttttacaat	ataaaaaagg	5220
aaggatgcca	cagctcagcc	agcaggacag	acagagatct	atgatgcttc	tgctgcacca	5280
ttgtttgtgg	tcaagaaagt	ctgttttcaa	tgatttatta	aattgtgggtg	ggagatggat	5340

10

20

30


```

gggtggcagtg gttaccagca acatgaatgt tcttaatgcc actgaacttc acacttacaa 5400
atggttacga cgataagtgt tatatgtatt ttaccacaat taaaaacagg taaatgcagg 5460
ccgggcacgg tggctcacga ctgtaatctc agcactttgg gaggccaagg caggcagatc 5520
acctgaggtc aggggttcga gaccagtctc gccaacacgg tgaaactctg tctctattaa 5580
aaatacaaaa attagccaga tgtggtgggt catgcctgta atoccagctt ctcaggaggc 5640
tgaggcagga aaatagcttg aaaccgggag gcagagggtg ccatgagctg agattgtacc 5700
attgcactcc agcctgggtg acaaaagcaa aactctgtct caaaaaata aaataaaata 5760
aaaataggta aatgcaaaca tatggtatag taatattatg ggctattatg agctacaaaa 5820
aagaatgact tgggactaca gttacagccc tcattcagga atttgtttta aatgtgggtt 5880
ggctcgtaag gcatgtacac aacattttga cgttcaaata ttccatagatt tggacagtga 5940
gcaccctctc aagctggctc ttctgtccca gaggtcccca ccagtcctcc agaacttctt 6000
tgctttctta cacaataaga tgcccatgc tcggcttgta cctttccttg cccagccct 6060
agaaccagct tcttcgtgga caagctctga ctcccttggg tggagaatgg tattcagaaa 6120
cccagacctg ggtctgggtg tgctcactgc tactgggggt cattgcttct aggcctctct 6180
gctgatggag gtaggatata cacgtacagt cttccctctt cccagattcc gtacttgagc 6240
tcgctactct gctaacattt atttatatcc cccaaattaa acctcacagc acttctgcaa 6300
tcactcactg acctgcagag tgtgaaaaaa ctgagtcacc atcacacgtt ccaaaactgag 6360
gtcaactgag gccacaacgc cccatcttct tgctccggct gtogagatgt aagcaagtgt 6420
ccttctctcg gtctagctag tgccatgctt tccacatcac tgtgcttttt gtgggcaatt 6480
ttgctgtata aaatgtcccc tgcacatatg ctgctgtgta gtgctcctag gtgcagtagg 6540
ctgccccacg ccttacagag agaatatgca tgagaggctt tattcaggta tgagttag 6600
cgtagttggc catgaattca atgttaatga atcaacaata tacagttaat aagggtcttt 6660
ttagagacag ggtctcactc tgtcacccag gcttttagagt ccagtgggtg gaccttggct 6720
cactgccgcc tcaacctcct gggctcaagt gatectccca cctcagcctc ccaaactgtt 6780
gggattacag gcgtgagcta ctgcactcag cctaaataag gtgtcttaga aacacacata 6840
agacaagggt atgggctgag tgcgggtggct catgcctgta atcccaacac tttgggaggc 6900
acaggtggga ggttcacttg aggcagaag tttgagacta gcctgggcaa catggcaaga 6960
cctcatctgt atattttttt aaatcagaca ggtgtgggtg tgcatgccta tagtccagc 7020
tactggagag gctgaggcag gaaaatggcc tgagcccagg aggtcaaggc tgcagtgacc 7080
catgattgta ccactgcatt ccagcctggg gtgacacagc aagacgctgt cttaaaaaaa 7140
aaaaaaaaaa aagccagggtc aggtatcgaa cagttggcaa aaacgttgtg acctgaggct 7200
cacaggaacc tagccogatg tttcccttag gagcaatggt tcagtattca ataattcagg 7260
gttcccagtg actttatgga gcataacttt caagaataac aagaaccaac tgtacgtgtg 7320
tatgtatact cacactttta ttttatttta ttttattttt tgagacagag tctcactctg 7380
tcaccacaggc tggagtaaaa tggcgtgac tcgactcact gcaacctcgg cctcccagg 7440
tcaagtgatt ctccagctcc caagtagctg ggattacagg tgtgccccca caaccggcta 7500
atctctgtat ttttagtaga gaaggagttt cgccacattg gccacgctgg tctcaaaactc 7560
ctaacctcaa gtgatccacc caccctcagc tcccaaagtg ctggaattac aggcagtagc 7620
tgccgtgctt agcctacata cacttttata cacacatgca tctatgacta tttctctatt 7680
tctgtgcatg tgtgcgtggc agtacctaca gtttcagcta tgtgtctggg tactgtctcg 7740
tccaagtttg taagcacctt ctccaaagt ctgagcctgg cttgtgttac tatccatag 7800
tttactttatt tgetcaatca atttacttat tagctccata accagcttcc catctgctcc 7860
agtagcctct gctgtcagtc acctctgcac cctacccacc cttgcttccg gatgctggat 7920
gccaatcacc cccgacacct ctacatagca ccacctcga catgctgctt ctttatttct 7980
tattttattt tttgagatgg agtcttactc tgttgccagg gctggagtgc agtggcacga 8040
tccaggctca ctgcaacgct cgctcctgg gttcaagtga ttctcctgcc tcagcttctc 8100
aaatagctgg gattacaggt gcccaaccac acgcccagct aatttttgta tttttagtag 8160
agatgggggt tcacatggtt ggccaggctg gtctcgaaat cctgacctca agtgatccac 8220
cttggcctct caaagtgctg ggattacagg tgtgagccac cgcgcctgggt ctgcttcttt 8280
aaatgccagg caccaacatt tgtgcaatgg ggtgggagga aagaacaggg aggagagcac 8340

```

10

20

30

actgcgggcc	cctgcactga	atccactgat	caatctgggg	gcaactgcc	tctccatctc	8400
ctgtcttcc	atccgtgaac	atctactgca	gtctctccca	atgtccttct	gtaaagtgtg	8460
attatgtttt	gcatacaggg	cttgcatatt	agttctcaga	tataatccat	atactttata	8520
taaaattcaa	accacattta	aaaaataaaa	actagcatga	ctataacgga	gtctgcaaca	8580
ttctcacaga	ctttatgata	aaacatgaaa	cttcaaagat	acttaggggtg	gggcaggggac	8640
aatgtttaag	gctgcttgga	agcctcccca	tccctgagcc	agaaagtcc	atctcccctt	8700
caaggggaaa	tgcttgaaaa	agcactgatc	aggctaaaa	gacagggatc	agggagtaat	8760
caaagtacaa	gtgagctggt	ctctccatt	ctgagcacag	caaagtccag	tctctccaa	8820
tccaagaatc	atacacctgt	ttgccaagaa	tgaagttcag	gtgtctacaa	gtggctgaaa	8880
atattcattg	ctgggcccatt	aacaacattc	ttggcaaaac	cataccttag	cttctcgtgg	8940
aaatttctta	aggtagaaga	aacaggaaac	accaggctc	gcttttatgt	agacagttcc	9000
atgaagccag	ggaccttccc	cacatccacg	tttcaattac	ctgcacgcag	ctcacagtgt	9060
attcaacatc	tacgcgtctc	tctactggg	gtggcggtgg	ccactcaaac	cctcatgcag	9120
ctacgatgac	cgcaattttg	gcaacataat	ttcatgtttt	tctctgggct	tttaccaca	9180
tcaagtacac	aattctgcag	ttgtctaaag	attcaaaatg	agggacttga	catttacaac	9240
ataataaaaa	ctctgggttt	cctttaacca	agcacatgtt	ctgcctttta	gagaaagctc	9300
tgcaaaactca	agctggagtg	ggatacttgc	tgacatcttc	aagcaccoca	ggaatagctc	9360
tactccccca	tttccacctt	ggctgaacca	tctatatccc	accaattccc	ccaacatccc	9420
tccatccgtc	catccatcca	cccaaggacc	tgctaagcca	ggagggtctc	cccactacc	9480
ccacagcctg	gcctcagccc	acaagggtc	tctctacatg	aatccaccg	caccagagta	9540
gaccaagtct	cccgtagact	ccaccctgac	cacctccatg	cctccagcca	ttcccacccc	9600
taaaaaccct	ccctgggtctc	tacacccagc	tgatgaatac	ttggctgaat	gtgacctggc	9660
ctcctggacc	caggtgaagc	ccacgtctc	cgtaagcccg	ccagctcacc	ctgcctctgc	9720
accttactg	gagagagccc	gcacttcacc	tcctcagggc	aggcatggct	gatgccaccc	9780
agtggaatct	ggtgcaaaag	agggcccgg	gcagagcagg	gctgcctgca	gagcaaggcc	9840
ctgggtgctg	ggccgagcac	ctccaatgct	ggcgtggaa	ccatccctcc	cattccaggt	9900
gctgtctcca	tcaagaatga	gagagctgct	gacatttgca	tgacaataat	gaataaatac	9960
catattttgc	ttcaaatcca	gaatagatgt	ggccagggtt	ggcatatgac	tgttgggaaa	10020
ggacagtittg	cctcttccca	aaaccaactg	gattataaaa	agcttttctt	aacgaccaca	10080
agagcggagg	agctcagggg	cagacaaaag	gaaggctggc	tgacagaaggc	gggagagtgg	10140
ggccttcagg	ggcgggtggg	gagagagaaa	gcctggagct	gcacccccaa	ggtctgtgta	10200
catcaggtgc	tacagaataa	caccacctct	tccagcttgg	ccccccactg	ccctctccca	10260
gcccagtcac	ccagacagca	ccccactccc	cacacacacc	tcacatctgc	ccgcctcaca	10320
ctcaccagct	tgggtctctc	atgcaacctg	gaacctgccc	ttggcctctc	agctcagcca	10380
ccccattcc	tgttggcccc	tggcccccca	togaattctc	tctaattcta	atgcacacac	10440
ttgcacactc	aaacacacac	acacacacac	acacacacag	cccagaggaa	aaccataatt	10500
gactgaggtc	caggcaagtt	tcccgagcag	ggaccacatt	tcaaaggcca	gggaagcagg	10560
cgaacaggaa	acatacaggg	ggcacgtttg	ggggtggagc	aggaaataag	aaatcacttg	10620
caaaagataa	aaagaaaatg	aggtagctgg	tttcagacac	ctcggagcac	acagaacagg	10680
acaggcgcc	ccgggtcttc	cctcaacagg	gagatggggc	aggcaggctc	ctgctgctcc	10740
accgcagagc	tgggggctat	ggccctgaca	ccaaggccct	ggggcaggcg	gggaggcagc	10800
tggtctcctg	cctgtgctcc	cgggcagggc	ctggccccc	aagggaactg	gccgaaggct	10860
ctgcttggtc	actccggaaa	gtcctgggag	acaagcaaa	gacttgctag	gtcactccaa	10920
acggcccaga	tgtgacaact	gtgaagaagc	cacaccaag	caagggtgaca	gaacaatgtt	10980
ggtgacgtca	ggttatcagc	ttacgctcaa	ctccacttac	ccggactcac	ccgtaacctg	11040
ccgtctcttc	ccaaccagta	aaggatgcct	aggtagaggg	gcacaaggcc	tggagcataa	11100
ttaccatttt	aaaggctctg	agaagtctctg	cggtagaggaa	gcctagttca	ctttctctcc	11160
cctaggattt	cccaactgcy	cctgatcaca	gaacattttt	tcatttccac	tcaggaaaca	11220
tattttgaaa	aacactggcc	tagaggcaga	agtgaatgg	aaaacacaaa	agtaaaactg	11280
aacaggaggc	actgggcaga	gaacggctcag	aggcgccctg	aatcctggac	cggtagagat	11340

10

20

30

ccccagcttg	gcatgctccc	ctccctgggc	ccagaccgcc	ccccccatt	tccctggataa	11400
gaaggctaata	gogcatcagg	gtgaagggct	tgcctgggct	acacccccag	gctcgccccca	11460
caccaatcgc	gctcctgcga	gagccagtga	ctttcttgat	ttggctactg	tggaattgtt	11520
tgcaactaac	caccccgat	acagatacaa	atgacaggat	gatcagatgt	aaaggaccca	11580
caggtctctg	tgatcggct	tcatgcagcc	agcatggcta	gtgccgtgca	gaatgagaat	11640
gacccaggc	aagtccttg	ctcccagacc	cagaacccca	tggagcccac	cagggtctggt	11700
tcacaagcac	tgtctgggtc	gggcagagat	tccagcaaga	ggagggaaca	tccatgcacc	11760
ggagccagtt	accagaagca	aatcgccctc	tccaaaaccc	aggctattaa	tggagtccac	11820
tgttgagtgg	agctggggtc	tagctatgga	atactgcaca	gcagagatct	tcctgagaga	11880
aagcagtttt	ccctgaaagc	catgtgtcct	ccactaactg	tgttttaatt	gggcgaacgt	11940
ctgtatctca	ttgcagtggc	cgcccatgtg	ctgacaaggg	gctggggggc	gggtggggag	12000
cagaagctca	ggggcctggg	agggaaaggaa	acaggccacc	agggctcccc	agaaggcatg	12060
tatctctctc	acaaacacac	gcatgcacac	acacgtgcac	acatactctg	caagccctga	12120
gttagcaact	gtggaatgtg	accagctcag	tgatcccagg	acaagctgct	agggaatatg	12180
acatttgatt	gatgtctgca	aatgtgcgtt	ttcactaatt	agaaggttta	gggcagagca	12240
gagaaaaata	tgtatttcag	agtcccagtt	tgacctgcca	gaaaccagcc	cattactaac	12300
attctttatt	tcaacaaaat	atagcattct	gattacatac	catcttgggt	ccacgcctcc	12360
tgccttgcca	agcccccgga	agcgccccaa	ggccatggca	aatagtgaga	gaaacagttc	12420
cagggtggag	actgactcag	gggtgtcagt	cagtgggggc	ctgatggccg	gtgggaggcc	12480
agcagtcac	accctctcct	tgggacagtt	gagtagctct	ccccagggt	catgtggcca	12540
ctcaggttca	tatgggaggc	gagaggagtg	gcagagtcca	ggagagtggc	tccgaagtca	12600
ctgttccctc	caggcctcag	tgtcttcac	cattaaatgg	gtaggctgag	gtctgggatg	12660
acaaggagg	cttgccactta	ctgaaaccca	tgggaggctg	ttcgccgatt	tcttttattg	12720
atggaagaaa	acactcgtat	aattcaagta	ccaattaaaa	ggcaggcact	ggaaccaccg	12780
tctgccaat	octagttttg	cctataccaa	atttgagcaa	gttaattgac	ctctcccagc	12840
ctcagtttct	tctctgttaa	aatgagggta	gggatggccc	ccagcccaca	gggcagctgg	12900
aaggattaaa	gaaatcaaac	atctcttaga	gcccacctgg	cacactgtga	tacacaacaa	12960
atgttagcta	tttttgtcta	tgaagtctag	attttatatc	ttgggtgttc	taaagcagga	13020
tacatttatt	taaaaacaag	gattttcatt	aaacacgtac	cccacagaca	gcaaccocat	13080
ggagactgct	cttaattcag	gccagtatcg	aaacgactct	aactacaagc	tttatacagg	13140
tctcttggt	gtccttcaaa	tccaaactaag	gtggtacttc	tgaagcactg	tgcacatgtg	13200
tgtgtgcatg	cacacgtgtg	ggaaggggcg	gctcacggat	ccctcaggta	ccccaccacc	13260
gcagtctcaa	gtcacaaaag	gacagagcag	ccgagggaagg	tctgtgcccc	actggaccct	13320
cgtgaagcca	ccaactctac	ctctgcgcgc	tgtcctgcag	actgggctac	cctttgggtg	13380
gggaccagca	tttgatgcaa	gaaaggcaga	cagaaaaggga	aaagggcaag	ttcgactcca	13440
gataaacacag	acagtaccaa	gcccagggt	ccataaatgc	cacgcagatg	gaagcattta	13500
ctgcgaggcc	acacagcaaa	cgcacggatc	cagggacgga	ggtgcagact	gcggtgcccc	13560
tgagccatga	ccctgcaaat	taccaccatg	ggaaaggagg	ctgccaaacc	ccccgacagt	13620
cggctgggct	ggcacagact	cgtggtttcc	atcgagggtg	gaggagggtg	gacgtcccag	13680
cccccccccc	atgccactg	cagagggaag	cggccgtttc	ccctgtgtgg	ttacaaaggt	13740
ctcattgttc	ttcctcacag	ggaggaaact	ggaggaccga	gctcagaacg	catttttagaa	13800
ctggcagaaa	agaacatctg	gggaaggaaa	cacatttcag	aaacaaacat	acctttgtac	13860
cagcttttat	tttctttaag	tgttgaaaaa	ataataataa	taaagacatg	ccaaatttat	13920
catcgctcta	caaaatccct	ttattgagca	aaacgtggca	gctctacttt	caaatgatta	13980
ctgttccctg	aaaattgcag	caacgtggat	gccaaaggcc	gaaggccgcc	atcagcagcc	14040
aaacaaaaga	tgccacctcg	ggctccgcga	cactgtacca	tgccagggaa	ctggacagat	14100
ttgggggaatg	ccacggtttg	cctttaaccc	cttgccctct	ggtctcctga	tgcactctcag	14160
aggctaacat	tctttgagga	actggcattt	cttagttgta	aatatgcagt	tgggtttggg	14220
agctgcctgc	aaagtccagt	gttgacgac	agctttgatt	tccttgggaat	caagtttaacg	14280
tgtcgagctc	ggaagttaag	aagaatttgg	agaagctgag	cactatgggt	ttgcaggccc	14340

10

20

30

tgggtgaact	cttccacca	gcattcattg	tggactgaca	gcgtgcgagg	ggctctgcag	14400
gcaggtgcac	aggacgaaac	acattccgtc	cgggggaaac	ctgcaggaaa	gctccctctt	14460
cttccctaagg	tgccgggocct	agcttcatgg	gtccctaccc	tccacgcctg	tcacactttc	14520
tgagtctcat	gtgggagctg	cttctgggtt	ctgacttcac	tcagtcctca	taggaggtgg	14580
aactactgtc	accccattht	acagatgggg	agactgggca	caaggggacc	aagaaacca	14640
tgcaaaagtc	cacttgtggg	atcagtgaca	ggggagatca	attcccaggt	tctttctgca	14700
agagttaa	tgttttcatg	ctgcctaagg	gggggcaact	gaaagaccac	tgcatactct	14760
tgccaaaagg	gtcaagcaca	ggagccgcag	ccagtggttc	agatccgcag	aggcgctggg	14820
gtgaccctcc	ccatacctgg	agggatgctt	gtccctccct	ggccttcaact	gggtccctcc	14880
atgaccgtgg	cctcccagga	cctcagcaca	atcccgggtc	tgtgctccag	gacaagccct	14940
ccgtcccca	gactgtgagg	aaatggaaac	aagaggggct	cgctgcagcc	cagcaccacc	15000
actgccccct	ctcaggggca	agaaccgtcc	tggaggtact	ggctttggag	ggggagcctg	15060
ggaggccagt	aagtcaacaa	gcctctactg	ctcatgggtg	ggatccacc	gcaggcccc	15120
acctgctggg	gcgggcaggg	acgggcggca	cagcttggcc	agggcagata	acccccacct	15180
tggccagggc	gaaggcagga	cacgtgggct	ccagcctggc	cccaccatcc	ctgcacaaca	15240
ctgggcaaa	tccacgtttt	cctcaactgg	gtgttgacat	ctgcaggaca	ggggcatgga	15300
ggtacagagc	gctgaagcca	cacagcaacc	taggagcgag	actccatgcc	tcccgggga	15360
ccccccccca	ccatgaggac	catgaaggct	tcccagtgc	cgcaaggact	ctgggttgga	15420
gacacacgtc	tccacacag	ccaggcctaa	cgctcttgta	actgggtggt	cccacctggg	15480
ctcagagctg	gagggccagg	agctcaaggc	tccgcagggt	ctgctctcat	cccagaggcg	15540
atggggagcc	acagcaggct	gcaggagaga	gggtggggcc	cctccacttc	agaggcccca	15600
tctggccac	agactggaga	gcacatctct	cagcaaccac	ggagcgccaa	ctgcgcacag	15660
ggctcggtcg	tcagagcggg	gcaaaaggac	tagccgtcac	ggccaggcg	agggaaagac	15720
gggtggcagg	gaccttgggc	agagggggaa	gaacctgggt	cccaggctgg	ccctgccttc	15780
agcagtgaag	ctgagtgggg	agggcgtgat	gcagggggcc	agaaagggt	gctggtcagc	15840
cgggaggagc	ccccacaga	ggaagcagcc	agcccagacg	cagatggcag	gggtccctca	15900
acaatgtcct	ctgaaaagga	gaggcgggga	ctgctctggt	gacacctaca	aatagatagt	15960
cagccctcag	ccccctgcca	tacttctgac	aaagcagagg	ccccagggg	ggcgccacc	16020
gaaggtacct	gcacctgtcc	cccagactcc	tagagccac	ctgaccccat	cccaccagg	16080
ctccagctac	aaaataaatg	cogaggccag	ctaggcaagg	acgcacactc	ggtaccgact	16140
gaataggctc	cacgttgtca	tgagcgcaac	ccacaggcca	ccaggccaca	ctatgcagag	16200
ctgagatggt	tccggccaag	cagcctctca	gctgagctga	acaagtccag	agtcccoggg	16260
gggtcgtcac	tatggagtaa	caattgcgat	gogattgtaa	ccctaacagc	taaccgtcac	16320
tgagccaggc	cctgagctag	gtacttttca	acgctgcctc	tctgcagcct	caggacgagc	16380
ctgtgggagc	ataaagatca	ttccctatca	cggatgggga	aaatgagctc	tgaagcagtt	16440
aacgtgcttg	tcccagaccg	cagagctagg	agcaggacac	aacagcagg	caggcaggaa	16500
cgggtgaggg	gggcttgcat	gggcttctct	ggaggctgct	catacacgca	acccccagg	16560
ccccgaccct	gcacctgcag	ctcgctactg	ccccctcagt	gactccagca	aacctcgggg	16620
taggggaagg	aggctgggaa	tacctcggtt	gtccgaaaca	gcagcttctg	cttgagggoc	16680
actgctgcat	aatgggttgc	gcccagcaca	ccccaaagca	cctgtgccac	ctgtggtgac	16740
cttccagcat	gccttgggtg	ccaagctggc	cttaggtgct	gtgggcagcc	aagaatagaa	16800
caggggccac	ccctcctctt	cacactaaca	caaagcaaga	ggcgggcact	tgcactgagt	16860
gcacccctct	agctcaaggg	cctcacggat	cacaggggct	agggcaagat	cccaattctg	16920
cattcccgte	tgcctttcat	cctgctctgc	caacaacagc	cagttaggct	ggggacatcc	16980
ctgaacctgt	ttctcacctg	aaacacatca	taccattgga	ccccagccct	ccgggagagg	17040
ccctaattcc	tgactgtggt	gagatcagat	cactggttaa	gtaccagaa	gggccttgg	17100
caggggctcc	aggggtgggg	ggtgatgggc	gtggtggtat	cccgctctgg	gctatagtcc	17160
acctgatgg	aggaggtctg	tggtcagaac	cgggctgtgc	agggcacagg	agcccagagg	17220
gacccccaga	gctcacctgg	tggctctctg	gcagggtccc	ctcaaccctc	agagaaaagc	17280
acagcaagga	ggccgcccag	agcccagcgc	ctagcaccca	gtggcggtgc	agacctgctc	17340

10

20

30

ggatcctgga	gatctctcat	cacccctcaa	gtcagtcag	cccaaccag	ggaccacag	17400
cccacggggc	cgtgaaggtg	tgtctgagtc	aagaaggcct	tgcacactgg	gaagccaagt	17460
ggcacctcct	ggtgtggagc	aggcggaaac	ccaccagcct	ctgctctgcc	agtgggcaca	17520
gctggacgat	gagcagaagg	ggctgttgct	taataaaagt	catttcctta	agaggataaa	17580
acctttcaaa	acagatggaa	attttttttt	aattaaaact	ggtggccaaa	gagatggaaa	17640
gcaccccttg	tgcctccctc	ccatcgtgac	ccatcctctg	cacacctcaa	gctgttcgct	17700
gcccaggtgt	ctcctgaggc	actgggggag	ggtgagaatc	cgtgagccct	cggccagccg	17760
tggctctctg	gagctctgcc	ccaggccatc	agggcacacg	ccgggcaccc	tggggggccac	17820
acagggcaga	gcccagctgg	gtcagcacac	agggccacac	tgggcacaca	agtctctgag	17880
cctccctctg	ggacgcagct	ctcactatcc	caccccaacta	ggtcccgggg	atctgtccca	17940
cagggttgata	tgtgttcaca	gaccactacc	agagccatgg	cctgctgttc	cggccgcagc	18000
caggtagtca	cttgctccac	agggacaggg	aacggccgac	ttgggggctg	ctctgcggca	18060
ggactagagc	tccagcagct	cagccctcct	gagaaggaga	actccatgct	ctaagaggca	18120
gacgcagcgg	acggcaccac	agccacacac	agccacgggg	gcccctgcatg	gcaggtcagg	18180
agtccttgac	cactcgtctc	ttgtaaccag	agctgcagtg	gagtcctaca	ggcaaggact	18240
gtggggcgga	gtggccacag	caaatgaatg	agtgtcccaa	gggagcaggg	ggctgcgggg	18300
aggcacagcc	gggacccagg	agtcctccgg	cactgcagca	aactccctgg	gccccctgag	18360
cagcgaccag	gtggcaagtg	catgaactcc	cgggggcata	acctgggagg	gtgacactct	18420
cttcgtgttc	aaattcttga	gaacgcatta	aaaatatcac	tcagtcacct	actctatagt	18480
tttaactcaa	aagtaccaaa	gtagccaggc	gcggtggctc	acgcctataa	tcccagttact	18540
ttgggaagct	gaggcaagag	gatcacttaa	gcccaggagt	tccaaatgaa	cctggggcaac	18600
atggagggag	cccatttcta	caaaaaaagt	gttttaaaaa	attacctggg	cctgggtggtg	18660
tgtgctgtga	ctcccagagg	ctcaggaggg	tgaggcgagg	gaaccacatg	aaaccagggg	18720
aggtagaggc	tgcagtaggc	tgtgatggca	ccactgcact	ccagcctggg	taacagagtc	18780
agactctatc	tcaaaataaa	tttaaaaagc	accaagccag	gcttggtggc	tcacacctgt	18840
aatccagca	ctcaggaggg	ctgaggcaag	tggatcacct	gagtcagaag	ttcgagacca	18900
gcccagccaa	catggtgaaa	ctccatctcc	actaaaaata	caaaaattac	ccaggcgttg	18960
tggcggtgct	ctgtaatccc	agctactcag	gaagctgagg	caggagaact	gcttgaacct	19020
aggagcgaga	ggttgagctg	agccaagact	gtgctactgc	actcaagcct	gggagacaga	19080
acgagactcc	atctcaaaaa	ataaataaat	caatcaaaac	caccaagact	ttttaatata	19140
aacatttatt	attccataat	tccctttttg	catgattaaa	aatgtttata	taaagtttcc	19200
tgaatatggt	aagaatgcc	agtgaaggct	gcaaatgccc	aagccccac	cgtggcatct	19260
cacggagctc	gggcccctagg	aggctggtgg	gtaccacgtg	gacccgagac	ttcacagtca	19320
agtccttttg	gggtacactg	ggtttcccac	accccagaaa	tatgggctct	tactgcagga	19380
ccatgggggt	cctcacactt	ggcccagaag	ctgtcacata	gccagacagg	tgttctacaa	19440
cctaggtctag	agggagctca	tgtccagca	gaattcgagc	cagaggaggt	aaaagatggg	19500
taagatctgc	tccctggaca	gatgaggcct	tggcctcaga	acagttactg	atcatctacc	19560
agacatcaca	ctagaggcag	agggggcgag	acgaagacag	cccctgtcct	caaggccctc	19620
ccaggttggg	tggaccatgg	aaggttccag	acagatctgg	caagagaagt	gcccacacca	19680
ggggcgagaag	atgggcaggt	ctgctcaggg	cggcacggcc	tggcaggcca	aaaagttcca	19740
acttcagatg	ctgggagaatg	ggcacgactg	tctgagaaaag	ggaaggatgt	gatgaaaact	19800
acttgagaga	aaattaatct	ggccagagca	taagataaat	gggcaaaggg	gaggttccag	19860
aaagcaagga	gaccaagtaa	aagctgatgt	cattggctct	gaatctaggg	tttactgaa	19920
tatgcaccgc	agggcctgta	ggtaaaagcct	cagagcccag	ggagtctgag	tggaggagag	19980
ggcaggggac	agagctgggg	cctgtgtcta	cagtgtctag	gaggaatagg	catggacgtc	20040
agctcggagg	ctccagctga	agtgaggagg	cggccagggc	agcacggcca	cggccggatc	20100
cagactcctt	ttgggaagca	agttcgctct	gggggaaagt	ttggagaaat	ggcctttacc	20160
cgcagaagca	agccccagaa	catatcttgc	tccaaaacta	tctcgtacag	tgaggacgtt	20220
aagcttcagg	tcccctagag	gagacagtct	gctccttctt	ggggcagaa	ccaagggtgg	20280
cagagcctgg	aaggcaccca	gcacccaggc	tgggtgtgtc	cagcccaggc	cacacgtcca	20340

10

20

30

gatagctatt	aatgccccgt	tgagcaatth	cctgagagct	ttgccaggca	ggtaccgcct	20400
ccccatctga	actaatacag	gggtacatcc	caaggaagaa	atgaaagggtg	cccacatttt	20460
gctctgggat	taactaggga	ggggagtgat	aattaactca	gtaattatat	ttgccatcgg	20520
gctaagtcta	aaattagtg	gcattagaat	ttctttctctg	agcagacacc	ggagtgagtt	20580
gggcagcagg	agtggctcgg	gcaagtccgg	acaaagggca	cctccagagc	cttccacaaa	20640
tgctagcaaa	accacaaaat	gtcaaggccg	gctccactgc	acccagcaga	tgaattcact	20700
tcacagccct	gagaccgcca	gctcatcgga	ggccatttaa	aatccagccc	tctgacacct	20760
gctggatatt	accatttacc	gtccccagat	caagagatca	aagggtggaa	cctgatagga	20820
cggctctgaa	gttcaccaca	aaagcataaa	cgtgcaagca	gagccaatac	gtcttttgaa	20880
aaggacaatg	agggtgggaat	ttacataact	gatcttaaaa	tatgttctga	tgcttcagag	20940
atggagacag	cagcattccg	gtacacaaag	acactcacag	gcagtggagc	acagtgaagg	21000
gtctggaatc	aggaccagg	tgtctgtgga	cactacacat	aaaagagcag	catttacaat	21060
gaatggatag	gatggaccat	cccaccaagg	tggtggacaa	ctccctattc	actggccaga	21120
cccctacctc	ataccatata	caaaaaaaaa	aaaaaaaaaa	aaacccagac	agaataatgt	21180
ctgaatgtaa	aacataaaac	agtaacagtc	ctggaagaaa	ataatggagg	atatatttat	21240
aatctggaga	tggagtaaca	agggatagga	aaaaagccat	agggaaaaag	tagagttatg	21300
attatatgaa	gttctttaat	atctttatga	taatgtacca	ccagaaacaa	ggatgaagga	21360
ctagctacag	accagcagtg	aaacctgaaa	caaacagaa	aaagaattaa	agtccatacc	21420
aaataaagac	ctcccacaaa	tctataagaa	aaagataaac	aggctggcac	cgtggcttat	21480
gtctgtaatc	ccagcacttt	gggaggcgga	gatgggtagg	tcacttgagg	tcaggagttc	21540
gagaccagcc	tggccaacat	ggtgaaaccc	tgtctctacc	aaaaatacaa	aaattagcca	21600
ggcgtgggtg	cgcatgcctg	tagtcccagc	tacttgggag	gctgagccag	gagaacagct	21660
ggaacccggg	aggcagaggt	tgcaagtgaac	caagatggca	atcgcgccac	tgcaactccag	21720
cctggaggac	acagcgagac	tctgtctcaa	aaaaaaaaaa	aaaagaagaa	gaagaaaaaa	21780
gaaaagaaaa	agacaacaga	aaaatgggcc	aaggataagt	gtaggcaatt	tcagaaaaag	21840
taaatcccaa	taaaccagaa	atgagggttg	tgcaaatcaa	aagggtgtat	aatttttaac	21900
caaactggac	caaagaaaac	acaaaaaac	aaaatcttgt	aattgccagc	atcagagagg	21960
atataggaaa	gtgtgtgttc	tcgtagatgc	ttgcaggtat	gaactgtctac	aggcttttag	22020
gagttatgta	tgtatgtatg	cttgtatgta	tgtatttgag	acagggtctc	gctctgttgc	22080
ccaggctaga	tctgttgag	tgctgtgatc	atggcttact	gcagccttga	cctcctgagc	22140
tcaatagatt	ttcccacctc	agcctttcaa	gtagctgaga	ctacaggagt	gtgcaatcat	22200
actcagctaa	tttttttaaat	tttttgtaga	catggggggt	ctcccaattt	tgcccaggct	22260
ggtctcgaac	tccctggactc	aagtgatcct	cctgcctcaa	cctcccaaag	tgctgggatt	22320
acctggatga	gccactgtgc	cgggcctcaa	tatctttaa	aacagaaatg	gacacactct	22380
ttgactagga	atgtatccta	taaaaacact	tatacacatg	cagagacaca	cgagcaagca	22440
tgctttgtaa	tagcaatgaa	ggctggaaaa	actcctcaat	caggtaaatg	ctgtcaagtg	22500
cacctgtgta	ctatgaaatg	gcacttggct	tttaacaaga	gcaaagacag	aaaagcaaaa	22560
gtacaaagta	gggtgtgatg	gcacatgcct	gcagtcaccg	ctactcagga	ggctgaggca	22620
ggaagatcct	ttgagcccag	gagttggagg	ccaggagctg	ggcaatagtg	agaaaaaata	22680
aaattaaata	ataataataa	taaaataggc	tgggcacagc	ggctcatgcc	tgtaatccca	22740
acactttggg	aggctgaggt	gggaggatcg	cttgatccca	ggagttcaag	gccagcctgg	22800
gcagcaaaag	aagacaccca	tctcaacgac	aaatttttaa	aaatcagcca	ggcaggctgg	22860
gcattggtgg	tcacgcctgt	aatcccagca	ctttgggagg	ccgaggcagg	cagatcactt	22920
gaggtcagga	gttcgagacc	agcctggcca	acgtggcaaa	accctgtctc	tactaaaaat	22980
acaaaaatta	gctgggcatg	gtggcagatg	cctgtagtcc	cagctactga	ggcacaagaa	23040
tcgcttgaac	caggggtggca	gaagttacag	tgagccgaga	tcgtgccacc	gcactccatc	23100
ctgggcgtga	gtgagactcc	tgtctcaaaa	aaaaaaaaaa	aaaaaaaaaa	aggagccagg	23160
cacggtgggg	tgaggggagg	cacagaagca	gcgcctcttc	tgggggcacc	cccaatctct	23220
agcgatccag	aggcctcagg	atcctgaagg	gagaaaaaac	gtgaagctcc	gtgctagaag	23280
agaccataga	gattggaatc	agctggttct	attttacaaa	aaaaggaaac	tgaggccctc	23340

10

20

30

agaaggtgag	tgcctctcaa	tgccccacag	ggaggcaggg	agagggctct	gagccctgca	23400
gggcccctgga	ttctttgcaat	ggggtggagt	ggagcctgtg	ccgcccccac	caggcacctt	23460
ctcaggagag	gagccgttgt	catatccttg	aaggggtcct	tgagcccttc	aaaaggctaa	23520
aaaccacttt	cctccttgag	tgaaccttca	cctcagttta	accacaagaa	aaactacatt	23580
aaggcccagc	gcagtggctc	atgtctgtaa	tcccagcact	ttgggaggct	gaggtgggtg	23640
gacgccttga	gcccaggagt	tcaagaccag	cctgggcaac	atagtgaac	cctgtctcta	23700
caaaaaacaa	caaaatcagc	tgggcgtggt	ggtgcacacc	tgaggtccca	actacttgcg	23760
ggctgaggtg	agaggattgc	ttcagcccag	gaggtagagg	ctgcagtaag	cggtgactga	23820
atcactgcac	tccagcctca	gcaacagagc	aagactcaaa	aaaaaaaaaa	aaagcaggcc	23880
gggtgtggtg	gctcacgcct	gtaatcccag	caccttggga	ggccgagcgg	gaggatcagg	23940
aggtggagac	catcctggct	aacacggtga	aaccccgctc	ctactaaaaa	tgcaaaaaat	24000
tagccgggcg	tggtggcggg	tgccctgtagt	tccagctact	caggaggctg	aggcaggaga	24060
aaggcgtgac	cctgggagggt	ggagcttgca	gtgagctgag	atcacaccgc	tgactccag	24120
cctgggcgac	agagcaagac	tccatctcaa	aaaaaaaaaa	attaaatctc	aaaaaaaaatt	24180
acattaaggc	aaactaaaag	atgtttaaaa	tatatatatt	aaattaaata	cactccaata	24240
gagcaaatac	gaaaaatacc	agaaaaacaca	atccccgcac	ccccaggaca	acctcccagg	24300
gggtccacag	caagagaccc	caagcacgag	agacagagaa	cagtgtccct	gtggcggaac	24360
ctctggccca	tcaggctcta	ttagaaaata	aggctcctgc	cactgagaga	aaggaggaca	24420
gtcgcaccagc	agccacgggc	tctggcacac	cacgagtcag	gccagcaaag	tgtcaactgc	24480
ccccctacaag	gtgacaaaact	aggacaaaact	ggaaaccaga	ggctggacct	ggagcacagg	24540
gaccaccaca	tggggctggg	gaatgggcag	ggacctcaga	gcgccacca	catgcctaag	24600
agcagcgcgt	atgcgcagtc	ctctgcatgg	cttagggaca	cagggagctc	ccccaccccc	24660
caaccagga	aggcagcccc	cactacccag	gtagggaacg	gataggacca	gcaccccggt	24720
ctgctcgtaa	ctcagggctc	caggccccct	cgggggcaac	cagcacagag	ctcagacccc	24780
aaatatcttc	accacacctc	tggtcccat	ctggacaagg	gtgctgggga	ctggctctca	24840
gtcacaccct	cggggtaactc	ttcaaggac	agctggatgc	cccagggcag	gagcttttgg	24900
ccccagctc	cctcacccca	gacaccagct	cttgggaccc	caccagcatg	ggcaagggtg	24960
acaccatcgt	cccgaattttg	cagatgagga	aactgaggtc	gagggtcggc	acacggctct	25020
ccagagctga	agagaatgca	gagagcagcc	ggagccagcc	ggtgggtccc	tgaggccggc	25080
tcgtagcaag	ccacagctgc	ctccgcccac	cacacttgga	cctcactggc	cccaggacag	25140
ccctccaggg	cggcctggca	cagagcccac	accctgctgc	ttcctgaaca	aataagtga	25200
caaggccacc	aagccgagga	cctggatgta	gccccggctc	ccgccagggc	ctccccaaac	25260
gactcccat	ttggagagcg	cattaagtgt	ttccaaagcc	tcacaaacca	cagatgtccg	25320
gctgtctcac	ggcttctgta	acctgaactt	ggccctcact	ctgccctccc	agcactcctc	25380
tcaggggcca	ggccctcctc	ctgagatgcc	agcactgact	ccccacttg	tccccatcac	25440
ctggctcggt	cctgaacctc	ggcaggagag	tctcaggcca	gacccctcca	ccagccacct	25500
ccaccaggat	gcaggaggca	tgagacctgc	tcgtgccggc	tgggagatgc	aaccaaccac	25560
gatcaatcca	atcagcggat	gaactgacaa	atataatgtg	gtccctccac	acaatggaat	25620
attattcagc	cacaaaaagg	gctgaaatag	gcccggcggtg	atggctcaca	cctgtaatcc	25680
cagcactttg	ggaggccgag	gccggcagct	cacttgaggt	caggagtcca	agaccagcct	25740
ggccaacatg	gtgaaatccc	gtctctaact	aaaatacaaa	aattagctgg	gcgtgggtggc	25800
gggcacctgt	aatgcaagct	acttgggagc	ctgaggcagg	agaatcactt	aaacccaggga	25860
ggcagaagtt	gcagtgagcc	aagatcgcac	caccgcactc	caacctgggc	aacagagcaa	25920
gactccattt	caaaaaaaa	ataaaaggct	gaaacaccca	tacgtggtac	tacttggatg	25980
actcctgaaa	acgttacagt	aaccaaggaa	gtcagccacg	aagacgcatt	gtaagattcc	26040
cttcatgcaa	aatgcccaga	acaggcagaa	ccacagaggc	agaaagtcca	ctggtgttca	26100
ccaggggatc	cggggagagg	gaacgggaag	tcaccgtgta	atgggtatgg	gttttatttt	26160
ggggtgatgg	aaatctctta	taacttgata	gaagagaggg	ttgtaaacac	tgtgaatgta	26220
ccaaatgcct	gccttctata	ctttaatatt	ttatattata	taagtttcc	ctcaatttaa	26280
aaaaaaaaaca	actcgacacc	tttcacctag	gaaagatctg	gctttagctt	gcatttctctg	26340

10

20

30

taactcctgc	ctaaagcctt	ccagaagcctt	ccgctgcctt	gtggatcaca	accagactcc	26400
acaccatgat	ctggcctcta	agggcctctc	gcaggacacc	ccgaggggtga	aggagcacc	26460
gtgggcccac	ctctgcatag	ctgcaaaagct	tctttccctg	tccctcccctc	tacatgggaa	26520
gctctgcccg	cagggggcgg	gccttatctg	ccattctatc	gcactcaacc	ctagcacttc	26580
actcggtagc	agacaccaaa	gcaaaacagc	aacagcatta	taccgggcca	ggtgcacgtt	26640
aactcactga	attcatggta	ggaaggattc	tattcccat	ttacagggtga	gaaaactgag	26700
gcacacaaag	gtagcatcag	cttcctaagc	ctcccagcac	aggaagcggc	caggctggaa	26760
tcagaccctg	ggcgcagggg	ctctgtccac	agtgttaact	aactactcct	gcccccgagg	26820
gctgcagcgg	tgagtgagtg	agtttgtcag	tggaactggat	gtccaaggtc	atacaggaaa	26880
aatccagact	attgtaataa	cagcctctag	accggctggg	gccagaaaga	tcgaggacgc	26940
tgacacacaa	ctgcgtcac	tgagctctg	ccagggatgg	ggctaagggt	ctcacacagg	27000
gcagttaggg	ctccccatag	cctgggagag	gaacggggtg	agataacaga	aactaggtat	27060
ggtgcccgaa	gtcaaacagc	cactgagcat	gtaaacccag	gtgggtctga	ccccaaacc	27120
ctccaccccc	atcagccctg	caaccctctg	ctgcaaggga	gaaagcaact	cagaggcctc	27180
acctgcctac	atccccacc	cgtgtgtgtg	agttctacta	aatgcctgag	cagtgcaca	27240
gcacggctga	aattaaacgg	gttccaaaa	cgacaggaag	caogaagtga	atctcccag	27300
gaaagtgtctg	aacaaatgct	ggatcgggtt	caccggcgaa	ttctctggaa	ctgaagagg	27360
gagctaaaca	cacggggccc	tgttttgagg	gggactctct	cagggtgctc	cacacagcac	27420
ttggttaacc	ccactcagcc	cttctgggct	ctcccagagg	gcccggcctt	ggccttgggc	27480
atctacagga	ggaacctcca	gggggagagg	gggtgcctgg	acaggccggc	cctggaacaa	27540
gcacttgggc	cccaggagga	gaggactagg	gcttgggagc	tggggaagtt	ctcagcactg	27600
ggaccactag	aaacaaagcca	tttccgtgct	ttcacagctt	ccaattgcaa	caggaagcaa	27660
tcaggaaaaa	taattagcgg	cccacttact	ggcttcgctg	aggtccgagg	catgtatttc	27720
acacagtaaa	accagggata	taacatcaaa	accgttctgc	agaaagattc	ctccctttcc	27780
ttccatttta	ggcctggatc	accacattca	ctggggctcc	caggccttgc	tgccataatgt	27840
taaaataatc	aactctatct	ttgcctcaca	cacaactgaa	ctctacagct	ataattcttt	27900
ctctcaggg	gctcgaacca	catggacgac	aggcatttga	ctccagcaac	atcaccccaa	27960
aacgtgcaca	aaacccaaaa	ctgcaatgag	gtgaaaaggca	acgcggtcgg	cctagaaaacc	28020
ccccctttaa	aacaaacagt	ttccccaaaa	ccccttttgc	ctccttgacc	caggcatttc	28080
cggaaaaagg	agcggcgctg	gcctgtactc	cccagatact	gtcgtgtgtt	tgtcttcacc	28140
ttgttttgct	agctccagac	aaggccccac	aatgtaaaaca	cgtctctgaa	agaggcagat	28200
ttggggtgaa	actgtccata	gaatctctag	gcttgggtca	gaggcaggag	gacgtgaaac	28260
aaactccaag	ctctcctgt	tccccgtgt	cccccacacc	tccaagcaga	ggctgcagcc	28320
tgggggatct	gactacaggg	ccaccccgct	gcaccattca	cactggaaat	attcagggag	28380
acagctgttt	gccttaaggga	ggcccagaca	aaggggccc	aggtcctccc	cgctaaactg	28440
ccacaaacag	aacaggagcc	gcggcgtgca	caggcacttg	cggccgtgcc	acttggccag	28500
ccatactcca	gaaaaacaaa	acacgcacat	ccgaagagaa	tgatttaggt	agcaagaggc	28560
ttgcttga	aaaccaatgg	caatctccaa	attaaaagaa	catgtgtagc	gtttcacgac	28620
tgtttaagtt	tctgagttcc	tctgacctc	aactccaccc	cctgggaaac	accaaaagtt	28680
ggatgagaaa	gttccccccg	cctacctctc	cccacgggag	tgtacaactg	aggcacaagc	28740
ctgctcccc	cactgccccg	cgatctggga	ccacgtctcc	tccgcgtagc	cgacccgggg	28800
atggacacta	tctggggacc	cggcgccac	acggggcatt	cgggtcgcgc	gggcacctgg	28860
cagggtgtcag	tccgcttgga	aaaccacagc	cacgcgctc	acaggagcag	cgccaccggc	28920
taggcgcgcc	cgcgcgcggg	ctcagaactt	tctcgtgccc	acttcagccc	gtcctcggag	28980
cacgcggggc	ggcgcgcggg	ccgctggaaa	caggcttgcg	aaccggtccc	ccggggccagg	29040
ccgcctccg	cgcgccaaagt	ccccgctcgg	tgcgccggcc	ggggccacag	ggcccagcgc	29100
gggctcggct	cggtcccg	cttcccgcgg	gctcgggcag	gtgaggaccc	gcccgcgcgc	29160
cacctggcgg	agcgggcgc	ctcctcgcca	gcccgggacg	cagcgtcccc	ggggaggggc	29220
cgggtgggga	gacaaaagggc	ccgcgcgtgg	cggggacgc	ggggacggca	gggggatccc	29280
gggcgcgcgc	cccaactcgc	tcccaactcg	ccaagtgcgt	tccgagacgg	cggcggcgc	29340

10

20

30

cgcgcaacttg gccgcggggc cgcccggggc attgtccgag caacccggcg cccgtcttac 29400
acgccggggc cggaaggta tcgaatcagg 29430

<210> 8

<211> 33769

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> (33739), (33749), (33758)

<223> Identity of nucleotide sequences at the above locations are unknown.

<400> 8

cttccctta	caactgtct	tcgacccgc	tcggatgaa	actgaatgg	tttagcctta	60
gaggtctctg	gtctctaagg	gaggtgggtc	aggatgccg	ggacagggtc	ctcttcctgg	120
ggcaacgtgg	gggaacgagc	cacctacccc	tccactgaat	tgccctgggg	tgtgggtacc	180
gacggctcat	tcgggtgtcca	gggtctgaga	tgtgttgaca	ggaagaatga	aaggggatgg	240
gagggatggg	gcgaaagaag	ccacctgcag	ccccaggaa	tatctggcca	gcacaccgtc	300
accagcggc	ctgagccacc	cctgccagag	ccaggaggag	accctgcaa	tgggtcacca	360
gtgtgcagga	actcagaagg	tcatcacagt	taataccctc	catgcccaca	tgtgggaaaa	420
caggtttttt	cacaacaaac	aagataat	ttgttat	ggcaaaagga	ggcagggcag	480
ccccggacac	ctccatccca	cctcatcacc	cagccgcagg	gccccggcca	tccttcgaga	540
cagagtggat	gtcacaacct	cctgcaccg	aaccaagtgc	agctcccagg	ccacaggcca	600
cccaggaaag	gtccagtggc	ccccggaggc	tcccaccgca	ggcctcccac	cacagccggc	660
accaaccag	gatagctgtg	ttctcctggc	ttcttttcac	acgggttagca	gaaagctgag	720
atccggggaa	agctgagatc	cagggaagc	tgagaatcgg	cctctgctgc	cggagcgccc	780
accccagct	ctgctcccag	ctccaggggc	tccttctcag	gtgcccttac	aggaggcaga	840
gggcttgagc	cacctcctgg	gcctggggca	cgcaggatga	acgggggtcac	ggtgcaggcc	900
actgtccact	gcgcagatcc	caaggccata	aacagcctgg	ccacagtggc	ttcccagctg	960
gcaggcggcc	agattat	tggtgtttag	caattgatta	agtttctccg	ctgccccacg	1020
gggttaagtgg	tggggcaaat	gccgcaaccg	cagcatttga	cccgggatcc	tgtgccaaagt	1080
gaccataggg	tcacaaagca	caagggaagt	ggctggggccc	gatgetggct	ctgctggaac	1140
ctgaggccgg	ccactgtcac	ctgcacgggtg	cctgggacct	tccagcaagc	acagagaagc	1200
tatggccctc	caggagcagc	tggcaggcac	cctggcctgc	agtcaggggc	tctgtctgct	1260
cagctctaaa	acaggaaagt	cgctgctctg	cctggggtoa	gggcagccag	agagtgacca	1320
agtcagtgcc	ggcctcagga	agggacctgc	aggcgggtcc	cttctctctc	catccctcgg	1380
tgccagccag	cccctcctgt	ggccccccac	tgccctgcctc	tgcccccatg	ccccaccaca	1440
acctcaggcc	catggctgca	tggccactcc	ccaggcaggc	agtggggatg	ggatttcacc	1500
atgtttggcca	ggctggtctc	gaactcctga	cctcagggtga	ggagttccta	aagtgtctggg	1560
attacaggcg	tgagccaccg	cgccagccct	ccctgtggta	ctaaacactc	acacccctt	1620
gctggggacc	ctggtgaggg	aacacagcct	cacaagtga	gtgtggtttt	gttgagcaaa	1680
tgacgcctgg	gcagccctct	catctttgcc	taaaactgaa	gaatttaggg	gcgtggatgt	1740
ataaaacagt	tggtagctta	aatgaaaaag	aaggccacac	tcccccttt	aggcaggcgg	1800
cctaattctt	taaaagccag	cacagggtgc	ctttctgaac	ccaggcacac	agtaggtgtt	1860
caatggacag	cagcggttac	ttgtactgct	catgacaccc	tgtctgtggc	ctctgcagct	1920
ggctccagcc	tgacgcattg	ctgcgcccct	ccgcaaggcc	accccgttat	acatggaaac	1980
tctgtggaga	aggccttggg	ggccggccag	gacgcaggcc	ccagatccca	tctgcgccct	2040
tcctccatag	acctcagcga	gctctcggca	ccatgtgcct	caggcccat	taagaagtag	2100
ggccggccag	gcattggtggc	tcattgcctgt	aatccacgca	ctttgggagg	cccaaggctgg	2160

10

20

30

gtggatcacg	agatgggtcag	gagatcgaga	ccatcctggc	taacacgggtg	aaaccccatc	2220
tctactaaaa	atacaaaaaa	taagccgagt	gtgggtggcgg	gtgcctatag	tccaagctac	2280
tcgggagggt	gaggcaggat	aatcgcttga	gctcagcagg	cagagggttgc	agtttagcggg	2340
gacgcggcca	ttgcactcca	gcctaggtga	cagagagaga	ctctgtctca	attaaaaaaa	2400
aaaaaaaaata	aaaaaagaag	cagggccagc	cagggaacgac	ccctcacaca	gctcccaggga	2460
cgctgcctg	ggtatagggc	tcaggacccat	gaccgctgca	gtggccccc	agaaacgtta	2520
cttttgtcac	ccaccccgcc	tcagtggcag	tagccaaaat	aacggattag	aatggaacca	2580
tgtgacaatg	ccactgcccc	aactgacaga	agatggctat	cagcagttca	cgcggcccca	2640
ccatcacaca	gtgcagggca	ctctacaact	tatgcacact	tcccagaca	ccgtcccttc	2700
gaccctccca	ggtcagcaag	gcacacagg	cctacatttc	acagccacac	agcagagggc	2760
tgaggctgga	actcggatgc	tctgatttcc	gttcaatcac	atcccagag	gtggcacaga	2820
gacggggggc	ttctcttgac	aaagtcaaga	aagtcaactgc	cagctccact	gaagaccaaa	2880
gaacctcagc	tctcaaaccc	tcttgaaggt	gttacogaac	tctcccagcc	tgtttcctgg	2940
gtcccgatgt	tgggtcccggt	ggacacagga	agagggaaga	gctccctaga	gcagagcctg	3000
gtgcacctgc	cacactctca	gagggctgcg	cacgggaggga	ggagccgtgt	gcaggagtgg	3060
ggtctgggat	gaggggcgct	gtggccgggg	gcagggggca	ggggaagggt	gctccaggtg	3120
gtgggcacag	cacgagcagg	ggcagggagg	tccacactca	gatgtgcaca	gggagaaaca	3180
aatcgtgcat	ttccattgga	ataggcggtg	aaaggtagaa	aaacagagt	ggggccagga	3240
agggagtccg	agccttctag	tgtctctctg	caggtgagcg	gcagcccgag	gtgtcagctc	3300
agcagacttg	gggtccagg	gcggtgtctt	ctatcactga	ccccagggca	cacgggaactg	3360
gggaggggaga	gcagaggcac	agggcacggg	cagtgaacg	aaacaaggag	tcatacccaa	3420
atgccgaaa	ggcaaggagt	gcccgcagcc	gcacaagggt	tctgtctggg	caacgtgggc	3480
gtccaccacg	gccccgcacc	ctgcaagcgc	aaagctcgcc	actgaagata	aagggaagct	3540
gttggagctg	cggagctggt	ctgggggtccg	catggagctg	ggcttatgct	gcagtcacaa	3600
gggggacatg	gaagaggctg	caggggacaa	aaccagtgac	cacagctctaa	ctctgagcct	3660
gtggaaaggc	gcccacagca	ttcaccatc	ccagagatgc	cattccccct	gtgcccccg	3720
tccacggtag	cagcgttctc	caggaatatg	atgcgccctc	ctcctcttgc	atcagccctg	3780
acagtgaag	ttcaggccaa	aaagcagaag	agcacagctg	cgtgggttcca	tttccatgta	3840
gttctggaac	aggcaacgct	aatccaagg	gatagaagtc	aggagagtgg	tggagggggc	3900
gggggttgag	gatggcaaa	gggcacccgg	aactttccca	gtggtagaaa	tgttctctgt	3960
ctggaccgtg	tggtagttat	gcagacatat	gcagctgtca	aagttaatcc	aaatgtacac	4020
gttaaaatgt	gtgcgtttta	ttgcctgcaa	gttataacct	aattaaaaaa	ataaagttag	4080
cactcaggct	tcttccacaa	cttccctgaac	cgtgtgagct	gattttcttg	ctattaaaaa	4140
ttcacgggtc	atgggtgaga	acagcagctg	ccttctgttt	gcaaagtcaa	cgcacatcac	4200
tgcctggccg	cggcagaactc	ggccccacag	gacctccttt	cttttttccc	tttgacctac	4260
ttccctgata	agtgacaaga	cagccagact	ctgggaacaa	acgcccgtta	ttcgcccccg	4320
agctgagcgg	gcccctgcttc	ctgagctaat	ccgcccggac	agacggagg	acgtgagggg	4380
ctttgcgctc	ggctccagct	gtcagctctg	ccgtcagact	cgacagtggc	cccctctgtt	4440
cctcccgctg	cccccaactc	atccccgact	tctttttgtt	tctgttccct	gacagacgaa	4500
catctgttaa	aactctgtct	gggtgagctg	tggccagcgg	cccacaaatc	cccaagccgc	4560
accccgacct	catctgggag	ctgcggggag	cactgcctgg	ccaccctctg	gacatagctc	4620
tgagagccac	cggccagggc	acgtgtggcc	cgagtggcat	ggtgcacgcc	gctaagccca	4680
ctgcccacag	gcccccaagc	aggagggatg	tgcaggagac	aaaagtcaaa	agaacagggg	4740
cacgttccac	agaggatggg	gctggagggg	tggcagtgag	gaacagcagc	ttccgaggat	4800
ggcgttgagg	actcccaaat	aaggccctcac	tcctgtgtgt	tttagctcat	tccacataat	4860
tggaaaaaca	tggcagaaac	cgaagccagc	tgtgtgcttg	gtcctggggc	tgtgtggagg	4920
gggtggggag	gcccggaggcc	caggctctgc	actcgactgc	tggggatgag	agtgactctg	4980
agctgcagag	agcagcatcg	cagccgccat	ggtcccattg	agccccggcc	acgtggggcg	5040
gcagaggctc	gtgggatata	cctgccctgt	ctcatggggg	tcacttcagg	aggggggggg	5100
gagccaggac	acagcccagg	gctagcggtc	accctgcagc	tcagggggcca	cgtaaatagt	5160

10

20

30

gccaccttga	aggcacacag	cagtgcgggg	cccccccg	caccaacgca	tccctacctc	5220
taggaggccg	cctgtgtgcc	cctgggaacg	ctgctccctg	tcccttgggg	tcttgggtgtg	5280
accacccctc	cagcccttcc	cttggggaa	gcacctgact	ccctacaccc	agctgggtttt	5340
catttctctc	aatcagga	aaagcagaat	tcaagacatc	acagaaatgt	cttgcctctg	5400
aactccatga	aagataaacg	gtcagacacc	caggaggagg	tcccagggac	ccttgagtct	5460
cacctgaggc	tctggcttca	aacctcgaga	tgtttccagc	catgctagcg	ccgcccccca	5520
caacctgccc	cacacagtcc	tcccttggga	actcacagat	ttggccccca	cctgcccctg	5580
ttcttctggg	ggagtgggtg	cggtgggttg	gggtggggct	ggggactctg	gatgtgtctt	5640
aagagtctga	gtgattctga	cacagccagg	ccctgcccc	ctcctgacct	tgcgccccca	5700
ggaaaggagg	ccacacgcct	gaagcgccca	gcacaccccc	ctcctgcttc	cccagggtcac	5760
cgcctggccg	tgtgagccgt	gtccccact	gccccctcac	ccaccccgag	tccctcctggc	5820
agcaccagc	cttgggaagct	acttctgatt	acaaccgcg	aagggaagact	cgctccctcg	5880
gcactgaccc	agacagcctg	caccatcacg	ctgctcagca	caacccacac	agccttcttc	5940
caaaccocat	ggagcgggga	gtataatcac	cccccttcta	ccaacggaca	aactgaagca	6000
cagagagggt	aagtcacttt	cctaagctcc	caacacgatg	acaaaaata	gaagggtcagc	6060
cgcgaagtgg	aactaggtgc	tccaagtccc	cggtctgctt	gacactgcac	ctcctcgccg	6120
ccacgggtccc	gggtccgctt	gacactgcac	ctcctcgccg	ccacgggtccc	gggtccgctt	6180
gacactgcac	ctcctcgccg	ccacgggtccc	gggtccgctt	gacactgcac	ctcctcgccg	6240
ccacgggtccc	gggtccgctt	gacactgcac	ctcctcgccg	ccacgggtccc	gggtccgctt	6300
gacactgcac	ctcctcgccg	ccacgggtccc	gggtccgctt	gacactgcac	ctcctcgccg	6360
ccacgggtccc	gggtctgctt	gacactgcac	ctcctcaaca	ccaccaagggt	cccgggtctg	6420
cctgacactg	cacctcctca	ccaccacac	agtcccggtt	ctgcctgaca	ctgcatttcc	6480
tcatcaccac	agtcccggtt	ctgcctgaca	ctgcatttcc	tcatcaccac	gggtcccggtt	6540
ctgcctgaca	ctgcacctcc	tcaccgccac	gggtcccggtt	ctgcctgaca	ctgcatttcc	6600
tcaacaccac	tccttggccg	gtctcccaact	acaaaccag	ccatgtcttc	catcctgaat	6660
cctcttggcc	taaacatcac	tcacaatgcc	tccctcgga	acaggcaca	gtcccaccag	6720
cacagcctcc	tctgttaact	gcgtttccgc	tagccdaggg	ccagctccag	agccctcacc	6780
acagagcctc	tatccttcac	ccccggacac	tggacctcac	caacccatag	cctggaggag	6840
atccctgtgt	gacccagggt	cctcctctgc	cagactctga	atttctactgc	ccaaagtgc	6900
acctcggaag	gctctctggg	cactggcagc	cctccatggg	caccgctcct	tctggccagc	6960
tctgacatcc	cggctgggtga	gggtgcctgc	acgaggcctc	tggccactgg	gacctcacag	7020
cogtgcgtgc	agctgaacaa	agcgacagaa	tttcaagttt	tcttcaagtt	gcccctgggt	7080
gagcagctcc	aggtagtttt	cagtcgagge	gaggcgctcc	gtcagcagcc	aggcggcaca	7140
gctaattcat	gcccgcgggg	cgcacggccg	caataccaat	gggcacctgc	agcctggaaa	7200
gccacagagg	aaccgagaac	agcgactgtg	ctcagggtgac	aggactgtgg	tcttttaaca	7260
aaacattttc	ctttaacgtg	atattttacg	gcaagggaatg	aaacctggag	ggcaggacat	7320
ttggatacta	aagccccagg	ctgcgcgctg	gtctgctttg	tgaagtctga	agccccgcgc	7380
ccattctggc	cccgttcaca	ggtccggctc	tgactcacca	gcttcaatgc	taggcctgtc	7440
ctgtcctcca	accagaacat	gaattcctta	aggacaaagc	cgtttctcgc	ccatcccat	7500
ctccctctgg	attaagaaat	atgggaagat	cttctagaac	caacctcaat	ttgcagagag	7560
ccatcctggg	gacaaacctt	tgaatgctt	ctaagaagag	tttaggtttc	ttctcaactc	7620
taaaacctct	agaaaactct	atttccacac	cagctgcccc	tggaaacttt	cagcttcaaa	7680
agggcccagg	gcaggggagc	ggaggagcca	gcacccacac	cgagcaccag	cctgttaatt	7740
aacgggaagc	gggtggggcc	catctccagg	cagctctgag	gtcagactgg	ggaaccatgc	7800
ttacaaaaaa	aagtgaactg	aaacgctcac	gtcctcatgc	aaaaaccagc	tcccagttgc	7860
atctttctgt	ctcattgagg	agcttttttc	tccctttgac	agaacacctt	acacacggca	7920
tctggaacca	aagcagaaag	attcaggctc	agagtataac	agtccccaca	ctggctgcat	7980
gtggacgttc	cgggccagga	gtctcgccca	agcagggcct	ataaatgaca	caaatgtttt	8040
ttctcctgcg	tgccagtcac	gtctcaactg	agttatgtgt	aaaagtgcct	ctcacggctg	8100
agggcaaaaa	cagttcccac	aagactagag	aaagggtgac	cctgacggct	gagtctctag	8160

10

20

30

ggagcgtgga	gctgcgtgct	cagccctgcg	gccctgacgg	ctctggaatg	gaaaagctat	8220
ccaactggaa	gggcagggct	cgctgctagt	ccagcgggtcc	aacccacacag	gtgtctgtgg	8280
tgtcagctcc	atgccacaga	gcccagggtct	ggggccagag	ccaccaggcc	ccctgccagc	8340
ctgcaggggc	ctcctcctct	gggtagccta	accacccctt	gtgagcgcag	gcagcctcct	8400
ctaatacaca	cagggcctgt	ccccccctct	cccccgcttg	caggaaaatg	agccctgagg	8460
actccccagg	gctgctcttg	gcctggacat	ggagactggg	aattacattt	gcagaaggag	8520
cgcaatgccc	ttgaagggtc	cagccaagag	cagccagtcc	ccagggctca	gaaggcccag	8580
ctgttagaac	cctggggagc	agcaaaagag	caggggtctcc	acctaaagtct	atagcccctg	8640
cctcttcttg	ttgggaaaga	aatcaacgcc	cctttactgg	ctcccactga	cagcccactc	8700
ccccagggtat	gggaggattc	tgggaogatg	caggcaaac	tggaccctga	gtgaacctgc	8760
cccagctctc	acgggccttg	ccccagccac	agcacctaag	gcgcgggtca	tggtgacaac	8820
atgaagggtga	taagggtcatg	gacagtggac	atggcagctg	gacactgggc	accactgga	8880
tgcaggcac	ccagcacggc	tccgtcacc	ctggatgagc	agtggccctt	tgcaaggccag	8940
ggtagccttg	gcaagttatt	tgggggtctc	caagcttgct	cagctgtgcg	acttctactga	9000
gccatgagtc	tgggatttta	tcaggggcca	caccogttcc	tggaaactctg	atacgtgagg	9060
gagccacaca	gggaccctta	acaaaagctc	ccagggcaac	atgttctctt	gcctcagtct	9120
cccaaatagc	tggtgattaca	ggcgcacgac	taccgcccgg	ctaatttttg	tatttttagt	9180
agagacaggg	tttcaccatg	ttggccaggc	tggctctgaa	cccctgacct	caaagtatcc	9240
ttcactgtt	agggcaaggc	acctgacagg	cacgactgca	cgatctgctt	gttgggggct	9300
gtgtccattc	cccactcctt	cgacaaatgt	ccacaccagg	ccttgctttg	acaccccaag	9360
aacagagatg	gtgacacctg	cttctacat	gcccattgct	ctcccaaggc	agacatcccc	9420
agcagatgca	acacagctgt	taggcagaca	tcaccaatcg	atggtggcaa	cagacaccag	9480
gcccgtctcc	ctctaactcc	agtggccagg	ccccaaagcca	gctctcacct	gcccactccc	9540
aaccocagc	agcaagactc	agaaatggca	aaaacacaaa	gagaacagaa	acgcccata	9600
gcgggaggat	gactaaaaga	catgtcttga	taagatattg	ttcaggcata	ggccaggcac	9660
agtggctcat	gcctgtgatc	ctagaacttt	aggaggttga	ggtagggtga	tcacctgagg	9720
ttaggagttc	aagaccagcc	tagccaacat	ggtgaaaccc	catctctact	aaacatacaa	9780
aaattagcca	gacatagtag	cgggcgccctg	taatcccagc	tgcctggggag	gctgaggcag	9840
gagaattgct	tgaacctggg	aggtggaagc	tgctgtgagc	cactgtactc	caacctggac	9900
aacagagcaa	gactctgtct	caaaaaaaaa	aaaaaaaaaaa	gatatccttc	actaaaactc	9960
atgtctttga	tacatatatta	cctcctgcaa	tcgcaaatgc	ttctgcagtg	cataaagtga	10020
aataaatagc	aggaagcctt	acgggttogat	cacccacaca	gacacacagt	cacatacagg	10080
aaaaacgcag	ggagggtctg	ggaacaaaaa	aacagaagat	aaaatgtgga	gacagacaca	10140
ccaagagagt	aagagaccac	ctccagacct	cccttcagct	tctcaaacac	acgagccggg	10200
cccgttacag	aattttcggg	gaccgctgca	aatgggaagt	gcagacagcc	ccttactcaa	10260
aaggtaggaa	tttcagggtca	acaacagagc	tcacctcata	tgactacaca	ggtcacacag	10320
cccgtgaagt	cggtcccaac	accagcatgc	tcctgcctca	aagccgctgc	acgtgctggt	10380
ccttctogcc	tttccctctt	ttagtccctc	agatctcagg	cctcctgaga	gagacctctg	10440
acctgcgggc	tcaggcgggc	acacccccag	tacaggagtc	tcgggctcag	cccctgctgt	10500
gttcogtacc	cgatccagggt	ctgtccctatg	tccatctgtg	tgcgggcttg	cttcttgaca	10560
tggcccccac	cacacgtgtg	cctcggggca	ggggaacagg	cccgtctcat	taactgcttt	10620
cttctcagat	attttctgga	atatttgttg	atatttggga	acatatatgc	tccacctttt	10680
tcagactagc	caggacagagc	tgcatttttt	tttttttttt	tttgagacag	ggtctcactc	10740
tgttgcccag	gctggagtat	agcggcatga	tcttggtcca	gtgcaacctc	cgctccctag	10800
gctcaagcaa	ttctctctgc	tcagtctccc	aagtagctgg	gattacaggc	ccgtgccact	10860
actgcccagc	taatttttat	attttttagta	gagatggagt	ttcaccatgt	tggccaggct	10920
ggtcttgaa	tcctgacctc	aaatgatcca	cctgccttgg	actcccaaat	tgttgggatt	10980
acaggcgtga	gccactgcgc	ccggcccagag	ctgcctgttt	tacacctttg	ccatattccg	11040
gtgattctct	ctccctctcg	tcctccgggc	ctgactgtgg	tggccactcc	ctgccgtcat	11100
gagccogtat	gtcctcaactc	tttccctttc	cgcaggagact	tcaaccaaca	ctgcagagcg	11160

10

20

30

cagggtccag	ctccagcact	gagttcagcc	totttctcacc	aacagacagg	caggaaagaa	11220
aacaaactct	gagaaggcca	aggttcccg	gcagccagca	agccaagcat	ccttctccgc	11280
tgaggttgt	gcagccgagg	cacccctcc	tccagggagc	aggcagcgtc	ctggggcagt	11340
ctgcgaggga	gaccagggcc	cttgctccac	cagggcccca	ggatbgggg	cagcagcaaa	11400
ctcatggctc	tggaagccag	acccacctg	ctagaacctc	ctatgccacc	tgctgtgggc	11460
aaccccaggc	tggtgacttg	ccctggcctc	ctctgtaaac	aaagggtcca	tccaacctgg	11520
tcaaacact	cctcccttc	aagggtctat	aatcctccct	taacctgctt	ggtccaaacc	11580
cctggtgtcg	ccaggtcact	caggagcgag	ctcatctgga	ctccttccct	gggtccagtt	11640
tctctctcaa	cattgccttt	gaggccgagg	tgaacggtca	acagcgaagg	gccccagagg	11700
tgatggagga	gcgggtgtcc	aagacactca	ccctttctaa	tgactgact	ccctcgtgga	11760
ctcacttgtg	ccgtctcccc	caccacccca	gccccagagc	ccagagtgcg	agcgccagag	11820
gccccggatt	ctgtctgcac	cgcggggtcc	ccagtgcctc	ggagcaatgc	cagcaccggg	11880
caagtgttcg	acaaatgcct	gctgaatgag	caaatggatg	gatgaacgaa	tgaatgagca	11940
agcagatgaa	tgaatggggg	gctgtccaga	gccgtgagga	ctaggccggc	caagtcccca	12000
tttctcaaat	tctccttctc	ccgacttggg	aaacaagatg	cttggtcggg	gaggtctctc	12060
aaccatcccc	tgccagcagg	ggcacagcgg	acagaccctt	tgatgtaaca	gccatgtctt	12120
cattaaagat	gccttgctct	cagaaagaga	aagacaaata	caaacctgga	aaatcctcac	12180
caaacgcagg	acccctgcc	gggagcagag	aaaagaccca	cacgccacgg	gcgccacgac	12240
cacacacaca	ccccagccgc	tgacacacaa	cacagaccct	agccagcaag	aacaggggga	12300
ccaggaaact	gttcttaaa	tcaggacccc	catgtgtcca	gacagcagtg	agagcaagga	12360
cacttctcca	tccaccggat	gccaggagag	tccttttagg	gggccccaca	ccgagactct	12420
gcccttagga	ctgttcctga	gtgtggaagc	cagccactt	ggaagcccc	tgccctcccg	12480
agtgggacac	cggcacagga	agcaggccct	gtccccacc	actttctgca	agctgggccc	12540
catcacgcta	cagaaacggg	gaggactggg	cccagggatg	gcgctttcct	gacacctctc	12600
gttaccacct	cgcttgccag	gccccagggt	cagccccaga	ggccagactg	gctatcccag	12660
gccccggagc	atccccgaag	gcgagctgca	tcctgaacgt	gtgtgatttc	ccgaagggcc	12720
cgccccgaac	cgacacctgg	aaagaaagat	cctcagccgg	tgccccagag	gagaagagcc	12780
atgcctcact	gcaacacagt	cccagggaagc	accaagtgc	tgaggaccaa	ggcggagagt	12840
aaaaaagtgg	aaaatatctg	gggcaaaaat	aaaacaaaac	aaaacaggat	tgacctcctg	12900
ggctcaagca	atcctcccaa	ctcagcttcc	cgagtagctg	ggaccacaga	cttgaatcac	12960
cacaccggcc	aagtggatca	tttcgaacgg	gtttgcagag	gttccttctg	gggcaccccc	13020
ggggggccga	acccattccc	gccaggcccc	gccccgccc	cccgccccgt	cccgccccac	13080
cgctcacct	gccttacacg	tcctgcccgt	gtcctgcagc	tgacacccc	tggggcaggc	13140
gcatgtgtag	aaaggctcgc	ttggggacag	caggcacagg	tgggagcagc	cgccattgtc	13200
ctcctcacag	cgagtgtgga	ctgagaaaac	caggacagac	tgagagaagg	ttccagaaga	13260
ggaccgtcac	ttgtttctga	atgagtcaca	tcctgcctcg	tcctccgtga	cagcctccag	13320
tgtgtccctc	tgcccaaaac	tcggcctcaa	gtggcatcag	ggacctcccc	gcgggcacca	13380
ttccacctgc	ctcatcgctg	gccccgtcca	catggggccc	tcagcctggc	cagacggcct	13440
gcaatttccc	caaaaccagc	cgtgaccttc	ctggccaccc	tcacaccag	atgtgacctg	13500
cccatggagt	gacatcctcc	ccatctgctt	cctccaccca	agctcctatg	actagaacac	13560
cctccccagc	tcctcgagag	ccccaaagga	cacccctctg	caaaggctgc	ccccacgct	13620
ccaatggccg	gggtcaggac	ctgcctgtgt	ggtagtgcg	ggaaccccag	agacaatggg	13680
ctcctgggca	aaaggcttgt	cttgtctttg	tgctatgtgt	ggacccagca	gcttccatag	13740
gaacactgtc	cttctgtctg	ggatggccaa	gcttgtcact	ctcccaagcc	ctcctatgac	13800
caacagcaat	tgaacggaac	tcgataaatg	cttcacgac	ctcattcaaa	ccaggggaaa	13860
gctgggtgta	gcagccccaa	aatacggata	taactggaac	aacaaactca	tcaaaatgaa	13920
cctctccctc	cctcatgctg	ccccaaagtg	agatgggttt	tgtgaccacg	actttctcac	13980
caggaaacag	ctccagagag	ccccaccctc	ctgtgtcctg	ctctgggaac	agctggcacc	14040
cctaggcccc	acatttcaat	tcaaagtcca	aaccttccat	aatggcctgg	ccagaaatct	14100
ccatccctgg	tcctgtgtgg	agtgggccc	tgccccaga	gcgcagccc	cactgtcaca	14160

10

20

30

gaagctgggtg	cattttcccca	tcagggaacct	ctgtcacaaac	ccagcgtggc	ccccaggctg	14220
agaactgctg	attctgggca	gattattcat	tgataaatac	gcgacttgca	gggccaagca	14280
tggtgggtca	tacctgtgac	cccagcactt	tgggaagtca	aggtgtgagg	atcactggag	14340
cccacgagtt	tgagacaagc	ctgggcaaac	tggcaaaatc	tctcatctct	attaaaaata	14400
catacacaca	cacacacaca	cacacacaca	cacatatata	tgtatatata	aataaccata	14460
tatatatata	cacacatacg	tgtatgtgta	tataaatata	tatacacaca	cacacagaca	14520
actttctctg	ggccttgaaa	acgaggcaac	cttccttgga	aatcccttg	ccactgctga	14580
gcctgaaata	gccccatga	gctctgcaga	gggtctctct	gcaggcccg	gtcccccagc	14640
cagccacaca	cctccctcca	ttgcagcagg	taccccttta	gagagggggc	cccccagagc	14700
atgggcttct	gcaggggagg	gtcacctgcc	ccccaccccc	acccacgccc	gcgcaccccc	14760
acgccccgcg	atcctccccc	tcccctgccc	cgcgcccccg	ctccccccag	ccccctcacc	14820
ctctcccccg	tgcccccaacc	ggcactcaca	aaaaggctgc	cgctcctggc	tcagcacctg	14880
gatgtccatg	gggtgagtata	gggcactcag	gatctccttc	ctcttcccc	cagtgcgctt	14940
gttgagggca	tggatggagc	gggtctgcca	gtctgtccag	tacagagtgt	ccccgggag	15000
cgtcaggggc	aaggggtgcg	tcaggctgcc	ctccaccacc	ttctgcctgc	agtcagggaa	15060
gcgggggtgga	ggagccatca	ggagggtccc	cgcacagtca	ttgtgctgta	cccaattaat	15120
ttcttttttt	tttttttgaga	tggagtctcg	gtctgtcgcc	caggctggag	tgcagtgatg	15180
taatctcagc	tcactgcaac	ctccgcctcc	cgggttcaag	caattatcct	gcctcagcct	15240
cccagtagtc	tgggatcact	gatgcccacc	actacgcccc	gatgattttt	gtatttttag	15300
tagagacagg	gtttcatcat	gttgccaagg	ctggctctga	actcctgacc	tcaggtgatc	15360
caccacacct	agcctctcaa	agcgtgagg	ttacaggcgt	gcgccaccat	gccaggcttc	15420
ccatttgctt	tcaaccagac	aagttagggc	aggccaagag	ccccaggagc	tggcgccctc	15480
gtacatttct	cccgcgctgc	acagggcacc	tcccaaacac	agcctgtgat	ggtgacacac	15540
gggtcccccc	aggccaagtgc	gcaaagtctc	ccccagggaa	gaaaggagga	agccatgcct	15600
ggcaaaaagc	acacctctcc	tgcccacagc	tttaacctct	gtatacaaat	caggccatgt	15660
gcactcgctc	cttcttcaaa	tgctcataat	ttatactttc	agagtaaatg	aaacttggca	15720
tcaacccgag	aaacagctat	tcttttctag	atgcttacag	tgcccagcaa	atgaggactc	15780
gggtgtaaatg	agattatgga	cactggaaac	aggatcataa	tgtgacgtgg	tcggtaatgt	15840
gcagttttat	ttgcttaaatg	accctcgccc	cgtgacaggc	tccctgaggg	tgggctgggg	15900
ggcagaggtc	cccgccacgt	ccccagccct	cagcacagtt	gccaggagag	ggtgacactc	15960
atgaagtggc	acagggaaga	tgggagctgt	gggtctgtga	gatccaccac	ctcttctgtt	16020
cattttttgt	gatgtgtgtt	tttaagaaaa	ttattgaagt	aaaattcaca	ggacatacgt	16080
ttactttttt	tttttttttt	ggagatgggg	tctcactctg	tcacccagg	tggagtgcag	16140
tggtgtgatac	tcagctcact	gcaacctctg	cctcccagg	tcaagcgatt	ctcccacctc	16200
cgctccagga	gtagctggga	ccacaggcgt	gcaccaccac	acccagctaa	tttttggggg	16260
gtatcttttt	ggtagagaca	gggtttcgcc	atgttgcccc	aggctggtct	tgaagccctg	16320
agctcaggcg	atccaccgcg	cttgccctct	caaagtgtctg	ggattacagg	cataagccac	16380
tgcaccagc	ctaaattttac	cactttaaag	tgaatagtgt	tacctagtgc	attcgcaagg	16440
cggtgcagcc	tccactctctg	tctagtccca	aagcacttcc	attgccccac	aggcaaaccc	16500
cacacccggc	agcagtcctg	ccccagtcct	cgcccccagc	ccgggcaaac	acttttgatg	16560
gacttaacta	cacacattct	caacatctca	tataaacgga	atcaccaat	acagcctctg	16620
atgtctgtct	tctttgaact	ggcaccatgt	tttcgagggt	catccaggct	gtagcatgtc	16680
agtgtctcat	cccgtttttag	gggtgaacca	tattccagtg	tgcagacaga	aaccaatctg	16740
tgcattccatt	caccacttgg	gggacctttg	tgtcatttcc	acccctcggt	gttgtgcaca	16800
gtgctgctac	ggacatttact	gtccattcac	attttgtgtg	aagacctgtt	tlogattctt	16860
aagagtatac	agctaggagc	ggaattgtctg	ggtcatacgt	aaatcaatgt	ttacgtctca	16920
aggaatcaac	aaactgtttt	ccacaatgtt	gtcttttttg	tttgttttct	gagacaggg	16980
cttgctctgt	caccaggtct	ggagtgcgg	ggtgtgatca	tggctcactg	cagcctcaat	17040
ctcctaagct	caatccatcc	tctgtcctca	gcctcctgag	tagctgggaa	cacaggatatg	17100
taccaccatg	gccagctaat	tttctaattt	tatttttttt	tgtttttgtt	tttttgagac	17160

10

20

30

agagtctcgc	tctgtcgcgc	aggctggagt	gcagtgggtgc	catctcagct	cactgcaage	17220
tctgcctccc	gggttcacac	cattctcctg	cctcagcctc	ccgagtggct	gggactatag	17280
tcacccggcca	ccacgcctgg	ctaatttttt	tgtattttta	gtagagatgg	ggtttcacccg	17340
gtttacccag	gatggctctg	atctcctaac	tccatgatcc	acctgccttg	gcctcccaaa	17400
gttctgggat	tacaggcgtg	agccaccacg	cccgacctta	cttttaattt	tttaatttta	17460
ttattttatt	ttattttttt	tttttttgag	acagagtctc	gctctgtagc	ccaggctgga	17520
gtgcagtggc	gggatctcag	ctcactgcaa	gctccacctc	ccaggttcac	gccattctcc	17580
tgcctcagcc	tcccagtag	ctgggactac	aggtgcccac	cacgatgccc	ggctaatttt	17640
ttgtattttt	agtagagaca	gggttttact	gtgttagcca	ggatgatctc	aatctcctga	17700
cctcgtgac	cgcccgctct	agcctcccaa	agtgtggga	ttacaggcgt	gagccaccgc	17760
gcccagccct	tttttttttt	tttttttttt	ttttgagata	gagtcttgct	ctgtcgccca	17820
ggctggagtg	cagtggcggg	atctcagctc	actgcaagct	ccgcctccca	ggttcacgcc	17880
attctcctgc	ctcagcctcc	cgagttagctg	ggactacagg	caccaccac	cacacctggc	17940
taatgttttg	tatttttagt	agagacgagg	tttcaccgtg	ttagccagga	tggctctgat	18000
ctcctgacct	cgtaatccgc	ccgcctcggc	ctcccaaagt	gctgggatta	cacgcgtaag	18060
ccatggcgcc	cagcccatgt	ggccattttt	cagttagaga	agccagaggc	ccatcactct	18120
cggttgctcc	ctgggcatg	ctctgcctca	gccagaagca	ctgagggaag	gtcagcctcg	18180
gcccttgccc	cagccacagt	cacagataaa	ggggcctgca	caggctctgtg	tggctccaga	18240
gctcgtcacc	caacacacga	cgcttccatg	tgaatagccc	cagggtgcatc	atgaagagcg	18300
atggccgctg	cagaggcaga	agaatcccg	ggggaagcag	gtgggagaga	ggctgagaac	18360
agaccagacc	ctggagctac	agaccctatg	ttccaacct	ggctgggact	agctgtgttg	18420
cacgcctgta	atccacatgc	ttctctgtgc	acaggggato	aaaatagcaa	acacaggcta	18480
ggcacagtgg	ttcacacct	taatccagct	gctttgagag	gccgaggtgg	acacatggct	18540
taagctcagg	agtttgagac	cagcctgggc	aacatggtga	aacctcgtct	ctacaaaaaa	18600
aataccaaat	aaattagcca	ggcgtggtgg	tacgtgcctg	tggctctcagc	tacttggaag	18660
gctgagggcg	gaggaacact	tgagcccaag	aagtcaaggc	tgtggcccg	tgtgggtggct	18720
cacgcctgta	atccacatgc	tttgagaggc	tcaggtgggt	ggatcacttg	tgatcaggag	18780
ttcaagacca	gcctggccaa	catggtgaaa	ccccgtccct	actaaaaaaa	tacaacaatt	18840
tgcaggcgct	ggtggcgggc	acctgtaatc	ccagctactt	gggaggctga	ggcaggagaa	18900
tagttagaac	ttgggagggtg	gaggtttag	ttagccaaga	tgggtgcgct	gcactccagc	18960
cagggggaca	gagcaagact	ccatcccaaa	aaaaaaaaa	acaaacaaac	aaacaaaaaa	19020
agaggtcaag	gctgcagtga	acctgatgtg	tgccaatgca	ctccagcctg	ggtgacaaag	19080
tgagaccctg	cctcaaaaaa	ataaaaaat	aaataaaaa	aaaacataat	agcaaacgtt	19140
tcatagaggt	ggtatgagca	ttaaatgaac	tgataaacgt	ccctggaaaa	cagtaagtgc	19200
tatggaagga	ttcgctgcgc	ccaccgccac	caccattagc	atgtttcaac	ctccatcacc	19260
ctcactgtcc	cctgtcacca	tcctttgacc	agggcactcc	cagctgcagc	ctttctatcc	19320
tcttgtccac	ccttcataac	tgtaagatca	ctcagctccc	aagaaccaca	gtctacaggg	19380
taaccacatt	tccaaatctc	aaaccagacc	cgctggtctg	cacttccagg	gacaacagga	19440
tattttcaaa	ccagcccaaa	agagatgtgt	ggctcagcat	aagaggaaaca	ggagaaactg	19500
aggcctcttg	ccctgagaat	gagcttggaa	gtggatgtcc	cggcctcact	caaaccttca	19560
gatgactgag	gcccagccag	gagcttgagt	gtaccctcag	gtcataccct	gagccagaag	19620
caccacgcta	atccactcct	catcactgac	tcctcccca	taaaaaacct	gtttgctgtt	19680
tcaggctgtt	aagttgtggg	ctgttttggt	acacagcaat	ggataactaa	cacacgaggc	19740
ctggcaagtg	tggagcaaa	ctgcccgaag	cctcaagtct	gttcattgtg	gtgttgccct	19800
gtgtttcag	aaatccagcc	actgagtcct	cccatgcagt	cactactgcc	ctctgcacag	19860
acacctgcca	catccctgoc	tgggccagga	gtccactag	tgcaggaaatg	gggtctgccg	19920
tcccaggagg	atccctgaca	cctagcacag	ggctagcagc	aggcagcact	tggttagtga	19980
ataaactgcc	cttcacctgt	acacagaagg	gatgtttcta	taaggggtaa	ttaagtacag	20040
agctgggaag	ctatgctgac	cagaaggctc	taaaagcaat	taaccaacga	ggggaaaacc	20100
cttcctactc	attctcgccc	cattttattg	agcactgacc	atgtggaagg	ccccctgggtg	20160

10

20

30

agactgggga	atgcaccaat	aactgagaca	gottccggct	gttgccctca	ggatgcctga	20220
gctgggtag	ggccagggtg	gggtgggtgc	gtgtgacagg	gttactgttc	acaaccctgc	20280
cgggccataa	gcccctccca	acaattccaa	aatccaaaac	gctctgaaga	tggaaagctt	20340
ttgttgetca	tctggtgaca	aaacctcatt	tgggtgcatgg	gccgggtgcg	gtggctcacg	20400
cctgtaatcc	cagcactctg	ggagccgagg	ggaaggatcc	cttgagctta	ggagtttgag	20460
accagcctga	gcaacatgtg	agaccocgtc	tctaccaaaa	atacaaaaat	tagccagggtg	20520
tgggtggcgca	ctcctgtagt	cccagctact	cgggaggctg	aggcgggagg	atcgcttgag	20580
cctgggaggt	gggggctgca	gtgagctgag	attatgacat	tgcactccag	cctgggtgaa	20640
agagtgagac	tctgtctcaa	aaaaacaaag	ttaaaaaaaa	aaaaactgtg	catgggtgtg	20700
ggctacagat	agtcttttct	gccctactta	gaatgaacgt	gccacatttg	ctatagaaat	20760
attcaagggc	tgggtggcaaa	tgccacacag	accctgacgc	tgttccaagt	tctgagaagt	20820
cctgcattcc	tcagggcccc	agagtttcag	agaagagtct	gtaggcctga	gttaagaagg	20880
aacgccttca	aaagccctgg	ggacaaaggg	gaaaggggtg	ccccaggact	gcgtgggtac	20940
ctaccggaac	gagccgtcca	ggttggcacg	gtggatgaag	ctgagcttgg	cgtcagccca	21000
gtagagcttc	tgctcctcca	ggtcgatggt	cagtcatttg	ggccagtaaa	tgtccgagtc	21060
cacaatgatc	ttccgggtgc	tgccatccat	ccctgccgcg	tcaatccggg	gcgtctcacc	21120
ccagtctgtc	cagtacatgt	acctgtgacg	ggggcagggc	aagagaagca	gctaacacag	21180
atctgttttt	tgtttttgtc	tgcatagatg	cagacatgaa	acaacagaca	gtgaacttgc	21240
cctaaaatct	cacccatcgg	aaataaccaa	caggatgggt	ttcaggattt	cctgccttaa	21300
gctgggcaat	caaaaataac	tatttccaac	ttgttctcag	ttaacagtaa	attctgggca	21360
ccttcccttc	ttgtggatag	aaagattcct	tgttcttttg	atgattgcct	agtgtactct	21420
gctgtaagtt	ttttaagaaa	cttcagggtta	tttctgattt	ttttgttacc	atgaaaatgc	21480
tgtaaatgaa	cctctaaaag	gcaattcaaa	acactcagga	tggaaatatta	tttagtggtta	21540
taaagaaatg	agctatcggc	tgggcccagt	ggctcacacc	tctaattcca	gcactttggg	21600
aggccaaggc	gggtggatca	cgaggtcggg	agatcaagac	catcctggct	aacacagtga	21660
aaccccgctc	ctactaaaaa	tacaaaacat	tagccaggcg	tggtagtgag	cacctgtagt	21720
ccagctact	taggagctg	aggcaggaga	atcatttgaa	cccgggaggg	ggaggttgca	21780
gtgagcagaa	atcgacccat	tgcactccat	cctgggcgac	agagcgagac	tccatctcaa	21840
aaaaaaaaaa	aagaaaagaa	aagaaatgat	ctatcaagcc	atgaaaagac	atggagggaaa	21900
cttaaatgca	tgttagtagg	tgaagagacc	aatctgtatg	agtcaggttc	taaacactct	21960
ggaaaaagca	aatacacaga	gacagttaaag	catcagtggt	tgccaggagt	tggagaggag	22020
agggatgaat	gagtgaggca	cagaaaatca	gggcagtgga	actatcctgt	atgacatgga	22080
atggtgggtg	catgtcctta	ctcatctgtc	taaaccaaga	atgtacaaat	caagggcgaa	22140
ccctcgtgta	aacgtggatt	ttgggtgatg	gtgcgtcagc	cagctttcat	cagttgtaac	22200
aaatgtacca	ccctgcacag	gatgctgaca	gttgggaagg	ctgtgtgggt	gtgaggacag	22260
ggatgtatag	gaactcagta	cctgtctgtc	atcaattttg	ctgtgaacct	acaactgttt	22320
gaaaaaatta	agtctattta	aaaacaacaa	aacatggcca	ggcacgatgg	cttgacactg	22380
taattccagt	acttcgggag	gctgaggtgg	gtgggtcact	tgagccaccc	tgggcaacat	22440
ggcaaaaatc	cacctctaca	aaaaataaaa	attaaaaaaa	agttagctgg	gcatggtggc	22500
acactcttgt	agteccagct	acttgggagg	ctgacgtggg	aggatccctt	cagccctggg	22560
aggtcgaggc	tgcagtgagc	tgtgactgta	ccactgcact	ccagcctgga	tgacagagtg	22620
agaccctgcc	taaaaaaaaa	aaaaaaaagg	ctgggtgcgg	tggctcatgc	ctgtaattcc	22680
agcgcttttg	gaggccgaga	tgggcggatc	acgaggtcag	gagatcgaga	ccatcctggc	22740
taacacggtg	aaaccocgtc	tctactaaaa	gtacaaaaaa	aaaaattagc	cgggcacgtg	22800
ggcggaaccc	tgtagtccaca	gctactcggg	aggctgaggc	aggagaatgg	cgtgaacccg	22860
ggagggcgag	cttgacagtga	gccaaagatca	caccactgca	ctctcagcct	gggagacagc	22920
aacactccgt	ctcaaaaaaa	aaagaataaa	acccatggct	gggatggacc	ctgaacctgc	22980
agctgcagct	gttctctgggt	aggtctgtgg	gcgacgtggc	tttgcctctc	catgttccca	23040
agagacaagc	atcaccatcc	catgagaaac	aagcacatcc	tcagggcgcc	cttacgtgat	23100
ctctggccaa	tgaaccaaga	caaagtgagc	agacaccagg	tctgggatgg	caggctccac	23160

10

20

30

ccccaccagt	gcccagtggt	cctgttttgg	aggtgaccac	aggggtgtgtg	cccagagggt	23220
gggctgtgact	ctcagcggag	accagagggg	aaccacacca	gcttgaggga	ctcagttccc	23280
atcccagcca	gctgggatga	gccacaggac	acaagggctg	gcagacctat	tgtgttttgt	23340
ccacccttca	cagcagagaa	aggggacagt	gcccagaatg	tccctctgagg	agcctcctcc	23400
cactcttgggt	cottgtaaaa	tgggtgctgac	tcccttgctc	ccttcttctc	ggggtgggcg	23460
gcaaacccca	ttccctcag	ccttagcaag	tgatttagaa	acaggcagct	cgcccaagcc	23520
aggcatgaga	gtgatcccg	gacacaggga	gaacaagccc	cgctttgccc	tctgggggtc	23580
tccattcagc	agaagaggca	aatgacagac	acacagccgc	ctcctccccc	accatggtgc	23640
tctgcagcct	caggagcctc	aggtgcacca	agggccaccc	catccagggg	gcatgtcttc	23700
cttgagtgggt	atcgttcctg	agcgagtacc	atctccacct	tccagagggg	ctgtgacaag	23760
atcaacaaga	atgaggcat	aggagcctcg	aaccaaacad	gccctcttcc	ctgcagaggc	23820
tgactgcgcc	cagctcttat	caccaagccc	ctgctcctcc	ggccccgtgg	ggacagggtg	23880
agaggggtgt	cacatggaac	agctctccaa	acagtcctcc	tcaagctgct	gtctcctgtg	23940
catctagtga	gaaccaaac	aacaaaggga	aggtgggaat	tgctattccc	attaggcaga	24000
tgagaaaact	gaggcccca	aaggctggcc	tgttccaggt	tacaggcgct	gagcggtctg	24060
tctgggaaca	cacttgggtg	ctgctgagg	cccgagcccg	gccatcatat	gactcaccct	24120
tccccagcaa	agcccgggtg	tgggtgaact	tttccctggc	gcttgggaact	ccaaggtgct	24180
ggcagccagc	ccagggaagg	ctcccgcgtg	cctgcggcag	acgccttgc	ttacctgcac	24240
gtccccaccc	ctaggagcct	ggacagagcc	cagaccctcc	gccacctcct	gagaaggtat	24300
caggggcac	agtctggact	tgggggggaa	tccacacagg	ccttccccaa	atgctccacc	24360
gtggcccatg	gaaaaggctg	gaaaacgtgc	aggagcagga	gcttccgcat	ggagcataat	24420
tcacattcct	tcccagagtt	tcataacaga	ggcctgctgg	tttcccttaa	tggggaattt	24480
gcgagccagt	cgggtgaccg	agactggttg	gcgtggagct	gctcttgcat	agtctcaaac	24540
gctaccacaa	gcccagccaa	attccacgga	ggaaaatcga	cttccgaaga	aaagagctgc	24600
agcatggcct	tccgtgcagag	ccagctgcgg	ttgtgtgtgt	gtgttatttt	agggaggggc	24660
cattttgcat	tttaaaggag	gggttgggtt	tcacctgtgc	tttaatttga	gacccggggg	24720
ccactgcagc	ccttgttcag	gctggtacag	gccggggact	cctcccatgc	taagccagtgc	24780
tctttctggc	cccagatcct	caggggccag	aggggtcatcc	ccagagcccg	ctctgccacc	24840
cacatgggta	ccttgggctc	gggagggatg	tgccttccct	caacctgtcc	tggatgtccg	24900
cacggggcca	cctgcattgc	tgaactgca	acgaagtcca	gtctcaggag	gggccccctc	24960
ggctgcagg	ctcttgatcc	ttttggccac	gtgcacactg	aggtggacgc	tgggacccag	25020
agacccctt	catgatgatg	gccggggcag	gaacccctcc	ctctgaggaa	ggacctgggt	25080
gggggacagc	actgcaggag	ggcacaggag	atgacggggg	ctctagcagg	gccgggagga	25140
aggccaagat	gctcctcgca	accgtgtgcc	tgtggccagg	acagaggaca	aaacccacct	25200
ccactgtccc	cactctcagg	acagcagtc	tgccccagga	ctcagcgccc	acacttatgc	25260
ctgaggacca	ctattcaagt	cagtatttgg	cgagcagggg	ttgctgccgc	ggcgctgtg	25320
acaggctgga	atcctctccc	tctccctctc	cctctccgga	gacatggagc	ctacaggggac	25380
agagtcagca	cctcagggtg	ggaccatggc	tggcgctcat	agcatcactg	gatctgatga	25440
gtgggagccg	gcatctcact	gttttccact	tctcattcaa	atgactggag	caaaggggag	25500
gtgtggggag	aggcccagga	atcaacacta	aggtcaactt	tgcccccagg	ggcaggggtg	25560
ggagtgaaca	gccacagggt	tgatcctggg	gagggcttct	gggagagaat	tcagagggcaa	25620
gcatgtagag	gaaccatttc	aaatagttaa	gaaaagccag	agccaaacag	ggacagttgg	25680
ctcgcagaga	tgatgcaggc	aaagccagct	cagatctgag	catgggaaag	actactccca	25740
accaaggggc	cagcatctcc	caaccaagca	ccaagtaacct	cccaaccaaa	tgccaagcac	25800
ctcccaatca	aatacctccc	aacccaagcac	ctagcacctc	tcaactggac	accaactact	25860
cccaaccagg	caaccaagtac	ctcccaacca	agtgcgaagc	acctcccaac	caagtaccaa	25920
ttacctccca	accaagcgcc	tagcacctcc	caactgagca	tcatgcacct	cccaacagag	25980
catctagcac	ctcccaactg	atcacctccc	aacctagcac	cgagcacctc	ccaaccaagt	26040
gcagagcacc	tcccaaccaa	gtgccaaagca	cctcccaatc	aaatacctcc	caaccaagca	26100
cctagcacct	ctcaactgga	caccaacaac	tcccaaccaa	gcgccaaagca	cctcctaaca	26160

10

20

30

aagtaccaat	caocttccaa	cogagcacct	agcacctccc	aactgagcat	catgcacctc	26220
ccaacaaatc	acctagcacc	tcccgactga	tcacctccca	acctagcact	gagcacttcc	26280
caaccaacat	agcaaaagcc	ataaagaagt	aaaaagacaa	aaccacgtag	gcatggagac	26340
tggacttctg	gtggcgagga	aagggcattt	ttattataac	gacagctaac	atttgttgaa	26400
ctcacaaaact	gttcttggtg	ttttcctcat	gacatgcagc	atggtcacgc	ctctgtacag	26460
acaaggatac	tgaggcacag	agtggcaccc	tgccaacctt	gtctcatctt	tttatcgaac	26520
ctacatgcag	agtggcagca	aatccagctg	tcttttctct	tcagaacaga	tcccaaatct	26580
cgccactoct	taccccacac	agtgaggtgt	ccccgctgct	gctttctgtc	gccaggatcc	26640
cggtataaac	cgtggagagg	gctcctgccc	ccacgccacc	cacccacacg	ctcactctcg	26700
ctccagccac	caggggatgc	cttccagcac	gagtcagagc	tggcacctcc	tctgctcgag	26760
acctcatgtg	tcctctcctc	acaccttggg	ccctgtttcc	ctacattctg	ctacagcccc	26820
tcaaacaggc	cccgccccaa	accagcccag	ggcctttgca	ctggctgac	cctctgcctg	26880
gaccgcgctg	ccccagaca	gccacacggt	tctcagcctc	atctgcttcc	agtctcgact	26940
caaaagtcac	caagaggcct	tcccagcacc	tgagctccga	cggaagcccc	tgcgccacagc	27000
acccaagcac	tgctttatcc	ccctacgcac	acgtcccttt	caaatactat	tcattttacca	27060
tctcctccca	ctcactgaaa	gggcccagaga	ctgggctata	cccgctcgct	ggggagcagg	27120
accaggcgca	agggctcaca	aatgcagtgg	atgcctggtt	gggaggtgag	ggagctgcag	27180
cgacccacgc	tgggagggaa	cgcaatgaca	ggaggagcgc	aggtcctggc	gacacgatgg	27240
ccatggcagc	cgctggtgag	caaccgcagg	ccggccctgg	gagagggtct	ctagcaagct	27300
gctatcttca	gcctctccga	ctactgcaga	tgcctccctc	tagccagaga	cactgctaca	27360
ccagccgac	cttccaaaaa	gaaggtcagt	aaccccgcga	ctcctggagc	cacagtgcag	27420
ggggagaggg	ctgagagggc	aacagtccac	caagcggaac	agaggctgcc	ccggaggtca	27480
gctggctccc	cggcagctgc	aggggtggct	agccactcgc	gagggcagcg	agggcatacg	27540
aggggctcca	gggatgagtg	gttgcccagc	acagcacccc	tgggagggcg	ggggcacttc	27600
tcaggtagtg	ggggcacgag	gctgctctgg	cctgacctoa	gggactcaaa	atactttggc	27660
gataaattcc	accgtgtccc	accctgctg	gtaccccata	cttacacaca	gactggttca	27720
gatgcagaca	ctctcgcgca	catactcgct	cacacgggca	catacacgtg	cacacacagt	27780
cacatgcgca	cactcataca	cacacaaata	tccactcaca	cgcattgcag	cacacacacg	27840
gacacacaca	ggctcacacg	tatgcacgca	tatgcgtgca	cacgcacaca	cacacacaca	27900
cgtctacatc	ctcccactcc	cacactcagt	tgtctcagaca	cacacacgcc	tggctctcac	27960
acaaacctgt	tgggtctctga	aaggctccag	cccttcccat	gctcgtcaga	agccagtcaa	28020
tggcttctca	agtccaccaca	cagatcaaaag	aggtgaactt	ggccacatgg	cactctgctt	28080
cctgagctcc	caaacaccag	ccttggtgag	gacagacctt	cacccacac	cctcattccc	28140
actaccctgg	gcaggcccag	aggaggggca	tctgcaggat	ctggcaacca	gccccctccg	28200
cccggtccct	gcagccggca	ccatgggagt	cagggggagg	tcactgcaaa	gggcaacagc	28260
aagttggtgg	ccccaggact	agagcccagg	ggtcttcagt	cctactccag	agcttggaca	28320
ctgtccacac	gggcatggcc	aaggggaagg	cttccagagc	cctgacttca	gggaggaggg	28380
caggcgggct	cctgtggcag	gcttgatgc	atggccgccc	actcctggga	ctttctaacc	28440
tagaatatct	aggtcaggct	gggtgcagtg	gctcacgcct	gcaatcccaa	cactttggga	28500
ggccgaggag	ggtggatcac	ttgagggttag	gagtttgaga	ccagcctggc	caacatggcg	28560
aaacctctgt	tctactaaaa	atacaaaacc	tagccagggtg	tggtagtgca	cgcctgtaat	28620
cacagctact	caggaggctg	aggcaggaga	atcacttgaa	ctcgggaggt	ggagggttgc	28680
gtgagctgag	atcgtgccat	tgcgcaaaaga	agatctaggg	cggccctcca	accggtgagg	28740
tccaggctgg	gagtgtctgag	agactgtggt	gacactgaat	gaactaacag	gcaaagggtc	28800
tccaactgag	cctgggggtg	gtgggaaatg	gctcttgtgt	tctagtcaag	acctctgcca	28860
accagttctg	acactgaccc	agcacagaa	ctgacaggtc	agcaagggcc	agggcttagc	28920
acagccaggg	taagggtgtg	tgtacggccc	ccagagtcac	tcccaggctg	caagaaaagg	28980
gacaaaggag	ggacaagggg	tggccaagca	aactgttccc	tctgctcggg	agtctgggg	29040
gacctggcct	agctggccag	tggagctggg	ccacctcccc	ttaaactctc	caccccgga	29100
ttcgactcca	aagctttcct	gccacccacg	ctctccccac	ctgggatcac	ggccaggccc	29160

10

20

30

tgagcettca	agggcccagg	tgaactcagc	cagactagga	gctgaggagg	acacagggca	29220
gcttccagaa	cggacccgag	aaccactccc	agcaggttct	gcttccagac	aaggagctgc	29280
actttttcag	ccaatgcaat	tagaaaagcca	ggagaaagggtg	caaattccac	ctgcctgagc	29340
gtccgcactt	cccaggccgc	ccaccatata	cacagcaaag	atgtgtttaa	ccattcaaac	29400
ccatggccaa	ccacatcggt	tgccctcagac	atgcaagttt	taaaaaggaa	cataactatg	29460
ggccaggcac	ggtggttcac	gtctgtaatc	ccagcacttt	gggaggccga	ggtgggtgga	29520
tcacctgagg	tcaggagttc	gagaccagcc	tagacaccat	ggtgaaaccc	catctgtacc	29580
aaaactacaa	aaattagctg	ggcgtggtgg	tgggcgcctg	taatcccagc	tacttgggaa	29640
gctgaggcag	gagaatcact	tgaacccggg	aggcgaagggt	tgcagtgcgc	cgagattgtg	29700
ccactgcact	ccagcctggg	caacaaggga	gactccatct	caattaaaaa	aaaaaaaaaa	29760
aaaaagggaac	ataactatgg	agtctcaagg	ggaagtaatt	ccttcaacaa	taacaaatct	29820
tgaagagctga	gctctttttt	tttttttgaga	caggatctcc	tcactttgtc	gcccaggctg	29880
gagtgcagtg	gtgggatcac	agctcactgc	agcctcgatc	tcccaggctc	aaatgatcct	29940
cctacctcag	cctcccaaga	agctgggatt	acaggtgcat	accatcacac	ccgattcatt	30000
tttgtatact	ttgaagagat	gggggtctcac	catgttgccc	agtgtggtct	tgaattcctg	30060
gactcagggtg	atctgcccgc	cttggcctcc	cagagtgcctg	ggattacagg	cctgagccaa	30120
cacccccacg	ggttcatttt	cagagtgcga	ccgagtgcctg	gggttacagg	cctgagccaa	30180
ccccccacg	ggttcatttt	aagagtgcga	ccgagtgcctg	gggttacagg	cctgaaccaa	30240
ccccccacg	ggttcatttt	cagagtgcga	ccgagtgcctg	gggttacagg	cctgagccaa	30300
ccccccacg	ggttcatttt	aagagtgcga	ccgagtgcctg	gggttacagg	cctgagccaa	30360
cacccccacg	ggttcatttt	cagagtgcga	ccgagtgcctg	gggttacagg	cctgagccaa	30420
ccccccacg	ggttcatttt	cagagtgcga	ccctttttct	gaaaaacaa	ttgggctcat	30480
gcaaatctga	gagagcatgt	gtgacactcc	ccgccccctg	gacccagggtg	gagtcgcagc	30540
aggggtttacc	cgtgagcggg	gtccaaggcg	atggccctcg	gctgggtcaag	gtcctgccag	30600
aagagcacct	tccgggatgt	gccattgagg	ttggccacct	cgatgcgggt	ggtctctgag	30660
tccgtccagt	acagcttctt	gcccacccag	tcgcaggcga	ggcgcctggg	agagaccagg	30720
ccggagatga	ccacgttctg	caaggcggcc	ccgctctggt	tcaggtaggt	ctgcttgatg	30780
gctcctctgc	tcacgtctgt	ccagtacacg	gctcccttgg	aaaactggaa	gtccactgcg	30840
gcccgcctct	ccaggccgct	gaccacgatg	gtggactcca	gcttgactcc	gccggcgctc	30900
accagccgta	cgtcccgggc	gttggcaaat	agcaggagcg	gcgaggctgt	ggggcagaag	30960
caaaccgtga	gggccaactg	ctaagccagc	aagatacaca	gccctgggat	ggagcactat	31020
gcccagagca	ctcctggtac	tgccctgccc	atgcccaaga	cctccagttc	cttccctcca	31080
ccctaaggc	gttgtcagga	agttgcctgg	gcagccccgg	cccgcctcat	tcagaggctc	31140
ctgcagcgca	gcaaacagcc	ttcttcccac	attcgggtgac	agcacctggt	tggttaccaa	31200
ctgttacgtc	tggtccccca	gatattgggtg	acccttctctg	ccatgcccaa	aacctcccac	31260
atcgtcctcc	agaggctaca	ggggccctgt	cctgttctgc	agagaagcca	catccctttt	31320
gttggcctga	cacaggggat	ggggacatgc	aggcacagca	ctggccatgc	tgcctgcctac	31380
agaccagcc	acagggccac	attttttgag	gggttcagag	cccaggccag	acagagcctc	31440
aagattccct	tacaagtctt	tgaccactgt	ccaagctcag	gcccgtttcc	ttggccgtgg	31500
catcagcttc	ccatccaccc	ctgtattcca	tggttctccc	accctgcttc	tggacattcc	31560
tacatttaaa	gggtcactct	ggaatgccac	cccttggtctc	agacaccttc	cacagctccc	31620
tgtgccagtg	ccatgcagaa	caaggtcaga	ccccctagcc	tgccctccaa	ggccttgccc	31680
tctggcctca	cctacacttc	tctccaccac	cccaccccaa	gcattcctga	tctgcctggg	31740
gccaggctgg	ctccctcacc	tccctgtgca	cgcagccctc	cagcccttcc	tgccctgtga	31800
agaagcctca	tctcacagac	aacggtctca	ttcccacaac	gggtctcaatg	agaaatcagg	31860
agaggccttc	agaccatcac	cccaccagac	acctcagacg	tcggaccagg	aggggtccagc	31920
aacccccaac	acagactcag	agggactaag	aagccacatg	aggagtgaac	acaagatgtg	31980
gacaggagga	ggttaagggc	ctccaggggag	ctccatcagt	ccgtgttctg	ctgtcagcag	32040
ggttaggctg	ggctggccac	aaacaccccc	aaaaaacatc	tgaagccttg	gcttgaacaa	32100
gctgacattc	ctcatgaaaa	ctgcagaccc	ctgggtcctc	ctgcgcagat	gggggagccc	32160

10

20

30

```

agccaacccc acactcccac cttcaccacg aaagagaaaag ccaaaacaaa ctcaactcag 32220
ccaatgacaa tcacagaact gaatcctgta gttagtctcag ttgggttcat ttcagcaggg 32280
gaaagatttg cagcctctat gagggtagct ggggaacacaa agggccagag catggcccag 32340
gagaccccag cgcagtgggg tagatgggtc cgagcacagg cctccctgcc aagacaagca 32400
ctgggtcaaa tccctggccc tccattccc agggagacatg ctccacagga tgggaggaca 32460
cacagaggac ctgaggccag gaaaatgaca gcggcgccctc cgccgcccc aacgtgctgt 32520
catcatctta ggtctacagt tctttgtggc aacgagggac actgtgaaag tcaaacaca 32580
ggaaggcata ggccacaaat aaagacaaac gggacttcat gggaagctaa agatcttctg 32640
catcaaaaga cactatcgag agagtaaaaa ggcaaccac agaagtagag aaaatatttc 32700
caaatcatag atctactaag agattaatat ccatgaaata cagagaactc ctaaaactca 32760
acaatgagaa aacaactaag ccaactcaaa aatgggcaaa caacttgaac agacatttct 32820
ccaaagatga catataaatg gccataaaac acatcaaaac aggtttaata tatccctaata 32880
catcagggaa atgcaaatca aaactacaat aagataccat cttgcacca ttaggacggc 32940
tactatcaaa aaaacaaaat agcaagtgtt ggtgaggatc tggagcaact ggaacccttg 33000
tgcaccactg gcaaaaatgt gaaatgggtg agctactatg gaaaacagca tggcagttcc 33060
ccaaaaactt aaacacagaa ttaccatatg accagcaat ttogctttgg gttatatacc 33120
caaaagaact gaaaacaggg acacaatcag atatgcatac accttggtac acagcagcat 33180
ccttcccaac agctaaaaca tggaggcagc caggcatggt ggctcacgac tgtaatecca 33240
gcactttggg aggtctgaggc ggggtggatc cctgagggtc ggagttcgag accagcctgg 33300
ccaacatggt gaaacccctt ctctactaaa atacaaaaat tagctgggag tagtgacggg 33360
cacctgtaat cccagctact cacaagtctg aggcaggaga atcacttgaa ccctggaagt 33420
ggacgttgca gtgagccaag attgcgccac tgcattccag cctgggtgac acagcgagac 33480
tctgtctcaa aaaacagcaa aacaaaaaca aaaaaacaaa caaacatgga agcaaccaca 33540
gcgtccctct actgagggat gaatagcggg gcaaaatctg ctccatccac acaatggagt 33600
actattcagt ctcaaaaagg aaaaagattc tggtcaggca cgggtggctc tgctgtaat 33660
cccagcactt ggggaggctg agggcgggtg atcacctgaa gtcagggaat caaggcccg 33720
ctggccaaga ctggcaccna gctacacana aagtatangg ccccgga 33769

```

<210> 9

<211> 72049

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> (8356), (8385), (38585)

<223> Identity of nucleotide sequences at the above locations are unknown.

<400> 9

```

tataccttgc ggggaccttc ggctcctgtg gtgaagacaa tatgaagaaa atagaaatta 60
cccataatth tgccacacag acttagttgt gtccatgtat cttgtgcacc ttttttctgt 120
ttacggatca aaatcgactt ttagggtcag gcgcggtggc tcacacctgt aatcccaaca 180
ctttgggagg ctggagttag ggttgggggg tggatcactg aagatcagga gtttgagacc 240
agcctggcca acatggcgaa actccatctc tactaaaaat aaaagattag ccaggcgtgg 300
tgggtgggtg ctctaattcc agctactccg gaggtctgag caggagaatc gcttgaacct 360
aggagacaga ggttgcagtg agccaggatc acgccactgc actccagcct ggcaacagag 420
cgagactctg tctcaaaaaa aaaaataaaa ataaaataaa taaatacata aattgacttt 480
taggagattg gttcaaacaa tgtgtgtaat gttgtgtctg agtggttttc atttatcggt 540
catgcaaat cggacatcat tcaactcttc ccagagtgtg ctgttttctt gcctgtgtca 600
tcaccogtca ccttgaatgc cctcgtttag gtaaaataag tacattttat tcaaaaatat 660

```

10

20

30

ttgaggacat	ttgggttgtc	tccaggttct	tggtcttgag	ttttgctgtt	cttgtggagc	720
catggtgggtg	tctgggttgca	ggaacctcca	tgcgttccag	ctgctgcttc	tgctgtgtt	780
cttagagagg	aaatgctggg	gtccgcggtt	cccggtctgc	tgaccaggaa	gcctgcgggtg	840
ctttacggcc	cttccagaag	cgggagatgc	ccccacttaa	gtgtcagaca	ggcctttcca	900
cctcactggc	agctcctgagc	ggctcccttc	tatttgacga	tgactgagaa	gttaccaatt	960
tccacgttta	ctgactgctg	tttctcctgt	taatttgat	ttatagtctt	cgctaattta	1020
ttgctagggt	tttgggtgtt	tccctattga	cttgtatgcc	ttttaatttt	ttaaacaaca	1080
ttaataatact	tcattttttt	agagcagttt	taagtttaca	ggaaaattaa	gggacaagta	1140
cagagagtto	cttccacctg	ctgtctctct	ctcctctctc	ccaccttccc	tccttccctt	1200
attgtaactt	tctttctgat	attataaaag	tcactcatgg	ctgggctgtg	tggtcacgc	1260
ctgtaatccc	agcacgttgg	gaggcagagg	caggcagatc	acctgaggtc	aggagttcca	1320
gaccagcctg	gccaacatgg	tgaaaccccg	tctctactaa	aaacacaaaa	agttagccag	1380
gcgtgggtgg	gggcacctgt	aatcccagct	actcaggagg	ctgaggcagg	agaatggcgt	1440
gaacctggga	ggcagagggt	acagtgcgtc	gagatcgctc	cactgcactc	cagcctgggc	1500
aataagagtg	aagcttcgtc	tcaaaaacaa	agtcacacac	gcttcttgta	cgagggtcat	1560
ttggccgagg	ggccagatgg	ctcaccatct	agttgggaca	ggccatgagc	tgggaatgct	1620
ttttacatat	ttacatgggt	gagaagaaaa	tcaggagaat	aatgttttgg	gacatgggaa	1680
aatgacatgg	aatttgcatt	ttagtgtcca	taaatgaagt	tttgtttgct	cccagctgtg	1740
ttgactgagg	caggctggct	tctacagct	gcggcagagc	tgaggaggcg	ggaaggagac	1800
cgtgcaggcc	gcagcaccga	aaatatattg	tctctggccc	ttcccagagt	gcttgccgac	1860
ctctgtccga	cagctagaag	gaaggatagg	acccgtccga	cgataaccac	tgttgacatt	1920
tgagcgcggt	tccttcccg	cttttgtgtg	agagtggcag	tctgtttgct	tttgtggtcg	1980
ggatctgctg	cacgcacggc	gggtgttttg	catgaggctt	cctggaggat	agggctgggc	2040
tccgagctgc	acgcagtggt	gcgtgtcctg	catgcagtgg	ggcctcagaa	gagagctgtg	2100
gtggcgggg	cagtgcacac	gctgtgtgg	gccaggcctc	cacgctcaga	tcagcccg	2160
cgacagggtt	gggccacctt	ctctctggcc	tctgtgcagt	ggcccaggcc	gtctgtctctg	2220
cctggcacac	ttgcctctgt	ccttccactg	aagcgctcct	cttaccctct	gctcccggt	2280
gggtacgttg	aattgtgtcc	ctcaaggaga	tatgctaaag	gtctaaccct	aggaacctgt	2340
gtatgtgatc	taatttggaa	acagggctct	ggctgatgta	atcaagcgag	gatgaggtca	2400
ccctagagta	gggggacctat	atccacgggt	ctgggtgtct	catgagagca	ggtgagcaga	2460
cactgacact	caggggtgaa	ggctgcagtg	agtcagaaca	gggcttagtg	cgatggcggc	2520
cacaagccaa	ggaactccaa	gtatttctct	caacaccaga	agctggaaga	tgccaggaa	2580
gatcctgccc	tggagccttc	ggaggagatc	tgtccctgca	gacgtcttga	cttttgattg	2640
cagggatgca	tgtcttaggg	tgtgtggggg	gggtgcatttc	tgatgttaga	agccacctgg	2700
ttgggtggcga	tgtgtcacgg	gagccctctg	caggttctgc	gtgtccatgt	ggtcggggac	2760
agaggtgggc	agggacggac	ggtgtcgagc	tggacatgtc	catgacgtcg	gccatccctt	2820
gggatggctt	ttttgttttg	aggataaggc	tgcctgcag	gaagctgtgc	cctgcctggc	2880
ccttgcccca	agccctggc	ctgtgcttgg	cctcgcgga	gggatgtgc	ccttctctct	2940
tgcatgctg	cagggaggaa	ggggagaggt	cagcagcccg	cctggaggag	gctcgggcga	3000
ggggaagggt	tcactttcag	gcaatgttgt	ggggctgttt	aaacaacccc	aaagaaaacc	3060
atttggccaa	actgttagtt	tccaaacatt	ttacttctct	ggtgttttaa	ttaaattccta	3120
ccaagactct	gtagctgggt	ccagggaagg	agttggcctc	tcttctttat	agcccgccac	3180
agtcagtccc	ctgcacctgc	ccttcccaac	cccaggcctg	cttccccgtg	gccatggctg	3240
ctgcccggag	ctctctacac	acagaacctc	ctggaggcca	gctgtgggca	ccagccttgg	3300
cagggctgtg	gcggagccca	ggctgtgtgt	actctctctg	cagctgtctc	ctgctggcct	3360
ggctggacag	cgtccccacc	accactgggg	tcacctctgt	gctgggtcaca	gctcactcag	3420
accttcaggc	aaatgggttg	gacctgcct	ctctcccagg	tgtctcagtc	tctgcaaac	3480
tcaaaaacct	cagaggcctt	gcagcctgag	gggtgtcaga	gacacctcct	tgaatcagt	3540
aaacacctac	agattcacc	cagcagtgaa	aggactgctt	cgccacagag	gtttgattta	3600
ctcctaagta	attggaagg	atgcccagaa	taggttctct	atgggtgggac	tagaggccct	3660

10

20

30

ctgctgacct	agttaacaga	gggctagggc	tgggtgtgct	cagccctga	aggttctagg	3720
cccatttggg	accccccgcc	agaacctgcc	acaacctgcc	atgtgggtgac	agctacctaa	3780
atcccagagg	ctcttgagct	ggagagcaga	cctctcaatc	tcagcaggcc	ccccacacag	3840
accccataac	cctagtctgc	cttcacagta	cagttcgtgg	ctatgtgttc	acggatgggtg	3900
ttgttccct	aaggtctctg	cctgtgacc	ccaagggcgt	cctgagggca	gattccaagt	3960
ctgtttcgtc	cacccctcct	tccctagcag	cgggtccagg	gcctggcctg	aactagcttc	4020
ccacagagat	actgggtgga	tgatgaaggc	agccaggcgg	caagtgaaaa	acgcacttcc	4080
tgcatgtgct	ggctcctggg	attgaagtgt	ttgaggaagc	aaagtgaagt	gagctttcct	4140
cttgccgctg	tgtgtccttg	ggccgggagc	ctacctctc	tgagcgttgg	ggtccttgtc	4200
agtagaatgg	ggcatcctca	tagctcaagg	ggtgggtgtg	gaaaattgtg	ctattgtgtt	4260
actttaatga	tttttttttt	ttcgagacaa	agtctcacc	caacgcgcag	gctggagtgc	4320
agtggcgcga	tctcagctca	ttgcaacctc	tgcctcctgg	gttcaagtga	ttctcctgcc	4380
tcagcctccc	aagttagctg	aattacagga	gtgcgccacc	aggcccgcca	tatttttcta	4440
tttttagtag	agagggggtt	ttacctgtt	ggctaggctg	gtcttgaact	cctgacctca	4500
ggtgatccac	ctgcctcgcc	ctcccaaagt	gctgggatta	caagcatgag	ccaccgcgcc	4560
cgccctactt	tagtgatttc	ttaggaggac	agagggaaag	ggctggcaag	acaggcttgg	4620
aatgtgtttt	gggatcaagt	gcgggtttct	gtctggcaact	ggcgttctct	gtggggccat	4680
gatggacaca	ctgctgaggt	caagcgtgat	tgcctctggc	ctgtgcctgg	cagtctcatt	4740
ggaaagtctt	gtagacatcg	tgtggatggg	gctcttcccg	gccaaagcct	tggggacctt	4800
ccaggactgt	gatctcccca	cagtggctgt	taagcaggga	cctttcgtga	agtggagtct	4860
ctggctccct	ccaagtcata	gctagacagg	gactcgggca	tcgccaagcc	tggctgatta	4920
ttcactggat	gaggagacag	gcccagagag	gggcaggaac	ctgcccagg	tcaccagca	4980
ggccccagag	gtttcggctc	cggattctcc	ctgctcatcc	ctggatgtag	tgctgctgtg	5040
gatgtggttc	tgtgctgggg	gctgtggaga	gcagggggct	tgtgccagga	ccccagttag	5100
ggtggcgccc	tcgccatgag	gcgcactgtt	ggtatggggc	ggccatccac	tggggtgtgg	5160
ggaggaaacg	ctttcctgag	gaggaggtgg	cgggaggaac	agcttccctg	aggaggaggt	5220
ggcgggtgct	tgtgacctgg	gccttgaagg	acagggtccat	tgtcaacaga	acattttggg	5280
agtggagcct	agagggagaa	aatttgttga	aattcagatt	ccctccccc	taccaatata	5340
caaccaatca	gatgcccctg	accagatcta	aatttggctc	tcagagattt	ccattgttagc	5400
tggggacttg	gggaaccttc	taagtgcctg	ctctgcctct	ccccagcctg	cctgcctcag	5460
tttcccagc	cctgggcccc	tgtcgctgtt	gcatcacgt	gggcgcctc	tagtggagga	5520
atcagattat	gcactccggg	gcttggagca	ggagttagga	ggggctcctg	tctttccttg	5580
aaacgttgga	tgccgggata	ctggaacagt	ctctgcattc	ctcctggcga	gaaccagagc	5640
ctgggacacg	gggacctatc	gttggttgaa	ggctgcagcc	tggcagggca	ctcaggagat	5700
ctggcagttg	gctgcagggc	caggctctagg	ggccagggca	tcagggaggc	tctgggctgg	5760
ttcagccccc	ggcccttttg	cagatttgtga	cctgggcccc	tgtgcagggg	catggccaca	5820
ggatgctggg	aggggtctct	gacctgacc	ttcttggctc	tgtgcacct	tgagaccaga	5880
aaggtctgga	acaaatgagt	agacgatgcc	ctaacctggg	gaggagagca	catcctgac	5940
ccagcaacct	cgggaaggat	ctgtcaggat	tatggggcac	cctgggggccc	ccaagtctgc	6000
atgggtctcc	acttgcaatt	tctgtaggaa	gctctgataa	atccaaactg	ggggctctag	6060
gacacagtca	gaaatgctga	taccgttgtg	tgtggagcct	cgggcccctgg	gggtcaggag	6120
catgtggagg	gtgggcccacg	gggggttcaga	agagaatcct	gtaaccccc	accccccaaa	6180
ctgaagccca	cttgaggggc	atggctgaaa	ggttgggggg	tctccgtgcg	tctgtggag	6240
tgggtggtga	ggagtccctg	ggtttgcacg	cctctgggoc	tgagcggcgg	gaccccgctc	6300
acagcggatc	cttgggcccc	gttgcctcaga	tgtctcaga	gtgttgctgt	ggccacggag	6360
ggagcctgag	ttaagcttct	cttgtgcggg	ttgtacgctg	tcagggtcaca	ctggtgagtt	6420
aggcagggca	cagatgccca	gagcagaggg	aactttcctt	ggggattcaa	cactgcaag	6480
tcttaggggc	tggcaaatcc	tgccctcagc	tagagagggg	gcttttattt	gagaccagaa	6540
tcacctgagc	atcctcctgt	ccccagctgt	gtccagcctg	tctgcaggga	catcctgaga	6600
ggaccaggct	ctccctcat	ccacctgcct	aagtgcact	ctgaacctg	tcacctgtg	6660

10

20

30

ccgtggaggg	gcgtgacctc	aagctgetca	gccagcagca	ggcttggccc	tggggggcag	6720
cagagaccca	gggtggtgtg	gggtgggtgc	ttcgtggcgt	ggttctgaaa	cttcgttgga	6780
agtgtgtgga	cagtgccttg	cctgtttctc	gtgggacctc	atttagaaac	gaggtctgag	6840
ttactggggg	tcactactgt	gttctgatgg	cccagctgtg	tggaggccgc	ggtgcagccc	6900
catccaagga	gccaggggcc	tgggtctagc	cgtgaccaga	atgcatgccc	cggagtgtgt	6960
tctcatctcg	cacctgtgtt	gcctgggtgt	tcaagtggtc	gtgaaactct	gtgttagctc	7020
ttggtgttcc	tgaagtgtcc	cccgggtctc	aggcctcaga	accagggttt	cccttcatct	7080
cggtagcctg	ggagcatctg	ggcagttgag	caaagagggc	gattcacttg	aaggatgtgt	7140
ctggccctgc	ctaggagccc	cccggcacgg	tgctggggcc	tgaagctgcc	ctcgggtggt	7200
ggagaggagg	gagcgatgaa	gtggcgtcga	gctgggcagg	aagggtgagc	ccctgcaagg	7260
tgggcatgct	ggggacgctg	agcagcatgg	ccagcagctg	ggtctgcagc	ctggtagccg	7320
gcgggacttg	tgggtggggc	tggttgtgtg	ccaggagagg	ggctggcagg	agacaagggg	7380
gactgtgagg	cagctcccac	ccagcagctg	aagcccaatg	gcctggctgt	gtggctctca	7440
gctgcgtgca	taacctctca	gtgcttcagt	tctctcattt	gtaaaatgag	gaaacaaaca	7500
gtgccagcct	cccagaggtg	tcatgaggat	gaacgagtga	ccatgtagca	tgggctgggt	7560
gcgtgtcacc	taacatcacc	agcctttgca	aggagagccc	tgggggcctg	gctgagtatt	7620
tcccttgccc	ggcccacccc	aggcctagac	ttgtgcctgc	tgcaggccct	tgacccctga	7680
ccccattgca	cctgtctcca	caggagccga	ggaggtgctg	ctgctggccc	ggcggacgga	7740
cctacggagg	atctcgtctg	acacgccgga	cttcaccgac	atcgtgctgc	aggtggacga	7800
catccggcac	gccattgcca	tgcactacga	cccgtctagag	ggctatgtct	actggacaga	7860
tgacgagggt	cggggccatcc	gcagggcgta	cctggacggg	tctggggcgc	agacgctggt	7920
caacaccgag	atcaacgacc	ccgatggcat	cgcggtcgac	tgggtggccc	gaaacctcta	7980
ctggaccgac	acgggcacgg	acgcctcga	ggtgacgcgc	ctcaacggca	cctcccgcaa	8040
gatcctgggt	tccgaggacc	tggacgagcc	ccgagccatc	gcactgcacc	ccgtgatggg	8100
gtaagacggg	cgggggctgg	ggcctggagc	cagggccagg	ccaagcacag	gcgagagggg	8160
gattgacctg	gacctgtcat	tctgggacac	tgtcttgcac	cagaacccgg	aggagggctt	8220
gttaaaacac	cggcagctgg	gccccacccc	cagagcggtg	attcaggagc	tccagggcgg	8280
ggctgaagac	ttgggtttct	aacaagcacc	ccagtgttcc	ggtgctgctg	ctgggtccat	8340
gcgtagaaag	ccctgnaaac	tggaggggagc	cctttgtccc	cctgncttca	gtttcctcat	8400
ctgtagaatg	gaacgggtcca	tctgggtgat	ttccaggatg	acagtatgta	cagtaagggc	8460
agcctctgtg	acactgacca	cagtacaggc	caggcctctt	tttttctttt	tttttttgag	8520
atggagtctc	actctgtcgc	ccaggctgga	gtgcagtggg	gtgatctcag	ctcactacaa	8580
cctctgcctc	ctgggctcaa	gtgattctcc	tgcctcagcc	tcctgagtag	ctgggattac	8640
aggtgcctgc	cactgtgtct	ggctaattgt	tgtatttttg	gtagagatgg	ggtttcaccg	8700
tcttggccag	gotgggtcga	aactcctgac	ctcagggtgat	ccacctgcct	cagcctccca	8760
aagtgtctgg	attacaggca	tgagccacca	cgcccgggtc	ggccaggcct	cttttgaaaca	8820
ctttgcacac	catgggtctt	ttcatccagg	ggggtaggta	cagttgtaca	gttgaggaca	8880
ctgaagccca	gagaggctca	gggacttgcc	cagggtcaca	cagcaggatg	tggcaggtgt	8940
ggggctgggc	ctggcagcgt	ggctccagct	ttccagcata	gaaatctgtg	aaagcagata	9000
gtttgtcggg	cggtagggga	gactttctga	gacccgcccc	agcggctcag	agggtagtag	9060
ccaggggcct	tcctgggggc	tcataaacca	gaacactgaa	tgggaaaacc	ctgatggagg	9120
aggcgagtg	gagctgtggg	tgccgatggg	aagtcccaga	ggagctggga	ggtcagtagc	9180
ggtgctgccc	tctgtggagc	acttagtggg	caccaggtgt	gtttccagg	tcattggcct	9240
gggaacctgaa	gctcagaagg	tgaagtaact	tgccagggc	acccgtcggg	cagcggcggg	9300
cagaggattt	gtgggctgtg	gagcctgtgc	tctgggcccc	gccctggggg	ttgtgagtgt	9360
gctggccggg	gagcttttcc	tgcaagtggg	ctggtgtcta	ggagccagca	tgtcaggcag	9420
caggcagcgg	gagtgacgca	ggcagcggga	gcacagcagg	cagagggcgg	ggctcagagc	9480
gccatccgtg	gacctggggg	cacggaggca	tgtgggagag	ggctgtctca	tggcagtggc	9540
tgaagggtctg	ggttgtgtccc	cagggagggt	ggatgagggt	aagaagtggg	gtccccaggg	9600
gctttagcaa	gaggaggccc	aggaaactggt	tgccagctac	agtgaaggga	acacggccct	9660

10

20

30

gaggtcagga	gcttgggtcaa	gtcactgtct	acatgggcct	cggtgtcctc	atctgtgaaa	9720
aaggaaagga	tggggaagct	gactccaagg	ccctccctag	ccctgggttc	atgagttctga	9780
ggatcccagg	gacatgggct	tggcagttctg	acctgtgagg	tcgtgggggc	cagggagggg	9840
caccgagctg	gaagcgggag	gcagaggggc	tggccggctg	ggtcagacac	agctgaagca	9900
gaggtgtgta	cttggggcct	cagaaccttc	acccctgagc	tgccacccca	ggatctgggt	9960
tcctcccttg	gggggcccc	gggaacaagt	cacctgtcct	ttgcataggg	gagcccttca	10020
gctatgtgca	gaaggtttctg	ctctgcccct	tcctccctct	aggtgtctcag	ctcctccagc	10080
ccactagtc	gatgtgaggc	tgccccagac	cctgggcagg	gtcattttctg	tcactgacc	10140
tttgggatgg	gagatgagct	cttggcccc	gagagtccaa	gggtgtgtgt	ggtgaaacct	10200
gcacaggggtg	gaagtgggca	tcctgttccc	aggggagccc	ccagggactc	tgggtactgg	10260
gcttgccgct	ggcatgtctca	gtcctccagc	acttactgac	accagcatct	actgacacca	10320
acatttacaa	acaccgacat	tgaccgacac	cgacatttac	cgacactgac	atttaccaac	10380
actgtttacc	aacactgaca	tctactgaca	ctggcatcta	ccaacactga	catttaccca	10440
cactgacatt	taccaaacct	atttaccaac	actgacatct	actgacattg	gcactctacca	10500
acaccaacat	ttaccgacac	caacatttac	caacactgaa	atttacggac	accgacattt	10560
accgacacog	tttaccacaa	ccgacgttta	ccgacaccca	catttaccca	cactgatatt	10620
taccaaacct	gacatctact	gacgtctggca	tctactgaca	ccgatgccag	catctaccaa	10680
caccgacatt	taccaaacct	gacatttacc	aacactgaca	tttaccgaca	ttgacattta	10740
ctgacactga	catctactga	cactggcatc	tactgacact	gacgtttacc	gacactagca	10800
tctactgaca	ctgacattta	ccaacaccag	catctaccaa	caccgacatt	taccaaacct	10860
gacattttact	gacactgata	tctactgaca	ctggcatcta	ctgacaccaa	catttaccaa	10920
caccagcatc	taccaaacct	gacatttacc	aacaccgaca	tttaccacaa	ccgatgttta	10980
ccaacgcoga	cgtttaccga	cgccagcatc	taccaaacct	gacatttacc	gacaccgaca	11040
tttaccgaca	ctgacattta	ctgacactga	catctactga	tactggcatc	taccgacact	11100
gatattttacc	aacgccagca	tctactgaca	ctgatgttta	ccaacacoga	catttacgag	11160
caccgacatt	tactgacacc	aattatttact	gacatcaaca	tttagccatg	tgatgggggc	11220
cggcttgggg	gcaggccttg	ctcttggcac	tggggatgct	gcagagacca	gacagactca	11280
tggggtcatg	gacttctgct	tcttctccag	cctcatgtac	tggacagact	ggggagagaa	11340
ccctaaaatc	gagttgtcca	acttggatgg	gcaggagcgg	cgtgtgtctg	tcaatgcctc	11400
cctcgggtgg	cccaacggcc	tggccctgga	cctgcaggag	gggaagctct	actggggaga	11460
gcocaagaca	gacaagatcg	aggtgaggct	cctgtggaca	tgtttgatcc	aggaggccag	11520
gcccagccac	cccctgcagc	cagatgtacg	tattggcgag	gcaccgatgg	gtgcctgtgc	11580
tctgtctatt	ggccacatgg	aatgcttgag	aaaatagtta	caatactttc	tgacaaaaac	11640
gccttgagag	ggtagcgcta	tacaacgtcc	tgtgggttacg	taagatgtta	tcattcggcc	11700
aggtgcctgt	agacacagct	acttgagagc	tgaggtggga	ggatcgctgg	agtccaagag	11760
tttgaggcca	gcccggggca	aggggacaca	ggaatcctct	gcactgcttt	tgccacttac	11820
tgtgagattt	aaattatttc	acaatacaaa	attaagacaa	aaagttaatc	acatatccac	11880
tgccttgctt	aagacagaaa	acatgggtgt	tgttgaaagc	agaggcagct	gctggcctga	11940
gtttgggtgat	tggttcctaa	gcagttgaag	gcagttttgt	ttttccatag	atgtctgttc	12000
tccttttctg	gggtgcagcc	tcgcccctgt	gctgtggtcg	ggtttcagtg	gcctcgtccc	12060
gtggacgcag	cctcgccctg	ccgctgtggt	cgggtttcag	tggcctcgtc	ccgtggacgc	12120
agcctcgccc	tgcgctgtgt	gtcgggtttc	agtggcctcg	tcocgtggac	gcagcctcgc	12180
cctgcgctg	tggtcggggt	tcagtggcct	cgtcccgtgg	acgcagcctc	gccctgcgcg	12240
tgtgttcggg	tttcagtggc	ctcgtcctgt	ggacgcagcc	tcgcccgtgc	gctgtggtcg	12300
ggtttcagtg	gcctcgtccc	atgggcgtgc	tttggcagct	ttttgtcac	ctgtggagcc	12360
tctcttgagc	ttttttgttt	gttgtttgtt	tttgtttgat	tttgtttgat	tgtttgtttt	12420
tgttgtcgtt	gttgttgccc	aggctggagt	gcagtggcgc	gatctcagct	cactgaaacc	12480
tctgcctcct	tgggttcagc	ccattctcct	gcctcagcct	cccacatagc	tgggattaca	12540
agtgcgccgc	accagccctg	gctaaatttt	gtatttttag	tagacagggg	gtttcaccat	12600
gttggtcagg	ctggtctgga	actcctggtc	tcacatgatc	cacctgcctc	ggcctcccaa	12660

10

20

30

agtgttggga	ttacaggcgt	gagccaccgc	gccagccct	ctgttgagca	tattttgagg	12720
ttctcttggg	gccagtata	tgtacatgtg	tcccatcgc	accatcgtca	ccattgagg	12780
tgacattggg	gcctctctc	ggggtggatg	cctccctctg	ttccagcaa	cttctgaagg	12840
attttcttga	gctgcatcag	tcttgttga	cgccaccatc	ggggtcacct	ttgtctctct	12900
cagggctccc	aggggaggcc	cgaatcaggc	agcttgccag	gcagggcagg	atggagaaca	12960
cgagtgtgtg	tctgtgttgc	aggatttcag	accctgcttc	tgagcgggag	gagtttcagc	13020
accctcaggg	tggggaaccc	agggatgggg	gaggctgagt	ggacgccctt	ccacgaaaa	13080
ccctaggagc	tgacagtggt	gccatttcct	gctggagctc	cttgtaaatg	ttttgttttt	13140
ggcaaggccc	atgttttcgg	gccgctgagg	atgatttgcc	ttcacgcac	ccgctacccc	13200
gtgggagcag	gtcagggact	cgcgtgtctg	tggcacacca	ggcctgtgac	aggcgttgtt	13260
ccatgtactg	tctcagcagt	ggttttcttg	agacagggtc	tcgctcgcct	accaggcga	13320
gagtgcagtg	gcgcaatcac	ggctcgtctg	agcctcaatc	tccttgggct	caggtgatcc	13380
tctgcctca	ccctctgagt	agctgggact	acagacacat	accaccacac	ccagctagtt	13440
tttgtgtatt	ttttgtgggg	ggagatgggg	ttcgcgtgtg	gtgcccaagc	tgatctcaaa	13500
ctcctgaggg	acaagcgatc	cacctgcctc	ggcctcccaa	agtgtctgga	tgacaggcat	13560
cagccgtcac	acgcagctca	atgattttat	tgtggtaaaa	taaacatagc	acaaaattga	13620
tgattttaac	catttttaag	tgaacagtcc	aggctgggcg	tggtggctta	tgcttgtaat	13680
ccagtgactt	tgagaggctg	agggtggcag	atcacctgag	gtcaggagtt	tgagaccagc	13740
ctggccaaca	tgatgaatc	cagtctctac	taaaaataca	aaaattagcc	ggcatgggtg	13800
gcaggtgcct	gtaatccag	ctactcggga	ggctgaggca	ggagaatcgc	ttgagcccg	13860
gaggtggagg	ttgcagtgat	ctgagatcat	gccactgcac	tccaatctgt	gtgacagagc	13920
aagactctgt	cttgaaaaat	aaataaataa	aaaaaatttt	aaaaagtga	caattcaggg	13980
caatttagtat	gaggacaatg	tgggtgcagg	atctctgcta	ctatctactt	ctagaacact	14040
ttcttctgcc	ctgaaggaaa	ccccatgccc	accggcactc	acgcccattc	tccctctctt	14100
cccagcctct	gtcaaccact	aatctacttt	ctgtctctgg	gggttcactt	cttctggacg	14160
ttttgtgtga	ctggaatcct	gcaatatgtg	gtccctcgct	gtggcttctt	tccatagcat	14220
tgtgttttcc	agattcacc	acacattgtc	gcacgttatc	agaattctcat	tctgactgg	14280
gtgcagtggg	ttaggcctgt	aatcctaaca	ttctgggagg	ccaaggcggg	acgatcactt	14340
gaggcaggag	tttgagacca	gcctggccag	cctagcaaga	ccccagctac	caaaaaattt	14400
taaaagttaa	ctgaacgtgg	tgggtgggtg	caottgtggt	tcccagctac	ctgggagggt	14460
gaggttgagg	gatcgttaa	gccaggagg	tcaaggctgc	agtgagctat	gatcgacca	14520
ctgcactcca	gcctggacaa	cagagcaaga	ccctgtctga	aaaaaaaaac	aaaaaaaaaa	14580
gttcctttct	ttttgtgggt	ggatgacatc	ccattgtatg	gccacagcac	attttgtttg	14640
tctgtttatc	gggtgtgtgg	cagtgtgttc	caccttttgt	ctcctgtgaa	taatgctgct	14700
gtgaacattt	gaattcaagt	ttttgtttga	acacctgttg	tgaattattt	ggatatatgt	14760
gtaggggtag	gattgctgag	tcctatggta	atgttaggtt	tgacttactg	aggaaccatt	14820
aaactgtttt	caacagtggc	tgcgccttcc	tgcattccca	ccggcagtg	gtgagggttc	14880
tgactttacc	tcctcacaaa	cgttcttttt	ccatttaaaa	aaatattcag	ccagggtgct	14940
tggctcacgc	ctgtaatccc	agcacttttg	gaggccgtgg	cgggcggatc	acctgaggtc	15000
aggagttcga	gacgagcctg	gccaacatgg	tgtaacccca	tctctaccaa	aaatataaaa	15060
attagccggg	tgtggcagcg	ggcgcctgta	atcccagcta	cttgggaggg	tgaggcagga	15120
gaatcacttg	aaccggggag	gcagaggttg	cagtgagcca	agatcgcgcc	actacactcc	15180
agcctgggtg	acaagagtga	aactccatct	aaaataaaac	aaaaataaaa	ataaataaaa	15240
atltattaaa	acattcatca	cagccagcct	agtgggtgtc	ccatgtggct	ttgcctcgca	15300
tttccctgat	aactaggatg	ctgagcgtct	tgccccaggc	ttgccacacc	tcagcacttt	15360
gagatacgtc	gcacagtccc	catttgcgaa	cgagaaatga	ggtttaggga	acagcagctg	15420
tgtcatgtca	cacagcgagc	aggggtgtct	tgagccgtct	gacccacag	ccgaccaagc	15480
tccaatcctt	accgcctcct	agtgttggg	atgtagccca	gggtgctccc	acatttttca	15540
gatgagaaca	ccgaagctca	aaacaggagc	gttttgtcca	cattggatac	acgatgtctg	15600
tggtttggtc	ctgaagtcac	tttatatctc	agtggccag	actggagtag	gacagggggg	15660

10

20

30

tctggggaat	ggggaagggtg	tctcaggtga	aagggaaggaa	ttccagattc	tccatactgt	15720
ccttgggaag	ttagaagact	cagaggggtct	ggcaaagtca	gacaaagcaa	gagaaatgca	15780
gtcaggagga	agcggagctg	tccagggaaca	gggggggtcgc	aggagctcac	ccccaggaac	15840
tacacttgct	ggggccttcg	tgtcacaatg	acgtgagcac	tgcggtgtga	ttaccactt	15900
tttttttttt	tttgaggtgg	agtctcgctc	tcttgcccag	tctggagtgc	agtggcacga	15960
tctcggtcca	ctgcaagctc	tgccctcccg	gttcacgtcca	ttctcctgcc	tcagcctccc	16020
gcgtagctgg	gactacaggc	gcctgccacc	gcgcccggct	aatttttcta	tttttagtag	16080
agatgggatt	tcactacatt	agccaggatg	gtctcgatct	cctgacctca	tgatccgccc	16140
gtctcggcct	cccaaagtgc	tggtgattaca	ggcgtagacc	accgcccgcg	gccccatttc	16200
ccactttaag	aatctgtctg	tacatccctca	aagccctata	cacagtgtctg	ggttgctata	16260
gggaatatga	ggccttacagg	ccatgggtgct	ggacacacag	aagggaaggga	ggtcaggagg	16320
tagaaggggc	gagagaggga	acaggcggag	gtcacatcct	tggtcttcaa	aatgggcccag	16380
ggagagacac	cctctgagca	tggtaggaca	ggaaagcaag	attggaacac	attgagagca	16440
accgaggtgg	ctgggcgtgg	tggtcttacgc	ctgtaatccc	aacactttgg	aaagctgagg	16500
tggtgggatt	gcttgaggcc	aggagttaa	gaccagcctg	gccaacatgg	tgagaccccc	16560
tctctactaa	atatacaaaa	attagccagg	cgtgatgggtg	catacctgta	atcccagctg	16620
ccttgggaggc	tgaggcaggga	gaattgctta	aacctgggag	gcggagggttg	cagtgcagccg	16680
agatcccgcc	actgcactcc	agcctggggc	acagagtggag	actccatctc	aaaaaaaaaa	16740
aaaaaaaaaga	taaaaagacc	aaccgaggaa	ttgaagtggg	ggggcgtcac	agtagcagaa	16800
gggggatcgt	ggagcaggcc	acccgtgggt	catgcactgg	aagctcatta	cctgacgatt	16860
tgtagctcat	cactgggggc	ctaaggagaa	tagatactga	aggatgagga	gtgatggcgc	16920
ggggcacggg	tgtctttggg	ggccagaact	tggggactgc	tggggtgcct	cactgcaggc	16980
cctctcagcg	ccctttatat	gcttacacag	gctgtttcta	agagggggat	acattgcata	17040
agcgttttca	gactacctca	tcatgggtcc	ctttctttac	cctctgtggc	cctggtggcg	17100
cactctctgg	gaagtgagag	gtggatgccc	agaccggccc	tgccatccac	ctgcacgtcc	17160
agagctgact	tagcctcgag	attgctgctg	gcacctcctg	ccccgggaca	cctcggatgt	17220
gcccgtggag	atgctggctc	tgtgttttct	gctggagtgt	ggtgcgtctt	ttcctcctgc	17280
aagtggccac	cgctcttggt	tatgtcctca	ggcctctgcg	agtcattggct	gcttctcagg	17340
tccttgccca	ggccaggagg	caaaccctcc	tggtcactttg	ttcagggggtg	gatgcgccag	17400
tgttctctgt	gtggacggcc	atctcacatg	aggggtcttg	gcctgcaggc	tcgttcaggga	17460
aacaccgcct	gagtatgcag	tgtgtgccag	ctgtgtccca	ggcaatggcg	gggacagtgg	17520
ctgctgctgg	gggtgtgggtg	gcttctgggg	actctgggga	cagctgaggt	gcaaggagcc	17580
acggctcctt	gaggatgcag	ttggactcca	ggtggaagg	atgggtgggg	gaggtataaa	17640
tggtgtcagg	gaggagacac	atgttgaaca	atgggaacat	ttttaagatg	ctatgtcggg	17700
aggcaacaag	gtggccaacc	cagggtgctga	ggagcccaca	ccagccctgg	acgtgttttg	17760
ccgctcacct	ttgctgggga	gtggtgggag	agaggattcc	gttccacgtg	gtggtgtgcg	17820
cagctgggct	gtgtggagct	gggcgttagg	aggaagggtg	tttctgcggg	gctagccggg	17880
ctctgccttt	gaacacaatc	aggctccagg	ttttcagcat	ccagtgcattg	agaggacttc	17940
acgggcagct	gtggtctgatc	ccttgatgaa	ttgggagaag	aacaaaggtc	tatgaaatga	18000
gggttctatgt	agatggcatt	agagacgccc	acaacagatt	tacagagtgg	agcggagacg	18060
gcggatgggt	ctgggaggcc	cctcctgctg	gccttgactg	tgacagctgt	cctgggaatc	18120
agcttccagg	ccgccccagc	agcctgactg	acacacacag	gggttttagc	cccatcctgc	18180
gaccagctgt	tgccatcctc	agtgcagctc	gggagtggcg	gtgggtccag	ccctgggccc	18240
cctccccacc	tgctggggcc	caccagggcc	agtcctgaca	cctacagggt	gcttggagcc	18300
gcatccgagt	cctgccccac	cacgtgtgaa	gcccagtggt	tcgtgggctg	aggctcccctg	18360
attgcacccc	cacttccctt	ctgcttcaca	tagctgcctc	ttctcaccgt	ttttccagcc	18420
tcctgggcta	ggaattccag	tgtgtgtgctg	gctttgcccc	aggacacctc	cttagccctc	18480
ttcctgagtc	tagagccccg	gggttgggaa	gtcctggccc	ctgggacacc	tgccagccaca	18540
ctcagcttct	cctgtgagcc	tcagacatgt	cccctcagga	ccaagccctc	acgttcttgc	18600
ctccccgccc	acctgggctc	agccaggggga	aggcctggct	gggagcgtct	cccctctgcc	18660

10

20

30

ctgcccctct	ccctccctac	cctgcccctc	tctcctctgc	cccgccatgg	cttttatatc	18720
ctgtgccaca	agacatggct	gtgtgtgaaa	gtggcagggg	ctggcatctc	tgtgggtctc	18780
tgaggcccac	gctccagtgc	cactcttccc	accgctgggc	cgtgccctca	tgctggaggg	18840
acagcccagc	cctctcccga	accccagccc	catgtgccca	gctgcccccg	gcccctctcc	18900
ctggaagccg	gggtcactcc	agccgtatgc	catgggtggg	acatcctgct	tccttggcct	18960
tccagggaag	gtcctctttc	caaatggcga	cacctgggtc	ctgcccggag	gctggaagct	19020
gtggcccctg	tatgcccctc	cagggtctgt	gcgctcggtt	ggcccaggtt	cccatcaccg	19080
tcacatcac	catcatcatt	gtcatttctc	ttgtctgtga	gccggcctgg	tctcccagag	19140
cagagaccct	ctgaggtcca	gcctgagttg	gggtctccgt	gctgacccct	gacggggact	19200
caggacgtac	caggtctggg	tcaggagtga	cccccaaac	tcgtgccctt	tgacaggcac	19260
ccctgacttt	tgctaagtgg	gtggaggtga	catcacttac	agcgggagtg	atgggacagg	19320
gtctgttggc	tgcactgtgc	tcccagggat	ctggggagag	gctatatccc	tgggctttgg	19380
cactgcagag	ctgtgtgtgt	ttgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	19440
gtgtgtgttt	gcgtgcgcgc	acatgtgtat	aagatctttt	tttattacat	gaagcaagat	19500
aactgttgct	gtttcccttt	gggttttgtg	ttcaacagag	tggggtactt	cttccctcag	19560
acaacagaac	tctccccttt	aaacaogtgc	tgtcagaggg	tgggtcttgg	gctcatgtct	19620
gtttgcacag	ccgagtcaga	ggaaaacacag	ggttcttcat	aaaaaactcg	cacagcaggc	19680
gactgtccag	agtcagcctg	caggacggca	gcagccctgc	ccctcagagc	acagctaggg	19740
tgggctgctt	tgggatctcc	cgtcattccc	tcccagctgg	cagccggcgg	ccggccattt	19800
ccctgggtgtg	ctggctcagg	gggcgtgcgc	ctgctctgct	cacctgggga	atgggacaga	19860
agctggcagc	tcggagagga	cagggtctgga	cccttgggtg	gcctctgggt	ggaccatctc	19920
attgtcctca	gacacagcct	ctcgggtcta	gtttcatttc	ctgaaaaaca	agtgcacaga	19980
actagagcag	gagtcgagag	ctacggcccc	cgggccagat	ccagccctgc	cacctgtttt	20040
cacaccatgc	tcaagctgag	tgggtttttac	attttttaat	tacttgaaaa	aaaaaaagcc	20100
aaaggagggtt	tcatgaccca	tgaaaatttat	atggaattca	aaaaaaaaaa	attatatgga	20160
attcaaattt	cagtgtccat	aaataatttc	ttgagacagg	gtctcgctct	gtcaccagg	20220
ctggagtga	gtgctatggc	atggctcgct	gtacccttga	cctcccaggc	tcaagcgatc	20280
ctcctgtctc	agcctcctga	gtagctggga	ctacgggtgt	gtgccaccaa	gcccggttaa	20340
ttttttttta	atttttagtaa	agacagggtc	tttctatggt	gcccgagctt	ttctggaaact	20400
ccatcttggc	ctcccacaa	gctgggatta	caggctcgag	ccacggagcc	cagcctgttt	20460
ttgttttttc	actgataaa	ttttgcgggg	tgtggtagt	tgtgcctcta	gcgatttggg	20520
aggctgaggt	gggaggatcg	cttaagccca	ggagtttgag	gctgggctca	agtgatcagg	20580
agggtgaacta	tgatcatgtc	attgcattcc	agcctgggtg	acagagcaag	aacctatctc	20640
ttaaaaatat	atatttaaaa	agtattgggt	gtggtggctc	acgcctgtgg	tcccagctac	20700
ttaggcacat	gagggtggag	gatggcttga	gcccaggagt	ttgaggttgc	agcgagccaa	20760
gatcgtgtca	ctacactcta	gcctgggtga	cagagcccag	accctgcctc	tttaaaaaaa	20820
aaaacccaaa	aacatgtatt	ggaaacacag	catgcctggt	cagtcaogtg	ctctccatgc	20880
tgctttctgc	tccagagacc	ottatggcct	gaaagctgaa	aatattttct	atcctttaca	20940
aaaaagtttg	ctgacctctg	tcctggaaaa	ttcatctccc	aagttctctt	ccggcactgg	21000
cgttcctggg	tgtcctaaat	ttggcccctg	ttatttctga	actctgtttt	ggctctgttc	21060
cctcccagga	gccaggacag	gcacgttctc	tgcattctgt	cccctgacgc	ccagaggctt	21120
ggctcggctc	aggcattctt	ggaaatatct	ggctccagga	aaggcagagg	cctcctgagt	21180
gggcccagag	ggaacctgcc	ccaggctctg	gggaggcctg	acccagcaga	gtggcttttg	21240
ccgatgggtt	gggcccgtca	agatgtgctg	aaagtgtgct	tcagaaggcc	actttgggat	21300
tccttccctc	agtattagag	caactgagag	ctgctcattg	caagcctgat	gttttcccag	21360
ttggccgggt	ccaccgggtg	ccctgggatt	ctgggatctg	ggtggaaggt	agggggcttg	21420
ggggagtgtc	ctgggttctg	gaatccaggt	ggcaagtggg	gaggttcagg	gagtggcttc	21480
tgagccacca	taggggtctc	tgtgggaggc	tctgcccato	caggagattc	cgcaggccct	21540
gccggcccag	agccagcgtc	ttgcgcttgc	cagagctaca	gccagcccca	gccgggtgga	21600
acagcccgtc	gcctcctctc	actttgtttt	ggggccacct	gggagtgtgg	agcaagggtg	21660

10

20

30

gagagggagg	aagtggctgc	eggccgctgc	ccagcaccct	tgtttgctt	gggcccctcg	21720
tgggtccctt	tttattgtc	ttcaatgaag	ccagggaat	ggacttcctt	gcctcacttc	21780
agttcaacat	gtctggaagt	ttggtattaa	aattaagaaa	gtgtggaaat	agagcaagaa	21840
gagaaaaatc	tctccaagag	ataatagtga	cctctgagct	gggcgcggtg	gctcaccgct	21900
gtaaatccca	gtacttttgg	aggctgaggg	gggcagatca	cctgaggtcg	ggagtttgtg	21960
accggcctga	ccaagatgga	gaaaccccg	ctctactaaa	aataaataaa	taaataaata	22020
aataaataca	aaattagcca	ggcatgggtg	cgcttgccca	taatcccagc	taaggcagga	22080
gaatcgcttg	aacctgggag	gcaaagggtg	cagtgaagca	agatcacgcc	attgcactct	22140
agtctgggca	acaagagtga	aactccgtct	caaaaaaat	aaataaataa	aaaataaaaa	22200
tagtgacctc	tggccagggtg	tggcagctca	taccggtaat	cccagcactt	tgggaaggag	22260
gocgagatgg	gcagattgct	ttagcacagg	agtttgagac	cagcctggcc	aacatgggtg	22320
aaacccatct	ctacaaaaat	agaataaaat	ttaagaggta	atagtgaact	tttggtagat	22380
cgaaacctgg	attgctttct	ttttctaaat	gctgattctt	ttctttgttg	tgtttgtgtt	22440
ctgtgccgat	gtccctcccc	cagccctgtt	attgtgagt	gaagaagggg	aaagggttcg	22500
cccgctactg	tgagccctc	ctctcagct	gggtgtcctt	ggagaagcct	gcacttcttc	22560
attgtacgcc	agggtcgggt	ccctccctgg	agtgttctg	tgctgctggg	atggggccaa	22620
ccctcagat	gtttctgag	tgtcacacac	aggtgtgtgc	attcatggcc	tttgcgtgtc	22680
ttctgttgt	ggaggcaaaa	atgtgaagaa	ccctagatga	ttttgggacc	agggtccat	22740
caactgctgt	tcattgcaca	ccggagcatc	caggcatggg	tggagagctc	agacttccag	22800
gcacggctgc	aggggctggt	ctaaccatgt	tcccgccgc	ctgctcgtca	gaaccgcctg	22860
ttgggagctg	ttatcatgat	accatacctg	ggccctgggc	tatccgattc	tgacttaatt	22920
gctccaggtt	ggggccaggc	cgttgtttgc	tgttttgttg	ttcttctgt	gacgttagcc	22980
actgggctaa	tctgagcccc	tcagttacag	gtggagaaac	tgagaccat	gggggtgcaa	23040
ggacttgccg	aggaccacga	gccccttggg	ggcagagctg	aggcggggcc	tggctttggg	23100
tcccagagct	tccagtcccc	ttcccgctct	cctaacagct	tttttttttg	agacaagatc	23160
tcacctgtc	accaggctg	gagtgcattg	gcctgatctc	ggctcactgc	aacttctgct	23220
agctgcttc	cagcgattct	cctgcctcag	cctcccgagc	agctgggatt	acagggtgtg	23280
gcgcctatgc	ccaagctcgt	tttttttgta	cttttagtag	agatagggtt	tcaccatggt	23340
ggccaggctg	atctcgaaat	cctgacctca	aatgatccgc	ctgcctcgcc	ctcccaaagt	23400
gctaggatta	caggctggga	tcacactgtg	cctggcccta	gcagctttgt	cctgtgccat	23460
ccaacaacag	atgaccgaag	tctttgtttc	ttaacatgca	ttccatctgc	cttacagttt	23520
tgccacctgc	aaaacagagg	acttgtcgct	tttctggtaa	gctggaaatg	taatctggta	23580
gcaggaggcc	tgtggaagct	tgcccttaat	ggccttgtgt	ctctttcatc	ctgtcctgag	23640
agccggagaa	cttggatggt	gcacctaaat	caaccttcct	gttaacatac	agttctgcag	23700
gctcatggat	catcagaacc	acgtectatc	tcacggggct	gtatgcttcc	gttggttcag	23760
gtgtttttac	cttgacagta	ttttctctc	ggtggctttt	gcggtgggtg	cttttaatca	23820
gcattgactc	ttcaagaaaa	atatttagct	gctacatctc	agaggagaca	gggtggaaaag	23880
catctgagac	ctgcaggctc	agacttagaa	ccagaagtgc	cctcagagtt	catccggccc	23940
tgccccagcg	ggaatgaggt	tcacagagaa	gcgggagaa	tttgcccag	gccctgcccgt	24000
tgctcataac	tgccccagg	ccttacattt	gctccagggt	ctgccccagg	ccctgcagtt	24060
gctcataact	gccccagggt	cttatatttg	ctccagggtc	tgccccagg	cctgcagttg	24120
ctctgtgtgg	tgggtgtgat	ctggagccct	ccgcccattg	ctgcacctgg	ggcaggcatt	24180
gctaattgat	cccaggactc	cttccctggg	agcacgccct	ggttctccag	gcagccgctg	24240
cctgtcagcc	tgcagtgggt	cgggagagga	cacctgcttg	cctgggtctgt	tccaaatctt	24300
gcttctcatc	ccagcacagg	taggggggtg	tatgggaaag	ggatcctcag	ttggccctgt	24360
cactgctcta	tcagctgggg	acgtggcatc	ctagtgaaaa	catcatggcc	gggcgcgggtg	24420
gctcacgcct	ggaatcccag	cactttggga	ggctgaggag	ggtggatcac	ttgaggtcag	24480
aagtttcgaga	ccagcctggt	caacatggtg	aaacccatct	ctactaaaaa	tacaaaaatt	24540
cgcagggtgt	ggtggcgggt	acctgtaatc	cgagctactc	gggaggctga	ggcaggagaa	24600
tcgcttgaa	ctgggagggtg	gagcttgacg	tgagccgaga	tcttgccact	gcactccagc	24660

10

20

30

ctgggcaaca	gagtgagacg	ctgtctcaaa	atctcaaaaca	aacaaacaaa	caaaaaacaa	24720
acaaacaaag	cgtcatttat	ccagcaccoc	tggggaacca	tgctacctgg	tgttttatgg	24780
tacctggcaa	gggtcagggtg	aagttgctgc	tcttgggcat	tgaaccogtc	ttgtttgggg	24840
cagctcaggc	cccaggcagg	gtccgggttg	gctctcggtg	gtgtggccct	ggcccatcca	24900
gacctatatt	tctgccgtcc	tgacagtgat	caatgttgat	gggacgaaga	ggcggaccct	24960
cctggaggac	aagctcccg	acattttcgg	gttcacgctg	ctgggggact	tcactacttg	25020
gactgactgg	cagcgccgca	gcatcgagcg	gggtcacaa	gtcaaggcca	gccgggacgt	25080
catcattgac	cagctgccc	acctgatggg	gctcaaagct	gtgaatgtgg	ccaaggtcgt	25140
cggtagtcc	gggggggtcc	aagccatggc	tcagccatgc	agacttgcat	gaggagggaag	25200
tgacgggtcc	atgctgggc	ataagtgttg	agctcagggtg	ccccgacctg	gggaaggggca	25260
ggacaggaaa	ggtagacagta	tctggccaag	gacagatggg	aagggaacca	gggagctgat	25320
tagggagtg	ttatggacta	ggaatgtcgg	taacaatggt	tagaaagtga	ctaaccatttg	25380
ttgagcacct	gctgtgtgccc	cggccctggc	cggggagcctt	cgtgcccaca	gtgaccccg	25440
ctgcaaatgt	agtcccttgc	cctactcgca	ctggggagca	ggacgcagag	ccgtgcaact	25500
cacaggtgcc	aagctcaggga	ctccctcctg	ggtctgcctg	ggctgggctg	tgcttggtgc	25560
ccctgtggcc	cacgcgatgtg	caccttcac	ctgaaagcca	ggatcttcag	gacgtccccc	25620
gaggaggctg	ttgtctggca	caatgatttg	tcctctcctg	aaaagggtgac	agagttacac	25680
tgagagagac	agcatccagg	tgccggcagg	acaggccctgg	ggctcggcgg	cagggacctct	25740
gtgtcctgcc	gggggtccac	actgcacctg	cttgtcagag	gcactcagtc	aatctttgct	25800
gatgaaggat	gagaggacag	aggacgtgat	gcttgctgct	gcattgcctg	cagtcctggg	25860
tgagatgcc	gggttgactc	tgctgccctg	cgggtggatg	tgatgtcaga	tccccggctt	25920
taaaatacga	gggagctggg	aattgagggga	gcaggttggg	gcagaaagca	cagccccgtg	25980
gaagccttga	gctgaggcag	tgtagggcgac	ccctggagca	gtgagtgcctt	ccttcatggc	26040
cttcategca	cctgcagtc	ctcatgtagg	ggatgccatc	catgaattta	gttttccag	26100
cctcctttaa	aaacgcgttc	atgctggggc	cggggcagtg	cagtggctca	catctgaaat	26160
cccaccactt	tgggagggcg	aggcgggttg	atcatgaggt	caggagatcg	agaccatcct	26220
ggctaacaag	gtgaaacccc	gtctctacta	aaaatacaaa	aaattagccg	ggtgcggtgg	26280
cgggcgcctg	tagtcccagc	tactcgggag	gctgaggcag	gagaatggcg	tgaacccggg	26340
aagcggagct	tgcatgagc	cagatttgcg	ccactgcagt	ccgcagtcgc	gcctgggcga	26400
cagagcgaga	ctccgtctca	aaaaaaaaaa	aaaaagtaca	aaaaaaaaaa	aattagctctg	26460
ggtgtggtat	cacgcgccta	taatctcact	actcgagagg	ctgaggcgga	gaattgcttg	26520
aaccaggag	gtagagggttg	tagtgagccc	gtatcgtacc	actgcctcc	acctgggcaa	26580
tagagcgaga	ctctgtctca	aaaagaaaaa	aaaaaaaaga	acatttatgc	caggtgtggt	26640
ggtcctatgc	tgaaatccca	gaactttgga	agactgaggc	aggaggatca	cttgagccca	26700
gaaattttgag	agtgtcttcc	ctgggcaaca	tagagagacc	tcactcttac	cagaaaaaaa	26760
aaaattagcc	cggcatgggtg	gcatatccct	gtggtccag	ctacttaggg	ggctgacgtg	26820
gcaggatcac	ctgagctcgg	aggcagaggt	tgaagtga	tgagatcatg	ccactgcact	26880
ccagcctggg	tgacagacag	agaccctgtc	tcaaaaaaaa	aaaaaaaaaa	aagcatttac	26940
tatccaccat	ggaagggtgag	actgacctgt	gagtgttggt	tcaaagaaca	aaaaataaac	27000
ccagagata	agacaaaagg	gtgcctccat	gggggtgtga	tttaaagctg	agaaattggg	27060
cttcttcccc	ctccccctctc	accccggtgt	ttgtctaaagg	agatgggaaa	aaggattctt	27120
tttttggtg	aaatatatta	cactaaatta	aagccaattt	taacagcact	ttggttgatg	27180
agtgaatta	acagactggc	caaaaaataa	cgaacgggtc	gtactatgtg	aaaaagaggc	27240
agctttggcc	atgctggggc	aatgtgagtt	ttcagggttg	ctgggaatgt	ctgtgaatcg	27300
gagggaaggc	ctagctggga	ctctcaggag	ccaaggccct	gaggggcaac	ttgcctggtc	27360
cctgccctga	ggcgttcact	gctttcttcc	tgggccagat	cacaggcccg	gaggctggac	27420
cactgggctg	gcactcttgc	cgagctgtct	cctgacttcc	tgaccatgct	cctttcagca	27480
gccttgctgc	actttagttt	ccttgaatga	aaaatgggga	tgagaatagc	tcctacctcc	27540
aagggtgaatg	gagtgagttc	ggacaggtga	ctccctggga	ccagtgcctg	gcgcctgaca	27600
aggtccagtc	agagcccgca	ctgtgtttac	tgataccctt	ggctgtacca	ggggagaact	27660

10

20

30

tggttgccat	tgccaggtgt	tctcccacca	cccccaactac	tgtccctggt	tgatgtgtgg	27720
cgggaataaa	gctgtgcaca	ttggagcttt	tggcacatcc	tggttttcag	gtgaaaggtg	27780
cgtgtgtgtt	tgaggggtta	gcctggccaa	cccagccatg	aggtcggacc	tgacctgggg	27840
gtgagtcctg	agctcggcac	ccctgagctg	tgtggtcac	ggcagcattc	attgtgtggc	27900
ttggggcggca	cccccttccc	tgttgggtg	ttgatgttta	gactggagcc	tctgtgttcg	27960
cttcacaggaa	ccaaccgtg	tcgggacagg	aacgggggtg	gcagccacct	gtgcttctgc	28020
acacccacag	caacccgggtg	tggctgcccc	atcggcctgg	agctgctgag	tgacatgaag	28080
acctgcacog	tgccctgagc	cttcttggtc	ttcaccagca	gagccgccat	ccacaggatc	28140
tcctctgaga	ccaataacaa	cgacgtggcc	atcccgtca	cggcgctcaa	ggaggcctca	28200
gccttggaact	ttgatgtgtc	caacaaccac	atctactgga	cagacgtcag	cctgaaggta	28260
gcttgggcca	gaacgtgcac	acaggcagcc	tttatgggaa	aaccttgcc	ctgttccctgc	28320
ctcaaaggct	tcagacactt	ttcttaaagc	actatcgtat	ttattgtaac	gcagtccaag	28380
ctaatacaat	atgagcaagc	ctatttaaaa	aaaaaaaaa	tgattataat	gagcaagtcc	28440
ggtagacaca	cataagggtc	tttgtgaaat	gcttgtgtga	atgtgaaata	tttgttgtcc	28500
gttgagcttg	acttcagaca	ccccaccac	tccttgctcg	gtgcccgttt	gctcagcaga	28560
ctctttcttc	atttatagtg	caaatgtaaa	catccaggac	aaatacagga	agactttttt	28620
tttttttttt	tgagacagag	tcttactctg	ttggccaggc	tggagtaacc	tagcgtgagc	28680
tcagctcact	gcaacctccg	cctcccaggt	tcaagcgatt	cttctgctc	agcctcctga	28740
gtagctggga	ctacagacat	gcaccaccac	accagctaa	ttttttttat	attttttagta	28800
gagacagggg	ttcatcatgt	tggccaggct	ggtcttgaa	tcctgacctc	aggggaacag	28860
acgggggttg	cctcccaaa	ggcggaata	acaggggtga	gccaccgttc	cgggcctagg	28920
aaaacttttt	gccttctaaa	gaagagttta	gcaaaactag	ctgtgggctg	gccttctgat	28980
tctgtaaaga	aagtttgatt	ggtggctggg	tgcggtggct	cacacctgta	atcccagcac	29040
tttgggaggc	cgaggtgggc	agatcacctg	aggtcgggag	ttcgagacca	gcctcaccaa	29100
cgtggagaaa	ccccgtctct	actaaaaata	caaaaaaaaa	attaaccggg	catggcggcg	29160
cctgcctgta	atcgcagcta	ctcaggaggc	tgaagcagga	gaattgcttg	aacctgggag	29220
goggaggttg	tggtagctgt	agatggcacc	attgcactcc	agcctgggca	acaaaagtga	29280
aactccgtct	cagaaaaaaa	aaagtttgat	tgggtgtaac	aaagcgcat	tgttttatgga	29340
ttgtctgtgg	cagcttttgt	tctgccgaga	tgagttgtga	cagatctgta	tgggctctaa	29400
agcctaaaa	atgtgccatc	cgccccctta	cagaaaaagt	gtgctgacct	ctgttctaaa	29460
gtattggaca	actacaatgt	ttgtctattt	attattctat	gatttgtttt	ctgctttttg	29520
ttgttggtgt	tgttggtgag	atagggtttc	cctctgtcac	tcaggctgga	gtgcagtggt	29580
gtaatttcag	ctcactgcag	cctcgacctc	ctgggctcta	gtgatcctct	catctcagcc	29640
tcctctagtag	ctgggactac	aggcacacac	cacctctcct	ggctgatttt	tttttttttt	29700
tttttttttt	gtggagacag	ggtttccgca	tgttgcccag	gctgggttca	aactcctagg	29760
ctcaaacacc	cacctcagcc	tcccaaaagt	ctgggattac	aggcgtgagc	cacctgccc	29820
agcctattct	actgtttgta	ttacatagct	ttaaaagatt	ttttatgact	ttaagtcaca	29880
aggggttctt	gtagaaaaaa	atatatatat	aggaaagtat	aaaaagaaag	taaaaattgt	29940
ccataacctc	tccagccaga	gacgaccgtt	gctgacacct	cagcatattg	cctttaagtc	30000
ttttttctct	aagatagcat	ttctcttcat	cacagtcata	tgctacgcag	aattctgtat	30060
cctgattttt	tcacttgaca	ttacaacagg	tatttgatgg	cgctgtgaca	aactctttgg	30120
cacaatcttt	taaatgtatg	aaatactcca	ctgcacagat	gtttgctttt	aggcttaact	30180
gttcttttat	tttgctgtgtg	ctggttacag	ccgggcacag	tggctcatgc	ctgtaatcac	30240
aacactttga	gagggtgagg	caggaggatc	acttgagccc	agaagtttga	gaccggcctg	30300
ggcaacatag	tgagacccca	tctctacaaa	aaactttttt	aataagtcgg	gcgtagtggg	30360
gcatagtctgt	agtcccagcc	accaaggagg	ctgagttggg	aggattgctt	gagccccagg	30420
aggttgatgc	tgacgtgacc	tgagattact	ccactgtact	ccaacctgag	cgacagagca	30480
agacttgtct	ggggaaaaaa	aaaaaaaaaa	tatatatata	tatatatata	tatatacata	30540
tatacataca	cgcacacaca	cataataata	aaatatatat	ttataaatat	ataatatata	30600
atataaaaat	atataatttat	aaataaaaat	tataaattat	attataaagt	aaatatataa	30660

10

20

30

tatataatat	aaaaatatat	attatataat	atataataaa	atatataata	taaaaatata	30720
tattttataaa	taatatataa	tacatactta	taagtatata	tttaaaatat	atgtaatgta	30780
tatttttttaa	tgtatgatat	ataatataca	tttataaata	cacattttata	ttatttttata	30840
taaaatatat	ataaaatctc	caagttgtctt	tttccaaaaa	ggtgtcttgc	tgcatttcaa	30900
acatttcattt	aaaaaacttg	atgctgggtga	tctgggtccag	aatgtgttca	gtagctgctg	30960
ccagtgggcca	agcatctcgg	gagatgtcta	caaaacacgc	tggttctggc	ctggcggtggt	31020
ggctcacgcc	tgtaatctca	gcactttggg	aggtgagggc	aggtggatca	actgaggtct	31080
ggatttcgag	accagccttg	ccagcttggg	gaaaccccat	ctctactaat	aatacaaaaa	31140
aatttagccag	gcgtggtggc	atgtgcctgt	aatcccacct	acttgggagg	ctaaggctgg	31200
agaatcgctt	gaacccaggg	ggcagaggtt	gcagtgaacc	gagatcgcac	cattgcactc	31260
caggctgggg	aagaagagcg	aaactccgtc	tcaaaaaaaa	aaaaaaagat	gctgggttcc	31320
aaaatgtggc	ccttttcttc	ctcacctgct	gccagaccat	cagccgcgcc	ttcatgaacg	31380
ggagctcggg	ggagcacgtg	gtggagtgtt	gccttgacta	ccccgagggc	atggccgttg	31440
actgtagggg	caagaacctc	tactgggccc	acactgggac	caacagaatc	gaagtggcgc	31500
ggctggacgg	gcagttccgg	caagtccctg	tgtggagggg	cttggacaac	ccgaggtcgc	31560
tggccctgga	tcccaccaag	gggtaagtgt	ttgcctgtcc	cgtgcgtcct	tgtgttcacc	31620
togtatgaga	cagtgcgggg	gtgccaaact	ggcaaggtgg	caggctgtcc	gtgtggccct	31680
cagtgattag	agctgtactg	atgtcattag	ccttgatggg	ggccaggact	ggtagggccc	31740
tcagaggtca	tggagtctcc	togtggagcg	ggtgctgagg	ctgtatcagg	cacagtgtct	31800
gctgctttca	cctgggcccgt	ctcacccaag	tgtccatgga	gcctgcgtag	ggtgggtatc	31860
tgtgtcgatt	ttacagatgc	agaaacaggg	tcagagaaac	cgagtgaact	ccctaaggtc	31920
acatacccca	ttagatcaga	gctggggccag	gaagtgtctg	ctcaggctcc	tgaccaggtc	31980
tccttgcttt	gtactcttgc	caaaacctat	atccagaact	gactttgagg	tccccggacc	32040
tcaggctcct	cggaaatggc	ctcttggagg	ctgctgagcc	acagcttagg	acccacctcg	32100
agaggcaaat	gtgctttgag	ctgccaggcg	tccctggggc	cctgccttgg	gcacgggggt	32160
cagacaggcc	ccagatgtgt	ggggcgtcct	tctggacttg	agttttcttt	tctgtgtggt	32220
ggacacagtg	ctcacccctt	aaagcacctg	tgatgtgtgc	agcagcccaa	tccctgcctg	32280
tgcctgttcc	tgctagggaa	ggaagggaata	cttcaggatg	gcaggacaac	agaaagaggt	32340
ccagggtttta	gagcaagggc	aggtcaaaact	tagaaaaatc	tggaaatgag	atgtgcattt	32400
cctcttctgg	atctgtctaa	agaagaggga	aggaggggct	gctgggggag	gagcccagag	32460
ccgagtttac	atccggatcc	cgcaaggcct	cccctgccct	gaggctctgt	tttgtgatgt	32520
gcttgtgtcc	atcctgggtt	ctgccgtgtc	cccaacatcc	ggccaagcct	aggtggatgt	32580
tccagcacac	actcacccct	tctgtgcacc	tgtttttgtg	tccgtaagtg	ggtatttact	32640
caccttacga	gtgagccact	gtgggaatcc	agggaggtgg	cgcagtgaac	acccctggag	32700
ggatatgtgt	gtggcagggg	tcgaggggtc	cgcccttccc	tgttctctgc	gcgtggcctt	32760
ctccaggacg	gggagggctg	agctgaagag	gtggggacag	ttgcgtcccc	ccgccaccca	32820
ctgtcctcgc	gtgagagcag	actcactgag	cctgcccttc	tcccttgtgc	cttcacagta	32880
catctactgg	accgagtggt	gcggcaagcc	gaggatcgtg	cgggccttca	tggacgggac	32940
caactgcatg	acgctgggtg	acaaggtggg	ccggggccaac	gacctacca	ttgactacgc	33000
tgaccagcgc	ctctactgga	ccgacctgga	caaccaatg	atcgagtctg	ccaacatgct	33060
gggtgagggc	cgggctgggg	ccttctgggc	atggagggcg	gggcagccgg	gcgttggcca	33120
cctcccagcc	tgcgcgcacg	tacctgtgtg	cctgcaagtt	ccccaacctg	gcaggagctg	33180
tggccacacc	cacgactgcc	cagcagcctc	accctctgct	gtgggagtgg	tccccgtcca	33240
ccccctgggtg	cctttgtctg	agttatgtcg	ggagaggctc	tggtagacgc	tgtttcctgt	33300
gcacctgctg	ggcactaggt	cccagctaata	ccctgtgcca	ggactctaat	ttcacccctaa	33360
cacacatggt	ggttttcatt	gctggggaag	ctgaggcctg	agcacatgac	ttgccttagg	33420
tcacatagct	ggtgagtcca	ggatccccc	gagataccag	ggccagcact	cgatccccac	33480
ccagccctga	accccacat	gtgctgggat	tgtgctggga	gtgtccacac	gcctgggacc	33540
ccagggctgg	tgtctctatc	tcccttttcc	agatcatgag	aatgaggctc	agggaagttt	33600
gaaaaaaacc	tatcccaagt	cacacagcaa	caggagcagg	atttgaaccc	agaaaagggg	33660

10

20

30

accgcacact	ctgttctgct	agagtagtta	gctgtcctgg	gtgatatggc	aggtgacagg	33720
ggcaactgtg	cttaacaaag	gaacccocat	ccccctgcc	aagttgggag	actagaaggt	33780
caggggcaga	agctctgaag	ggccagggtg	agtggctgac	acctctaata	ccagcacttt	33840
gtgaggccaa	ggcgggcaga	tgatttgagc	ccaggagtgc	aagatcagcc	tgggtaatgt	33900
agtgagacgc	catctctaca	aaaaaatttt	ttaaaaatta	gctgggcatg	gtggttcatg	33960
cctgtagtcc	aagctacttg	ggaggctcag	gtgggaggat	tgcttgagcc	caggagggtg	34020
aggttgttgt	gagctgtgat	catgccactg	cactccagcc	tgggcaatag	agtgagaccg	34080
tctccaaaaa	aaaaaaaaga	agaagaaaaa	gaagctctga	ggctccaagt	ccccaggcac	34140
cccttggtct	gagggcagac	aagggaggag	agggtcacct	gggcagccct	gacttttgtc	34200
ccctggcaaa	gggaccttca	gtgaccttgg	ccctaggaga	gcctctgagc	acgtcagcca	34260
tgtcgaaccg	ctcagggaag	gcagcaagaa	tttggcttct	gacctctgcc	tctcctaata	34320
gccatctgca	ctgggtgtgg	ttgtgcccat	tttacagatg	aggaggctgg	ggcatcgacc	34380
agctgaatgc	cttgtccag	gtactgcgta	ggcagagctg	gcagttgaac	cccgtgtcct	34440
ggttgtcgct	gggggtgggc	tgcacctga	cttgtgaggc	cagttagcaag	gtttgcacgt	34500
gacttcgtga	ccgtcaccca	gctctgcagc	acatcccggt	acccagctca	tccaggccgc	34560
atgcaaacct	ggttgccagg	gagaaaccag	tcacgcgaca	gctgtgggtg	cctgaaatga	34620
ttaagctcat	taatcacccc	ggagtggagg	cagactcaga	tgaaaaccag	caaaagccct	34680
ggaaactcat	gtgaccttgc	caatgagggc	ggccatgtgc	attgcagcct	ggccgtcact	34740
cctcggtaag	tgttttggac	ttaaacgctc	cggatgttta	ctgagtgcct	gattaataac	34800
atggaaggcc	tggtctcatt	gctgtgggag	tgaaggatgc	acagccaggc	ctgacatgat	34860
gagaacaaga	acctggagtc	tcgctgcctg	ggtggtaatc	ctggccctgc	cacttagcaa	34920
ctgtgtgact	gtagccaggc	cacttaattt	tgctagatcc	tgccctgcgt	tcagtggatc	34980
ttgtctggtt	tccaagggtg	ccaaacactt	taaggcattc	atgtggtcgc	taggctgcag	35040
ggttgaaccc	tggctcacc	cgcaggggcg	cgtgtgctct	gtggcctggc	tgtgcctttg	35100
ctgacaccgt	gcccggtgtg	gttcatgcag	gtcaggagcg	ggtcgtgatt	gccgacgac	35160
tcccgcaccc	gttcggctcg	acgcagtaca	gcgattatat	ctactggaca	gactggaatc	35220
tgcacagcat	tgagcgggcc	gacaagacta	gcggccggaa	ccgcaccctc	atccagggcc	35280
acctggactt	cgtgatggac	atcctgggtg	tccactcctc	cgcgcaggat	ggcctcaatg	35340
actgtatgca	caacaacggg	cagtgtgggc	agctgtgcct	tgccatcccc	ggcggccacc	35400
gctgcggctg	cgcctcacac	tacaccttgg	acccagcag	ccgcaactgc	agccgtaagt	35460
gcctcatggt	cccccgaccc	tcactccctc	gttagatcag	gctggttctg	ggagctgacg	35520
ctgaaaggag	cttctcatct	ggggttcctg	ggtgtacata	gatggttggg	taggttgtgc	35580
actgcacaag	ctgcatgatg	ctacctgggg	gtccagggtc	aggctggatg	gacttgttgc	35640
ttcatcagga	catagataaa	tggccaaaac	tcctcagctg	gaaggctcct	ggcaggatct	35700
ttgggtgtga	aaaccagtca	caggggaagg	gtgcttgctc	atactgccag	cacagtgtct	35760
agtgttttcc	atagcgctcg	tttactcctc	aagcctggag	ggtggggagt	agcatggtcc	35820
catttcacgt	acaaggaacc	cgatgcacag	agagggtgtg	caaccatcc	aaggccatac	35880
aaactggggtg	ggttgagccg	gggttgactg	tggcaggctg	gctcaagagt	ccctgtcctc	35940
gaacccttgc	caggcagcct	ggcatcagct	cggggaattt	ttgccctgac	ccttggaagc	36000
aagtgggcct	ctttgttctc	atgtcagtga	tgagaagagt	gactttccta	tggccctctc	36060
ggagtacagg	tgtttcctgt	tggcgggctc	ttcccccata	acatcagcag	cgagctgggt	36120
atgattccct	acgcagaact	tgatagttta	taaagctcct	tgtcatccag	gccccgttgg	36180
agtctcacgc	agacctgggc	gcaggcgggg	ctggtcttgc	ctgtcccagc	tgcattggatg	36240
gggaacttga	ggcttgcaaa	ggttaagggg	ctggttcgag	cccacgctgg	caggagatgg	36300
gcctggggcca	gagtctggga	cttcccatgc	ctgggctgtc	tttggctcct	ttgtccacca	36360
tccctccctg	gggcatgac	cttagagagc	caaattggag	tgcaggtaac	ccacggcaag	36420
gaggggttgc	catgactcag	agtcctccgt	ctgtggccgg	cagtacctgg	tgcacagact	36480
tggatttcag	accagccact	gtagcccgct	gacggtgcgc	tcgaagtgcc	acagcttctg	36540
aagccaggca	ggactcaggc	caggagactc	tgttagctgt	tgagaggggg	aggccaacgg	36600
atgttctggg	tctgctagag	agctgggtct	tcggatcctg	gtaccagtgc	actgagagga	36660

10

20

30

ggcccagctt	gattctgggg	ctgccttggt	gtggcatgtg	ctgctcactg	acaccctcga	36720
ggagtgtctt	ctctcgggct	tgttgactgt	gcccggtttt	cgcagttca	ctggtgcaca	36780
catagggcaca	tagcaaacgg	cacacacagt	cgtgggtatg	agtttcacta	cattccacca	36840
ccagtgttca	ctaccattac	ctgccttcgg	tcttaagtgt	tcatcattta	aaaataaatt	36900
tattgggctg	gacgcggtgg	ctcatgactg	ttatcccagc	actttgggag	gctgaggcgg	36960
gcagatcacc	tgaggtcagg	agttcaagac	cagcctggcc	aatatgggtga	aactccatct	37020
ctactaaaaa	tacaaaatta	gctgggcatg	gtggggcatg	cctataatcc	cagctactca	37080
ggaggctgag	gcaggagaaat	ggcgtgaacc	cgagaggcag	agcttacagt	gagcccagat	37140
agcaccactg	cagtcacagc	tgggcaacag	tgcgagactc	catctcaaaa	aaaaaataaa	37200
taaaataaag	aaaaataaat	ttatgatcta	tttcaaaaat	aacacatgta	ctttgaaaca	37260
gcagagacac	atatgacacg	gagaatgaaa	ttcccocatag	cgcaccccca	agagacagcc	37320
ctggtccccc	cgtctttccc	gtggacctcc	agcggggcag	atgctgagcc	gcctgttgct	37380
gagtggcatg	ctatcccgtc	ctccagctcc	tctgtggctt	acagacaccc	acctgcagcc	37440
ctgtctttgc	ctcctctagc	gcccaccacc	ttcttgctgt	tcagccagaa	atctgccatc	37500
agtcggatga	tcccggacga	ccagcacagc	ccggtatctc	tcctgcccct	gcattggactg	37560
aggaacgtca	aagccatcga	ctatgaccca	ctggacaagt	tcactactct	ggtggatggg	37620
cgcacgaaca	tcaagcgagc	caaggcacgc	gggaccagg	caggctgccct	gtgggaaggg	37680
tgcggggtgt	gcttcccaag	gcgctcctct	tgcgtggttc	caggctgctg	cccctgtcct	37740
tagcagaggg	aggaacaga	ggatggctct	gggtgaatga	tgacttgggc	ttcgattatg	37800
tagtcacagg	gtatgacctt	gagatgcgtg	gaaccccgag	actgtgatta	tatgtagaaa	37860
ctgggtttcc	cgttgtttta	agtagtcatg	gtggggtcag	acccacagg	acttttgtct	37920
tttcaagaaa	gaaaatgggt	gtgtgtcatg	caggggtagt	tggtactggt	taatccaggt	37980
ttatccttta	ttttgtggga	actgtacagt	cattttctgt	acaatgctgt	atatgctctt	38040
ctgaaagaca	cctatgcaaa	atcgacagct	aaaaatgaca	caactcatag	ggaaagcggg	38100
gccagggcac	agccctcaaa	atctccatca	atgacatgta	agaaaagaga	ggaaacctggg	38160
aaatagcaaa	gtgccttttg	cacattaaat	ggtagcttat	atcccacaat	actgtgcatt	38220
cgtaaacgtt	aatgtgtgca	taaatacggc	acttcaacct	gggaagatct	ggagtgggt	38280
tatgagtgtg	gaagggtgta	gcgcagagt	ttttgtgaaa	cactggaaag	aggattgtgg	38340
gaaatcaaat	ggaaagtctt	caccccaggc	gtggagaaga	gtgggtcatg	gccccagcag	38400
tgagcccagg	gaggtcagag	acggaggtgt	gtgtgtgggt	gtgacctctc	gcagttccct	38460
gccggtctga	gttttttgca	ttcgcttaat	gtttctcgtg	gaggaaattg	tgcattgagca	38520
aatgtgaaac	cgtgctgtgc	tcaaattgtc	ctaatacatc	attgcattgg	aacagattgg	38580
ctttnttttt	tttttttttt	tttttttttt	tttgaaatgg	agtctcactc	tgtcaccagc	38640
ctggagtgc	gtggcatgat	cttggctcac	tgcaaccttt	gcctcctatg	ttcaagtgat	38700
tttctgtcct	cagcctcctg	agtaactggg	attacagggc	atgagccacc	goggccggcc	38760
agatttgcat	ttttgaaaca	actgctaggc	tgggcccggg	ggctcacacc	tgtaatccca	38820
gcactgtggg	aggccgaggc	aggtggatca	cctgaggtca	ggggttcgag	accagcctgg	38880
ccaacatggg	gaaaccccg	ctctactgaa	tatacaaaaa	tcagctgggt	gtggtggcgg	38940
gtgcctgtaa	tcccagctac	tcaggaggct	gaggcaggag	aattgcttga	accaggagg	39000
cagaggttgc	ggtgagccga	gatcacacca	ttgcactcca	gcctgggcaa	caagagcaaa	39060
actccatctc	aaaaaataaa	aaatagaaaa	acaagtgtct	tagcgggaagt	gagcactttg	39120
cggagtccagg	cttgtgtggc	ctgttccaca	aatgatgtgc	tcacggtggc	ctcaggccca	39180
cctggagtct	gcagcatggg	gcacaacagg	ttcatttagt	tagaattcca	ggacaggcct	39240
ggctcctaag	cagccttctt	ttacaaaaac	tgcagagccc	gcctgtatcg	tagcactttg	39300
ggagggccgaa	gtgggtggat	cacgaggtca	ggagtccaag	accagcctgg	ccaacatggg	39360
gaaaccccat	ctctactaaa	tatacgaaaa	ttagctgggt	gtggtggcac	gcgcctgtag	39420
tcccagctac	tggggaggct	gaggcagaat	tgcctgaacc	tgggaggtgg	aggttgcagg	39480
gatctgagac	catgtcattg	cactccagcc	tgggcaacag	agcgagacgc	catctcaaaa	39540
aaaaaaaaac	tacagagcca	cacggcctct	ttctccaccg	agtgttgggt	tgggagcttg	39600
tgttattgtg	gtgaaatctt	ggtactttct	tgaggcagag	agaggctgag	cgcctggaga	39660

10

20

30

gactttcaca	tgggtcgcca	tgctcgccgt	cggtttcgct	gttgtgtctcc	ccatctgaag	39720
gctgggtgcg	tccagacagg	ctggacgcgc	ctttccacca	gatccttcct	cccgcagcag	39780
tttctagtta	cgttgtactg	tgaggtctgt	gtccttggtt	gatggcaaaa	gtcagccgaa	39840
ttgaaattca	gagccatgcc	tggtccctcg	gagcttctct	cctgggcagc	tgtgatcatt	39900
gcctctgctg	tgggtgtggg	ggtggaaatg	gattcctttc	atccttgcttg	ctacagggtga	39960
ctgtcacgtg	gagtcctttg	gagagaggga	cgtgttaatt	gatggatgtg	gctcccatgc	40020
tgagaaagct	cctgggcgta	cattgcctta	gagtttcatt	ggagctgcgt	tcttttatgg	40080
tgtctgctag	gcagaagtga	tgaagacttg	gaagaaaacc	cagaagggtt	tccacttaat	40140
ttggaaaatg	tgcttttccc	ctcctgtgtc	ttttgctaag	gtccagcctc	ctgcagcctc	40200
ccgctctctg	ggactctggc	tttgattctt	tattaggagt	ccccctgctc	ccccaaaaga	40260
tgggtgtctaa	attatcatcc	aattggccga	ggttttgttt	tctattaatt	gtttttatatt	40320
tttattgtgg	taaatttata	taacataaaa	tttgccattt	taattgtttt	gttattgttg	40380
tttttgagac	aggtgtctac	cccagtgcgc	aggtctggagt	gcagtgggtg	gatcatggct	40440
cactgcagcc	tcagcctcca	gggtctccagt	gatcctctca	cctcagcctc	tctagtagcc	40500
gggactacag	gcatacacta	ccacatctgg	ctgatttttt	gtattttttt	tttattgtag	40560
agaccgcgta	tgttgcccag	gctggctctca	actcctggac	tcaagccatc	ctcccacctc	40620
accctcccaa	agtgtctggga	ttacaggcat	gagccacaac	accagcccat	tttaattttt	40680
tttttttttt	ttgagatgga	gtctcactct	atcgcccagg	ctggagtgcg	gtggcgtggg	40740
atcaactcac	tgcaacctct	gcctcccagg	ttcaagcgac	tctcctgcct	cagcctcctc	40800
ccgagtagct	gggattacag	gtgcccattca	ctatgcctgg	ctaatttttt	tatttttttag	40860
cagagacggg	gtttccaccat	gttggccagg	ctggctctga	actcctaacc	tgggtgatccg	40920
ccgcctctgg	cctcccaaaa	tgctgagatt	acaggtgtga	gccaccgtgc	ccggcctttt	40980
tttgtttttg	agacagggtc	ttgcctgttc	accagactg	gagtgcattg	gtgggctctt	41040
ggctcactgc	agcctccgcc	tcccaggctc	aagttgtgca	cctccacacc	tggctaactg	41100
tatttttatgt	agagacagat	ttcaccatgt	tgcccaggct	gggcttgaaa	tggactcaag	41160
cagtccaccc	acctcagcct	cccaagtgct	tgagattaca	ggcgcgagcc	accgcaccca	41220
gcccatttta	cctattctgc	agttgacagt	tcagtggcat	tcagtcaagt	cacgaggtaa	41280
ccatcactgc	cattcatctc	cagactactt	caccttctcg	gcagatgtcc	gaaactgtcc	41340
gcattgaaca	cactcctcat	ctccctctga	cagccaccat	tctactttgt	atctctctct	41400
gccttctcta	ggtacctcat	gtaagtggaa	ttataccaat	atltgcccct	gtgtgactgg	41460
cttctttcat	gtgacatggt	gtctccaagg	ttcatctgtg	ttatagcctg	tgtcagaatt	41520
tccttcccta	aagcctgaat	aataaccctg	tgtaaaggct	gggcgcgggtg	gctcacaccc	41580
tctaatacca	gcattttggg	agtccgaggt	gggcagatca	cttgagggtca	ggagtttgag	41640
accagcctgg	ccaacatagt	gaaaccctgg	ctctaactaa	agtacaaaat	tagctgggtg	41700
tgggtggcgc	cacctgtaat	cccagttact	caggaggctg	aggcaggaga	atcgcttgta	41760
cccgggaggc	agaggttgca	atgaaccaag	attgtgcctc	tgcagtccag	cctgggtaac	41820
agagtggagc	ttcctgtctc	aaaaaaaaaa	aaaatcatcg	gatggatgga	cggaccactt	41880
cttgttattt	atccatccac	gggtgctagg	ttctctccac	ctttggttgt	cgtgaataag	41940
gccactatga	acatttccct	ccgtggtgaa	ggttttgtac	tagtgaggaa	aaggcgtggt	42000
tgtggtgttg	cataggattc	tggttaagaa	gtttgcacta	accataagta	tttgtactac	42060
attaaaatga	aagctcaggg	gcccggcgcg	gtggctcacg	cctgtaatcc	cagcactttg	42120
ggaggccagg	gcccggcgat	catgaggtca	ggagatcaag	accatcctgg	ccaacatggt	42180
gaaaccccg	ctctactaaa	aataccaaaa	aactagccag	gtgtggtggc	gggcacctgt	42240
agtcccagct	acttgggagg	ctgaggcagg	agaatggcgt	gaacccggga	ggcggagctt	42300
gcggtgagcc	gagatcgctt	cactgcactc	gagcctgggc	aacagagcaa	gactccgtct	42360
cacgcaaaac	tctgtctcac	gcaagactcc	gtctcaaaaa	aaaaaagagt	tcagggttta	42420
tgaaactggc	cagccgcgta	aagtttgctg	tgttgttttt	gtgcccggga	ggagtgtggc	42480
cagggtgtca	cgtcacacag	tacacgtttc	tcagatgggt	gttctccaga	ctgctgtccc	42540
aaagtctgtt	tttgcatctg	gttcccacag	acccaccctc	cacgggtgagc	ctgattttgg	42600
ccagggtagc	tggaaatctt	cttgtctttc	agcccggcag	ctgtaccagt	ccagggtcca	42660

10

20

30

cagctagtgg	cttttaggaa	ggaatttgtt	cagttggcct	tgacacatgg	ccccctaggg	42720
tccacagctc	tgtagtgatg	tggatgttgt	tatctacaaa	gacacatgat	ccttcgtgtc	42780
cagatgaaag	tgatgatgtc	tttgcagctg	cccagcaagg	ctgtgtgtgt	gtgtgtgtgt	42840
gtgtgtgtgt	gtgtgtgtgg	tgtgtgtgtg	gtgtgtgtgt	gtgtatgggg	gagggaggca	42900
ccctttccat	ctgggggtgt	gtgtgtgtgg	gggtgtgtgt	tgtgtgtgcg	cgtgtgtgtg	42960
gtgtgtgtgt	tgtgtgtgtg	tatgggggag	gcacccttcc	catctgggtc	caagagactg	43020
ggcctgggga	agacgcttct	ttttatctac	ttagagactt	tgttttattt	gtattttttt	43080
gagacagggg	ctcactctgt	caccacaggc	gggggtatgt	gatatgagca	tagctcactg	43140
cagcctcggc	ctcccaggct	gaagcgatcc	tcccacctca	gccttctgaa	tagctgggac	43200
tgtaggcgtg	cgtcaccata	ctgagctatt	gttttttttg	tttggttggg	ttaatttttt	43260
ttgatacaga	tggagtcttg	ctatgttgcc	cagactagtc	tcaaactcct	gaactcaagt	43320
gattctccca	cctcagtttc	ccgacattct	gggatcacag	gtgtgagcca	ctgctgtctc	43380
cctgttttat	taactgctga	aagacctaga	taaagaaagt	ctgaaaagac	ttactatcag	43440
agcaccatcc	taagatgatt	ccctctgact	caatggagag	ggaggggagc	ttttccttca	43500
ggcctgggtg	gcaggagccc	aggtgtctca	ggccccattt	gccccaggcc	aaatcactcg	43560
ggaacttggg	tgcagctgtc	tttcagggtg	acccaaagga	accagatccc	cgcaggcagt	43620
aggcttctgg	gctctcctct	cctcctacgt	cagctcagta	agagcccttc	gaagggatgc	43680
tgtgtcggag	gccccaaaag	cccaggctca	tccctgagat	gcacagggtg	ggctgggctt	43740
aggcagcgct	cgagcatctc	ctggacgggtg	accccagaga	gtgtggagac	ggagagtccct	43800
tgagagtcc	tgagagacgt	ggctgcccctg	ccttcccagg	aggggctctg	agtcattccc	43860
cacactcacc	tgcccctacc	caccctcacc	tggcccccag	cctcacctac	ccccacatct	43920
gtaccgatcc	ctttaccocg	accttcccta	cccacccctc	cctcccctgt	accttcacct	43980
ccccactcga	cccgcccctg	caccctcacc	tgtccccccac	cttcacctaa	ccccaccctt	44040
cacctgcctt	ccctcaccct	ggcctccttc	cgttggggaa	gggggtgtta	ggggcgggccc	44100
ccaaactgtc	tgtcctgggtg	ccttcgagag	aaaacagtac	gtgaggggccg	cagtcacaaa	44160
gctttagctcc	tggaaagggtg	aggagacagg	gatgtgttgg	gaaggggccc	atgggtcttg	44220
atcccttctc	gactgtcaat	ggggccttca	tgggagcgcc	agtctagtga	tgacacagctg	44280
ggtgcccggg	gggtgggtga	ggaggcctaa	agtcagaggg	ggcaagagct	ctccagagg	44340
ctgttgtcct	aatcgctctg	gcatactcag	gcgggcacgt	agttaggagc	tgattggaga	44400
ggagagaccc	ccacaccaat	actgggattt	gactttcagg	ctaaacttga	gaagtgtggc	44460
ctctgtgtgc	ctgccagagc	tctccagcca	gtgcccaggg	ctctccagcc	agtgccgggg	44520
ggtctccacc	agtgccgggg	ggtctccgcc	agtgccaggg	gtctccgcca	gtgcccaggg	44580
gtctccgcca	gtgctcagga	gtcttggttt	ctttgtctta	cagccctttg	ttttgacctc	44640
tctgagccaa	ggccaaaacc	cagacaggca	gccccacgac	ctcagcatcg	acatctacag	44700
cgggacactg	ttctggagct	gcgaggccac	caataccatc	aacgtccaca	ggctgagcgg	44760
ggaagccatg	gggggtgggtg	tgcgtgggga	cgcgcacaag	cccaggggcca	tcgtcgtcaa	44820
cgcggagcga	gggttagggag	ccaacgggtg	ggtgggggtg	ctgcccgtcc	aggcgtgccc	44880
gccgtgtctt	ctgccgaatg	ccagcctctc	acaggctggg	gagactttcc	accctgggga	44940
tccaatgggt	ggctttccag	ggtcccaaaa	gcaaacacag	gctctttcac	agccctccca	45000
ggaaaagcaga	aagccccaag	ggctggaagg	gaagggggag	ctctgctgag	aggttacaag	45060
gcagcgtctg	ccgacggggg	ttgcagttga	taggttttgt	atcatccttg	ttaaacttga	45120
accctgtgca	gaaatccctt	ccacggcatg	ggggctgcct	gttgactcgc	tcctgttcca	45180
ccacagggag	ctcctgggct	tcttccctcc	agaggccccc	gacgtctcca	cctgttggtc	45240
gtcagagctt	ctggttgggtg	ggaaggcacc	caggaccttg	aggtctccag	agagaaaagc	45300
cagggaaaga	gggagacoga	aacctatgtg	acatgaaact	caggctccaa	actgagcacg	45360
ggaacgtttg	gggacaggag	cgcgatggcc	ttcctcagat	agctgggggg	ctggcatgaa	45420
gacgggagct	acagccagca	caggtcctgg	gccgggagcc	cagagattga	gccctgactc	45480
tgtcacttac	tggccacgtg	accttggggc	ggtggcatag	cctcttgagg	actcagtttc	45540
ctcattggta	ggagtgcagg	ccacagtggg	gcggcctctg	cagcacacgg	ggggctcggg	45600
gggcggaagc	cccggtctca	taaggcggct	gtgcaggagc	cagccgagct	ggtctcccaa	45660

10

20

30

cagccagggc	tccgggggtcc	ttagcagctg	tgggggggct	gcacctgttt	cccatggctg	45720
ctgtcagaaa	ttaccagaag	ccaggtggct	gagagtaatg	gacacttggt	ctctcacagt	45780
tccctgagggc	tgaagcccga	gatcgagggt	tgggcagggc	cctgcgccct	ctgaaggctc	45840
tgagggaacc	tttgggcttc	tgggtggctcc	aggcaccctt	tgacttggtg	tcctgtcact	45900
ccagtctctc	tgtctggctg	cacatggcgt	ggcctcttct	gtaccattga	aggacacttc	45960
agttggattt	agggcctacc	ctcaccattt	gtggtcgtat	cttgatcctt	catgacattt	46020
gtaaagacct	tgcttccaaa	taagctcaca	ttctgaggtt	ctggggtgag	cggaattttg	46080
gagagcattg	ttcaactagt	atagaatgtg	acctgtcagc	ctcgggcagc	cctgagaggc	46140
aggggctttc	cacagcccag	ctgggtgccc	tgggtccctg	gctgtccgag	gagacgccat	46200
ccccacacct	gtccttcacc	cgccaccctc	cgcaggttac	ctgtacttca	ccaacatgca	46260
ggaccgggca	gccaagatcg	aacgcgcagc	cctggagcgc	accgagcgcg	aggtcctctt	46320
caccacccgc	ctcatccgcc	ctgtggccct	ggtggtggac	aacacactgg	gcaagctggt	46380
ctgggtggac	gcggaacctg	agcgcattga	gagctgtgac	ctgtcaggta	cgcgccccgg	46440
ggcctgccct	aaccgcagac	acccggcctt	cattgtcagt	aatggcagca	gctgccacat	46500
tgctcgagac	ctgcccgtgag	cccagtgcgc	cgcacggggc	tttgtgtgta	gcgtgttttg	46560
tcctcacact	gacagctgta	ggctgggggt	ctgagtgagc	cccacagggc	agaggcagaa	46620
aatgagtctc	agagaggggt	agcagctgac	ttggggcccc	acagcaggag	atggagcagg	46680
actgcagcct	agcctctgcc	cccagcacct	gcgcaagaag	ctgctctgct	ctggactgtg	46740
ttaggctgct	agggctggag	agaaatgaga	ggttggtgct	agagaggggg	cgcaggtccc	46800
catggtcttt	cctcttatga	tgaggtagat	gggtgaaggg	aggggccatg	cttgcaaggg	46860
ccagtgacct	aggcccgcgc	ttggaactga	tggccttcat	cccagaccca	gcccagggtg	46920
gaccagggct	ttccgagggc	ttgtcttggg	tggcctgctg	tccagggaact	ctgctgcagc	46980
tcccaccctc	tgccaaagca	tggaatcccc	caggctccct	ggcagtcctg	tcaacctctg	47040
tcctcccaag	ctgagtgtgg	ggcaagttct	ggaggtcagc	actgctcagg	ggggcccaag	47100
ggctgcttgc	agggggccaac	cgcctgacct	tggaggacgc	caacatcgtg	cagcctctgg	47160
gcctgacctc	ccttgccaag	catctctact	ggatcgaccg	ccagcagcag	atgatcgagc	47220
gtgtggagaa	gaccacccgg	gacaagcggg	ctgcgcatcc	gggcccgtgc	gcccacctca	47280
ctgcgcatcc	tgcagtggag	gaagtcagcc	tggaggagtt	ctgtacgtgg	gggctggcag	47340
tggggtgggc	aggggtggct	ctaaacccga	cccctggagg	aggctggagg	ccagtgcgaag	47400
atcctgtgtg	gcctcagcca	ggcgggtggc	tctgccagat	gccaactggt	gcccgcctgg	47460
gttcagcgac	atgtccgaat	gtcccagggc	ctctgagggt	gttttctttt	gccgcagaac	47520
aaatcaccac	gaacagcggt	ttaagacaac	accaactott	tttttttttt	tttttttttg	47580
gtcaggatct	tgtctgtgtg	cccaggctgg	ggtgccctgg	tgcacacaca	gttcactgca	47640
gcctcgacct	ctgggcttaa	ttaagtgaac	accttgccct	agcctcccag	gtagctggga	47700
ctacaggtgg	gcaccaccac	acctggctaa	tttttttttg	tagagacggg	gtttcccat	47760
gttgccagg	ctggctctgca	actcctgggc	acaagctatc	tgctgtgtgt	ggcctcccaa	47820
agtgtcagga	ttataggtgt	gagccactgg	cctgacaaca	cccacggatt	gtctctcagt	47880
tctgtaaggc	aaagtccagg	cacagcgtgg	ctcacctggg	ttctctgtct	agggctctcac	47940
ggggccagaa	tcaagtggtc	aggaacgctg	ggccctcagc	ggaggctctg	tggagaaatt	48000
agcttccttg	ctcactcagc	aggtagcagt	tgtgggatcg	aggttctggt	ttctctctgg	48060
ttattggctg	gggaccactc	tcagctccca	gaggccaccc	caggctcctg	cccctgggct	48120
ctctctgcct	cagcagtggt	ggctccctgc	gtcagtcctt	cccgcacctt	gagtcctctt	48180
gatttgcctc	ttaaaggccc	tgtgattcgg	ctcagccacc	tttagattag	gttagcctcc	48240
cctttgatag	actccaagtc	ggctgattaa	taaccttact	cacatctgca	gaatcccttc	48300
tgccacataa	ggtcagtacg	ccgtgctggg	gactgggggt	ggaaattacg	gggtcattta	48360
ggattctgcc	tgccactgcc	ttgctgtgtc	ccagggtctg	ggggaggggc	ctccacagct	48420
gggaccacag	tccttccctc	cctccatggt	aaocatctga	ggattacttg	agaccagcct	48480
gggcaacatg	gtgagaaccc	atccctacaa	aaaatacaaa	caaaaaggga	ccaggctggg	48540
cttgggtggc	catgcctata	atcccagcac	tttgggagac	caagggtggc	tgatcacttg	48600
aggttgggag	ttcgagacca	gcctgcccac	catagtgaac	tcccgctctt	actaaaaata	48660

10

20

30

caaaaattag	ctgggtgttg	tggcaggcgc	ctgtattccc	agctactggg	gaggctgagg	48720
tgggagaatt	acttgaacct	gggaggcgga	agttgcagtg	agccaaaatt	acgccactgc	48780
actccagcct	aggcaataga	gtgagactcc	gtctcaaaaa	aaaaaaagg	ccagggtggg	48840
tagtgacaaa	gagaccctat	cccaaaaaaa	ccgaacactg	aatccttgag	actgagtaag	48900
gacctgtga	aatttttctg	ggtggggcag	ggaacagagc	gtcttctgtc	atttcttcca	48960
cctgggtgtg	gtcagctctc	cctccaagct	gcctcctctt	cttctcattg	tcgggtgttt	49020
ggacacattt	ggttaactgg	atagaataac	gcgagttccc	agggacttgg	tccatttget	49080
attttatttt	attttatttt	attttatttt	attttatttt	ttatttattt	atttatttat	49140
tgagatggag	tttcgttttt	gtcgccagg	ctggagtga	gtggcgcgat	ctcggttcac	49200
tgcaacctct	gcctcccagg	ttcaagtgat	tctctacct	cagccttcca	agtaactggg	49260
attacaggca	cccaccacca	taccaggcta	atttttttgt	attttttagta	gagacgggtt	49320
ttcgccattt	tgcgccaggct	ggtcttcaac	tcctagcctc	aggtgatcca	cgcacctcgg	49380
cctcccaaag	tgctgggatt	acaggcatga	gccaccacgc	ctggcaccat	ttgctatttt	49440
aattcccatg	tgtatttagt	ttcccaggct	gctgtaacaa	atgaccacaa	actggatggc	49500
ttaaagcaac	agaaatggat	ttcccacatg	tgctggagac	cagaagcctg	cgaacaaact	49560
gttggggagg	ctgtgcttcc	tctgggggct	ccaggaggga	tctattttgt	ggcccttcca	49620
gtgtgtgtgg	tgcagcgtt	ccacacttgt	ggatgcgcgc	cctcaacctc	tgcctatctt	49680
catgtgtcca	tctcctttgt	gtctgcgtct	ttacctcttc	ttcttgtctg	tgttgctctt	49740
tataaggacg	tttgtcattg	ggttttagggc	ccaccacaaat	catccgagat	gacctcgtct	49800
tgagatcctt	aacctgcaaa	gacctttttt	ccaaaaaaag	gttatgctca	cagattctag	49860
gccttaagac	atgggtgtat	ctttctgggg	ggcactatcc	aacctcttat	acaatgaaag	49920
acgggaagag	ggccagggtg	ggtagttcac	gcctgtaatc	tcagcacttt	aggaagctga	49980
agcgggagga	tcacttgagc	ccaggagttt	acaagtagct	aggcaacatg	atgagacccc	50040
atttctacaa	aaagtataaa	aaaaaaaaaa	aaaaaaaaag	ccagggtgtg	tggctcacac	50100
ctgtaatccc	agcacttttg	gaggctgagg	caggcagatc	acgaggtcag	gagattgaga	50160
ccatcctggc	taaacagggtg	aaaccccgtc	tctactaaaa	atacaaaaaa	ttatggccgg	50220
gcgcagtggc	ttccgcctgt	aatcccagca	ctttgggagg	ccgagggtgg	tgaattacaa	50280
ggtcaagaga	tcgagaccat	cttggctaac	acggtgaaac	cccatcaaga	tcacaagggtc	50340
aagagatgga	gaccatcctg	gctaacacgg	tgaaccccg	tctctactaa	aaatacaaaa	50400
aattagccgg	gcactggtagc	gggcgcctgt	agtcccagct	gctcgggagg	ctgaggcagg	50460
agaatggcgt	gaacccggga	ggcggagcct	gcggtgagcc	gagatcgctc	catgccattg	50520
cactccagcc	tgggtgacag	agtgcagctc	cgtctcaaaa	aaaaaaaaaa	aaagaaaatt	50580
agccaggcac	agtggcagggt	gcctattgtc	ccagctactt	gggaggctaa	ggcaggagaa	50640
tggcatgaac	ccgggagggtg	gagtttgcag	tgagccgaga	tcattgccact	ggcctccagc	50700
ctgggcgata	gagcaagact	ctgtctcaaa	aaaaaaagcc	aggcatgggtg	gtgcatgcct	50760
gtagtcccag	ctactcaaga	ggctgaggca	ggagggttgt	tcgacccacg	gagatcaagg	50820
ctacagtgg	ccatgatcgc	accactgccc	tcagcctgg	gtgacagagt	gtgacctgt	50880
ctcaaagtaa	gtaaatagga	ggagagacaa	gtgggcagtt	cagactgatg	gtatgggcac	50940
agtagagact	ggtgcagaca	ggctggcctg	tgatgtcaag	caacttctgt	aactgtttcc	51000
ggcatccatt	tgtgtgtcaa	tttcogtgtc	agtaggaaga	ctctgtaggc	tgcacaggag	51060
aataagtggt	aggatcctcc	cagagaggcc	gggcctgcag	gaggggccagt	tctcatgagt	51120
tcttatttgg	ccctaccct	ccaggctgtg	gttctgaggt	gggagacaga	gcctgacctc	51180
tgtttgtctt	gttttgtctt	tgacgcagcc	cacctatgtg	cccgtagcaa	tgggtggctgc	51240
ttccacatct	gtattgccaa	gggtgatggg	acaccacggg	gctcatgccc	agtcacacctc	51300
gtgtcctgtc	agaacctgct	gacctgtgga	ggtaggtgtg	acctagggtgc	tcctttgggg	51360
tgatggacag	gtacctgatt	ctctgcctgc	taggtgtctg	cctggcatcc	ttttaaaatc	51420
acagtccctg	tggcatccag	tttccaaagc	tgattgtgtc	ttcctttgcc	ctcctttctt	51480
ttctactatg	tgcattcggg	gctatgaatt	ttcctctaag	tactgcgttt	cctgcacctc	51540
acaaattttg	ttacattttc	attttcagggt	agtttgaata	tttttacct	tctcctgaga	51600
tgacatcttt	ggctcatgtg	ttatttagaa	gtgttgctta	gtttctaaag	agttggggct	51660

10

20

30

tttccagctg	tctctctgca	actgatttct	aattttaatto	tactgtagtc	tgagagctta	51720
ttttatatga	tttctgttat	ttttaaagtgtg	ttgggtgtgg	tgtttttgtt	gttattgttt	51780
ttgtgtcttt	ttgttttgtt	ttgcttcgtt	tgtttttgtt	ttgagacagt	gtcttgtctt	51840
gtcactcagg	ctggagtgtca	atggcgcgat	ctcagctcac	cgcaacctct	gcctcccggg	51900
ttcaagtgtat	cctcttgccct	cagcctcctg	agttagctggg	attacaggtg	cacgccacca	51960
tacccagcta	atftttgtat	tttttagtaga	gacgggggtt	caccatgttg	gtcaggctgg	52020
tctcgaaactc	ctgacctcgt	gatccgcccc	cctcgggcctc	ccaaagtgtc	gggattatag	52080
gcgtgagcca	ctgtgcctgg	ccattaggtg	tgttttatca	cccagcatca	tgagtttat	52140
cttgggtgaat	gttctgtgta	ctcttgaaaa	gaatgtggat	tctgtctgtt	ttgggtggag	52200
tgttccagaa	acatcaatta	gatccagttg	gttaatagt	ctcatcaggt	tgctctctac	52260
cttccctcct	gactgcctgc	ttgagctgtc	agttattgac	aggggtgtgg	agtctccaac	52320
totaatgggtg	gatttgttta	tttctcctag	tagttctatc	ttttctctc	cttctaccct	52380
tgatcctctt	ctccccctag	ggcttctctg	tgttgggtgg	gggagagtgg	ggtagtgaag	52440
aacctggact	ttagggccaa	agaggccagg	gttcaaatcc	tggtctctgtc	acttcccagt	52500
tgagtgaacc	tggtctgtgc	ctgaatctct	gtgagcctcc	acttctctct	ctgtgaaatt	52560
gagagcaact	acctggcagg	ctgtcatggg	catcaagtaa	cagggcactc	caactggacc	52620
atgacacgtg	atgcacagga	atgccagctg	ctatgccatg	ggtgtggcag	tagtaataaa	52680
gtgaccatct	gtatcctcac	cacagtgaag	cctgtccagg	gctttctctc	ctatgcccc	52740
atgcctccag	gtggccttgg	atcctgttgg	ttctgtgtct	tgctcagoga	cctttctccc	52800
gtgggagttc	ctgggggttc	agcttcatcc	tacagacagc	agcacacact	ggctgtgcac	52860
cctttttttt	tttttttttt	tttttttttga	gatggagtct	cgcttttttc	gcgcaggctg	52920
aagtgcagtg	gtgtgatctt	ggctcactgc	aacctctacc	tctgggttc	aagtgtttt	52980
cctgcctcac	cctcccaagt	agctgggatt	acaggctccc	accaccagc	ccggctaatt	53040
tttgtatttt	cagtagagat	gggtgtttcac	catgttggcc	aggatggtct	tgaactcctg	53100
acctcaggtg	atccgcccac	ctcagcctcc	caaagtgcag	ggattacagg	cgtgagccac	53160
cacaccogga	gtgcoggttg	tttttagcag	tttgtcttgt	tctggagag	actggctcct	53220
gcccaggagc	tggggagata	gggcccgggg	gtgtgtcctc	acacctcgag	tttggccgta	53280
agcagagggg	acattttgtg	actgtccccc	tctgtagctt	cccagcagct	tttctccaag	53340
ttacagccca	aaagctcagg	tggatttgca	acccaacggt	gtctgtgcac	ctcccactga	53400
tgcccgaact	gccctggcca	agaaacgggg	ccgtcagaac	gtgcactaa	ctgcagcctt	53460
gggcctccat	gccagaggcc	atgcccttcc	atccaccacc	ccttggtcctg	ggcctggcc	53520
ctcctggctc	gggaactcca	ggccccctcc	tcaoggatcg	agagacgtgt	atttaaccga	53580
caggtgcttg	tcattctctt	gtggcctctt	ctccaggag	atcacagaag	gacagggcct	53640
cactgaggtc	tgggacatgg	acctttgat	agtggcagga	gccaggctgg	gcaagaggcg	53700
gccacagtca	cctcagcagt	gccatcacca	ccgccattca	gcccttccct	gagccgggcg	53760
cgcccctggc	tctggcccca	gtgtcccagt	tacagctcac	aggagcttgt	ggtgcccagc	53820
ggctgcttct	gattgagagt	cgaggtcgga	ggctttggga	ggctgagagg	ctgtcgggtt	53880
tcacaaactgc	tgaggagagc	ttgggtctca	tctcaggtct	gccccatgtc	gccctcaacc	53940
tccagccacc	ggctcctcgt	gtcccccatg	gccaggcacg	gcttgacagc	atctgtctgt	54000
ggctcctctc	agcctgtctg	ggctgacct	ggcacgtct	cctgtggctg	agcccagtg	54060
ggacagctgc	ttccttttat	tacctagaa	ctctcgtctt	tgatcaggcc	ccctccccta	54120
tgccacacag	tccctgtcac	tgggtgagc	ccagtagtca	tggggaaggc	ctgcgggttc	54180
caaacatcca	aaggcttgcg	tgcagcatga	cagcttgaaa	ccgatgtttt	ttaccttgat	54240
cagatttcag	cttggcgggg	gctttgtctc	gctttcagtg	aggcctgggc	cgatttccca	54300
gcatccctcc	ctgaggccag	cctctgtttc	ctgtgatttt	ctgcacaaag	tgggaggagg	54360
gagtcttagg	aaatgggggg	ccacctcgaa	acctaggcct	cctctggctt	ctctgtgcca	54420
gtgccccccac	gctttgtgtc	tgtgtcccca	gcccattggg	ctgtgttatt	ccctgagtgc	54480
tgccgcatgc	ccagcccgc	ctgaggacgt	ggagcccga	ggggcaggat	ggcctccatg	54540
gtcacacgta	ggaagtggcc	tccacctctc	gatgatcctc	tccccccctc	cctttcagcg	54600
ccttccccgg	gggtgtcatc	agcctcctg	cctgtgcttt	gtcccgtctt	ctgcaggcgc	54660

10

20

30

atgggacgtg	ctgacagggtc	ctctgcgggg	ttctgcctt	gctatgcgca	cgctggtcac	54720
cacagaggcc	tggcccttct	tctgtagcag	tcccacccc	gcaacagggtg	tggctgctga	54780
ccacctgctt	tctgcccctc	tgggtcctgag	gagggcgag	tgggcactca	ggcgtggctg	54840
agcagatgtg	tgttgccggg	aggaggaagg	actgtccag	tcagggtctga	atttcccacc	54900
cggagcattt	ctgctgtatt	tgggtgtagcg	cctgctgctt	aaagctctga	ttcccagttg	54960
gcaccttttc	ccttctgcat	tgaaaaacat	acggatgcat	gtcttcttgc	agtgaatgtg	55020
tattctccca	gcctctcttc	tgggttgggg	ctggagggtg	agcggcacac	aggagccgca	55080
gcgatggagg	atgtgcgggt	gcagcacccc	gtacagcagg	gatgccaaac	ccgcgctgag	55140
tccctctcaa	ccttctgctt	gaagcccagt	cacgccattg	cctgggtttt	gctgggcggg	55200
gctgcatgtg	atgttctcct	ctgtccctcc	cccagagccg	cccacctgct	ccccggacca	55260
gtttgcatgt	gccacagggg	agatcgactg	tatccccggg	gcctggcgct	gtgacggctt	55320
tcccgagtgc	gatgaccaga	gcgacgagga	gggctgcccc	gtgtgctccg	ccgccagtt	55380
ccctgcgcg	cggggtcagt	gtgtggacct	ggcctgcgc	tgcgacggcg	aggcagactg	55440
tcaggaccgc	tcagacgagg	tggactgtga	cggtagggcc	ctcccgtca	aggctctgcc	55500
aagaccctgg	ccttgccttc	cgggatacga	gcttggggct	gcctccggcc	tcacaggagt	55560
aggggctctg	aaaacctttg	cttgacggga	gattgccaa	tctgtctttt	aggcccaaca	55620
aggaaaactc	tgcagttcca	cccacctctg	cccaccaggt	agtgtggctt	gaaggcagac	55680
tgtgagggtc	tatctcacct	tcctgcatta	ggtcaggagt	ttcacagaaa	cctgaggcac	55740
attcaggggt	gggctgcaga	ggtccatggc	tcacacctg	gaaaatccgc	ccccaaaaga	55800
cagtgtctgc	tccactgacc	agtctgtggg	atagtgtcta	agcctgagtg	gtttctatca	55860
acatgtagaa	tcaggaggta	taaaagagatt	tgtcaggcca	tcctgggccc	tctctgacca	55920
gcaggatctt	cctttagatc	ttgacagtga	aacacatctc	ttctgtgccc	cctgtgagtt	55980
ttctttcatt	cattcattca	ttcattcatt	cattcattca	ttcgagacag	agtcttgcct	56040
tgtcacccag	gctggagtgc	cctgggtgtaa	tctcggtcca	ctgcaacctc	tgcctccagg	56100
gttcaatcga	ttctcctgcc	tcagcctccc	gagtagctgg	gatgacagg	ggcaccacc	56160
atgcctggct	aatttttcta	tttttagtag	agacagggtt	tcacctggtt	ggccaggctg	56220
gtctcgaact	cctgaacctc	ggtgatccgc	ccgcctcagc	ctcccaaagt	gctgggatta	56280
caggcatgag	ccaccgcgcc	cggcctgagt	ttctctttta	tgaaggacct	gcttggttgg	56340
ttgcctgcca	catgttgtca	gcaccatggg	cccaggactg	ctgaggagct	gttgatgccc	56400
tcgctctccc	agagccaccg	gctctgttag	ataatcaca	tgcagtctgg	ccactgtcct	56460
acgtcctcat	tcacaaagag	cagacatttc	gtagaagatg	agggcctggg	agtaacctcc	56520
ctgcatgttt	ttctataaag	gcatagtgg	taagtccctc	cagctcattg	accattggag	56580
aatttttatgg	aggctgtaga	ctaggggctg	gtaaaactaag	ggcccagggg	ccaaatccag	56640
cctgccacct	acttttgtaa	ataaagtttt	cctgggtgcac	agccatgccc	attcattcat	56700
ttgcacaatg	tctgtggctg	ctttcatgcc	aaaagcagga	gaactgagtg	gttatgctgg	56760
agacctacgg	ccttcaaagc	cccagacctc	acgtctggcc	cttgacagac	agagcttccc	56820
cagccctgct	gcgcacccg	gcccagcatg	tgtgtgtgt	gtgatttcag	cctgcaggag	56880
ccgtgggttag	gaattgtccc	tgtgttggtc	cattttgcat	tgcatagaag	gagcacctga	56940
ggccgggtag	attatgaagg	aaagaggctc	gtctggctca	tggttctgta	ggcagacca	57000
gtatggcacc	cgcactgctc	cagcttctag	tgaggctcca	ggaagctttg	actcatgggtg	57060
gaagtogaag	cgggagcagg	tgcatacat	ggtgagagag	ggagcaacgg	agagagagag	57120
agagagagag	agagcgccct	tccctcttgc	cctcaccttg	agaggagatg	ccaggctcct	57180
ttaagtaacc	agctcccatg	tgaactcaca	gtgagagccc	atttgctact	gcccagagg	57240
caccaggcat	ctgctcccat	gacccaaaca	ctgccacca	ggccctacct	ccaaccttgg	57300
ggctatattt	tattctgttc	tatgctatgc	tatgctatgc	catgccatgc	catgccatgc	57360
tattctctatt	ctattatttg	agacagaatc	tcgctctgtt	gcccaggetg	gagtgacgtg	57420
gcgatgatct	ggctcactgc	aacctccacc	tcccagggtc	aagcgattct	cctgcctcag	57480
cctcccgagt	agctgggatt	acaggcacac	accaccacac	ccgggttaatt	tttgtatttt	57540
caatagagat	ggggtttcac	catgttggcc	aggctggctc	caaaactcctg	gctcgaagtg	57600
atccacttac	ctcggcctcc	caaagtgcac	tgattacaga	tgtgagtcac	tgcgccagtc	57660

10

20

30

gagggtcaca	tttccgttga	gatttgagg	ggcagacgtt	ggagccatct	gagccccctc	57720
gtcccgctct	agcttctctc	cccgtgtgcc	ccgcggtgct	ggtggcaggc	ccttacgcgc	57780
gttctggctg	cacgctctgt	tccagaagct	ttcttccctg	cttggttacc	agaaaatcat	57840
cccatccatt	acaaggacag	ggtcccccta	tctcccatte	ccagggcagg	acaccggggg	57900
cagggcaggt	ggggaaactga	gcaagtctct	tgggggcagg	cgtggctatg	gctccctctg	57960
ggtgggcgtc	tggggagggg	tggaggcagc	cgtcagcgcc	ctggcttgct	cttctccctc	58020
ggccagagac	tgtggccttg	tgtgtctccc	gtgtgggctg	cctgcacctc	cagtgggttg	58080
tgtccctcc	cctccctccc	cctcaagctc	tgtgagcac	cactgccttc	cacagccccc	58140
actctcggga	ggcgaggctc	ctcgtggcca	ttctctgctc	tggcacccac	ccccccacca	58200
acctggtaga	gccttgggcg	gggtctgtta	ctccttgcat	ggcgtagacc	tccccacagt	58260
aggcacctga	cacatacctc	ctggggggca	ggcaggaggt	gcgttgaggt	ctcagccctg	58320
gcagtccctc	cctcgcgttg	cataggcctc	gccacaggg	catcgagggt	gggtggagac	58380
tgtactagac	cactccccgc	tggctctaga	aagggtccca	tctgtctgct	ctctgttttg	58440
agtcacagac	ttggttgctg	tgccttgcat	ggtgggctgg	ggggcaccct	ccagcctctc	58500
tgagtgcatt	gcctctcctt	gcagccatct	gcctgcccac	ccagttccgg	tgtgcgagcg	58560
accagtgtgt	cctcatcaaa	cagcagtgcg	actccttccc	cgaactgtatc	gacggctccg	58620
acgagctcat	gtgtggttag	ccagcttctg	gcacggggaa	ggggcgctcc	ggctgggttc	58680
ccccaggaac	gtggagttta	ggggaggaga	cgtgcctttc	cagcggggct	gggggctgtg	58740
tgggagactc	aggcggtctg	gaggctcctt	gcgggaggca	gggaagcctt	tcccagggca	58800
gcggccaggga	ggacagactg	tgagctgttg	gctcggcggc	tacagagtct	gcctcagtg	58860
gcggggctga	tgggttccag	gtgcctgcag	cacgcaccca	cccacgggac	cttgcctgagc	58920
agcgtctgtc	aggcagcaag	attaccogag	ggctgcagtg	gtcctgttcc	ctggcagctt	58980
actgtctggc	tgaggaggag	tgtatgtcac	atatgcacac	atgtcatgtg	cacacacatg	59040
tacatgacaa	catcccatat	gctcctcaaa	tagcatgacc	tgtacagtca	cggatatagg	59100
gcctagggga	taggaggcca	agacagtcag	ggaagacttt	ccagaggcag	tggctcctga	59160
aaggctgtct	gattcaggca	ggaaggagc	tgagttcaga	taggaagtag	caatgagtca	59220
ttgtgtctgg	ggacatggcc	actccttcgc	tgacagggga	cctgggctga	gagctcctct	59280
cttatggctg	cagtcgggag	agaagtctgt	tggggggaga	agggggcttc	ctcaagggac	59340
tccctgtgcc	ctttggcacc	ttcgtgccag	gtcaggcttg	aggcctgaag	gcagtgggtg	59400
ggggcaccac	gggtcgccct	ctctgtctgg	caagttccca	gtctgacggg	cctgtgcccgt	59460
ggggccccagc	tgtgggggcg	ctgttgatgc	gcagccaggc	ctcgccgcga	gagcccgac	59520
gcttccattc	cgtgaacttc	atcgacgccc	tcaggatcgc	tgggcccggc	ctgtgggaga	59580
gtgaatgtgg	cttttgccaa	agttgagtct	ggagcctgga	aacttcccta	tgggcagcct	59640
tgtatgtgga	gtggcccaag	gagccacccc	agccgacctt	gccccctccg	tggctgggtg	59700
gcggcaccag	gggctgcctg	gctttgtctg	ttcaccaaca	tcaccggggc	tggccagggc	59760
gcgctcactt	ctgccaccac	cgagggccct	gggcgaagga	gtgaatacca	ggctgccttg	59820
gcagggatgt	ggttagggct	gtggggagtc	ggacagcggc	gggggtcaga	ggaggaggag	59880
ggtgcaccgt	gcaggctgaa	gggcccagtt	accctgaggt	tggccaggct	ccccaggcct	59940
agcctcccag	ctccccactt	ttctccccac	cctccaccag	tggcaaaagcc	agccccctca	60000
gggcgcacgg	tgtctgcccc	caaggagggc	ccattccggt	gggggttaatg	ttggccacct	60060
ctttctgttt	gtctctggca	gaaatcacca	agccgccttc	agacgacagc	ccggccccaca	60120
gcagtgccat	cgggcccgtc	attggcatca	tcctctctct	cttcgtcatg	ggtgggtgtc	60180
atthtgggtg	ccagcgcgtg	gtgtgccagc	gctatgcggg	ggccaacggg	cccttcccgc	60240
acgagtatgt	cagcgggacc	ccgcacgtgc	ccctcaattt	catagccccg	ggcgggttccc	60300
agcatggccc	cttcacaggt	aaggagcctg	agatatggaa	tgtatctggag	gaggcaggag	60360
agtagtctgg	gcagcttttg	ggagtggagc	agggatgtgc	taccccaggc	cctcttgcac	60420
atgtggcaga	cattgctaata	cgatcacagc	attcagcctt	tcccactgag	cctgtgcttg	60480
gcatacagaat	ccttcaacac	agaggcctgc	atggctgtag	caaccacccc	tttggcactg	60540
taggtgtgga	gaaagctcct	tggacttgac	cttcatattc	tagtaggaca	tgtgtgtgtg	60600
tgtccacaaa	tcctcatgta	ccctagaaat	gaatgtgggg	gcggctgggc	tctctccaga	60660

10

20

30

gctgaaggaa	tcactctgta	ccatacagca	gctttgtctt	gagtgcagct	gggatttgtg	60720
gctgagcagt	tacaattcct	acgtggccca	ggcaccagga	acgcaggctg	tgttttaga	60780
tggtctggga	gcccgcaccg	agagctgcac	catgctgggt	tgatcacat	gggtgaccat	60840
ggtatgtcta	agaaggtgga	gtccctgtga	ggtctgcagg	tgccccaca	gctccaggcc	60900
accttgagga	ttgcctctgc	ctgcccagcc	ctgagttccc	tctcccctgt	cctgtcccac	60960
tgtcacccca	agccggcctc	attgggagcc	tggtggatgg	cagggtatag	atgtaacctg	61020
attctctctg	gggagcgggg	ttatctggct	tctcaagagc	tcttaggagc	ccacagtggg	61080
ggcaccatca	cagtgcagc	agccccaga	gaacgcggcc	ctgtctgttc	ctggcgtgct	61140
ctgtgctgcc	cgcctctggg	tccctgcccc	agtcgcaggc	cccttgagg	aggtaacctg	61200
tgtctccctg	ttcacagatg	agccccgggg	agctcactct	agtagtgagg	agagaggcct	61260
gcggtctcagg	gagcggggga	catttccaac	aggacacacc	gccttggtct	gagtctcgtg	61320
ggtagtggga	gcagaggaga	gcgcctatg	tctgtggggc	ggcttggtg	agcctggaag	61380
ccacctgacc	tccccgtcc	cttccctgcc	aggcatcgca	tgcggaagt	ccatgatgag	61440
ctccgtgagc	ctgatggggg	gcccggggcg	ggtgcccctc	tacgaccgga	accacgtcac	61500
aggggctctg	tccagcagct	cgccagcac	gaaggccacg	ctgtaccgc	cggtaggggg	61560
cggggccggg	gagggggggg	gcgggatggg	gctgtggggc	cctcccaccg	tcagtgtctg	61620
ccaccggagg	cttcccgggt	tcttgggggc	tgtagccacc	cctctgaggc	atgcttgctt	61680
tottcccttt	tcaaaccttt	ctgcttcctt	ctttaatgac	attgttgatt	gtggataatc	61740
tgaaaactac	acaaaaatat	aaagagccaa	aatctcacc	aatccacct	cctagagtgg	61800
ctgttgggct	cgtgcagcat	ccaggcgccc	gtctgtgttc	cgacggccc	agcccatcga	61860
tagccgcctg	caccaggcct	gtctgccttc	tgtgagcctc	cccacagggt	tccctccaca	61920
aacacctgt	tctcccacc	agggtctggt	gcttcttgga	aaacagctgg	atggttttgt	61980
gcatagcaga	caaacacagg	gtgattttcg	tgggtataat	actccctgga	gcttttggca	62040
gggtgagggg	ctggctccag	ctgagccacg	ccttgagtga	aatgactgtg	aggagaataa	62100
actgcogctg	ccctccagga	tcactggggc	tggctgggga	gaacccccgt	ttctgggagc	62160
acagtccag	gatgccagg	cgagcttggg	gcccagatgt	gaactcctga	gtgtaaacag	62220
cgggggctga	cttgacatgc	tttgtatgct	tttcatttgt	tctgcagct	gtatgcccct	62280
aagggtagtc	cagccccctt	ctgcttcctc	tggggcctcg	ccagttagcc	ccacctgtct	62340
ggggctgggt	cctcctgcc	ttctgggtat	ccctcacatc	tggggtcttg	tcttctgtt	62400
ttatttttct	tttttttttg	agacggagtt	tcacttttgt	tgcccaggct	tcagtgcaat	62460
ggtgtgatct	ctaggctcac	cgcaacctct	gcctcccagg	ttcaagcagt	tctcctgctt	62520
cagcctccct	agtagctggg	attacaggca	tgtgccacca	cgcccagcta	attttgtatt	62580
tttagtagag	atggggtttc	tccatgttgg	tcaggctgat	cttgaactcc	ctacctcagg	62640
tgatccgccc	accttggcct	cccaaaagtgc	tgggattaca	ggcgtgagcc	accgcacctg	62700
gcttttttct	tttcttttct	tttctttttt	ctgagacagg	gtctcgctct	gtcaccaggg	62760
ctggagtgc	atggtgtcat	catggctaac	tgagcctct	accttctagg	ctcaagcaat	62820
cctcccatct	cagccccctaa	gtagctagga	ctgcacgcct	gcaccccat	gcccagctaa	62880
tattttacatt	ttttgtagag	atgaagtctc	actatattgc	ccaggctggg	ctccaactcc	62940
tggactcgag	cgatcctcct	gcctcgccct	cccagggtgc	tgggattaca	ggcgtgagcc	63000
accgtgcctg	gcctggggta	ttgtcttctt	atggcactgc	actgtgggtg	gcctgggaa	63060
ggaagtagca	gaagagggtt	cttcttgggt	tcttgagacg	taactgagtg	ttctggaggc	63120
cccagggcct	ggctttgttt	agggacaaag	ggaactggta	accagaagcc	gagagtttaa	63180
acacccactg	cccttcttcc	ctgctcctgc	tgtgcacacc	cagcttaacc	agccaggagt	63240
gctaggaacc	caagcagggc	ccccgagcac	acagcaggca	gctcacgaat	tctcttttcc	63300
tggtctccct	tgggagctgg	gaggatctta	atcaggcaat	aagagatggc	actgagcagc	63360
cagctaattt	tttaaatcac	tttattgttt	aaccatattg	ctcaccact	taaaaaaggg	63420
tacagttcag	tgggttttag	tgtattcaca	gatgtgtgca	accctcacca	cagttaattt	63480
tagaacattt	tcttgcctcc	aaaagaaact	ctgcatgaag	ccagctgttt	ttaaatttag	63540
aaagttattt	tgcctccttt	aaatatatgt	tcatggtaca	aaattcaaaa	gatacagaag	63600
agtctgcagt	ccaaagagac	tccgccccca	tgacgccaag	caggactccc	tgggaggcat	63660

10

20

30

ggcctcctgc	agtgtgtttc	ttctatgtcc	ccccaggggt	catctgtaca	tatgcaagca	63720
tacaagagcg	tggactttgt	tttccaagcc	agaagataat	tgtagattta	tgtgcagttg	63780
tgagaaagag	cacagaccca	tttatcctct	gcctgggttc	ccccagtgct	gcctgccatc	63840
ttgcatgact	tccattccta	tcataagcaa	gacactgata	acgattcttt	caccttatte	63900
agattgacat	aagtgttttt	tgtttgttct	tgagacaaac	ttcctctgtc	acccagtggg	63960
agtgcagtg	cacaatcaca	gctcactgca	gcctcaaac	cctgggctca	agcgattctc	64020
ctgcctcagt	cccctcaagt	agctcagatg	gcaggtgtgc	accatcatgc	caggctaatt	64080
tttaaatctt	ttgtggaggt	gaggcctcac	taaatcttct	gggctagtct	tgaactcctg	64140
agctaaagt	atcctcctgc	ctcagcctcc	caaagtggta	ggattacagg	catgagccac	64200
tggcctcggg	ctgacatatg	tgttttcgta	agcccgaaag	atagcatctg	aagagtcaac	64260
attgagcctt	gccttttgct	gctaattgatg	tataaaagct	gctgttctga	gcatttcgga	64320
ggctccagc	tgccgtgtgc	accctgccta	gagctctacc	gtaacccatc	tccgggagga	64380
gggtgctattg	ttttcctcat	tttgcaacaa	ggaggctgaa	gaactgagca	tgaaccactg	64440
gcctgggtcg	ttcggttggt	aggcagtggg	gccaggccat	ccaactcaca	accaccttct	64500
actctgcttc	ccccgcaccc	tgaagtttgt	tctgttttga	ggacacagcc	gtcacattct	64560
ttggtgctga	acagcactcc	ttgtcaggtg	tggctggggc	cccactggag	ggcatcatgg	64620
tcctctctcc	tgtgcgggtt	gaaccttggc	tgtttcaacc	actcctgcca	agtggccctc	64680
tgaaggggac	agtccatctt	ttctcagcag	agggccacac	tggcaaaaag	gtccctggca	64740
ccctttctct	ccacctgtct	aatatagagt	aaaaatggta	tcatgttaag	atcttcattt	64800
atatttattt	tatcatgaat	gatgtaagca	tcatcttctg	tgtttaagaa	cctttgggct	64860
cagcgtgatg	gcttgagctt	gtaatctcag	cacttttaga	ggctgagatg	agcggatcac	64920
ttgagcccg	gagtttgaga	ccagcctggc	caactgggag	aaaccccgtc	tctagttaaa	64980
atttaaaaat	tagccgggta	tgggtgatccc	agctacttgg	gagtcctgaag	catgagaatt	65040
gcttgaacat	gggaggcgga	ggttgagctg	agccgagatc	gcgccattgc	actccagcct	65100
gggagacaga	gagagactct	gtctcacaaa	aaaaaaaaaa	aaagaaaaga	aaagaaatta	65160
tcaatctcct	cttttatggc	atatatatat	atatatatat	atatatatat	ttatttccct	65220
ttcttgggta	tgttcataaa	ggcctccctt	gctctgatca	taaaaaacaa	cttattttca	65280
cactctctct	cttttttttt	tgagacagag	ttttgctcct	gttgcccagg	ctggagtgcg	65340
gtggcgcaat	ctcagctcac	tgtaacctcc	gcctcccggg	ttggagtgat	tctcctgcct	65400
taccttcccg	agtagctggg	attataggca	tgcaccacca	tgcctgggta	attttgtact	65460
tttagtagag	acgggggttt	ctccatgttg	gtcaggctgg	tctcgaaact	gcgacctcag	65520
gtgatccacc	cacctcggtc	tcccaaagt	ctgggattac	agacgtgagc	cacctatgcc	65580
agcccaact	ctctttctta	acgtcctcct	cctttcgttt	tacgttcaca	tctttaattc	65640
ttctgggatg	taattagatt	tgtatgagca	ggtgggcatc	cagcttgttt	cctggctgat	65700
ggcttatggg	tggcgtgaat	tagtcggggg	ctatcaggag	gcagaaactc	tatgagaatt	65760
tgaacagaga	aagttccgtc	tacaggctta	ttaccaggga	ctggaatagc	agaaattgaa	65820
cagttagatg	tacagagaac	tctaagaatg	caggaatagg	ccaggcatgg	tggctcacac	65880
ctgtcatccc	agcactttgg	gagaccaagg	cgggtggatc	acctgaggtc	aggagttcga	65940
gaccagcctg	gccaacatag	tgaaacccca	tctctactaa	aaatacaaaa	aaattagctg	66000
ggtgtggtgg	cgcatgcctg	taatcccagc	ttctcgggag	tctgaggctg	gagaatcact	66060
tgaacctggg	aggcagaggt	tgtagttagc	cgagatcatg	ccattgtact	ccagcctggg	66120
caacaagagc	gagactcagt	caaaacaaca	acaacgcagg	aatagcagat	gagccgaggt	66180
ggggcctccc	cagcccccac	ccccaccctc	gcaccctggg	ccgagatcca	gtcctctttg	66240
aatagggcct	gggcgttggt	caogggacat	ctgagacatt	gocgaggcgc	tgcactggtg	66300
gatcttgcca	gaagtctgcc	cagtgcagat	ttgggcagaa	tctcaaaact	ccttgggatg	66360
taggagagaa	accaggcctg	gtcaagttca	tgggaagagg	tggaaacaga	ccccataggc	66420
tggggcttgg	gcagctgtag	gaagccctct	ctgctgcctc	cctgcctgct	ctctgctttg	66480
aagcatcttc	cccagtgcct	ccagtctcat	gccctctcaa	cgttgggggt	aaatcctgag	66540
gaataccacg	actggctctc	tgggccaag	aggaccctct	ccagaaagag	cagggccccg	66600
tgcggcttcc	taaagggcag	gggaagggcc	tggccactcc	ccagaggcta	ctcaccagcc	66660

10

20

30

atcaggatag	ccccaggaag	caggccttct	cgagccatt	ttattacttt	attttattat	66720
tttattttaat	ttttaaattta	ttttttgaga	cagagtctca	ctctgttgcc	caggctggag	66780
tgcagtgggtg	cgatctcaac	ccactgcagc	ctctgcctcc	agggttcaag	ggattctccc	66840
acctcagcct	cccaagtagc	tgggattaca	ggtgcccgc	accacacccg	gctaattttc	66900
atatttttag	tagagacgag	gtttcaccat	gttggccagg	ctggtctcga	actcctgacc	66960
tcaagtgatc	cgcccgccctc	ggcctcccaa	agtgtcaggt	caagccatt	ttaaagttga	67020
agaaactgag	gctgaggtaa	attccctccc	cagggatcct	gctgcagcca	gaaggtggta	67080
aaacaggact	tcacccggt	ctgtctggcg	tgaaaggcag	tgttcttgta	ccaccctagg	67140
gggcctgaga	gaactgagtc	cctcgggcat	aactgacagt	tctgttccca	ttattccgca	67200
ggggctcgga	tctggctgta	tgctttccag	gatggccttg	gagaccaca	taagccctac	67260
accctttggg	aagctgcatg	ttgggttggg	gtgccgtcag	tggcacttgt	ggaaggtgca	67320
gacctgtgtg	ggtgtgtggg	cccaggcccc	ctggtccctt	cctccctttg	tagggctggg	67380
tgtgtgctgc	ctggacctgg	ggggcacgtt	cacgtggtga	atttgtotat	ttactatccc	67440
cgctttgggg	ctggtgcccag	cacaggccct	tgtgaagggg	gtgcctttgt	ctggagtggg	67500
actgtggccc	ctccctcagc	gtggtgactt	ctgtgtcagg	gcttcagcag	ggacgcagag	67560
ccctgagtg	ttcggaaaca	gggcgtcatt	gcaggagtta	gactgtgtgt	gatggaggga	67620
ggaggggcag	gaggaaaggt	cagaaggaga	gttccctggga	aggtccctga	ggagcctggg	67680
gaggtgctaa	ctggtgtgga	ggacactcag	ggcctgtggg	gacatctcct	actgctgggg	67740
gccagccaca	aagggaactg	gccgaagtcc	tgtcccccgc	ttcacagccc	agcatctggg	67800
cacaaggcag	gtacttggaa	gggcgcgggc	acctggggcca	aaagtgcctg	ggttcccttt	67860
gcctttcaat	gagatgacct	tgggggcagg	tggctgctgc	ctccctcct	gtcccaggt	67920
tttgcccaact	ggccagagga	agggttcctg	ggaagcaggg	ggccagaaag	ccctctctgc	67980
aaggaaagcc	cgagggtgtg	gggaggaagg	aaggaaatgcc	caggctggcg	aggctctaag	68040
tcacccctggc	ttggtctctc	tcagatcctg	aaccgcgcgc	cctcccgggc	cacggacccc	68100
tcctctgtaca	acatggacat	gttctactct	tcaaacattc	cggccactgc	gagacgttac	68160
aggtaggaca	tccctgcag	ccctccatgg	ccattgggtt	cccggccagcc	cgtggtggag	68220
gggcctaatc	cccattgccac	tgatgagggg	aggatattctg	ggtgctagt	ggcaggtgcc	68280
gggcccagcc	ctgctccctc	ctgctctgcc	aaccacacta	ggctgcctcc	ccagacaagc	68340
tcagcgggca	ctgcatgttg	ggttcagaaa	tcagcagaac	tccacgttct	gagctgctct	68400
tcaagttgct	cctatggggg	ttactttttaa	gctgggaaat	ggctgtggcg	tcgagggggc	68460
gggggcttgg	gctccaaact	ctgactgtgt	gtttgagtcc	ggctgtggaa	acctagccat	68520
tgagatgccc	cctcttggtg	gctctgtcct	cttaggatgg	gacaagtctg	tgaaggctgc	68580
tgcagcacc	accgtagacc	cctaatcgtg	tgacgtcacc	aggatgggtcc	gggctgctca	68640
cttgccacag	tggcctgttt	gagcccggga	agccaaaggg	gctgctcagc	tggacaccag	68700
ccccccagc	tgcccatggt	ggggtcacag	gccccacctc	cctgggttggg	gaggggcaac	68760
tgagagtgtg	gagagggtggg	accaggtgt	gctggtctcc	gcaggggctg	gatcagagcc	68820
tgggatgggc	agggtgagcc	tccctgacct	taaccagtg	gtgtcaggca	acgtggccca	68880
ccgcgccagc	gcaccaggcc	ccaccccgc	aggatgaagg	gtgggatagg	ctgggcctgg	68940
gccaggacac	ctctggacca	cgcattccctc	attgcttggg	tccctggagc	agcagggcct	69000
cccgagtgtg	gtgccgctg	ccacctagt	gccatttcca	cgaactccca	ggcctggctg	69060
gggagccgga	actgcagcct	ccatttccac	ccactccgg	gtcggggccac	ctccctgatg	69120
cctcagttat	atatcaaaact	gtcacagtct	gtcccacagc	cttacagacc	actgtctcca	69180
gaatggtcac	atccacactg	ggcagcccag	tctcgctagt	tccctgctcc	acctcctgce	69240
tttgctcatg	ccggtcctgc	tctggggcca	ccgcggacac	atcttcccc	cgcgcgcgt	69300
ctgacctcac	agcagctggg	ccccaagagg	agtatcctgt	cctgctgcac	ttttctcaac	69360
accoggtgtt	ggctgcacct	tcccacccat	tgcaggcccc	tctgtgacag	gacgggggct	69420
cctaaacaca	ccacagttcc	gagtctgaac	tcacacagt	ggatgoggcg	tttctggggc	69480
acagttgggt	gcaggttagcc	tctgggagga	tgggaggtca	ggagccatct	tgcgagttag	69540
gttgcttgaa	ctcaggatgg	aagtgttccg	ggcccatgg	ttgctgtatt	agcctgttct	69600
cacgctgcta	ataaagacat	acccaagact	gggtaattgt	aaaggaaaga	ggtttaacgg	69660

10

20

30

actcacagtt ccacctgcct ggggtggcct cacaatcatg gtagaagaca aggaggagca 69720
 agtcacatct tacatggcct caggggaacag acagcatgag aaccaagcga aaggggtttc 69780
 cccttgtaaa accatcaagt ctatggagat ttattcacta ccacgagaac agtatggggg 69840
 gaaccacccc catgattcaa tcatctccca ctgggtccct cccacagcac gtgggaatta 69900
 tgggagtaca attcaagatg agatttgggt ggggacacag ccaaacccta tcggttgcca 69960
 acatttacag taacagtgtt aggtgaacag ttgtccagtc tcctgttttg tcggacactg 70020
 tttctagcac cttccaggca gaatctcatg tatccttcac tttcgaaatg ggtactattt 70080
 catccccact tttatcaatg agaaactaaa gctcgaagag gtcaagtaag ttcttgacca 70140
 aggtcagcta gcaggctcta gaggcctcgt tctccttaga ggcagccttg ccagggccca 70200
 ggcttggcag gctgcagggc aggtgcgggc atgcccattg tagaggtggg accattgagg 70260
 ctacagagag gtaagtgatg agccttggcg acacagcggg gtgggtccag agtccggcct 70320
 gcatcttctg gagctggcca gtggacaggc ctttcccggt caccagcccg gggctgctgt 70380
 gccaccagg gcggatgtgc ctaccgaatc ccactcctct gtgtgtgtcc ctttcaggcc 70440
 ctacatcatt cgaggaatgg cgcctccgac gacgcctgc agcaccgac tgtgtgacag 70500
 cgactacagc gccagccgct ggaaggccag caagtactac ctggatttga actcggactc 70560
 agacccctat ccacccccac ccacgccccca cagccagtag ctgtcggcgg aggacagctg 70620
 ccgcctctcg ccgcgcacag agaggagcta cttccatctc tcccgccccc ctcctgcccc 70680
 ctgcacggac tcatcctgac ctgcggcggg ccactctggc ttctctgtgc ccctgtaaat 70740
 agttttaaat atgaacaaag aaaaaaatat attttatgat ttaaaaaata aatataattg 70800
 ggattttaaa aacatgagaa atgtgaactg tgatgggggtg ggcagggctg ggagaacttt 70860
 gtacagtga gaaatattta taaacttaat tttgtaaaac agaactgcc a ttcttttgtg 70920
 cctgtgtgc atttgagttg tgtgtcccg tggagggaat gccgaccccc ggaccacccat 70980
 gagagtcctc ctgcaccccg gctcctcctc gtccggctcc tgcagggaag ggcctggggc 71040
 ttgggcagag gtggatatct cccctgggat gcacccctga gctgcaggcc gggccggcctt 71100
 tatgtgcgtg tggcctgtgc cgtcagaaag ggccttggc ttcatcacgc tgttgcctgt 71160
 cgtcttctc agattcttag tctttttttt tttttttttt ttttgagacg gactctttct 71220
 ctgtcatcca ggtgggagtg cagtggtaga atctcagctc actgcaagct ccgactccca 71280
 ggttcaagtg agtctcctgc ctacgcctcc cgagtagctg ggactacagg tgcgcgccac 71340
 cacaccgcc cagctaattt ttgtattttt agtagagatg gggtttcacc atgttgacca 71400
 ggatgatctc gatctcttga cctcgtgatc cgcacacctc ggctcccaaa agtgctggga 71460
 ttataggcat gagccactgt acccagctga ctcttagtca cttttaagaa ggggactgtg 71520
 ccttcatttt tcaactgggc ctgcagaata tatgcctggg ctctgggctc ttctgaacct 71580
 gtgttgctt ccatctgacc tctctgtgcc agcccaaggc tgcctgctctt cctgagggca 71640
 aggagcccca tgactgcgtg ttgactcgtt ggatggggct gctgagccca ctctgccaca 71700
 ccacgtgcc ctggcaggga gggaaatccct gggtcctcac aggaacagtc agcaagccac 71760
 acctgacgco tgcctgtggc ccatccctgc ggtgctggag aagacagaca aggcctggtc 71820
 actgcctctg caggggtccc agtccgtgga aggagacagt aatctaggca ttttcgggtg 71880
 ggaagctgag ctgttctcgt gtctgaagg ccaggcggga acagccgtct tcagagggaa 71940
 gggagaaaaa gcacatcgca tcagtggaga agggcctgac tccctcagc atggtggagg 72000
 gaggtcagaa aacagtcaag cttgagtatt ctatagtgtc acctaaata 72049

<210> 10
 <211> 8705
 <212> DNA
 <213> Homo sapiens

<400> 10
 ggactcaggg gcagcaggga ggtacacca tggtagtggt gcggaccata gggggtaattg 60
 agaggggtgaa tcgatgggaa ctgggggaca caatcgaagt ggttccagag tcgggctgta 120
 ctaattaaaag agacggggca gtggacaggc attttcagtt gactgccagc ggagtgttct 180

10

20

30

gccaacagg	gaggatatgc	gtacagaatc	atactcgatc	agcatgagtc	caattcagac	240
cgtacatcag	tggagatatg	ggtcccccga	tgactccgtg	gaacactgat	gtttgtgaca	300
ggggagtaga	gcaccagcca	tcagcaggcc	agtaaatacat	accggcctgc	gaaattggac	360
tcagaccggg	atccaccctg	accgacgtcc	caagccccca	ccccccaccc	cccaccatgg	420
gccgagatcc	agtcctcttt	gaatagggcc	tggccgtggt	tcacgggaca	tctgagacat	480
tgccgaggcg	ctgcattggt	ggaatcttgc	agaagtcttc	ccagtgcaga	tttgggcaga	540
atctcaaaact	gccttgggat	gtaggagaga	aaccaggcct	ggtcaagttc	atgggaagag	600
gtggaaacag	accccatagg	ctggggcttg	ggcagctgta	ggaagccctc	tctgctgcct	660
ccctgcctgc	tctctgcttt	gaagcatctt	cccagtgcc	cccagtctca	tgccctctca	720
acgttggggg	caaatacctga	ggaataccca	gactggctct	ctgggccaaa	gaggaccctc	780
tccagaaaga	gcagggccca	gtgcggcttc	ctaaagggca	ggggaagggc	ctggccactc	840
cccagaggct	actcaccagc	catcaggata	gccccaggaa	gcaggccttc	tcgagcccat	900
ttttattactt	tattttatta	ttttatttaa	ttttaaat	attttttgag	acagagtctc	960
actctgttgc	ccaggctgga	gtgcagtggg	gcgatctcaa	cccactgcag	cctctgcctc	1020
cagggttcaa	gggattctcc	cacctcagcc	tcccaagtag	ctgggattac	aggtgcccg	1080
caccacaccc	ggctaatttt	catattttta	gtagagatga	ggtttcacca	tgttgccag	1140
gtgggtctcg	aactcctgac	ctcaagtgat	cggcccgctc	cggcctccca	aagtgtcagg	1200
tcaagcccat	tttaaagtgg	aagaaactga	ggctgaggta	aattccctcc	ccagggatcc	1260
tgctgcagcc	agaagggtgg	aaaacaggac	ttcaccgggg	tctgtctggc	gtgaaaggca	1320
gtgttcttgt	accaccctag	ggggcctgag	agaactgagt	ccctcgggca	taactgacag	1380
ttctgttccc	attattccgc	aggggctcgg	atctggctgt	atgtcttcca	ggatggcctt	1440
ggagaccac	ataagcccta	caccttttgg	gaagctgcat	gttgggttgg	ggtgcctgca	1500
gtggcacttg	tggagagtgc	agacctgtgt	gggtgtgtgg	gcccaggggc	cctggctcct	1560
tcctcccttt	gtagggtctg	ttgtgtgtctg	cctggacctg	gggggcacgt	tcacgtggtg	1620
aattttgtcta	tttactatcc	cggcttttggg	gctgggtgcca	gcacaggccc	ttgtgaagg	1680
ggtgcctttg	tctggagtgg	gactgtggcc	cctccctcag	cgtgggtgact	tctgtgtcag	1740
ggcttcagca	gggacgcaga	gcccctgagt	gttcggaaca	agggcgatcat	tgacaggagt	1800
agactgtgtg	tgtaggaggg	aggaggggca	ggaggaaagg	tcagaaggag	agttcctggg	1860
aagggtccctg	aggagcctgg	tgagggtgta	actgggtgtg	aggacactca	gggcctgtgg	1920
ggacatctcc	tactgtctgg	ggccagccac	aaagggaact	ggccgaagtc	ctgtccccgc	1980
cttcacagcc	cagcatctgg	tcacaaggca	ggtacttgga	agggcgcggg	cacctggggc	2040
aaaagtgcct	gggttccctt	tgcttttcac	tgagatgacc	ttcggggcag	gtggctgctg	2100
cctccctccc	tgctccacag	ttttgccaa	tgccagagg	aagggtccct	gggaagcagg	2160
ggggccagaa	gcctctcttg	caaggaaagc	cggaggggtg	tgggagggaag	gaagggaatgc	2220
ccaggctggc	gaggtctctaa	gtcaccctgg	cctggctctc	ctcagatcct	gaaccgcgcg	2280
ccctccccgg	ccaaggaccc	ctccctgtac	aacatggaca	tgcttctactc	ttcaaacatt	2340
ccggccactg	cagagaccgt	caggtaggac	atccctcgca	gccctccatg	gccattgggt	2400
tcccgcacag	ccgtggtgga	ggggccta	cccatgcca	ctgatgagg	gaggtattct	2460
gggtgcta	gggcagggtg	cgggcccagc	cctgcctccc	tctgctctgc	caaccacact	2520
aggctgcctc	cccagacaag	ctcagcgggc	actgcatgtt	gggttcagaa	atcagcagaa	2580
ctccacgttc	tgagctgtctc	ttcaagttgc	tctatgggg	gttactttta	agctgggaaa	2640
tggtctgtgg	gtcgaggggc	cgggggcttg	ggctccagag	tctgactgtg	tgtttgagtc	2700
cggctgtgga	aacctagcca	ttgagatgcc	ccctcttggt	ggctctgtcc	tcttaggatg	2760
ggacaagtct	gtgaaggctg	ctgcagcacc	caccgtagac	ccctaactgt	gtgacgtcac	2820
caggatggtc	cgggctgtctc	acttgccaca	gtggcctgtt	tgagcccggg	aagccaacgg	2880
ggctgctcag	ctggacacca	gccccccgag	ctgcccattgt	tggggtcaca	ggccccacct	2940
ccctgggttg	ggagggggcaa	ctgagagtgt	ggagagggtg	gacccagggtg	tgctgggtctc	3000
cgcaggggct	ggatcagagc	ctgggatggg	cagggtgagc	ctcctgacct	ttaacccagt	3060
ggtgtcaggc	aacgtggccc	acccgccagc	cgcaccaggc	ccccccccc	caggtgaagg	3120
ggtgggatag	gctgggctcg	ggccaggaca	cctctggacc	acgcattcct	cattgcttgg	3180

10

20

30

gtccctggag	cagcagggcc	tcccagagtgt	ggtgcccgcct	gccacctagt	ggccatttcc	3240
acgaactccc	aggcctggct	ggggagcccg	aactgcagcc	tccatttcca	ccccactccg	3300
ggtcggggcca	cctccctgat	gcctcagtat	tatatcaaac	tgtcacagtc	tgtcccacag	3360
ccttacagac	cactgtctcc	agaatgggtca	catccacact	gggcagccca	gtctcgttag	3420
ttctctogtoc	cacctccctgc	ctttgctcat	gcccgctctg	ctctggggccc	accgcggaca	3480
catcttcccc	ccgcccgcgc	tctgacctca	cagcagctgg	gccccaaagag	gagtatcctg	3540
tcctgctgca	cttttctcaa	caccgggtgt	tggctgcacc	ttcccaccca	ttgcaggccc	3600
ctctgtgaca	ggacgggggc	tcctaaacac	accacagttc	cgagtctgaa	ctcacacagt	3660
gggatgcggc	gtttctgggc	cacagtttggg	tgcaggtagc	ctctggggagg	atggggaggtc	3720
aggagccatc	ttgcgagtc	ggttgcctga	actcaggatg	gaagtgttcc	gggcccattg	3780
gttgctgtat	tagcctgttc	tcacgctgct	aataaagaca	tacccaagac	tgggtaattg	3840
taaaggaaag	aggtttaacg	gactcacagt	tcacctgcc	tgggggtggcc	tcacaatcat	3900
ggtagaagac	aaggaggagc	aagtcacatc	ttacatggct	tcagggaaca	gacagcatga	3960
gaaccaagcg	aaaggggttt	ccccttgtaa	aacctcaag	tctagtgaga	tttattcact	4020
accacgagaa	cagtatgggg	ggaaccaccc	ccatgattca	atcatctccc	actgggtccc	4080
ttccacagca	cgtagggatt	atgggagtag	aattcaagat	gagatttggg	tggggacaca	4140
gccaaaccct	atcggttgcc	aacatttaca	gtaacagtgt	taggtgaaca	gttgtccagt	4200
ctctgttttt	gtcggacact	gtttctagca	ccttccaggc	agaatctcat	gtatccttca	4260
ctttcgaaat	gggtactatt	tcctccccc	ttttatcaat	gagaaactaa	agctcgaaga	4320
ggtaagtaa	gttctcggcc	aaggctcagct	agcaggctct	agaggccctcg	ttctccttag	4380
aggcagcctt	gccaggggccc	aggcttggca	ggctgcaggg	cagggtcggg	catgcccattg	4440
gtagaggtag	gaccttgag	gctcagagag	ggtaagtgt	gagccctggc	gacacagcgg	4500
gggtgggtcca	gagtcgggcc	tgcattcttct	ggagctggcc	agtggacagg	cctttcccg	4560
tcacagcccc	ggggctgctg	tgcaccaccag	ggcggatgtg	cctaccgaat	cccactcctc	4620
tgtgtgtgtc	cctttcaggc	cctacatcat	tcagggaatg	gcgcccccca	cgacgcctcg	4680
cagcaccgac	gtgtgtgaca	gcgactacag	cgccagccgc	tgggaaggcca	gcaagtacta	4740
cctggatttg	aactcgggact	cagaccccta	tcacccccca	cccacgcccc	acagccaagta	4800
cctgtcggcg	gaggacagct	gcccgccttc	gcccgcacc	gagaggagct	acttccatct	4860
cttcccgccc	cctcogtccc	cctgcacgga	ctcatcctga	cctcgcccg	gccactctgg	4920
cttctctgtg	ccccctgtaa	tagtttttaa	tatgaacaaa	gaaaaaata	tattttatga	4980
tttaaaaaat	aaatataatt	gggattttta	aaacatgaga	aatgtgaact	gtgatggggt	5040
gggcagggtc	gggagaactt	tgtacagtgg	agaaatattt	ataaacttaa	ttttgtaaaa	5100
cagaactgcc	attctttcgt	gcccgtgtgt	catttgagtt	gtgtgtcccc	gtggagggaa	5160
tgcgcagccc	cggacaccaca	tgagagtcc	cctgcacccg	ggcgtccctc	tgtccggctc	5220
ctgcagggaa	gggctggggc	cttgggcaga	ggtggatatt	tcacctggga	tgcattccctg	5280
agctgcaggc	cgggcctggc	ttatgtgcgt	gtggcctgtg	ccgtcagaaa	gggcccctggg	5340
cttcatcacg	ctgttgcgtg	tgccttccct	cagattctta	gtcttttttt	tttttttttt	5400
ttttttgaga	cggagtcttt	ctctgtcatc	caggctggag	tgcagtggta	caatctcagc	5460
tcactgcaag	ctccgactcc	caggttcaag	tgagtctcct	gcctcagcct	cccgagtagc	5520
tgggactaca	ggtgcgcgcc	accacacccg	cccagcta	ttttgtattt	ttagtagaga	5580
tgggggtttca	ccatgttggc	caggatgac	togatctott	gaacctcgtga	tcggcccacc	5640
tcggcctccc	aaagtgcctg	gatttataggc	atgagccact	gtaccagct	gactccttagt	5700
cacttttaag	aaggggactg	tgccttcatt	tttcaactggg	ccctgcagaa	tatatgcctg	5760
ggctctgggc	tcttctgaac	ctgtgttggc	ttccatctga	cctctctgtg	ccagcccaag	5820
gctgctgctc	ttctgagggg	caaggagccc	catgactgcg	tgttgactcg	ctggatgggg	5880
ctgctgagcc	cactctgcca	caccacgtgc	ccctggcagg	gagggaaatcc	ctgggtcctc	5940
acaggaacag	tcagcaagcc	acacctgacg	cctgctgtgg	gcccattccct	gcgggtgctgg	6000
agaagacaga	caaggcctgg	tcactgcctc	tgcagggttc	ccagtcctgtg	gaaggagaca	6060
gtaatctagg	cattttcggg	ggggaagctg	agctgttctc	gtgtcctgaa	ggccaggcgg	6120
gaacagccgt	cttcagaggg	aaggagagaa	atgcacatcg	catcagtggg	gaagggcctg	6180

10

20

30

acttcctca	gcatggtgga	gggaggtcag	aaaacagtca	agcttggtgc	tgggtgacag	6240
tgcattta	aatcaaaata	taggctgggt	acggtggctc	atgcctgtaa	tcccagcact	6300
ttgggaggct	gaggcagggtg	gatcacttga	ggccaggagt	ttgagaccgg	cctggccaac	6360
atggcaaac	ctcaactact	aaaatacaaa	aactagccgg	gcgtgggtgt	gcacgcctgt	6420
aatcccagct	acttgggagg	ctgagggcagg	agaattgctt	gaacctggga	ggcggagggt	6480
gcagttagcc	gagattgtgc	cactgcactc	cagcctgggc	aacagagcaa	gactctgtct	6540
caaaaaaaaa	aaaaaaaaaa	gcaatacaaa	atacaaatat	cactttcact	aaaagaaggg	6600
atggaagacc	caaaacaaac	agaaaacaac	aaaatggcag	gagtaagtcc	ccacttatca	6660
ataataacat	tgactgtaaa	taggctaagc	tctgcaatca	aaagagtggg	ccaggagcgg	6720
tggctcacgc	ctgtaattcc	aacgcttttg	gaggttgagg	cggatggatc	atttgatgtc	6780
acgagtttta	agaccagcct	ggccaacaag	gtgaaacccc	atctgtacta	aaaatacaaa	6840
aattagccag	ggcgtagtgg	cacgcacctg	taatccacgc	tacttgtgag	gctgaggcag	6900
gagaatcact	ggaggctggg	aagcggaggt	tgtgtgagc	caagatggag	ccactgcact	6960
cccacctggg	cgacagagtg	agatcctgtc	ttaagaaaaa	aaagagtggg	tgaatggatc	7020
aaaaaacaa	acccaaccat	ctcttgcata	caagaaacac	actttacct	taaaaacaca	7080
ctaggccagg	tgtggtggct	cacacctgta	atcccagccc	tttgggaggc	ctgactggca	7140
gatcacctga	ggccaggagt	ttcagaccag	cttgaccgac	atggcaaaac	cccatctctc	7200
ctaaaaatac	aaaaaaaaaa	aaaaaagaaa	aaggctggaa	gtagtgtgtg	gtgcctgtag	7260
ccccagctac	ttgggaggct	gaggcaggag	aattgcttga	atccgggaag	tggaggttgc	7320
agttagccag	gatcgtgcca	ctgcactcca	gcctgggtga	cagagcgaga	ccctgtcata	7380
aaaaaaaaaa	gaaaagaaaa	gaaaaacgag	aaaaacaaac	acaaaattag	tagaagaaaa	7440
gaaataataa	agatcagaac	aggccaggct	catgggcaca	gtggtcaca	tctacacctg	7500
tcaggagttt	tgaccagctc	ggccaacat	ggcaaaaccc	catctctcct	aaaaatatga	7560
aaaaaaaaaa	ataggctgga	tgtggtgatg	tgtgtgtgcc	tgtagcccca	gctacttggg	7620
aggctgaggt	gggagaatca	cttgagccca	ggaagtggag	gctgcagcga	gtcatgaatg	7680
caccctgcac	tctagctggg	taactggagt	gagattctgt	ctcaaaaaag	caaagaccag	7740
agcagaaata	aatgaaatgg	aaatgaagga	aacaatgcaa	aatgatacaa	aaagtttttt	7800
cgaaaagata	aacaaaatca	acaaaccttt	agccagatta	agaaaaaaag	agagaagacc	7860
caaataaata	aatccgaga	ttaaaaagga	gacattacca	ctgataccac	agaaattcaa	7920
aggatcatta	gaggcaacta	tgtgcaacta	tatgctaata	aactggaaaa	cctagaagaa	7980
ctgggttaaa	ttctagacac	atacaacct	tcaagattga	accatgaaga	aatccaaaac	8040
ctgaacaggc	cgggcacggt	ggcttacgcc	tgtaatccca	gcacttttga	aggcctgaga	8100
tcaggagttc	gagaccagcc	tggccaacat	ggtgaaaccc	catctctact	gaaaaaatat	8160
aaaaattagc	cgggcgtggt	ggcgggtgcc	tctaattgtc	gccactcggg	aggctgaggc	8220
aggaaaatca	cttgaacctg	ggaggcatag	gttgcagcga	gccgaggttg	caccactgca	8280
ctccagcctt	ggcgacagag	ccagactcca	tctcaaaaaa	attaaaataa	caaaaacctg	8340
aacagaccac	taacaagtaa	tgcgatgaaa	actgtaataa	aatgtttccc	aacaaagaaa	8400
gcccaggaac	aaatggcttc	actgctgaat	tttaccaaac	attttttttt	ttttgagacg	8460
gagtctcgct	ctgtcgccca	ggctggagtg	cagtgggtga	acctcggttc	gctggtaact	8520
tatgcctctc	aggctgcaag	tgattttctc	gcttcaggcc	ccccgagtgg	ctggaaaatta	8580
gatgggtact	gtcaaacaa	gcctggctaa	atttctatat	ttccttcaag	tagaagatgt	8640
gcttccaaca	aagggttggg	tacggctggc	ttctgaaaat	cttggatttc	aaggctcccc	8700
aaaag						8705

<210> 11
 <211> 66933
 <212> DNA
 <213> Homo sapiens

<400> 11

10

20

30

tataatcaag	cggttcogt	ccagtcoggt	gggaagattt	tcgatatgct	tcgtgatctg	60
ctcaagaacg	ttgatcttaa	agggttcgag	cctgatgtac	gtatttttgc	taccaaatac	120
agcaatagta	atgggtctca	gtccocgtgg	atggaggagc	aaattcggga	tgccctggga	180
agcatgggtc	taaaaaatgt	tgtacgtgaa	acggatgaag	ttggtaaagg	tcagatccgg	240
atgagaactg	tttttgaaca	ggccattgat	caacgctctt	caactgggtg	ctggagaaat	300
gctctttcta	tttgggaacc	tgtctgcaat	gaaattttcg	atcgtctgat	taaaccacgc	360
tgggagatta	gataatgaag	cgtgcgcctg	ttattccaaa	acatacgtc	aatactcaac	420
cggttgaaga	tacttcgtta	tcgacaccag	ctgccccgat	gggtggattcg	ttaattgogc	480
gcgtaggagt	aatggctcgc	ggtaatgcca	ttactttgcc	tgtatgtggg	cgggatgtga	540
agtttactct	tgaagtgtc	cgggggtgata	gtgttgagaa	gacctctcgg	gtatggtcag	600
gtaatgaacg	tgaccaggag	ctgcttactg	aggacgcact	ggatgatctc	atcccttctt	660
ttctactgac	tgggtcaacag	acaccggcgt	tcggtcgaag	agtatctggg	gtcatagaaa	720
ttgccgatgg	gagtcgcctg	cgtaaagctg	ctgcacttac	cgaagtgat	tatcgtgttc	780
tggttggcga	gctggatgat	gagcagatgg	ctgcattatc	cagattgggt	aacgattatc	840
gccaacaacg	tgttatgaa	cgtggtcagc	gttatgcaag	cggattgcag	aatgaatttg	900
ctggaaatat	ttctgcgctg	gctgatgcgg	aaaatatatt	acgtaagatt	attaccgcgt	960
gtatcaacac	cgcocaaattg	cctaaatcag	ttgttgctct	tttttctcac	cccgggtgac	1020
tatctgcccg	gtcaggtgat	gcacttcaaa	aagcctttac	agataaagag	gaattactta	1080
agcagcaggc	atctaaccctt	catgagcaga	aaaaagctgg	gggtgatattt	gaagctgaag	1140
aagtatcac	tcttttaact	tctgtgctta	aaacgtcatc	tgcatacaag	actagttaa	1200
gtcacgcaca	tcagtttgct	cctggagcga	cagtattgta	taagggcgat	aaaatgggtg	1260
ttaacctgga	agggtctcgt	gttccaactg	agtgtataga	gaaaattgag	gccattctta	1320
aggaaactga	aaagccagca	cctctgatcg	accacgtttt	agtctacgtt	tatctgtott	1380
tacttaatgt	cctttgttac	aggccagaaa	gcataactgg	cctgaatatt	ctctctgggc	1440
ccactgttcc	acttgtatcg	tcggctctgat	aatcagactg	ggaccacggg	cccactcgta	1500
tcgtcggctc	gattattagt	ctgggaccac	gggtccactc	gtatcgtcgg	tctgattatt	1560
agtctgggac	cacgggtccca	ctcgtatcgt	cggctctgata	atcagactgg	gaccacgggc	1620
ccactcgtat	cgtcgggtct	attattagtc	tgggaccatg	gtcccactcg	tatcgtcggg	1680
ctgattatta	gtctggggacc	acgggtccac	tcgtatcgtc	gggtctgatta	ttagtctgga	1740
accacgggtc	cactcgtatc	gtcgggtctga	ttattagctc	gggaccacgg	tcccactcgt	1800
atcgtcgggc	tgattattag	tctgggacca	cgtatccact	cgtgttgctg	gtctgattat	1860
cggctctggga	ccacgggtccc	acttgtattg	tcgatcagac	tatcagcgtg	agactacgat	1920
tccatcaatg	cctgtcaagg	gcaagtattg	acatgtcgtc	gtaacctgta	gaacggagta	1980
acctcgggtg	gcgggttgat	gcctgctgtg	gattgctgct	gtgtcctgct	tatccacaac	2040
atthttgcga	cggttatgtg	gacaaaatac	ctgggttacc	aggccgtgcc	ggcacgttaa	2100
cggggctgca	tccgatgcaa	gtgtgtcgtc	gtcgacgagc	tcgcgagctc	ggacatgagg	2160
ttgccccgta	ttcagtgctg	ctgattttgta	ttgtctgaag	ttgtttttac	gttaagttga	2220
tcgagatcaa	ttaatacgtat	acctgcgtca	taattgatta	tttgacgtgg	tttgatgggc	2280
tccacgcacg	ttgtgatatg	tagatgataa	tcattatcac	tttacgggtc	ctttccgggtg	2340
atccgacagg	ttacggggcg	gcgacctcgc	gggttttcgc	tatttatgaa	aattttccgg	2400
tttaaggcgt	ttccgttctt	cttcgtcata	acttaatggt	tttattttaa	ataccctctg	2460
aaaagaaagg	aaacgacagg	tgctgaaagc	gagctttttg	gcctctgtcg	tttcctttct	2520
ctgtttttgt	ccgtggaatg	aacaatggaa	gtccgagctc	atcgctaata	acttcgtata	2580
gcatacatta	tacgaagtta	tattcgatgc	ggccgcgaagg	ggttcgcgtc	agcgggtggt	2640
ggcgggtgtc	ggggctggct	taactatgct	gcacagagc	agattgtact	gagagtgcac	2700
catatgcggg	gtgaaatacc	gcacagatgc	gtaaggagaa	aataccgcac	caggcggccat	2760
tcgccattca	ggctgcgcaa	ctgttgggaa	gggcgatcgg	tcggggcctc	ttcgtattta	2820
cgcagctgg	cgaagggggg	atgtgctgca	aggcgaattaa	gttgggtaac	gccagggttt	2880
tcccagtcac	gacgttgtaa	aacgacggcc	agtgaattgt	aatacgaact	actatagggc	2940
gaattcagag	tcggtaacccg	gggacccctc	agagtcgacc	tcgaggcacg	caagcttctc	3000

10

20

30

ttgtgcccgt	tgtacgctgt	caggtcacac	tggtgagtta	ggcagggcac	agatgccag	3060
agcagaggga	actttccttg	gggattcaac	acgtgcaagt	cttaggggct	ggcaaatcct	3120
gccctcagct	agagaggggg	cttttatattg	agaccagaat	cacctgagca	tcctcctgtc	3180
cccagctgtg	tccagcctgt	ctgcagggac	atcctgagag	gaccaggctc	ccccctcatc	3240
caactgccta	agtgccactc	tgaacctgtg	ccacctgtgc	cgtggagggg	cgtgacctca	3300
agctgctcag	ccagcagcag	gcttggccct	ggggggcagc	agagacccag	gtggctgtgg	3360
gggtgggtgct	togtggcggtg	gttctgaaac	ttcgttggaa	gtgtgtggac	agtgccttgc	3420
ctgttctctg	tgggacctta	tttagaaacg	aggtctgagt	tactgggggt	catcactgtg	3480
ttctgatggc	ccagctgtgt	ggaggccgcg	gtgcagcccc	atccaaggag	ccagggccct	3540
gggtctagcc	gtgaccagaa	tgcattgccc	ggagggtgtt	ctcatctcgc	acctgtgttg	3600
ccgtgtgtgt	caagtggctg	tgaacctctg	tgttagctct	tggtgttcct	gaaagtgtcc	3660
cgggtctca	ggcctcagaa	ccagggtttc	ccctcatctc	ggtggcctgg	gagcatctgg	3720
gcagttgagc	aaagaggcg	attcacttga	aggatgtgtc	tggccctgcc	taggagcccc	3780
ccggcacggg	gctggggcct	gaagctgccc	tgggtgggtg	gagaggaggg	agcgatgaag	3840
tggcgctcag	ctgggcagga	aggggtgagc	cctgcaaggt	gggcatgctg	gggacgctga	3900
gcagcatggc	cagcagctgg	gtctgcagcc	tggtaccggg	cgggacttgt	ggttggggct	3960
ggtttggggc	caggagaggg	gctggcagga	gacaaggggg	actgtgaggg	agctcccacc	4020
cagcagctga	agcccaatgg	cctggctgtg	tggctctcag	ctgcgtgcat	aacctctcag	4080
tgtctcagtt	ctctcatttg	taaaatgagg	aaacaaacag	tgccagcctc	ccagagggtg	4140
catgaggatg	aacgagtgc	catgtagcat	gggctgggtg	cgtgtcacct	aacatcacca	4200
gcctttgcaa	ggagagccct	gggggcctgg	ctgagtattt	cccttgcccg	gcccacccca	4260
ggcctagact	tgtgcctgct	gcaggccctt	gacccctgac	cccattgcac	ctgtctccac	4320
aggagccag	gggctgctgc	tgtggcccg	gcggacggag	ctacggagga	tctcgctgga	4380
cacgccggac	ttcaccgaca	togtgcgtga	ggtggacgac	atccggcagc	ccattgccat	4440
cgactacgac	ccgctagagg	gctatgtcta	ctggacagat	gacgaggtgc	gggccatccg	4500
caggcgctac	ctggacgggt	ctggggcgca	gacgctggtc	aacaccgaga	tcaacgaccc	4560
cgatggcatc	gcggtcgact	gggtggcccg	aaacctctac	tggaccgaca	cgggcacgga	4620
ccgcacogag	gtgacggccc	tcaacggcac	ctcccgaag	atcctgggtg	cggaggacct	4680
ggacagagccc	cgagccatcg	cactgcaccc	cgtgatgggg	taagacgggc	gggggctggg	4740
gcctggagcc	agggccaggc	caagcacagg	cgagaggagg	attgacctgg	acctgtcatt	4800
ctgggacact	gtcttgcatc	agaaccggga	ggagggtctg	ttaaaacacc	ggcagctggg	4860
ccccaccccc	agagcggtga	ttcaggagct	ccaggggcgg	gctgaagact	tgggtttcta	4920
acaagcacc	cagtgggtccg	gtgctgctgc	tgggtccatg	cgtagaaagc	cctggagacc	4980
tggaggggagc	cccttgttcc	cctggcttca	gtttcctcat	ctgtagaatg	gaacggtcca	5040
tctgggtgat	ttccaggatg	acagtagtga	cagtaagggc	agcctctgtg	acactgacca	5100
cagtacaggc	caggcctctt	tttttctttt	tttttttttg	agatggagtc	tcactctgtc	5160
gcccaggctg	gagtcgagtg	gtgtgatctc	agctcactac	aacctctgcc	tcctgggctc	5220
aagtgtattct	cctgcctcag	cctcctgagt	agctgggatt	acagggtgct	gccactgtgc	5280
ttggctaattg	tttgtatttt	tggtagagat	ggggtttcac	cgtcttggcc	aggctggctg	5340
caaactcctg	acctcaggtg	atccacctgc	ctcagcctcc	caaagtgtctg	ggattacagg	5400
catgagccac	caegcccggt	caggccaggc	ctcttttgaa	cactttgcac	accatgggtc	5460
ttttcatcca	ggggggtagg	tacagtgtga	cagttgagga	cactgaagcc	cagagaggct	5520
cagggacttg	cccagggtca	cacagcagga	tgtggcaggt	gtggggctgg	gcctggcagc	5580
gtggctccag	ctttccagca	tagaaatctg	tgaagcaga	tagtttgtcg	gtcggtaggg	5640
gagactttct	gagacccgcc	ccagcggctc	agagggtagt	agccaggggc	cttcctgggg	5700
gctcataacc	cagaacactg	aatgggaaaa	ccctgatgga	ggaggcgcag	tggagctgtg	5760
gggtccgatg	ggaagtcccc	gaggagctgg	gaggctcagta	gcgggtgctgc	cctctgtgga	5820
gcacttagtg	ggcaccaggt	gtgtttccag	gttcatggcc	ctgggacctg	aagctcagaa	5880
ggtgaagtaa	cttgcccagg	gcacccgtcg	ggcagcggcg	ggcagaggat	ttgtgggctg	5940
tggagcctgt	gctcgtggcc	cagccctggg	ggttgtgagt	gtgctggccg	gggagctttt	6000

10

20

30

cctgcaagtg	gactgggtgtc	taggagccag	catgtcaggc	agcaggcagc	gggagtgacg	6060
caggcagcgg	gagcacagca	ggcagagggc	ggggctcgag	cagccatccg	tggaccctgg	6120
ggcaccggagg	catgtgggag	agggctgtct	catggcagtg	gctgaagggc	tgggttgtgc	6180
cccaggaggag	gtgatgagg	gtaagaagtg	gggtccccag	gggctttagc	aagaggaggc	6240
ccaggaactg	gttgccagct	acagtgaagg	gaacacggcc	ctgaggtcag	gagcttggtc	6300
aagtcaactgt	ctacatgggc	ctcgggtgtcc	tcactctgtga	aaaagggaagg	gatgggggaag	6360
ctgactccaa	ggccccctcct	agccctgggt	tcactgagtct	gaggatccca	gggacatggg	6420
cttggcagtc	tgacctgtga	ggctgtgggg	tcaggggagg	ggcaccgagc	tggaaagcggg	6480
aggcagaggg	gctggccggc	tgggtcagac	acagctgaag	cagaggctgt	gacttggggc	6540
ctcagaacct	tcacccctga	gctgccaccc	caggatctgg	gttccctcct	tggggggccc	6600
caggggaacaa	gtcacctgtc	ctttgcatag	gggagccctt	cagctatgtg	cagaagggttc	6660
tgctctggcc	cttccctcct	ctagggtgctc	agctcctcca	gccactagt	cagatgtgag	6720
gctgccccag	acctggggca	gggtcatttc	tgtccactga	cctttgggat	gggagatgag	6780
ctcttgggcc	ctgagagtcc	aagggtctgt	gtggtgaaac	ccgcacaggg	tggaaagtggg	6840
catccctgtc	ccaggggagc	ccccagggac	tctggtcact	gggcttgccg	ctggcatgct	6900
cagtccctca	gcacttactg	acaccagcat	ctactgacac	caacattttac	aaacaccgac	6960
attgaccagc	accgacattt	accgacactg	acattttacca	acactgttta	ccaacactga	7020
catctactga	cactggcctc	taccaacact	gacattttacc	gacactgaca	tttaccacaa	7080
ctattttacca	acactgacat	ctactgacat	tggcatctac	caacaccaac	atttaccgac	7140
accaacattt	accaacactg	aaattttaccg	acaccgacat	ttaccgacac	cgtttaccac	7200
caccgacgtt	taccgacacc	gacattttacc	gacactgata	tttaccacaa	ctgacatcta	7260
ctgacgttgg	catctactga	caccgatgcc	agcatctacc	aacaccgaca	tttaccacaa	7320
ctgacattta	ctgacactga	tatctactga	cactggcctc	tactgacacc	aacattttacc	7380
aacaccagca	tctaccaaca	ccgacattta	ccaacaccag	cattttaccac	caccgatggt	7440
taccaacgcc	gacgtttacc	gacgccagca	tctaccaaca	ctgacattta	ccgacaccga	7500
cattttaccga	cactgacatt	tactgacact	gacatctact	gatactggca	tctaccgaca	7560
ctgatattta	ccaacgccag	catctactga	cactgatggt	taccaacacc	gacattttacg	7620
agcaaccgaca	tttactgaca	ccaattattta	ctgacatcaa	cattttagcca	tgtgatgggg	7680
gccgggttgg	gggcaggcct	tgtctttggc	actggggatg	ctgcagagac	cagacagact	7740
catgggggtca	tggactttctg	cttctttctcc	agcctcatgt	actggacaga	ctgggggagag	7800
aaccctaaaa	tcgagtgtgc	caacttggat	gggcaggagc	ggcgtgtgct	ggtcaatgcc	7860
tcctctgggt	ggcccaacgg	cctggccctg	gacctgcagg	aggggaagct	ctactgggga	7920
gacgcccaaga	cagacaagat	cgagggtgag	ctcctgtgga	catgtttgat	ccaggaggcc	7980
aggcccgagc	accccttgca	gccagatgta	cgtattggcg	aggcacogac	gggtgcctgt	8040
gctctgctat	ttggccacat	ggaatgcttg	agaaaatagt	tacaatactt	tctgacaaaa	8100
acgccttgag	agggtagcgc	tatacaacgt	cctgtggtta	cgtaagatgt	tatcattcgg	8160
ccagggtgcct	gtagacacag	ctacttgag	actgaggtgg	gaggatcgct	ggagtccaag	8220
agtttgaggc	cagcccgggc	aaaggggaca	caggaaatcct	ctgcactgct	tttgccactt	8280
actgtgagat	ttaaattatt	tcacaataca	aaattaagac	aaaaagttaa	tcacatatcc	8340
actgccctgc	ttaagacaga	aaacatgggt	gttggtgaag	ccagaggcag	ctgctggcct	8400
gagtttggtg	attgggttct	aagcagttga	aggcagtttt	gtttttccat	agatgtctgt	8460
tctccctttg	ctgggtgcag	cctcgccctg	ctgctgtggt	cggttttcag	tggcctcgtc	8520
ccgtggacgc	agcctcgccc	tgcctgtgtg	gtcgggtttc	agtggcctcg	tcccgtggac	8580
gcagcctcgc	cctgctgctg	tggctggggt	tcagtggcct	cgtcccgtgg	acgcagcctc	8640
gccctgcgcg	tgtggtcggg	tttcagtggc	ctcgtcccgt	ggacgcagcc	tcgcctcgcc	8700
gctgtgggtg	gggttcagtg	gcctcgctcc	atgggcgtgc	tttggcagct	ttttgctcac	8760
ctgtggagcc	tctcttgagc	ttttttgttt	gttggtttgt	tttggtttgat	tttggtttgat	8820
tgtttggttt	tggttgctgt	gttggttgccc	aggctggagt	gcagtggcgc	gatctcagct	8880
cactgaaacc	tctgctcctc	tgggttcactg	ccatttctct	gcctcagcct	cccacatagc	8940
tgggattaca	agtgcgccgc	accacgcctg	gctaaatttt	gtatttttag	tagacagggg	9000

10

20

30

gtttcaccat	glttggtcagg	ctgggtctgga	actcctgggtc	tcacatgata	cacctgcctc	9060
ggcctcccaa	agtgttggtg	ttacaggcgt	gagccacccg	gcccagcctc	tgttgagcat	9120
atthttgaggt	tctcttggtg	ccagtgatat	gtacatgtgt	ccccatcgca	ccatcgtaac	9180
ccattgaggt	gacattgggt	cctctcctcg	gggtggatgc	ctccctctgt	ttccagcaac	9240
ttctgaagga	tttctcagag	ctgcacaggt	ccattgtgac	gtcaccatcg	gggtcacctt	9300
tgctctcctc	agggctccca	ggggaggccc	gaatcaggca	gcttgacagg	cagggcagga	9360
tgagagaacac	gagtggtgtg	ctgtgttgca	ggatttcaga	ccctgcttct	gagcgggagg	9420
agtctcagca	ccttcagggt	ggggaaccca	gggatggggg	aggctgagtg	gacgccttcc	9480
ccacgaaaac	cctaggagct	gcagggtgtg	ccatttcctg	ctggagctcc	ttgtaaatgt	9540
tttggttttg	gcaaggccca	tggttgccgg	ccgctgagga	tgatttgctt	tcacgcctcc	9600
ccgctacccg	tggtgagcag	tcagggactc	gcgtgtctgt	ggcacaccag	gcctgtgaca	9660
ggcgttggtc	catgtactgt	ctcagcagtg	gttttcttga	gacagggtct	cgctcgctca	9720
cccaggcgag	agtgcagtg	cgcaatcacg	gctcgtgtga	gcctcaatct	ccctgggctc	9780
aggtgatcct	cctgcctcac	cctctgagta	gctgggacta	cagacacata	ccaccacacc	9840
cagctagttt	ttgtgtatth	tttggtgggg	gagatggggg	ttcgtgtgtg	tgcccaagct	9900
gactcctaac	tcttgaggca	caagcgatcc	acctgcctcg	gcctcccaaa	gtgctgggat	9960
gacaggcatc	agccgtcaca	cgagctcaa	tgattttatt	gtggtaaaa	aaacatagca	10020
caaaattgat	gatttttaacc	atthtaaaagt	gaacagttca	ggctgggctg	ggtggcttat	10080
gcttgtaatc	ccagtacttt	gagaggctga	ggtgggcaga	tcacctgagg	tcaggagttt	10140
gagaccagcc	tggtccaaat	gatgaaatcc	agtctctact	aaaaatacaa	aaattagccg	10200
ggcatgggtg	cagggtgcctg	taatcccagc	tactcgggag	gctgaggcag	gagaatcgct	10260
tgagcccggg	tgagtgaggt	tgagatcatg	ccactgcact	ccaatctgtg	10320	
tgacagagca	agactctgtc	ttgaaaaata	aataaataaa	aaaaatttta	aaaagtgaac	10380
aattcagggc	atthtagtatg	aggacaatgt	ggtgcaggta	tctctgctac	tatctacttc	10440
tagaacactt	tctcttgccc	tgaaggaaac	cccatgcccc	ccggcactca	cgcccatctt	10500
ccctctcttc	ccagcctctg	tcaaccacta	atctactttc	tgctctctgg	ggttcaacttc	10560
ttctggagct	tttggtgtgac	tggaatcctg	caatatgtgg	tccttgctgt	tggtctcttt	10620
ccatagcatt	gtgtttttcca	gattcaacca	cacattgtcg	cacgttatca	gaatctcatt	10680
cctgactggg	tgagtgagg	taggcctgtg	atcctaacat	tctgggaggc	caaggcggga	10740
cgatcacttg	aggcaggagt	ttgagaccag	cctggccagc	ctagcaagac	cccagctacc	10800
aaaaaattht	aaaagttaac	tgaacgtggt	ggtggtgggc	acttggtggt	cccagctacc	10860
tggtgaggct	aggtgggagg	atcgcttaag	cccaggagg	caaggctgca	gtgagctatg	10920
atcgaccacc	tgactctcag	cctggacaac	agagcaagac	cctgtctgaa	aaaaaaaaaa	10980
aaaaaaaaag	ttcctttctt	tttggtgctg	gatgacatcc	cattgtatgg	ccacagcaca	11040
ttttgtttgt	ctgttttatcg	ggtggtgggc	agtgttttcc	accttttgct	tcctgtgaat	11100
aatgctgctg	tgaacatttg	aattcaagtt	tttggttgaa	cacctgttgt	gaattattht	11160
gatatatgtg	taggggtagg	attgctgagt	cctatggtaa	tggttaggtt	gacttactga	11220
ggaaccatta	aactgttttc	aacagtggct	gogccgttct	gcacccccac	cggcagtgtg	11280
tgagggttct	gacttttact	cctcacaaac	gcttcttttc	catttaaaaa	aatattcagc	11340
cagggtgctc	ggctcacgcc	tgtaatccca	gcactttggg	aggccgtggc	gggcggatca	11400
cctgaggtca	ggagtctcag	acgagcctgg	ccaacatggt	gtaaccccat	ctctacccaa	11460
aatataaaaa	ttagccgggt	gtggcagcgg	gogcctgtaa	tcccagctac	ttgggaggct	11520
gaggcaggag	aatcacttga	acccgggagg	cagaggttgc	agtgaaccaa	gatcgcccca	11580
ctacactcca	gcctgggtga	caagagtga	actccatcta	aaataaaaca	aaaataaaaa	11640
taaataaaaa	tttattaaaa	cattcatcac	agccagccta	gtgggtgtcc	catgtggctt	11700
tgctctgcac	ttccctgata	actaggatgc	tgagcgtctt	gtcccaggct	tgccacacct	11760
cagcacttht	agatcctgct	cacagtcccc	atthtgcga	gagaaatgag	gtthtagggaa	11820
cagcagctgt	gtcatgtcac	acagcgagca	gggggtctct	gagccgtctg	acccacagc	11880
cgaccaagct	ccaatcctta	cgcctccta	gtgtgtgtga	tgtagccccg	ggtgctccca	11940
catttttccg	atgagaacac	cgaagctcaa	aacaggagcg	ttttgtccac	attggatata	12000

10

20

30

cgatgtctgt	ggtttgggtc	tgaagtcact	ttatatctca	gtggtccaga	ctggagtagg	12060
acagggggtt	ctgggggaatg	gggaaggtgt	ctcaggtgaa	aggaaggaat	tccagattct	12120
ccatactgtc	cttgggaagt	tagaagactc	agaggggtctg	gcaaagtcag	acaaagcaag	12180
agaaatgcag	tcaggagagaa	gcggagctgt	ccaggaacag	gggggtcgca	ggagctcacc	12240
cccaggaact	acacttgtctg	gggccttcgt	gtcacaatga	cgtgagcact	gcgtgttgat	12300
taccagcttt	tttttttttt	ttgaggtgga	gtctcgctct	cttgcccagt	ctggagtgc	12360
gtggcacgat	ctcggtcac	tgcaagctct	gcctccggg	ttcatgccat	tctcctgcct	12420
cagcctcccg	cgtagctggg	actacaggcg	cctgccaccg	cgcccggtc	atttttgtat	12480
tttttagtaga	gatgggattt	cactacatta	gccaggatgg	tctcgatctc	ctgacctcat	12540
gatccgccc	tctcggtctc	ccaaagtgt	gggattacag	gcgtgagcca	ccgcgcccgg	12600
cccgaatttc	cactttaaga	atctgtctgt	acatcctcaa	agcctatac	acagtgtctg	12660
gttgctatag	ggaatatgag	gcttacaggc	catgggtgctg	gacacacaga	agggacggag	12720
gtcaggaggt	aaagggcg	agagaggaa	caggcgagg	tcacatcctt	ggctttcaaa	12780
atgggcccag	gagagacacc	ctctgagcat	ggtaggacag	gaaagcaaga	ttggaacaca	12840
ttgagagcaa	ccgaggtggc	tgggcgtggt	ggcttacgcc	tgtaatcca	acactttgga	12900
aagctgaggt	gggtggattg	cttgaggcca	ggagtccaag	accagcctgg	ccaacatggt	12960
gagaccccgt	ctctactaaa	tatacaaaaa	ttagccaggc	gtgatgggtc	atacctgtaa	13020
tcccagctgc	ttgggaggtc	gaggcaggag	aattgcttaa	acctgggagg	cggaggttgc	13080
agtgaagcca	gatcccgcca	ctgcactcca	gcctgggcca	cagagtga	ctccatctca	13140
aaaaaaaaaa	aaaaaaaaa	taaaaagacc	aaccgaggaa	ttgaagtggg	ggggcgctac	13200
agtagcagaa	gggggatcgt	ggagcaggcc	acctgtggt	catgcaactg	aagctcatta	13260
cctgacgatt	tggagctcat	cactgggggc	ctaaggagaa	tagatactga	aggatgagga	13320
gtgatggcgc	ggggcacggc	tgtctttggt	ggccagaact	tggggactgc	tggggtgcct	13380
cactgcaggc	cttctcaggc	ccctttatat	gcttacacag	gctgtttcta	agagggggat	13440
acattgcata	agcgttttca	gactacctca	tcatgggtcc	ctttctttac	cctctgtggc	13500
cctgggtggc	cactctctcg	gaaggtgcag	gtggatgccc	agacccgcc	tgccatccac	13560
ctgcacgtcc	agagctgact	tagcctcgag	attgctgctg	gcacctctg	ccccgggaca	13620
cctcggtagt	gcccgtggag	atgctggctc	tgtgttttct	gctggagttt	gggtcgctct	13680
ttctctctgc	aagtggccac	cgctottggg	tatgtctcca	ggctctctgc	agtcattggc	13740
gcttctcagg	tccttgccca	gcgccaggag	caaaccctcc	tggcactttg	ttcaggggtg	13800
gatgcgccag	tgttctgct	gtggaccccc	atctcacatg	agggctcttg	gcctgcaggc	13860
tcgttcagg	aacacccgct	gagtacgcag	tgtgtgccag	ctgtgtccca	ggcaatggcg	13920
gggacagtgg	ctgctgctgg	ggttgtggtg	gcttctgggg	actctgggga	cagctgaggt	13980
gcaaggagcc	acggctcctt	gaggatgcag	ttggactcca	ggtggaagg	atggttgggg	14040
gaggtataaa	tgggttcagg	gaggagacac	atttggaaca	atgggaacat	ttttaagatg	14100
ctatgtcggg	aggcaacaag	gtggccaacc	caggtgctga	ggagcccaca	ccagccctgg	14160
acgtgttttg	ccgctcacct	ttgctgggga	gtggtgggag	agaggattcc	gttccacgtg	14220
gtggtgtgcg	cagctgggct	gtgtggagct	gggcgctagg	aggaagggtc	tttctgccc	14280
gctagccggg	ctctgccttt	gaacacaatc	aggctccagg	ttttcagcat	ccagtgcctg	14340
agaggacttc	acgggcagct	gtggctgac	ccttgatgaa	ttgggagaag	aacaaaggtc	14400
tatgaaatga	ggtttcatgt	agatggcatt	agagacgccc	acaacagatt	tacagagtgg	14460
agcggagacg	gcggatgggt	ctgggaggcc	cctcctgctg	gccttgactg	tgacagctgt	14520
cctgggaatc	agcttccagg	ccgccccagc	agcctgactg	acacacacag	gggttttagc	14580
ccatcctgc	gaccagctgt	tgccatcatc	agtgcacgct	gggagtggcg	gtggttccag	14640
ccctgggcac	cctccccacc	tgctggggcc	caccagggcc	agtcctgaca	cctacagggt	14700
gcttgagacc	gcacccagct	cacgtgtgaa	gcccagtggt	tcgtgggctg	14760	
aggtccctg	attgcacccc	cacttccctt	ctgcttcaca	tagctgcctc	ttctcacctg	14820
ttttccagcc	tcctgggcta	ggaattccag	tgtgtgctg	gctttgcccc	aggacacctc	14880
cttagccctc	ttctgtagtc	tagagccccc	ggggttggaa	gttctggccc	ctgggacacc	14940
tcgagccaca	ctcagcttct	cctgtgagcc	tccagcatgt	ccctcaggga	ccaagccctc	15000

10

20

30

acgttcttgc	ctccccgccc	acctgggctc	agccaggsga	aggcctggct	gggagcgtct	15060
ccctcttgc	ctgcccttct	ccctcttacc	ctgcccttct	ctcccttgc	cggccatggc	15120
ttttatatcc	tgtgccacaa	gacatggctg	tgtgtgaaag	tggcaggsgc	tggcatctct	15180
gtgggtctct	gaggccacag	ctccagtggc	actcttccca	cccgctggcc	gtgccctcat	15240
gctggaggga	cagcccagcc	ctctcccgaa	ccccagcccc	atgtgcccag	ctgcccccg	15300
ccctctcccc	tggaaagccg	ggctcactca	gccgtatgcc	atgggtggga	catcctgctt	15360
ccctggcctt	ccagggaagg	tctcttttcc	aaatggcgac	acctgggtcc	tgcctggagg	15420
ctggaagctg	tggcccttgt	atgcccctcc	agggtctgtg	cgctcggttg	gcccaggttc	15480
ccatcacctg	catcatcacc	atcatcattg	tcatctcgct	tgtctgtgag	cgggctgggt	15540
ctcccagagc	agagaccctc	tgaggctccag	cctgagttgg	ggctctccgt	ctgacccctg	15600
acggggagct	aggagctacc	aggctctgggt	caggagtgac	ccccaaacct	cgtgccccct	15660
gacaggcacc	cctgactttt	gctaagtggg	tggaggtgac	atcacttaca	gcgggagtga	15720
tgggacaggg	tctgttggct	gcaetgtgct	cccagggtac	tggggagagg	ctatatccct	15780
gggcttttgg	actgcagagc	tgtgtgtgtt	tgtgtgtgtg	tgtgtgtgtg	tgtgtgtgtg	15840
tgtgtgtgtg	tgtgtgtgtg	tttgcgtcgg	cgcacatgtg	tataagatct	ttttttatta	15900
catgaagcaa	gataactgtt	gctgtttcct	tttgggtttt	gtgttcaaca	gagtggggta	15960
cttcttccct	cagacaacag	aactctcccc	tttaaacacg	tgtgtgcaga	gggtgggtct	16020
tgggtctcat	tctgttttga	cagccagagc	agaggaaaca	cagggttctt	cataaaaaaca	16080
ctgcacagca	ggcgactgtc	cagagtcagc	ctgcaggacg	gcagcagccc	tgccctcag	16140
agcacagcta	gggtgggctg	ctttggggtc	tccgtcatt	ccctcccagc	tggcagccgg	16200
cggccggccc	attccttggg	gtgctgggtc	ggggggcgct	cgcctgctct	gctcaccctg	16260
ggaattgggac	agaagctggc	agctcggaga	ggacaggsgt	ggaccccttg	gtggcctctg	16320
gctggaccat	ctcattgtcc	tcagacacag	cctctcgggt	ctagtctcat	ttcctgaaaa	16380
acaagtgcac	agaactagag	caggagtcga	gagctacggc	ccccggggca	gatccagccc	16440
tgccacctgt	tttcacacca	tgctcaagct	gagtggtttt	tacatttttt	aattacttga	16500
aaaaaaaaaa	gcaaaaggag	gtttcatgac	ccatgaaaat	tatatggaat	tcaaaaaaaa	16560
aaaattatat	ggaattcaaa	tttcagtgtc	cataaataat	ttcttgagac	agggtctcgc	16620
tctgtcaccc	aggctggagt	gcagtgtctat	ggcatggctc	gctgtaccct	tgacctccca	16680
ggctcaagcg	atcctcctgt	ctcagcctcc	tgagtacgtg	ggactacggg	tgtgtgccac	16740
caagcccgcc	taattttttt	ttaatlttag	taaagacagg	gtctttctat	gttgcccagg	16800
cttttctgga	actccatctt	ggcctcccaa	agtgtcggga	ttacaggctc	gagccacgga	16860
gcccagcctg	tttttgtttt	ttcactgata	aagtltttgc	gggtgtggta	gtgtgtgcct	16920
ctagcgaatt	gggaggctga	ggtgggagga	tcgcttaagc	ccaggagttt	gaggctgggc	16980
tcaagtgtac	aggaggtgaa	ctatgatcat	gtcattgcat	tccagcctgg	gtgacagagc	17040
aagaacctat	ctcttaaaaa	tatatattta	aaaagtattg	ggtgtgggtg	ctcacgcctg	17100
tgggtcccagc	tacttaggca	tctgaggtgg	gaggatggct	tgagcccagg	agtttgagggt	17160
tgcagcgagc	caagatcgtg	tcactacact	ctagcctggg	tgacagagcc	cagaccctgc	17220
ctctttaaaa	aaaaaaaaaa	aaaaacatgt	attggaacac	agccatgcct	gttcagtcac	17280
gtgctctcca	tgtgtcttcc	tgctccagag	acccttatgg	cctgaaagct	gaaaaatatt	17340
tctatccttt	acaaaaaagt	ttgtgtacct	ctgtcctgga	aaattcatct	cccaagttct	17400
cttccggcac	tggcgttctt	gggtgtccta	aatlttggcc	ctgttatctt	tgaactctgt	17460
tttggctctg	ttccctccca	ggagccagga	caggcacgtt	ctctgcatct	tgtcccctga	17520
cggccagagg	cttggctcgg	ctcaggcatt	cttggaata	tctggctcca	ggaaaggcag	17580
aggcctcctg	agtcagccca	gagggaacct	gccccaggct	tgggggaggc	ctgaccocagc	17640
agagtggctt	ttgcccagtg	gttgggccc	tcaagatgtg	ctgaaagtgt	tcctcagaag	17700
gccacttttg	gattccttcc	tccagtatta	gagcaactga	gagctgtctc	ttgcaagcct	17760
gatgttttcc	cagttggccg	ggtccaccgg	gtgccctggg	attctgggat	ctgggtggaa	17820
agtagggggc	ttgggggagt	gtcctgggtt	ctggaatcca	ggtggcaagt	ggtgaggttc	17880
agggagtggc	ttctgagcca	ccatagggtt	ctctgtggga	ggctctgccc	atccaggaga	17940
ttccgcaggc	cctgcccggc	cagagccagc	gtcttgcgct	tgccgaggct	acagccagcc	18000

10

20

30

ccagccgggt	ggaacagccc	gtcgccctct	ctcactttgt	tttggggcca	cctgggagtg	18060
tggagcaagg	gtagagaggg	aggaagtggc	tgccggccgc	tgcccagcac	ccttgtttgc	18120
cttggggccct	ctgtgggctc	cttttttattg	ctcttcaatg	aagccaggga	aatggacttc	18180
cttgccctcac	ttcagttcaa	catgtctgga	agtttggtat	taaaattaag	aaagtgtgga	18240
aatagagcaa	gaagagaaaa	atctctccaa	gagataatag	tgacctctga	gctgggcgcg	18300
gtggctcacg	cctgtaaatc	ccagtacttt	gggaggctga	ggcgggcaga	tcacctgagg	18360
tggggagttt	gtgaccggcc	tgaccaagat	ggagaaaccc	cgtctctact	aaaaataaat	18420
aaataaataa	ataaataaat	acaaaattag	ccaggcatgg	tggcgccctgc	ctataatccc	18480
agctaaggca	ggagaatcgc	ttgaacctgg	gaggcaagg	ttgcagtgg	ccaagatcac	18540
gccattgcac	tctagtctgg	gcaacaagag	tgaaactccg	tctcaaaaaa	aataaataaa	18600
taaaaaataa	aaatagtgac	ctctggccag	gtgtggcagc	tcatacccg	aatcccagca	18660
ctttggaagg	aaggccgaga	tgggcagatt	gctttagcac	aggagtgtga	gaccagcctg	18720
gccaacatgg	tggaaaccca	tctctacaaa	aatagaataa	aatttaagag	gtaatagtga	18780
ccttttggtg	gatcgaaacc	tggattgctt	tctttttcta	aatgctgatt	cttttctttg	18840
tgggtgtttgt	gttctgtgcc	gatgtccctc	ccccagccct	gttattgtga	gtggaagaag	18900
gggaagggtg	tgcgccgcta	ctgtgagccc	ctcctctcac	gctgggtgtc	cttgagaag	18960
cctgcacttc	ttcattgtac	gccaggcgct	ggtccctccc	tggagtgggt	ctgtgctgct	19020
gggatggggc	caacccctca	gatgttttct	gagtgtcaca	cacagggtgtg	tgcattcatg	19080
gcctttgctg	gtcttctctg	tgtggaggca	aaaatgtgaa	gaaccctaga	tgattttggg	19140
accagggtct	catcacctgc	tgttcattgc	acaccggagc	atccaggcat	gggtggagag	19200
ctcagacttc	caggcacggg	cgcaggggct	ggtctaacca	tgttcccgc	cgcctgctcg	19260
tcagaaccgc	ctgttgggag	ctgttatcat	gataccatac	ctgggcctcg	ggctatccga	19320
ttctgacttc	attgctccag	gttggggcca	ggccgttgtt	tgtctgtttg	ttgtttcttc	19380
tgtgacgtta	gccactgggc	taacttgagc	ccctcagtta	cagggtggaga	aactgagacc	19440
catgggggtg	caaggacttg	ccgaggaccc	agagccctct	gggggcagag	ctgaggcggg	19500
gcctggcttt	gggtcccaga	gcttccagtc	cccttcccgc	tctcctaaca	gctttttttt	19560
ttgagacaag	atctcaccct	gtcaccaggg	ctggagtga	atggcatgat	ctcggctcac	19620
tgcaatcttc	gctagctgcg	ttccagcgat	tctcctgcct	cagcctccc	agcagctggg	19680
attacagggtg	tgtgcgcgca	tgcccagctc	gttttttttt	gtacttttag	tagagatagg	19740
gtttcaecat	gttggccagg	ctgatctcga	actcctgacc	tcaaatgatc	cgcctgcctc	19800
ggcctcccaa	agtgttagga	ttacaggctg	ggatcacact	gtgcctggcc	ctagcagctt	19860
tgtcctgtgc	catccaacaa	cagatgaccg	aagtctttgt	ttcttaacat	gcattccatc	19920
tgccttacag	ttttgccc	tgcaaacag	aggacttgct	gcttttctgg	taagctggaa	19980
atgtaatctg	gtagcaggag	gcctgtggaa	gcttgccctt	aatggccttg	tgtctcttct	20040
atcctgtcct	gagagccgga	gaacttgat	gttgaccta	actcaacctt	cctgttaaca	20100
tacagtctctg	caggctcatg	gatcatcaga	accacgtcct	atctcacgcg	gctgtatgct	20160
tcctgtgggt	cagggtgttt	taccttgaca	gtattttctc	ctcgggtggc	tttgcggtgg	20220
ttgtctttta	tcagcattga	ctcttcaaga	aaaatattta	gctgctacat	ctcagaggag	20280
acagggtgga	aagcatctga	gacctgcagg	ctcagactta	gaaccagaag	tgcctcaga	20340
gttcatccgg	ccctgaccca	gcgggaaatg	agttcacaga	gaagcgggag	aactttgccc	20400
caggccctgc	cgttgctcat	aactgcccc	ggtccttaca	tttgctccag	gtcctgcccc	20460
aggccctgca	gttgctcata	actgccccag	gtccttatat	ttgctccagg	tcctgcccc	20520
ggctcctgag	ttgctctgtg	tgggtgggtg	gatctggagc	cctccgcca	ttgctgcacc	20580
tggggcaggc	attgctaatt	gatccagga	ctccttccctg	cggagcacgc	cctggttctc	20640
caggcagccg	ctgcctgtca	gcctgcagtg	gttcgggaga	ggacacctgc	ttgcctggtc	20700
tgttccaaat	cttgcttctc	atcccagcac	aggtaggggg	tgtatgggga	aagggatcct	20760
cagtggcccc	tgtcaactgct	ctatcagctg	gggacgtggc	atcctagtga	aaacatcatg	20820
gcccggcgcg	gtggctcacg	cctggaatcc	cagcaacttg	ggaggctgag	gagggtggat	20880
cacttgagggt	cagaagttcg	agaccagcct	ggtcaacatg	gtgaaaccca	tctctactaa	20940
aaatacaaaa	attcgccagg	tgtggtggcg	ggtacctgta	atccgagcta	ctcgggaggc	21000

10

20

30

tgaggcagga	gaatcgcttg	aacctgggag	gtggagcttg	cagtgaagcg	agatcttgcc	21060
actgcactcc	agcctgggca	acagagtgg	acgctgtctc	aaaatctcaa	acaaacaaac	21120
aaacaaaaaa	caaacaaaac	aagcgtcatt	tatccagcac	ccctggggaa	ccatgctacc	21180
tgggtgtttta	tgggtacctgg	caaggtgcag	gtgaagtgtc	tgtctctggg	cattgaaccc	21240
gtcttgtttt	gggcagctca	ggcccaggg	agggtcggg	ttggctctcg	ttgggtgtgg	21300
cctggcccat	ccagacctat	atttctgccc	tcctgcaggt	gatcaatgtt	gatgggacga	21360
agaggcggac	cctctcggag	gacaagctcc	cgcacatctt	cgggttcacg	ctgctggggg	21420
acttcatcta	ctggactgac	tggcagcgcc	gcagcatcga	gcgggtgcac	aagggtcaagg	21480
ccagccggga	cgtcatcatt	gaccagctgc	ccgacctgat	ggggctcaaa	gctgtgaatg	21540
tggccaaggt	cgtcgggtgag	tcgggggggt	cccaagccat	ggctcagcca	tgcagacttg	21600
catgaggagg	aagtgcagg	tcocatgctg	ggcataagtg	ttgagctcag	gtgcccgcac	21660
ctggggaagg	gcaggacagg	aaaggtgaca	gtatctggcc	aaggacagat	gggaagggac	21720
caaggagagct	gattaggggag	tgggttatgga	ctaggaatgt	cggtaacaat	ggtagaagag	21780
tgactaacat	ttgttgagca	cctgctgtgt	gcccggccct	ggccgggagc	cttcgtgccc	21840
acagtgaacc	cgtctgcaaa	tgtagtctct	tgcctactct	gcactgggga	gcaggacgca	21900
gagccgtgca	tctcacaggt	gccaaagctca	ggactccctc	ctgggtctgc	ctgggctggg	21960
ctgtgcttgt	tgcctctgtg	gcccacgcat	gtgcaccttc	caactgaaag	ccaggatctt	22020
caggacgctc	cccaggagg	tcgttgtctg	gcacaatgat	ttgtctcttc	ctgaaaaggt	22080
gacagagtta	cactggagag	agcagcatcc	aggtgcggca	gggacaggcc	tggggctcgc	22140
gggcaggggac	tctgtgtcct	gccgggggtcc	cacactgcac	ctgcttgtca	gaggcactca	22200
gtcaatcttt	gctgatgaag	gatgagagga	cagaggacgt	gatgcttgct	gctgcattgc	22260
ctgcagctct	gggtgagatg	cccgggttga	ctctgctgcc	cgtcgggtgg	atgtgatgtc	22320
agatccccgg	ctttaaaata	cgagggagct	gggaattgag	ggagcaggtt	ggggcagaaa	22380
gcacagcccc	gtggaagcct	ggagctgagg	cagtgtgggc	gacccctgga	gcagtgagtg	22440
cttctctcat	ggccttcctc	gcacctgca	gtcctcatgt	aggggatgcc	atccatgaat	22500
ttagttttcc	cagcctcctt	taaaaaacgc	ttcatgctgg	ggccggggca	gtgcagtggc	22560
tcacatctga	aatcccacca	ctttgggagg	cogaggcggg	tggatcatga	ggtcaggaga	22620
tcgagaccat	cctggctaac	aagggtgaaac	cccgtctota	ctaaaaatac	aaaaaattag	22680
ccgggtgctg	tggcggggcg	ctgtagtccc	agctactcgg	gaggctgagg	caggagaatg	22740
gcgtgaaccc	gggaagcggg	gcttgcaagt	agccgagatt	gcgccactgc	agtccgcagt	22800
ccggcctggg	cgacagagcg	agactccgtc	tcaaaaaaaa	aaaaaaaagt	acaaaaaaa	22860
aaaaattagt	ctgggtgtgg	tatcacgcgc	ctataatctc	actactcgag	aggctgaggg	22920
ggagaattgc	ttgaacccag	gaggtagagg	ttgtagtggc	cccgtatcgt	accactgccc	22980
tccacctggg	caatagagcg	agactctgtc	tcaaaaagaa	aaaaaaaaaa	agaacattta	23040
tgcaggtgt	ggtggctcat	gcctgaaatc	ccagaacttt	ggaagactga	ggcaggaggga	23100
tcacttgagc	ccagaaatct	gagagtgtct	tccctgggca	acatagagag	acctcatctc	23160
taccagaaaa	aaaaaaaaata	gcccggcatg	gtggcatatc	cctgtggtcc	cagctactta	23220
gggggctgac	gtggcaggat	cacctgagtc	tggaggcaga	ggttgaagtg	agctgagatc	23280
atgccactgc	actccagcct	gggtgacaga	cagagaccct	gtctcaaaaa	aaaaaaaaaa	23340
aaaaagcatt	tactatccac	catggaaggt	gagactgacc	tgtgagtgat	tgttcaaaga	23400
acaaaaaata	aaccccagag	ataagacaaa	aggggtgcctc	catgggggtg	tgatttaaag	23460
ctgagaaaatt	gggtctcttc	cccctcccct	ctcaccctgt	ggtttgctaa	aggagatggg	23520
aaaaaggatt	cttttttttg	ctgaaatatt	taacactaaa	ttaaagccaa	ttttaacagc	23580
actttgggtg	atgagtgaag	ttaacagact	ggccaaaaat	aaacgaacgg	tctgtactat	23640
gtgaaaaaga	ggcagctttg	gccatgctgg	gccaatgtga	gttttcaggg	ttgctgggaa	23700
tgtctgtgaa	tgggaggaag	ggcctagctg	ggaactctcag	gagccaaggc	cctgaggggg	23760
aacttgcttg	gtccctgccc	tgaggcgttc	actgctttct	tcctggggcc	gatcacaggc	23820
ccggaggctg	gaccactggg	ctggcactct	tgcagagctg	tcctctgact	tcctgacct	23880
gctcctttca	gcagccttgc	tgcactttag	tttctctgaa	tgaaaaatgg	ggatgagaat	23940
agctcctacc	tccaaggtga	atggagtgg	ttcggacagg	tgactccctg	ggaccagtgc	24000

10

20

30

ctggcgctg	acaaggtcca	gtcagagccc	gcactgctgt	tactgatacc	cttggttgta	24060
ccaggggaga	acttggttgc	cattgccagg	tgtttcccac	ccacccccac	tactgtccct	24120
gtttgatgtg	tggcgggaat	aaagctgtgc	acattggagc	ttttggcaca	tcctggcttt	24180
caggtgaaag	gtcgtgtgt	gtttgagggg	ttagcctggc	caacccagcc	atgaggtcgg	24240
acctgacctg	ggggtgagtc	ctgagctcgg	cacccctgag	ctgtgtggct	cacggcagca	24300
ttcattgtgt	ggcttgggcg	cacccctttc	cctgctgggc	tgttgatgtt	tagactggag	24360
cctctgtgtt	cgcttccagg	aaccaaccgg	tgtcgggaca	ggaaacggggg	gtgcagccac	24420
ctgtgcttct	tcacacccca	cgcaaccggg	tgtggctgcc	ccatcggcct	ggagctgctg	24480
agtgcacatga	agacctgcat	cgtgcctgag	gccttcttgg	tcttcaccag	cagagccggc	24540
atccacagga	tctccctcga	gaccaataac	aacgaagtgg	ccatcccgct	cacgggctgc	24600
aaggaggcct	cagccctgga	ctttgatgtg	tccaacaacc	acatctactg	gacagacgtc	24660
agcctgaagg	tagcgtgggc	cagaacgtgc	acacaggcag	cctttatggg	aaaacottgc	24720
ctctgttctt	gcctcaaagg	cttcagacac	ttttcttaaa	gcactatcgt	atttattgta	24780
acgcagttca	agctaataca	atatgagcaa	gcctatttaa	aaaaaaaaaa	gatgattata	24840
atgagcaagt	ccggtagaca	cacataaggg	cttttgtgaa	atgcttgtgt	gaatgtgaaa	24900
tatttgttgt	ccgttagact	tgacttcaga	cacccacccc	actcccttgt	cgggtcccg	24960
ttgctcagca	gactctttct	tcatttatag	tgcaaatgta	aacatccagg	acaaatacag	25020
gaagactttt	tttttttttt	tttgagacag	agtcttactc	tggtgcccag	gctggagtac	25080
cgtagcgtga	gctcagctca	ctgcaaccctc	cgctcccgag	gttcaagcga	ttcttctgcc	25140
tcagcctcct	gagtagctgg	gactacagac	atgcaccacc	acacccagct	aatttttttt	25200
atatttttag	tagagacagg	gtttcatcat	gttggccagg	ctggtcttga	actcctgacc	25260
tcaggtgatc	tgcccgcttc	ggcctcccaa	agtgtgaga	taacaggtgt	gagccaccgt	25320
tcocggcata	ggaaactctt	ttgccttcta	aagaagagtt	tagcaaaacta	gtctgtgggc	25380
tggccttctg	attctgtaaa	gaaagtgtga	ttggtggctg	ggtgcgggtg	ctcacacctg	25440
taatcccatc	actttgggag	gocgacgtgg	gcataccacc	tgatgtcggg	acttcgagac	25500
cagcctcacc	aacgtggaga	aaccccgctc	ctactaaaaa	tacaaaaaaa	aaattaaccg	25560
ggcatggcgg	cgctgcctg	taatgcgagc	tactcaggag	gctgaagcag	gagaattgct	25620
tgaacctggg	aggcggaggt	tgtggtgagc	tgagatggca	ccattgcact	ccagcctggg	25680
caacaaaagt	gaaactccgt	ctcagaaaaa	aaaaagtgtg	attgggtgta	ccaaagcgca	25740
tttgtttatg	gattgtctgt	ggcagctttt	gttctgccga	gatgagttgt	gacagatctg	25800
tatgggctct	aaagcctaaa	acatgtgcca	tccgcccctt	tacagaaaaa	gtgtgctgac	25860
ctctgttcta	aagtattgga	caactacaat	gtttgtctat	ttattattct	atgatttgtt	25920
ttctgctttt	tgttgttgtt	gttgttgttg	agatagggtt	tccctctgtc	actcaggctg	25980
gagtgcagtg	gtgtaacttc	agctcactgc	agcctcgacc	tcctgggctc	tagtgatcct	26040
ctcatctcag	cctccctagt	agctgggact	acaggcacac	accaccactc	ctggctgatt	26100
tttttttttt	tttttttttt	ttgtggagac	agggtttccg	catgttgccc	aggctgggtt	26160
caaaactccta	ggctcaaaac	cccacctcag	cctcccaaag	tgctgggatt	acaggcgtga	26220
gccaccatgc	ccagcctatt	ctactgtttg	tattacatag	ctttaaaaga	ttttttatga	26280
ctttaagtca	caaggggtct	ttgtagaaaa	aatatataat	ataggaaagt	ataaaaagaa	26340
agtaaaaatt	gtccataacc	tctccagcca	gagacgaccg	ttgctgacac	ctcagcatat	26400
tgcctttaag	tcttttttct	ctaagatagc	atttctcttc	atcacagtca	tatgtctacg	26460
agaattctgt	atcctgattt	tttcacttga	cattacaaca	ggtatttgat	ggcgtgtgta	26520
caaaactctt	ggcacaatct	tttaaagtga	tgaataactc	cactgcacag	atgtttgctt	26580
ttaggcttaa	ctgttctttt	attttgcttg	tgctgggttac	agccggggcac	agtggctcat	26640
gcctgtaatc	acaacacttt	gagaggggtga	ggcaggagga	tcacttgagc	ccagaagtgt	26700
gagaccggcc	tgggcaacat	agtgcagacc	catctctaca	aaaaactttt	ttaataagtc	26760
ggcgtagtgc	gtgcatagct	gtagtcccag	ccaccaagga	ggctgagttg	ggaggattgc	26820
ttgagcccca	ggaggttgat	gctgcagtga	cctgagatta	ctccactgta	ctccaacctg	26880
agcgacagag	caagacttgt	ctggggaaaa	aaaaaaaaaa	aatatatata	tatatatata	26940
tatatataca	tatatacata	cacgcacaca	cacataatat	aaaaatatat	atttataaat	27000

10

20

30

atataatata	taataataaaa	atatatatatt	ataaataaaaa	tttataaatt	atattttataa	27060
gtaaatatat	aatatataat	ataaaaaatat	atatttatata	atatataata	aaatatataa	27120
tataaaaaata	tatatattata	aataatatata	aatacatact	tataagtata	tattttaaaat	27180
atatgtaatg	tatatattttt	aatgtatgat	atataatata	cattttataaa	tacacatttta	27240
tattattttta	tataaaaatat	atataaaaatc	tccaagttgc	tttttccaaa	aaggtgtctt	27300
gctgcatttc	aaacattcat	ttaaaaactt	gaatgctggt	gatctggtcc	agaatgtggt	27360
cagtagctgc	tgccagtggc	caagcatctc	gggagatgtc	tacaaaaacac	gctggttctg	27420
gcctggcgtg	gtggctcacg	cctgtaatct	cagcaacttg	ggaggtctgag	gcaggtggat	27480
caactgaggt	ctggatttcg	agaccagcct	tgccagcttg	gtgaaacccc	atctctacta	27540
agaatacaaaa	aaaatttagcc	aggcgtggtg	gcctgtgcct	gtaatcccac	ctacttggga	27600
ggctaaggct	ggagaatcgc	ttgaaccacg	ggggcagagg	ttgcagtgag	ccgagatcgc	27660
accattgcac	tccaggctgg	gcaagaagag	cgaactccg	tctcaaaaaa	aaaaaaaaag	27720
atgctggttc	ctaaaatgtg	gccccttttc	tctcacctg	ctgccagacc	atcagccgcg	27780
ccttcatgaa	cgggagctcg	gtggagcacg	tgggtggagt	tggccttgac	taccccgagg	27840
gcatggcctg	tgactggatg	ggcaagaacc	tctactgggc	cgacactggg	accaacagaa	27900
tcgaagtggc	gcggctggac	gggcagttcc	ggcaagtcc	cgtgtggagg	gacttggaca	27960
acccgaggtc	gctggccctg	gatcccacca	aggggtaagt	gtttgcctgt	cccgctcgct	28020
cttgtgttca	cctcgtatga	gacagtgcgg	gggtgccaac	tgggcaaggt	ggcaggctgt	28080
ccgtgtggcc	ctcagtgtat	agagctgtac	tgtgttcatt	agccttgatg	gtggccaggga	28140
ctggtagggc	cctcagaggt	catggagttc	cctcgtggag	cgggtgctga	ggtctgtatca	28200
ggcacagtgc	tggctgcttt	cacctgggcc	gtctcaccca	agtgtccatg	gagcctgcgt	28260
aggggtgggt	tctgtgtcga	ttttacagat	gcagaaacag	gctcagagaa	accgagtgac	28320
ttccctaagg	tccacatacc	agttagagca	gagctgggcc	aggaagtgtc	gtctcaggct	28380
cctgaccagg	tctccttget	ttgcactctt	gccccaaaca	tgatccagaa	ctgactttga	28440
ggtccccgga	cctcaggctc	ctccgaaatg	gcctccttga	ggctgctgag	ccacagctta	28500
ggacccacct	cagagaggcaa	atgtgctttg	agctgccagg	cgtcctgggg	gcccctgcctt	28560
gggcaogggg	ttcagacagg	cccagatgt	gtggggcgct	tttctggact	tgagttttct	28620
ttctgtgtg	gtggacacag	tgctcacccc	ttaaagcacc	tgtgatgtgt	gcagcagccc	28680
aatccctgcc	tgtcgcctgt	tctgctaggg	aagggaaggaa	gacttcagga	tggcaggaca	28740
acagaaagag	gtccaggttt	tagagcaagg	gcaggtcaaa	cttagaaaat	tctggaatga	28800
ggatgtgcat	ttcctcttct	ggatctgcta	aaagaagagg	gaaggagggg	ctgctggggg	28860
aggagccag	agccaggttt	acatccggat	cccgaaggc	ctcccctgcc	ctgaggtcct	28920
gttttgtgat	gtgcttgtgt	ccatccctgt	ttctgcccgt	tccccaacat	ccggccaagc	28980
ttagggtgat	gttcacagac	acactcacc	tgtctgtgca	cctgtttttg	tgtccgtaag	29040
tgggtattta	ctcaccttac	gagttagcca	ctgtgggaat	tcaggggagg	ggcgcagtga	29100
ccacccctgg	agggatatgt	gtgtggcagg	ggtcgagggt	ctcgcccttc	cctgcttcc	29160
gcgcgtggct	ttctccagga	cggggagggc	tgagctgaag	aggtggggac	agttgcgtcc	29220
ccccgccacc	cactgtcctg	oggtgagagc	agactcactg	agcctgccct	tctcccttgt	29280
gccttcagc	tacatctact	ggaccgagtg	gggcggcagg	ccgaggatcg	tgccggcctt	29340
catggacggg	accaactgca	tgacgctggt	ggacaagggt	ggccgggcca	acgacctcac	29400
cattgactac	gctgaccage	gcctctactg	gaccgacctg	gacaccaaca	tgatcgagtc	29460
gtccaacatg	ctgggtgagg	gccgggctgg	ggccttctgg	tcattggagg	cggggcagcc	29520
gggcgttggc	cacctccag	cctcgccgca	cgtacctgt	ggcctgcaag	ttccccaaac	29580
tggcaggagc	tgtggccaca	cccacgactg	cccagcagcc	tcacctctg	ctgtgggagt	29640
tgtccccgtc	cacccctggg	tgcctttgct	gcagttatgt	cgggagaggc	tctggtgaca	29700
gctgtttcct	gtgcacctgc	tgggcactag	gtcccagcta	atccctgtgc	caggactcta	29760
atttcaccct	aacacacatg	gtggttttca	ttgctgggga	agctgaggcc	tgagcacatg	29820
acttgccctta	ggtdacatag	ctggtgagtt	caggatcccc	cagagatacc	agggccagca	29880
ctcgatcccc	accagaccct	gaaccccacc	atgtgctggg	attgtgctgg	gagtgccacc	29940
acgcctggga	ccccagggct	ggtgctctca	tctccttttt	ccagatcatg	agaatgaggc	30000

10

20

30

tcagggaggt	ttgaaaaaaa	cctatcccaa	gtcacacagc	aacaggagca	ggatttgaac	30060
ccagaaaagg	ggaccgcaca	ctctgttctg	ctagagtagt	tagctgtcct	gggtgatatg	30120
gcaggtgaca	ggggcaactg	tgttaacaa	aggaaccccc	atccccctg	ccaagttggg	30180
agactagaag	gtcaggggca	gaagctctga	agggccaggt	gcagtggctg	acacctctaa	30240
ccccagcact	ttgtgaggcc	aaggcgggca	gatgatttga	gcccaggagt	tcaagatcag	30300
cctgggtaat	gtagtggagc	gccatctcta	caaaaaaatt	ttttaaaaa	tagctgggca	30360
tgggtggttca	tgcctgtagt	ccaagctact	tgggaggctc	aggtgggagg	attgcttgag	30420
cccaggaggt	tgaggttggt	gtgagctgtg	atcatgccac	tgcactccag	cctgggcaat	30480
agagtggagc	cgtctccaaa	aaaaaaaaa	gaagaagaaa	agaagctct	gaggctccaa	30540
gtccccaggc	accccttggt	ttgagggcag	acaagggagg	agagggtcac	ctgggcagcc	30600
ctgacttttg	tccccggca	aagggaacct	cagtgacctt	ggccctagga	gagcctctga	30660
gcacgtcagc	catgtcgaac	cgctcaggaa	gggcagcaag	aattttgctt	ctgacctctg	30720
cctctcctac	tgcctatctg	cactgggtgt	ggttctgccc	attttacaga	tgaggaggct	30780
ggggcatcga	ccagctgaat	gccttgtccc	aggtactgct	taggcagagc	tggcagttga	30840
acccogtgtc	ctgggtgtcg	ctgggggttg	gctgcacct	gacttgtgag	gccagtagca	30900
aggttttgcac	gtgacttcgt	gaccttcacc	cagctctgca	gcacatccc	tgaccagct	30960
catccaggcc	gcattgcaaac	ctgttgccag	gcgagaaacc	agtcaccgca	cagctgtggt	31020
tgcctgaaat	gattaagctc	attaatcacc	ccggagttag	gacagactca	gatgaaaacc	31080
agcaaaagcc	ctggaaacct	atgtgacct	gccaatgagg	gcggccatgt	gcattgcagc	31140
ctggccgtca	ctcctcggt	cgtgttttgg	acttaaacgc	tccggatgtt	tactgagtgc	31200
ttgattaata	acatggaagg	cctggtctca	ttgctgtggg	agtgaaggat	gcacagccag	31260
gcctgacatg	atgagaacaa	gaacctggag	tctcgtctgc	tgggtggtaa	tctggccct	31320
gccacttagc	aactgtgtga	ctgtagccag	gtcacttaat	tttgctagat	cctgcctgct	31380
cttcagtgg	tcttctgtgt	tttccaaggt	ggccaaacac	tttaaggcat	tcatgtggtc	31440
gctaggctgc	agggttgaac	cctggctcac	cccgcagggc	gccgtgtgct	ctgtggcctg	31500
gtgtgacctt	tgtgcacacc	gtgcccgtgt	gtgttcatgc	aggtcaggag	cgggtcgtga	31560
ttgccgacga	tctcccgcac	ccgttcggtc	tgacgcagta	cagcgattat	atctactgga	31620
cagactggaa	tctgcacagc	attgagcggg	ccgacaagac	tagcggccgg	aaccgcaccc	31680
tcatccaggg	ccacctggac	ttcgtgatgg	acatcctggt	gttccactcc	tcccgcagg	31740
atggcctcaa	tgactgtatg	cacaacaacg	ggcagtgtgg	gcagctgtgc	cttgccatcc	31800
ccggcgccca	ccgctgcggc	tgccctcac	actacacct	ggacccagc	agccgcaact	31860
gcagccgtaa	gtgcctcatg	gtccccgca	cctcactccc	tcgtagatc	aggctggttc	31920
tgggagctga	cgtgaaagg	agcttctcat	ctggggttcc	tgggtgtaca	tagatggttg	31980
ggtaggttgt	gcactgcaca	agctgcatga	tgctacctgg	gggtccagg	ccaggctgga	32040
tggacttgtt	gcttcatcag	gacatagata	aatggccaaa	actcctcagc	tgggaaggctc	32100
tgggcaggat	ctttgggtgt	gaaaaccagt	cacaggggaa	gggtgcttgc	tcatactgcc	32160
agcacagtgc	tgagtgtctt	ccatagcgct	cgtttactcc	tcaagcctgg	agggtgggga	32220
gtagcatgg	cccatttcac	gtacaaggaa	ccgatgcac	agagagggtg	ggcaaccat	32280
ccaaggccat	acaactgggg	tgggttgagc	cggggttgac	tgtggcaggc	tggctcaaga	32340
gtccctgtc	ctgaacctt	gccaggcagc	ctggcatcag	ctcggggaat	ttttgccctg	32400
acccctggaa	gcaagtgggc	ctctttgttc	tcatgtcagt	gatgagaaga	gtgactttcc	32460
tatggccct	ctggagtaca	ggtgtttcct	gttggcgggc	tcttccccca	tgacatcagc	32520
agcgagctgg	ttatgattcc	ctacgcagaa	cttgatagtt	tataaagctc	tttgtcatcc	32580
aggcccgctt	ggagtctcac	gcagacctgg	tgcaggcgg	ggctggtctt	gcctgtccca	32640
gctgcatgga	tggggaacct	gaggcttgca	aaggtttaagg	ggctgttcga	ggcccaggct	32700
ggcaggagat	gggcctgggc	cagagtctgg	gacttcccat	gcctgggctg	tctttggtcc	32760
tgttgcctac	catccctccc	tggggccatg	accttagaga	gccaaatgga	ggtgcaggta	32820
acccacggca	aggaggggtt	gccatgactc	agagtcccg	tctgtggcc	ggcagtaact	32880
ggtgcaacga	cttggatttc	agaccagcca	ctgtagcccg	ctgacgggtg	gctcgaagtg	32940
ccacagcttc	tgaagccagg	caggactcag	gccaggagac	tctgttagct	gttgagaggg	33000

10

20

30

agaggccaac	ggatgttctg	gttctgctag	agagctgggt	cttoggatcc	tggtagccagt	33060
gcactgagag	gaggcccagc	ttgattcttg	ggctgccttg	tggtagccagt	tgcctctcac	33120
tgacaccctc	gaggagtgct	ttctctcggg	cttggtgact	gtgcccgggt	ttccgcagtt	33180
cactgggtgca	cacataggca	catagcaaac	cgcacacaca	gtcgtgggta	tgagtttcac	33240
tacattccac	caccagtggt	cactaccatt	acctgccttc	cgtcttaagt	gttcattcatt	33300
taaaaaataa	tttattgggc	tggacgcggt	ggctcatgac	tgttatccca	gcactttggg	33360
aggctgaggg	gggcagatca	cctgaggtca	ggagttcaag	accagcctgg	ccaatatggg	33420
gaaactccat	ctctactaaa	aatacaaaat	tagctgggca	tggtagggga	tgccataaat	33480
cccagctact	caggaggctg	aggcaggaga	atggcgtgaa	cccagaggga	agagcttaca	33540
gtgagcccag	atagcaccac	tgcagtcagc	cgtggggcaac	agtgcgagac	tccatctcaa	33600
aaaaaaaaata	aataaataaa	agaaaaataa	atttatgata	tatttcaaaa	ataacacatg	33660
tactttgaaa	cagcagagac	acatatgaca	cggagaatga	aattccccat	agcgcacccc	33720
caagagacag	ccctggctcc	cccgtcttcc	ccgtggacct	ccagcggggc	agatgctgag	33780
ccgcctgttg	tcgagtggtg	tgctatcccg	tcctccagct	cctctgtggc	ttacagacac	33840
ccacctgcag	ccctgtcttt	gcctcctcta	gcgcccacca	ccttcttgct	gttcagccag	33900
aaatctgcca	tcagtcggat	gatcccggac	gaccagcaca	gcccggatct	catcctgccc	33960
ctgcatggac	tgaggaacgt	caaagccato	gactatgacc	cactggacaa	gttcattctac	34020
tgggtggatg	ggcgccagaa	catcaagcga	gccaaaggac	acgggaccca	ggcaggtgcc	34080
ctgtgggaag	ggtgccccgg	gtgcttccca	aggcgcctct	cctgctgggt	tccaggtctg	34140
tgccccgtgc	cttagcagag	ggaggaaaca	gaggatggct	ctgggtgaat	gatgacttgg	34200
gcttcgatta	tgtagtcaca	gggtatgacc	ctgagatgct	tggaaacccg	agactgtgat	34260
tatatgtaga	aactgggttt	cccgttgggt	taagtgtgca	tggtaggggt	agaccccaca	34320
ggacttttgt	cttttcaaga	aagaaaaatg	tcgtgtgtca	tgcaagggtg	gttggtactg	34380
gttaatccag	gtttatcctt	tattttgtgg	gaactgtaca	gtcatttctg	ctacaatgct	34440
gtatatgtct	ttctgaaaga	cacctatgca	aaatgcgaca	gtaaaaatga	cacaactcat	34500
agggaaagcg	gggcccaggc	acagccctca	aaatctccat	caatgacatg	taagaaaaga	34560
gaggaaacct	ggaaatagca	aagtgccttt	tgacacattaa	atggttagct	atatcccaca	34620
atactgtgca	ttcgtaaacg	ttaatgctgc	aataaatacg	gcacttcacc	ttgggaagat	34680
ctggagttgg	cttatgagtg	tggaaagggt	tagcgcctga	gtttttgtga	aacactggaa	34740
ggaggattgt	gggaaatcaa	atggaaagtt	ctcaccaccg	gcgtggagaa	gagtgggtca	34800
tggccccagc	agtgcagcca	gggaggtcag	agacggaggt	gtgtgtgtgg	gtgtgacctt	34860
gcgcagttcc	ctgcggctg	tagttttttg	cattcgctta	atgtttctcg	tggaggaaat	34920
tgtgcatgag	caaatgtgaa	accgtgctgt	gctcaaattg	tcctaataca	tcattgcatt	34980
ggaaacagatt	ggcttttttt	tttttttttt	tttttttttt	tttttgagat	ggagtctcac	35040
tctgtcacca	gcctggagtg	cagtggcatg	atcttggtct	actgcaacct	ttgcctccta	35100
tgttcaagtg	attttcctgc	ctcagccctc	tgagtaactg	ggattacagg	catgagccac	35160
cgccggccgg	cagatttgca	tttttgaaac	aactgctagg	ctgggcgcgg	tggctcacac	35220
ctgtaatccc	agcactgtgg	gaggccgagg	caggtggatc	acctgaggtc	aggggttcga	35280
gaccagcctg	gccaacatgg	tgaacccccg	tctctactga	atatacaaaa	atcagctggg	35340
tgtgggtggc	ggtgcctgta	atcccagcta	ctcaggaggc	tgaggcagga	gaattgcttg	35400
aacccaggag	gcagaggttg	cggtagagcc	agatcacacc	attgcactcc	agcctgggca	35460
acaagagcaa	aactccatct	caaaaaataa	aaaatagaaa	aacaagtgtc	gtagcggaag	35520
tgagcacttt	gcggagtcag	gcttgtgtgg	cctgttccac	aaatgatgtg	ctcacgggtg	35580
cctcaggccc	acctggagtc	tcagcagatg	ggcacacacg	gttcattagt	gtagaattcc	35640
aggacaggcc	tggctcctaa	gcagccttct	tttacaacaaa	ctgcagagcc	cgctgtatc	35700
ctagcacttt	gggaggccga	agtgggtgga	tcacgaggtc	aggagttcaa	gaccagcctg	35760
gccaacatgg	tgaaccccca	tctctactaa	atatacgaac	attagctggg	tgtgggtggc	35820
cgccgctgta	gtcccagcta	ctcgggaggg	tgaggcagaa	ttgcttgaa	ctgggaggtg	35880
gaggttgca	ggatctgaga	ccatgtcatt	gcactccagc	ctgggcaaca	gagcgagacg	35940
ccatctcaaa	aaaaaaaaac	ctacagagcc	acacggcctc	tttctccacc	gagtgttggt	36000

10

20

30

gtgggagctt	gtgttattgt	ggtgaaatct	tggtactttc	ttgaggcaga	gagaggctga	36060
gcgcctggag	agactttcac	atgggtcgcc	atgtccgcgc	tcgggttcgc	tggtgtgctc	36120
cccatctgaa	ggctgggtgc	gtccagacag	gctggacgc	cctttccacc	agatccctcc	36180
ccccgcagca	gtttctagtt	acgttgtact	gtgaggtctg	tgctcctggg	tgatggcaaa	36240
agtcagccga	attgaaatcc	agagccatgc	ctggctccct	ggagcttctc	tcctgggcag	36300
ctgtgatcat	tgccctctgc	gtgggtgtgg	tggtggaaat	ggattccctt	catcttgctt	36360
gctacagggt	actgtcacgt	ggagtccttt	ggagagaggg	acgtgttaat	tgatggatgt	36420
ggctcccatg	ctgagaaagc	tcctgggcgt	acattgcctt	agagtttcat	tggagctgcg	36480
ttcttttatg	gtgtctgcta	ggcagaagtg	atgaagactt	ggaagaaaac	ccagaagggt	36540
ttccacttaa	tttggaaaat	gtgcttttcc	cctcctgtgt	cttttgctaa	ggtccagcct	36600
ctgcagcct	cccgcctctg	tggactctgg	ctttgattct	ttattaggag	tcctccctgc	36660
cccccaaaag	atgggtgcta	aattatcctc	caattggcgc	agggttttgt	ttctattaat	36720
tggttttatt	ttttatttgt	gtaaatttat	ataacataaa	atttgccatt	ttaattgttt	36780
tggtattgtt	gtttttgaga	caggggtctc	ccccagtgcc	caggctggag	tgcatgtgtg	36840
cgatcatggc	tcactgcagc	ctcagcctcc	agggtccag	tgatcctctc	acctcagcct	36900
ctctagtagc	cgggactaca	ggcatacact	accacatctg	gctgattttt	tgtatttttt	36960
ttttattgta	gagaccgcct	atgttgccca	ggtcgtctc	aactcctgga	ctcaagccat	37020
cctccacact	caccctccca	aagtgtctgg	attacaggca	tgagccacaa	caccagacca	37080
ttttaatttt	tttttttttt	tttgagatgg	agtctcactc	tatcgcccag	gctggagtgc	37140
agtggcgtgg	tatcaactca	ctgcaacctc	tgccctccag	gttcaagcga	ctctcctgcc	37200
tcagcctcct	cccagtagtc	tgggattaca	ggtgcccatc	actatgcctg	gctaattttt	37260
gtatttttta	gcagagacgg	ggtttcacca	tggtggccag	gctggctctg	aactcctaac	37320
ctggtgatcc	gcccgccctc	gcctcccaaa	atgctgagat	tacaggtgtg	agccaccgtg	37380
ccccgccttt	ttttgttttt	gagacagggt	cctgcctctg	caccagact	ggagtgcatt	37440
ggtgggctct	tggctcactg	cagcctccgc	ctcccaggct	caagttgtgc	acctccacac	37500
ctggctaact	gtatttttat	tagagacaga	tttcaccatg	ttgccccagg	tgggcttgaa	37560
atggactcaa	gcagtcacc	cacctcagcc	tccc aaagt	ctgagattac	aggcgcgagc	37620
caccgcaccc	agcccatttt	acctattctg	cagttgacag	ttcagtgcca	ttcagtcagt	37680
tcacgaggta	accatcactg	ccattcatct	ccagactact	tcaccttctc	ggcagatgtc	37740
cgaactgttc	cgcattgaac	acactcctca	tctccctctg	acagccacca	ttctactttg	37800
tatctctctc	tgcctctctc	aggtaacctc	tgtaagtggg	attataccaa	tatttgccct	37860
tggtgtactg	gcttctttca	tgtgacatgg	tgtcctcaag	gttcatctgt	gttatagcct	37920
gtgtcagaat	ttccttccct	aaagcctgaa	taataaccgc	ttgtaaaggc	tgggcgcggt	37980
ggctcacacc	ctctaactcc	agcatttttg	gagtcogagg	tgggcagatc	acttgaggtc	38040
aggagtgtga	gaccagcctg	gccaacatag	tgaaccctg	gctctactaa	aagtacaaaa	38100
ttagctgggt	gtgggtggcg	gcacctgtaa	tcccagttac	tcaggaggct	gaggcaggag	38160
aatcgcttgt	accggggagg	cagagggtgc	agtgaaccaa	gattgtgect	ctgcagtcca	38220
gctgggttaa	cagagtgcga	cttctctgtc	caaaaaaaaa	aaaaatcatc	ggatggatgg	38280
acggaccact	tcttgttatt	tatccatcca	cggtgtctag	gtttcttcca	cctttggttg	38340
tcgtgaataa	ggccactatg	aacatttcc	tccgtggtga	agggtttgtg	ctagtgcagg	38400
aaaggcgtgt	ttgtggtgtt	gcataggatt	ctggttaagaa	agtttgcact	aaccataagt	38460
atttgtaact	cattaaaatg	aaagctcagg	ggcggggcgc	ggtggctcac	gcctgtaatc	38520
ccagcacttt	gggaggccag	ggcggggcga	tcattgaggtc	aggagatcaa	gaccatcctg	38580
gccaaacatg	tgaaccctcg	tctctactaa	aaataccaaa	aaactagcca	ggtgtggtgg	38640
cgggcacctg	tagtcccagc	tacttgggag	gctgaggcag	gagaatggcg	tgaaccgggg	38700
aggcgaggct	tgcggtgagc	cgagatcgct	tcactgcact	cgagcctggg	caacagagca	38760
agactccgtc	tcacgcaaaa	ctctgtctca	ogcaagactc	cgtctcaaaa	aaaaaaagag	38820
ttcagggttt	atgaaactgg	ccagccgcgt	aaagtttgct	gtgttgtttt	tgtgcccggg	38880
aggagtgtgg	ccagggtgtc	acgtcacaca	gtacacgttt	ctcagatggg	ggttctccag	38940
actgctgtcc	caaagtctgt	ttttgcatct	ggttcccaca	gacccaccct	ccacgggtgag	39000

10

20

30

cctgattttt	gccagggtag	ctggaatctt	gcttgtcttt	cagcccgga	gctgtaccag	39060
tccagggtoc	acagctagtg	gcttttagga	aggaatttgt	tcagttggct	ttgacacatg	39120
gccccctagg	gtccacagct	ctgtagtgat	gtggatgttg	ttatctacaa	agacacatga	39180
tccttcgtgt	ccagatgaaa	gtgatgatgt	ctttgcagct	gcccagcaag	gctgtgtgtg	39240
tgtgtgtgtg	tgtgtgtgtg	tgtgtgtgtg	tgggtgtgtg	gtggtgtgtg	tgtgtgtatg	39300
ggggagggag	gcaccctttc	catctggggg	tgtgtgtgtg	tgggtgtgtg	gtgtgtgtgt	39360
gcgcgtgtgt	gtggtgtgtg	gtgtgtgtgt	gtgtatgggg	gaggcaccct	ttccatctgg	39420
gtccaagaga	ctgggcctgg	ggaagacgct	tctttttatc	tacttagaga	ctttgtttta	39480
tttgtatttt	tttgagacag	ggctctcactc	tgtcaccag	gctggggtat	ggtgatatga	39540
gcatagtctc	ctgcagcctc	ggcctcccag	gctgaagcga	tcctccacc	tcagccttct	39600
gaatagctgg	gactgtaggc	gtgcgtcacc	atactgagct	attgtttttt	ttgtttgggt	39660
ggtttaattt	tttttgatac	agatggagtc	ttgctatggt	gcccagacta	gtctcaaaact	39720
cctgaactca	agtgtattctc	ccacctcagt	ttcccgacat	tctgggatca	cagggtgtgag	39780
ccactgtctg	ctccctgttt	tattaactgc	tgaagagacct	agataaagaa	agtctgaaaa	39840
gacttactat	cagagcacca	tcctaagatg	attccctctg	actcaatgga	gagggagggg	39900
agcttttctt	tcaggcctgg	gtggcaggag	cccagggtgt	ccaggcccca	tttgccccag	39960
gccaatcac	tcgggaactt	ggatgcagct	gtctttcagg	gtaaccctaaa	ggaaccagat	40020
ccccgcaggc	agtaggcttc	tgggctgtcc	tctcctccta	cgtcagctca	gtaagagccc	40080
ttcgaaggga	tgtgtgtctg	gaggcccca	aagccaggc	tcctccctga	gatgcacagg	40140
gtgggctggg	cttaggcagc	gctcagcat	ctcctggacg	gtgacccag	agagtgtgga	40200
gacggagagt	ccttgagagt	cactgagaga	cgtggctgcc	ctgccttccc	aagaggggct	40260
ctgagtcatt	ccccacactc	acctgcccct	accacccctc	acctggcccc	cagcctcacc	40320
taccccaca	tctgtacoga	tccttttacc	cgcaccttcc	ctacccacc	tcacctcccc	40380
tgtaccttca	cctccccac	tcaccgcgcc	ctgcacctc	acctgtcccc	caccttcacc	40440
taacccccac	cctcacctgc	cctccctca	cctggcctcc	ttcctgtggg	gaaggggttg	40500
taaggggagg	ccccaaaact	gtctgtcctg	gtgccctgca	gagaaaacag	tacgtgaggg	40560
cgcagtccta	aaagcttgag	tcctggaagg	tggaggagac	agggatgtgt	tgggaagggc	40620
cccatggctt	tggatccctt	ctcgactgtc	aatggggcct	tcatgggagc	gccagtctag	40680
tgatgcacag	ctgggtgccc	ggcgggtggc	tgaggaggcc	taaagtccga	ggcggcaaga	40740
gctcttccag	aggctgttgt	cctaactcgt	ctggcatact	caggcgggca	cgtagttagg	40800
agctgattgg	agaggagaga	ccccacacc	aatactggga	tttgactttc	aggctaaact	40860
tgagaagtgt	ggcctctgct	gtcctgccag	agctctccag	ccagtgccca	gggctctcca	40920
gccagtgcct	gggggtctcc	accagtgcct	gggggtctcc	gccagtgcct	gggggtctcc	40980
ccagtgccta	gggggtctcc	ccagtgccta	ggagtcttgg	tttctttgtc	ttacagccct	41040
ttgttttgac	ctctctgagc	caaggccaaa	accagacag	gcagccccc	gacctcagca	41100
tcgacatcta	cagccggaca	ctgttcttga	cgtgcgaggc	caccaatacc	atcaacgtcc	41160
acaggctgag	cggggaagcc	atgggggtgg	tgtgtcgtgg	ggaccgcgac	aagcccaggg	41220
ccatcgtcgt	caacgcggag	cgggggtagg	aggccaacgg	gtgggtgggg	gtgctgccc	41280
tccaggcgtg	ccgcctgtgt	cttatgcoga	atgccagcct	ctcacaggct	ggggagactt	41340
tccacctggg	gatccaatgg	gtggctttcc	agggtcccaa	aagcaaacac	agggttttca	41400
cagcccgctc	gggaaagcag	aaagcccca	ggggctggaa	ggggaaagg	ggagctctgc	41460
tgagagggtta	caaggcagcg	ctggccgacg	ggagtgcag	ttgatagggt	ttgtatcatc	41520
cttgttaaac	ttgaacctg	tgcagaaatc	ccttccacgg	catgggggct	gcctgttgac	41580
tcgctcctgt	tccaccacag	ggagctcctg	ggcttcttcc	tcccagaggc	cccgcagcct	41640
cccacctgtt	ggtcgtcaga	gcttctggtt	ggtgggaagg	caccagagac	cttgagggtc	41700
ccagagagaa	aagccaggga	aagagggaga	ccgaaaccca	tgtgacatga	aactcaggct	41760
ccaaactgag	cacgggaacg	tttggggaca	ggagcgcgat	ggccttctcc	agatagctgg	41820
ggggctggca	tgaagacggg	agctacagcc	agcacaggct	ctgggcccgg	agcccagaga	41880
ttgagccctg	actctgtcac	ttactggcca	cgtgaccttg	ggcgggtggc	atagcctctt	41940
ggagactcag	tttctcatt	ggtaggagtg	acggccacag	tgggtcggcc	tctgcagcac	42000

10

20

30

acggggggct	eggtggggcg	aagccccggg	totataaggg	ggctgtgcag	gagccagccg	42060
agctgggtctc	ccaacagcca	gggctccggg	gtcccttagca	gctgtggggg	gcctgcacct	42120
gtttcccatg	gctgtgtgca	gaaattacca	gaagccagggt	ggctgagagt	aatggacact	42180
tgttctctca	cagttcctga	gggctgaagc	ccgagatcga	gggtgtgggca	gggccctgcg	42240
ccctctgaag	gctctgaggg	aacctttggg	cttctgggtgg	ctccaggcac	cccttgactt	42300
gtgggtcctgt	cactccagtc	tctctgtctg	gctgcacatg	gcgtggcctc	ttctgtacca	42360
ttgaaggaca	cttcagttgg	at tttagggcc	tacctccacc	cattgtgggc	gtatcttgat	42420
ccctcatgac	at ttgtaaag	acctgtcttc	caaataagct	cacattctga	ggttctgggg	42480
tgagcgggaa	tttgagagag	attgttcaac	tagtatagaa	tgtgacctgt	cagcctcggg	42540
cagccttgag	aggcaggggc	tttccacagc	ccagctgggt	gccctgggct	ccgtgtctgtc	42600
cgaggagacg	ccatccccac	acccgtcctt	caccgcgccac	cctcccgag	gtacctgtac	42660
ttcaccaca	tgaggagacc	ggcagccaag	atcgaaacgcg	cagccctgga	cggcaccgag	42720
cgcgaggtcc	tcttcaccac	cgccctcatc	cgccctgtgg	ccctgggtgg	agacaacaca	42780
ctgggcaagc	tggtctgggt	ggacgcggac	ctgaagcgca	ttgagagctg	tgacctgtca	42840
ggtaacgcgc	ccggggcctg	ccctaaccgc	agacaccgcg	ccttcattgt	cagtaatggc	42900
agcagctgcc	acattgtccg	agacctgccg	tgagccaggt	gcgcgcgcag	gggctttgtg	42960
tgtagcgtgt	tttgtctca	cactgacagc	tgtaggctgg	ggttctgagt	gagccccaca	43020
gggcagaggc	agaaaatgag	tctcagagag	ggtgagcgag	ctgcttgggg	ccccacagca	43080
ggagatggag	caggactgca	gcctagcctc	tgccccagc	acctgcgcaa	gaagctgtctc	43140
tgctctggac	tgtgttaggc	tgcgagggct	ggagagaaat	gagagtgggt	gcttagagag	43200
ggggcgagg	tccccatggc	ttttctctct	atgatgaggt	agatgggtga	agggaggggc	43260
catgcttgca	ggggccagtg	accgagggcc	gccgttggaa	ctgatggcct	tcaccccgag	43320
ccagcccccag	gtgggagcag	ggccttccga	gggcttctct	tgggtcgcc	tgcttccagg	43380
gactctgctg	cagctccac	ccctgtccaa	agcatggaat	ccccaggct	ccctggcagt	43440
cctgtcaacc	tctgtcctcc	caagctgagt	gtggggcaag	ttctggaggt	cagcactgct	43500
cagggggggc	cacgggctgc	ttgcaggggc	caaccgcctg	accctggagg	acgccaacat	43560
cgtgcagcct	ctgggcctga	ccatccttgg	caagcatctc	tactggatcg	accgccagca	43620
gcagatgac	gagcgtgtgg	agaagaccac	cggggacaag	cggactcgca	tccagggccg	43680
tgctgcccac	ctcactggca	tccatgcagt	ggaggaaagtc	agcctggagg	agttctgtac	43740
gtgggggctg	gcagtggggt	gggcaggggtg	gcctctaaac	ccgacccctg	gaggaggctg	43800
gaggccagtg	caagatcctg	tgtggcctca	gccaggcggt	ggtctctgcc	agatgccaac	43860
tgttgcccgc	tggggttcag	cgacatgtcc	gaatgtcccg	aggcctctga	ggttgttttc	43920
ttttgcccga	gaacaaatca	ccacgaacag	cgttttaaga	caacaccaac	tctttttttt	43980
tttttttttt	tgagtcagga	tcttgcctctg	ttgcccaggc	tggggtgccc	tgggtgcaaac	44040
acagttcact	gcagcctcga	cctctgggct	taattaagtg	aacaccttgc	ctcagcctcc	44100
caggtagctg	ggactacagg	tgggcaccac	cacacctggc	taattttttt	ttgtagagac	44160
gggggtttccc	catgttgccc	aggctgggtct	gcaactcctg	ggcacaagct	atctgcctgc	44220
tgtggcctcc	caaagtgtca	ggattatagg	tgtgagccac	tggcctgaca	acaccacggg	44280
attgtctctc	agttctgtaa	ggcaaaagtc	aggcacagcg	tggctcacct	gggttctctg	44340
ctcaggggtct	cacggggcca	gaatcaagggt	gtcaggaacg	ctgggcccctc	agcggaggct	44400
ctgtggagaa	attagcttcc	ttgtcactc	agcaggtagc	agttgtggga	tcgaggttct	44460
gttttctctc	tggttatttg	tcggggacca	ctctcagctc	ctagaggcca	ccacaggctcc	44520
ttgcccctg	gccctctctg	cctcagcagt	gggggctccc	tgcgtcagtc	cctcccacac	44580
cctgagctctc	tctgatttgc	ttctaaaggg	ccctgtgatt	cggctcagcc	acctttagat	44640
taggttagct	tcccccttga	tagactccaa	gtcggtgat	taataacctt	aatcacatct	44700
gcagaatccc	ttctgccaca	taaggtoatg	acgccgtgct	ggggactggg	gtgggaaatt	44760
acgggggtcat	ttaggattct	gcctgccact	gccttgcctg	gtcccagggc	ttgggggagg	44820
ggcctccaca	gctgggacca	cagtccttcc	tcccctccat	ggtaaccatc	tgaggattac	44880
ttgagaccag	cctgggcaac	atggtgagaa	cccatcccta	caaaaaatac	aaacaaaaag	44940
ggaccaggct	gggcttgggt	gctcatgcct	ataatcccag	cactttggga	gaccaagggtg	45000

10

20

30

ggctgatcac	ttgaggttgg	gagttcgaga	ccagcctgcc	caacatagtg	aaatcccgtc	45060
tctaactaaa	atacaaaaat	tagctgggtg	tggtggcagg	cgctgtatt	cccagctact	45120
ggggaggctg	aggtgggaga	attacttgaa	cctgggaggc	ggaagttgca	gtgagccaaa	45180
attacgccac	tgcaactccag	cctaggcaat	agagtggagc	tccgtctcaa	aaaaaaaaaa	45240
gggcccaggg	tggtagtgc	aaagagaccc	tatcccaaaa	aaaccgaaca	ctgaatcctt	45300
gagactgagt	aaggacactg	tgaatttttt	ctgggtgggg	cagggaacag	agcgtcttct	45360
gtcatttctt	ccacttgggt	gtggtcagct	ctccctccaa	gctgcctcct	cttctctctca	45420
ttgtccgggt	gttgacacac	tttggttaac	tggatagaat	aacgcgagtt	cccagggaact	45480
tggtccattt	gctattttat	tttattttat	tttattttat	tttattttat	tatttatttta	45540
tttattttat	tattgagatg	gagtttctgt	tttgtgcgcc	aggctggagt	gcagtggcgc	45600
gatctcggtt	cactgcaacc	tctgcctccc	aggttcaagt	gattctccta	cctcagcctt	45660
ccaaagtaact	gggattacag	gcacccacca	ccataccagg	ctaatttttt	tgatatttta	45720
gtagagacgg	gttttcgcc	ttttgccag	gctggctctc	aactcctagc	ctcaggtgat	45780
ccacgcacct	cggcctccca	aagtgtctgg	attacaggca	tgagccacca	cgcctggcac	45840
catctgctat	tttaattccc	atgtgtatta	gtgtcccacg	gctgctgtaa	caaatgacca	45900
caaactggat	ggcttaaagc	aacagaaatg	gattccccca	atgtgctgga	gaccagaagc	45960
ctgcgaccaa	actgttggga	gggtgtgtgt	tctctggggg	gctccaggga	ggatctatct	46020
gttgggccctt	ccagtgtgtg	gggtgccagc	gttccacact	tgtggatgcg	cgcctcaaac	46080
ctctgcccct	cttcatgtgt	ccatctcctt	tgtgtctgcg	tctttacctc	ttcttcttgt	46140
ctgtgttgcc	tcttataaag	acgtttgtca	ttgggttttag	ggcccaccca	aatcatocga	46200
gatgacctcg	tcttgagatc	cttaacctgc	aaagaccctt	tttccaaaaa	aaggttatgc	46260
tcacagattc	taggccttaa	gacatgggtg	tatctttctg	gggggcacta	tccaaccctt	46320
tatacaatga	aagacgggaa	gagggccagg	tgtggtagtt	cacgcctgta	atctcagcac	46380
tttaggaagc	tgaagcggga	ggatcacttg	agcccaggag	tttacaagta	gctaggcaac	46440
atgatgagac	cccatttcta	caaaaagtga	aaaaaaaaaa	aaaaaaaaaa	aagccagggtg	46500
tggtggctca	caactgtaat	cccagcactt	tgggaggctg	aggcaggcag	atcacgaggt	46560
caggagattg	agaccatcct	ggctaacacg	gtgaaacccc	gtctctacta	aaaatacaaa	46620
aaattatggc	cgggocagct	ggctcccggc	tgtaatccca	gcactttggg	aggccgaggt	46680
gggtgaatta	caaggtcaag	agatcgagac	catcttggct	aacacgggtg	aaccccatca	46740
agatcacaa	gtcaagagat	ggagaccatc	ctggctaaca	cgggtgaaacc	ccgtctctac	46800
taaaaataca	aaaaattagc	cgggcatggt	agcgggcggc	tgtagtccca	gctgctcggg	46860
aggctgaggc	aggagaatgg	cgtgaaccgc	ggaggcggag	cttgcggtga	gccgagatcg	46920
ctccatgcca	ctgcactcca	gcctgggtga	cagagtga	ctccgtctca	aaaaaaaaaa	46980
aaaaaaaaaa	aaaaaaaaag	aattagccag	gcacagtggc	agggtgcctat	tgtcccagct	47040
acttgggagg	ctaaggcagg	agaatggcat	gaacccggga	gggtggagttt	gcagtgaacc	47100
gagatcatgc	cactgcgctc	cagcctgggc	gatatagcaa	gactctgtct	caaaaaaaaa	47160
agccaggcat	ggtggtgcat	gcctgtagtc	ccagctactc	aagaggetga	ggcaggaggg	47220
ttgttcgacc	cacggagatc	aaggctacag	tgagccatga	tgcacccact	gccctccagc	47280
ctgggtgaca	gagtgtgacc	ctgtctcaaa	gtaagtaaat	aggaggagag	acaagtgggc	47340
agttcagact	gatggtatgg	gcacagtaga	gactgggtga	gacaggetgg	cctgtgatgt	47400
caagcaactt	ctgtaatgtt	ttccggcatc	catctgtgtg	tcaatttccg	tgtcagtagg	47460
aagactctgt	aggctgcca	gaggaataag	tgggaggatc	ctcccagaga	ggccgggcct	47520
gcaggagggc	cagttctcat	gagttctcat	ttggccccta	ccctccaggc	tgtggttctg	47580
agggtggaga	cagagcctga	cctctgtttg	tcttgttttg	tctttgcagc	agcccaccca	47640
tggtcccgtg	acaaatgggtg	ctgctcccac	atctgtattg	ccaagggtga	tgggacacca	47700
cgtgtctcat	cccagtcaca	cctcgtgctc	ctgcagaacc	tgtgacctg	tggaggtagg	47760
tgtgacctag	gtgctccttt	gggtgatgg	acaggtaact	gattctctgc	ctgctaggct	47820
gctgcctggc	atccttttaa	aatcacagtc	cctgtggcat	ccagtttcca	aagctgattg	47880
tgtcttctct	tgcctcctt	tctttctctac	tatgtgcatt	cgggtgctatg	aattttcctc	47940
taagtactgc	gtttcctgca	tctcacaat	tttgttacat	tttcattttc	aggtagtttg	48000

10

20

30

aatattttta	cacttctcct	gagatgacat	ctttggctca	tgtgttattt	agaagtgttg	48060
cttagtttct	aaagagtgtg	ggcttttcca	gctgtctctc	tgcaactgat	ttctaattta	48120
attctactgt	agtctgagag	cttattttat	atgattttctg	ttattttaaa	tgtgttgggt	48180
gtggtgtttt	tgtgtttatt	gttttttgtgt	ctttttgttt	tgttttgcct	cgtttgtttt	48240
gtttttgaga	cagtgtcttg	ctctgtcact	caggctggag	tgcaatggcg	cgatctcagc	48300
tcacccgaac	ctctgcctcc	cgggttcaag	tgatcctctt	gcctcagcct	cctgagtagc	48360
tgggattaca	ggtgcacgcc	accataccca	gctaattttt	gtatttttag	tagagacggg	48420
gtttcaccat	gttggtcagg	ctggctctga	actcctgacc	tcgtgatccg	cccacctcgg	48480
cctcccaaag	tgctgggatt	ataggcgtga	gccactgtgc	ctggccatta	ggtgtgtttt	48540
atcaccagc	atcatgcagt	ttatcttggt	gaatgttctg	tgtactcttg	aaaagaatgt	48600
ggattctgct	gttgttgggt	ggagtgttcc	agaaacatca	attagatcca	gttggttaat	48660
agtgtctcatc	aggttgtctc	tatccttcc	tcctgactgc	ctgcttgagc	tgtcagttat	48720
tgacaggggt	gtggagtctc	caactcta	ggtggatttg	tttatttctc	ctagtagtcc	48780
tatctttttc	tctccttcta	cccttgatcc	tcttctcccc	ctagggtctc	ctgggtgttag	48840
tgggtgggaga	gtgggttagt	gaagaacctg	gacttttaggg	ccaaagaggg	caggggttcaa	48900
atcctgggctc	tgtcacttcc	cagttgagtg	accctggctg	gtgcctgaat	ctctgtgagc	48960
ctccacttcc	tcctctgtga	aattgagagc	acttacctgg	caggctgtca	tgggcatcaa	49020
gtaacagggc	actccacctg	gacctgaca	cgtgatgcac	aggaatgcca	gctgctatgc	49080
catgggtgtg	gcagtagtaa	taaagtgacc	atctgtatcc	tcaccacagt	gaagcctgtc	49140
cagggctttc	tctcctatgc	ccccatgcct	ccaggtggcc	ttggatcctg	ttggttctgt	49200
gctctgctca	gcgaccttcc	tccctgggga	gttccctggg	gttcagcttc	atcctacaga	49260
cagcagcaca	cactggctgt	gcacctttt	tttttttttt	tttttttttt	tgagatggag	49320
tctcgtcttt	ttcgcgcagg	ctgaagtgtca	gtggtgtgat	cttggtctac	tgcaacctct	49380
acctcctggg	ttcaagtgt	tttctgcct	cacctccca	agtagctggg	attacaggct	49440
cccaccacca	cgcccggtca	atttttgtat	tttcagtaga	gatggtgttt	cacctgtttg	49500
gccaggatgg	tcttgaaact	ctgacctcag	gtgatccgcc	caacctcagc	tcccaaaagt	49560
cagggattac	aggcgtgagc	caccacaccc	ggagtgcggg	ttgttttttag	cagtttctct	49620
tgttcctgga	gagactggct	cctgcccagg	agctcgggga	gtagggccgc	ggggtgtgct	49680
ctcacacctc	gagtttggcc	gtaagcagag	gggacatttt	gtgactgtcc	ccctcctgag	49740
cttcccagca	gcttttctcc	aagttacagc	ccaaaagctc	aggtggattt	gcaacccaac	49800
ggtgtctgtg	cacctcccac	tgatgcccga	actgcctcgg	ccaagaaacg	gggcctgcag	49860
aacgctgcac	taactgcagc	cttgggcctc	catgcccagag	gccatgccct	tccatccacc	49920
acccctgggc	ctgggcctcg	ggcctcctcg	gctcgggaac	tccaggcccc	ttcctcacgg	49980
ctcgagagac	gtgtattttac	cgcacagggtg	cttgtcattc	tcttgtggcc	tcttctccag	50040
ggagatcaca	gaaggacagg	gcctcactga	ggtctcggac	atggaccctt	tgatagtggc	50100
aggagccagg	ctgggcaaga	ggcggccaca	gtcacctcag	cagtgccatc	accaccgcca	50160
ttcagccctt	ccctgagccg	ggcgccgccc	tggtcttggt	cccagtgctc	cagttacagc	50220
tcacaggagc	ttgtggtgcc	cagcggctgc	ttctgattga	gagtcgaggt	cggaggcttt	50280
gggaggctga	gaggctgctc	ggtttcacaa	ctgctgaggg	agacttgggc	tccatctcag	50340
gtatgcccc	tgtcgcctc	aacctccagc	caccggtoct	ccgtgtcccc	catggccagg	50400
cacggcttgc	agacatctgt	cgttggetcc	tctcagccgt	cgtgggctga	ccctggcaag	50460
tcctcctgtg	gctgagccca	gtggggacag	ctgcttccct	ttattaccct	agaactctcg	50520
tctttgatca	ggccccctcc	cctatgccac	acagtccctg	tcactcgggt	gagcccagta	50580
gtcatgggga	aggcctgcgg	gttccaaaca	tccaaaggct	tgcgtgcagc	atgacagctt	50640
gaaaccgatg	ttttttacct	tgatcagatt	tcagcttggt	gggggctttg	ctcagcttcc	50700
agtgaggcct	gggcccattt	cccagcatcc	cctcctgagg	ccagcctctg	tttccctgtga	50760
ttttctgcac	aaagtgggag	ggaggagtcc	taggaaatgg	ggggccacct	cgaagcctag	50820
gcctcctctg	gcttctctgt	gccagtgcct	ccacgctttg	tgtctgtgtc	cccagcccat	50880
gggactctgc	tattccctga	gtgctgccgc	atgccagccc	cgcactgagg	acgtggagcc	50940
ccgaggggca	ggatggcctc	catggtcaca	cgtagggaagt	ggcctccacc	ctccgatgat	51000

10

20

30

cctctccctc	ctccctttca	gcgcctctcc	cgggggtgtc	ctcagccctc	ctgcctgtgc	51060
tttgtccctg	cttctgcagg	gcctctggac	gtgctgacag	gtcctctgce	ggctcctgce	51120
ttgctatgct	cacgtctggt	accacagagg	cctggccctt	cttctgtagc	agtcccaac	51180
ccgcaacagg	tgtggctgct	gaccacctgc	tttctgcccc	tctggctcctg	aggaggcgcc	51240
agtgggcact	caggcgtggc	tgagcagatg	tgtgttgccg	ggaggaggaa	ggactgctcc	51300
agtcagggtc	gaatttccca	cccgagcat	ttctgtgtga	tttgggttag	cgctgctgc	51360
ttaaagctct	gattcccaat	tggcaccctt	tccctctctg	attgaaaaac	atacggatgc	51420
atgtcttctt	gcagtgaatg	tgtattctcc	cagcctctct	tctgggttgg	ggctggagggt	51480
ggagcggcac	acaggagccg	cagcagtgga	ggatgtgcgg	gtgcagcacc	ccgtacagca	51540
gggatgccaa	accgcgctg	agtcctctct	aactctgct	ttgaagccca	gtcacgccat	51600
tgcctgggtt	ttgctggggc	gggctgctgt	tgatgttctc	ctctgtccct	ccccagagc	51660
cgccacctg	ctccccggac	cagtttgcac	gtgccacagg	ggagatcgac	tgtatccccg	51720
gggctggcg	ctgtgacggc	tttcccgagt	gcatgacca	gagcgacgag	gagggtgccc	51780
ccgtgtgctc	cgccgcccag	ttccctgctg	cgcggggtca	gtgtgtggac	ctgcgcctgc	51840
gctgcgacgg	cgaggcagac	tgtcaggacc	gctcagacga	ggcggactgt	gacgggtagg	51900
ccctccccgt	caaggctctg	ccaagaccct	ggccctgccc	tcggggatac	gagcttgggg	51960
ctgcctccgg	cctcacagga	gtaggggtct	tgaaaaacct	tgcttgccagg	gagattgcca	52020
agtctgtctt	ttaggcccaa	caaggaaaac	tctgcagtct	cacctatcct	gtcccaccag	52080
gtagtgtggc	ttgaaggcag	actgtgaggg	tctatctcac	cttctctgcat	taggtcagga	52140
gtttcacaga	aacctgaggg	acattcaggg	gtgggctgca	gagggtccatg	gctcacaccc	52200
tggaaaatcc	gcccccaaaa	gacagtgtct	tctccactga	ccagtctgtg	ggatagtgtc	52260
taagcctgag	tggtttctat	caacatgtag	aatcaggagg	tataaagaga	tttgcctcagg	52320
catcctgggc	cctctctgac	cagcaggatc	ttcctttaga	tcttgacagt	gaaacacatc	52380
tcttctgtgc	cccctgtgag	ttttctttca	ttcattcatt	cattcattca	ttcattcatt	52440
cattcattcg	agacagagtc	ttgctctgtc	acccaggtct	gagtgcctctg	gtgtaatctc	52500
ggctcactgc	aacctctgcc	tccagggttc	aatcgattct	cctgcctcag	cctcccaggt	52560
agctgggatg	acagggtgct	accaccatgc	ctggctaat	tttgtatttt	tagtagagac	52620
aggggtttac	catgttggcc	aggtgtgtct	cgaactcctg	acctcagggtg	atccgcccgc	52680
ctcagcctcc	caaagtgtct	ggattacagg	catgagccac	cgcgcccgcc	ctgagttttc	52740
cttttatgaa	ggacctgctt	ggttgggttc	ctgccacatg	ttgtcagcac	catgggcccc	52800
ggactgctga	ggagctgttg	atgcctctgc	tctcccagag	ccaccggctc	tgttagataa	52860
ttcacatgca	gtctggccac	tgtcctacgt	cctcattcac	aaagagcaga	catttcgtag	52920
aagatgaggg	cctgggagta	acctccctgc	atgtttttct	ataaaggcat	agtgggttaag	52980
tccttcacgc	tcattgacca	ttggagaatt	ttatggaggg	tgtagactag	gggtgtgtaa	53040
actaagggcc	caggggccaa	atccagcctg	ccacctactt	ttgtaaataa	agttttcttg	53100
gtgcacagcc	atgcccattc	attcattttg	acaatgtctg	tggctgcttt	catgccccaa	53160
gcaagagaac	tgagtgggtta	tgtctggagac	ctacggcctt	caaagcccca	gacctcacgt	53220
ctggcccttg	acagacagag	cttccccagc	cctgtctgct	atcctggccc	agcatgtgct	53280
gtgtgtgtga	tttcagcttg	caggagccgt	ggtaggaat	tgtccctgtg	ttgggtccatt	53340
ttgcattgct	atgaaggagc	acctgaggcc	gggtagatta	tgaaggaaag	aggtctgtct	53400
ggctcatggt	tctgtaggca	gcaccagtat	ggcaccgcc	tctgctcagc	ttctagttag	53460
gtctcaggaa	gctttgactc	atggtgaaag	tcgaagcggg	agcaggtgca	tcacatggtg	53520
agagagggag	caacggagag	agagagagag	gcctctctcc	tcttgccctc	accttgagag	53580
gagatgccag	gctcctttta	gtaaccagct	cccatgtgaa	ctcacagtga	gagcccattt	53640
gctactgcgg	agagggcacc	aggcatctgc	tcccatgacc	caaacactgc	ccaccaggcc	53700
ctacctccaa	ccttgggggtc	atattttatt	ctgttctatg	ctatgctatg	ctatgcatg	53760
ccatgccatg	ccatgctatt	cctattctat	tatttgagac	agaatctctg	tctgttgccc	53820
aggtggaggt	gcagtggcat	gatcttggct	cactgcaacc	tccacctccc	aggttcaagc	53880
gattctcctg	tatcagcctc	ccgagtagct	gggattacag	gcacacacca	ccacacccgg	53940
ctaatttttg	tattttcaat	agagatgggg	tttcaccatg	ttggccaggg	tgggtctcaa	54000

10

20

30

ctctctggcct	caagtgatec	acctacctcg	gcctcccaaa	gtgccatgat	tacagatgtg	54060
agtcactgog	cccagtgagg	gtcacatttc	cgttgagatt	tggaggggca	gacgttggag	54120
ccatctgagc	cccctcgtdc	cgctctagct	tctctctccg	tgtgccccgc	ggtgctgggtg	54180
gcaggccctt	acgocggttc	tggctgcatg	ctctgttcca	gaagctttct	tccctgcttg	54240
gttaccagaa	aatcatccca	tccattacaa	ggacagggtc	cccttatctc	ccattcccag	54300
ggcaggacac	cgggggcagg	gcaggtgggg	aactgagcaa	gttctctggg	ggcaggcgtg	54360
gctatggctc	cctctgggtg	ggcgtctggg	gaggggtgga	ggcagccgtc	agcgcctggg	54420
cttgcctctc	ctccctggcc	agagactgtg	gccttgtgct	gctcccgtgt	gggctgcctg	54480
cacctccagt	gggttgtgct	ccctccctc	ccctccctc	aagctctgct	gagcaccact	54540
gccttccaca	gccccactc	tggggaggcg	aggtccctcg	tggccattcc	tgtccttggc	54600
acccaccccc	ccaccaacct	ggtagagcct	tgggcggggg	ctgttactcc	ttgcatggcg	54660
tagacctccc	cacagtaggc	acctgacaca	tacctcctgg	ggggcaggca	ggagggtcgt	54720
tgaggctctc	gccccggcag	tccctccctc	gcgtggcata	ggcctcgcca	cagggtcatc	54780
gagggctggg	ggagactgta	ctagaccact	ccccgctggg	cctagaaagg	gtcccatctg	54840
tctgctctct	gtttggagtc	cagaccttgg	ttgctgtgcc	ctgcatggtg	ggctgggggg	54900
caccttccag	cctctctgag	tgcattggcct	ctccttgcag	ccatctgctc	gccccaccag	54960
ttccgggtgtg	cgagcggcca	gtgtgtcctc	atcaaacagc	agtgcgactc	cttccccgac	55020
tgtatcgacg	gctccgacga	gctcatgtgt	ggtgagccag	cttctggcac	ggggaagggg	55080
cgtccgggct	gggttcccc	aggaacgtgg	agtttagggg	aggagacgtg	cctttccagc	55140
ggggctgggg	gctgtgtggg	agactcaggc	ggctgggagg	ctccttgcgg	gaggcagggg	55200
agcctttccc	agggcagcgg	ccaggaggac	agactgtgag	ctgtgggctc	ggcggtaca	55260
gagtcctgct	cagtggcgcg	ggctgatggg	gtccagggtc	ctgcagcacg	caccaccca	55320
cgggaccttg	ctgagcagcg	tctgtcaggc	agcaagatta	cccgagggtc	gcagtggctc	55380
tgttccctgg	cagcttactg	tctggctgag	gaggagtgtg	gttcacatat	gcacacatgt	55440
catgtgcaca	cacatgtaca	tgacaacatc	ccacatgctc	ctcaaatagc	atgacctgta	55500
cagtcacgga	tatagggcct	aggggatagg	aggccaagac	agtcagggaa	gactttccag	55560
aggcagtggt	tccctgaaagg	ctgtctgatt	caggcaggaa	gggagctgag	ttcagatagg	55620
aagtagcaat	gagtcattgt	gtctggggac	atggccactc	cttcgctgca	gagggacctg	55680
ggctgagagc	tccctctctta	tggctgcagt	cgggagagaa	gtctgttggg	gggagaaggg	55740
ggcttccctc	agggactccc	tgtgcccctt	ggcaccttcg	tgccagggtc	ggcttgaggc	55800
ctgaaggcag	tgggtggggc	caccaagggt	cgcctcctct	gctgggcaag	ttcccagctc	55860
gacgggcttg	tgcctggggc	ccagctgtgt	ggggcgctgt	tgatgcgcag	ccaggcctcg	55920
ccgcagagc	ccgcacgctt	ccattccgct	gacttcatcg	acgcctcag	gatcgctggg	55980
cgggcctctg	gggagagtga	atgtggcttt	tgccaaaagt	gagtctggag	cctggaaact	56040
tccctatggg	cagccttgat	agtggagtgg	cccaaggagc	ccacccagcc	gaccttgccc	56100
ctcccgctgg	tgggtggcg	caccaggggc	tgcctggctt	tgtctgttca	ccaacatcac	56160
ctgggctggc	cagggcgcg	tcaattctgc	caccaccgag	ggccctgggc	gaaggagtga	56220
ataccaggct	gccttggcag	ggatgtgttg	agggctgtgg	ggagtcggac	agcggcgggg	56280
gtcagaggag	gaggagggtg	caccgtgcag	gctgaagggc	cacgttacct	tgaggttggc	56340
caggctcccc	aggcctagcc	tcccagctcc	cccaacttct	ccccacctc	caccagtggc	56400
aaagccagcc	ccttcagggc	gcacgggtgc	tgcctccaa	gagggcccat	tccgttgggg	56460
ttaatgttgg	ccacctcttt	ctgtttgtct	ctggcagaaa	tcaccaagcc	gcccctcagac	56520
gacagcccg	cccacagcag	tgccatcggg	cccgctatgt	gcatactcct	ctctctcttc	56580
gtcatgggtg	gtgtctattt	tgtgtgccc	cgcgtgtgtg	gccagcgcta	tgcggggggc	56640
aacgggcccc	tcccgcacga	gtatgtcagc	gggaccccg	acgtgcccc	caatttcata	56700
gccccgggg	gttcccagca	tggccccctc	acaggtaagg	agcctgagat	atggaatgat	56760
ctggaggagg	caggagagta	gtctgggcag	ctttggggag	tggagcaggg	atgtgctacc	56820
ccaggcccc	ttgcacatgt	ggcagacatt	gctaatacgt	cacagcattc	agcctttccc	56880
actgagcctg	tgttggcatt	cagaatcctt	caacacagag	gcctgcabtg	ctgtagcaac	56940
ccaccttttg	gcactgtagg	tgtggagaaa	gctccttgga	cttgaccttc	atattctagt	57000

10

20

30

aggacatgtg	ctgtgttgtc	cacaaatcct	catgtaccct	agaaatgaat	gtggggcg	57060
ctgggctctc	tccagagctg	aaggaatcac	tctgtaccat	acagcagctt	tgtottgagt	57120
gcagctggga	tttgtggctg	agcagttaca	attcctacgt	ggcccaggca	ccaggaacgc	57180
aggctgtgtt	tgtagatggc	tgggcagccg	caccgcagag	ctgcaccatg	ctgggtttgt	57240
tcacatgggt	gacctaggta	tgtctaagaa	ggtggagtcc	ctgtgaggtc	tgcagggtgc	57300
cccacagctc	caggccacct	tgaggattgc	ctctgcctgc	ccagccctga	gttccctctc	57360
ccctgtcctg	tcccactgtc	accccagcc	ggcctcattg	ggagcctgtt	ggatggcagg	57420
gtatagatgt	aacctgattc	tctctgggga	gcgggggttat	ctggcttctc	aagagctcct	57480
aggagcccac	agtgtgtggc	ccatcacagt	cgcagcagcc	cccagagaac	gcggccctgt	57540
ctgttccttg	cgtgtctctg	gctgccccgc	ctgggttccc	tgccccagtc	gcaggccctc	57600
tggaggagggt	accatgtgtc	tcccgtttca	cagatgagcc	ccggggagct	cactctagta	57660
gtggccagag	aggcctgctg	ctcaggggag	ggggcacatt	tccaacagga	cacaccgccc	57720
tggctctgagt	ctcgtgggta	gtggggagcag	aggagagcgc	cctatgtctg	tggggcggtc	57780
tggctgagcc	tggaaagccac	ctgacctccc	cgtctccttc	cctgccaggc	atcgcatgog	57840
gaaagtccat	gatgagctcc	gtgagcctga	tggggggccg	gggcggggtg	cccctctacg	57900
accggaacca	cgtcacaggg	gcctcgtcca	gcagctcgtc	cagcacgaag	gccacgctgt	57960
acccgccggt	gagggggcgg	gccggggagg	ggcggggcgg	gatggggctg	tggggccctc	58020
ccaccgtcag	tgtgggcccac	cggaggcttc	ccgggttccc	gggggctgtg	ccaccgcttc	58080
tgaggcatgc	ttgctttctt	cccttttcaa	acccttctgc	ttccttcttt	aatgacattg	58140
ttgatttgtg	ataatctgaa	aactacacaa	aaatataaag	agccaaaatc	tcacccaaat	58200
ccacctccta	gagtgtgtgt	tgggtccgt	cagcatccag	gcggccgtct	gtgttccgca	58260
cggcccagcc	cacagatagc	cgcctgcacc	aggcctgtct	gccctctgtg	agcctcccca	58320
cacgggtccc	tccacaaaac	ccctgttctc	ccacccagg	ctggctgtct	cctggaaaac	58380
agctggatgg	ttttgtgcat	gacagacaaa	cacagggtga	ttttcgtggc	taaaatactc	58440
cctggagctt	ttggcagggg	gaggggctgg	ctccagctga	gccacgcctt	gagtgaatg	58500
actgtgagga	gaataaactg	cgcgtgcct	ccaggatcac	tggggctggc	tggggagaac	58560
cccgtttct	gggagcacag	tcccaggatg	ccaaggcag	cttggtgccc	agatgtgaac	58620
tcttgagtgt	aaacagcggg	ggctgacttg	acatgctttg	tatgcttttc	atttgttccc	58680
gcagctgtat	gcccctaagg	tgagtccagc	cccttctgc	ttcctctggg	gcctcgccag	58740
tgagccccac	cttgctgggg	ctgggttccc	ctgcccttct	gggtatccct	cacatctggg	58800
gtcttgtctt	cttgttttct	ttttcttttt	tttttgagac	ggagtttcac	ttttgttgcc	58860
caggcttcag	tgcaatgggt	tgatctctag	gctcacccga	acctctgcct	cccaggttca	58920
agcagttccc	ctgcctcagc	ctccctagta	gctgggatta	caggcatgtg	ccaccacgcc	58980
cagctaattt	tgtattttta	gtagagatgg	ggtttctcca	tgttggtcag	gctgatcttg	59040
aactccctac	ctcaggtgat	cgcgccacct	tggcctccca	aagtgtggg	attacaggcg	59100
tgagccaccg	cacctggcct	ttttcttttc	ttttcttttc	ttttttctga	gacagggtct	59160
cgtctgttca	cccaggctgg	agtgcattgg	tgtcatcatg	gctaactgca	gcctctacct	59220
tctaggotca	agcaatccct	ccatctcagc	ccctaagtag	ctaggactgc	acgcattgat	59280
ccccatgccc	agctaataat	tacatttttt	gtagagatga	agtttcaacta	tattgcccag	59340
gctggtctcc	aactcctgga	ctcagcagat	ccctcctgct	cggcctcccc	agggtgctgg	59400
attacaggcg	tgagccaccg	tgccctggcct	gggggtattgt	cttcttatgg	cacctgactg	59460
tgggtggccc	tgggaaggaa	gtagcagaag	agggttcttc	ttgggtttct	ggacagtaac	59520
tgagtgttct	ggaggcccca	gggcctggct	ttgttttagg	acaaagggaa	ctggtaacca	59580
gaagccgaga	gtttaaacac	ccactgcct	tcttccctgc	tctgtctgct	gcaaccacgc	59640
ttaaccagcc	aggagtgtca	ggaacccaag	cagggccccc	gagcacacag	caggcagctc	59700
acgaattctc	ttttcctgtt	ctcccttggg	agctgggagg	atcttaataca	ggcaataaga	59760
gatggcactg	agcagccagc	taatttttta	aatcacttta	ttgtttaacc	atatgactca	59820
cccacttaaa	aaagggtaca	gttcagtggg	ttttagtgtg	ttcacagatg	tgtgcaaccc	59880
tcaccacagt	taatttttaga	acatttttct	gccccaaaa	gaaactctgc	atgaagccag	59940
ctgtttttta	attagcaaa	ttatttttga	tcctttaaat	atatgttcat	ggtacaaaat	60000

10

20

30

tcaaaagata	cagaagagtc	tgcagtccaa	agagactccg	cccccatgac	gccaaagcagg	60060
catccctggg	aggcatggcc	tctgcagtg	tgtttcttct	atgtccccc	aggggtcatc	60120
tgtacatatg	caagcataca	agagcgtgga	ctttgttttc	caagccagaa	gataattgta	60180
gattttatgtg	cagttgtgag	aaagagcaca	gaccatttta	tccctctgcct	ggtttccccc	60240
agtgtgtcct	gccatcttgc	atgacttcca	ttcctatcat	aagcaagaca	ctgataacga	60300
ttcttttcacc	ttatttcagat	tgacataaagt	gttttttgtt	tgttcttgag	acaaacttcc	60360
tctgtcaccc	agtgggagtg	cagtggcaca	atcacagctc	actgcagcct	caaactcctg	60420
ggctcaagcg	attctctctgc	ctcagtcccc	tcaagtagct	cagatggcag	gtgtgcacca	60480
tcatgccagg	ctaatttttta	aattttttgt	ggaggtgagg	cctcactaaa	tttctctgggc	60540
tagtcttgaa	ctcctgagct	aaagtgatcc	tccctgcctca	gcctcccaaa	gtggtaggat	60600
tacaggcatg	agccactgcg	cctgggctga	catatgtgtt	ttcgttaagcc	cgaagatag	60660
catctgaaga	gtcaacattg	agccttgcc	tttgctgcta	acgatgtata	aaagctgctg	60720
ttctgagcat	ttcggaggct	cccagctgcc	gtgtgcaccc	tgccctagagc	tctaccgtaa	60780
cccatctccg	ggaggagggtg	ctattgtttt	cctcattttg	caacaaggag	gctgaagaac	60840
tgagcatgaa	ccactggcct	gggtcgttcg	gttggttaggc	agtggggcca	ggccatccaa	60900
ctcacaca	ccttctactc	tgcttccccc	gcaccctgaa	gtttgttctg	ttttgaggac	60960
acagccgtca	cattcttgg	ggctgaacag	cactcctgtt	caggcgtggc	tgggccccca	61020
ctggaggcca	tcttggtcct	ctctcctgct	gcggttgaac	cttggtgttt	tcaaccactc	61080
ctgccaagt	gccctctgaa	agggacagtc	catcttttct	cagcagaggg	ccacactggc	61140
aaaacggctc	ctggcaccct	ttctctccac	ctgtctaata	tagagtaaaa	atggtatcat	61200
gtaagatct	tcatttatat	ttattttatc	atgaatgatg	taagcatcat	tttgtgtgtt	61260
taagaacctt	tgggcccagc	gtgatggctt	gcagctgtaa	tctcagcact	ttaggaggct	61320
gagatgagcg	gtcacttga	ggccgggagt	ttgagaccag	cctggccaac	atggagaaac	61380
cccgctctcta	gtaaaaat	aaaaattagc	cgggtatggt	gatcccagct	acttgggagt	61440
ctgaagcatg	agaattgctt	gaacatggga	ggcggagggt	gcagtgagcc	gagatcgcg	61500
cattgcactc	cagcctgggc	gacagagcga	gactctgtct	caaaaaaaaa	aaaaaaaaaag	61560
aaaagaaaa	aaattatcaa	tctcctcttt	tatggcata	atatatatat	atatatatat	61620
atatatatat	atatatat	tttttttttg	gttatgttca	gaaaggcctt	ccgtgctctg	61680
atcataaaaa	acaacttatt	ttcacactct	ctctcttttt	tttttgagac	agagttttgc	61740
tctgtttgcc	caggctggag	tgagtggtgc	caatctcagc	tactgttaac	ctccgcctcc	61800
cgggttgag	tgattctcct	gccttacctt	cccgagtgc	tgggattata	ggcatgcacc	61860
accatgctcg	gctaattttg	tacttttagt	agagacgggg	gtttctccat	gttggtcagg	61920
ctggtctcga	actgcgcacc	tcagggtgatc	caccacactc	ggcctcccaa	agtgtcggga	61980
ttacagacgt	gagccaccat	gccagccca	cactctcttt	cttaacgtcc	tcctcctttc	62040
gttttacgtt	cacatcttta	attcttcttg	gatgtaatta	gatttgatga	gcaagggtggg	62100
catccagctt	gtttcttggc	tgatggctta	tgggtggcgt	gaattagtcg	gggtctatca	62160
ggaggcagaa	actctatgag	aatttgaaca	gagaaagtcc	cgtctacagg	cttattacca	62220
gggactggaa	tagcagaaat	tgaacagtga	gatgtacaga	gaactctaag	aatgcaggaa	62280
taggccaggc	atgggtggctc	acacctgtca	tccagcact	ttgggagacc	aaggcgggtg	62340
gatcacctga	ggtcaggagt	tcgagaccag	cctggccaac	atagtgaac	cccatctcta	62400
ctaaaaatac	aaaaaaatta	gctgggtgtg	gtggcgcatg	cctgtaatcc	cagctactcg	62460
ggaggctgag	gcaggagaat	cacttgaacc	tgggaggcag	aggttgagct	gagccgagat	62520
catgccactg	tactccagcc	tgggtggaag	agcggaaactc	tgtctgaaaa	aaaaaaaaaa	62580
aaacaagaagt	tcaacttgaa	gggaaaaatg	cgtattgttc	tttccctttg	ttatgtcacc	62640
agggcacagt	ccatcccagg	ctggcgctga	tccacgggct	ggagaggggc	tgccccagaa	62700
gaggacatgc	caggaagggc	ttggctgtgtg	ttcaggagcc	caggccaggt	caggtaaga	62760
ggtgttgagg	ctggacggga	gaggccagct	aggggctcat	gtaggatatg	aggggtcggc	62820
ccattttcaac	gtggaaaactg	agctcttctg	cttctcttcc	ttcttccactg	cattaagatt	62880
caataccgct	tgggaagcag	gtatttcctt	tctataaaag	gatgggtggg	agcctgagtg	62940
ttgggagaaa	gtgtagccgc	tgagttacta	acaactaggg	ctgccgtcaa	gcctatgggg	63000

10

20

30

aaagagagaa	gaggacat	ggaaggagag	agatcaagct	gtggcaccct	gggagaggac	63060
cacagaaaa	aggccagtga	gggggttccc	cggtggcatc	tgaagggtgtg	gcccaccag	63120
gaggtccaga	ggctgccagc	cgagtggccc	aggagaggga	acctcacagg	ggctgagtgg	63180
gacccaagcc	ctatccaccg	tcctaaccac	ccacatttct	cgggaaacaag	acctcccaca	63240
gtggcctccc	cggcagtggg	aatagccaaa	ctggcaacat	ggactttctt	caactgcccg	63300
ggcgatgctg	cctcagtgcc	ccagggcagg	caggaagctc	ccacacccat	tctgggaatga	63360
gggggttgag	gaaggctgag	ctgagcaaa	gacccatctc	tgtctgtggt	gggtggggagg	63420
gagcccatta	tacaagagac	ccctcagggc	tcagtggagg	gtgacagaga	cttggggagt	63480
agtggctgtc	actgcagagg	tgagagggtt	tggagagaag	gtacatgcct	ttttggccac	63540
attgagtagc	acctggtagc	cagttagtaa	cgtgtattgg	ataaacaaaa	gattaaacgg	63600
atgcaaaaa	aaatgttggc	tttgcttctt	tttaccctaa	cctcagttcc	ctcaagtaga	63660
ttctgggaac	accccttacc	tggctggact	gttgtgaagt	ttaaataagc	caggttaact	63720
tcacctcctc	ctttaagaca	cagctcagac	actgcctcct	ccaagaagcc	ccctctggct	63780
tcctgtgtga	atatgacggc	cctctgggct	ctagggtatc	ttagaacaat	gcttccttat	63840
ggcttttgaa	ccccgctgtc	tcctggattg	ggagcaaatg	caggggagga	gccacacctg	63900
actaatctct	gggtctccca	gcacataagt	ggcataaggg	cagggctgtg	cccgttccag	63960
gcacttactg	aaggatgtac	ttggcagagg	gtaggcagcc	ggcggatgag	ccctcactc	64020
tccccagctg	actgcgtggg	cgggaaaggc	gggttcagga	gacccagcct	ccctgggctg	64080
tcaccacctc	tgcacatcca	gccccattga	tcaagggttc	aatttttggg	gtcctgttgg	64140
gaggccagga	gactctctcc	aggcacttct	tcagggtctt	tgtgttaggg	tgtgtgtgtg	64200
tgtgtgtgtg	tgtgtgtgtg	tgtgtgtgtt	gttttatttt	atttatttat	ttatttattt	64260
atttatttat	ttatttattt	tgagacgcag	tctcgctctg	ttgccagggt	tggagggtgg	64320
tggcatgata	tcgctcactt	gcaagctccg	cctcccggtt	tcacgccatt	ctcctgcctc	64380
actcttctct	agtagccgga	ttacaggcgc	acgcaccatg	cctggctaatt	tattttgttt	64440
tttttagtaga	gacagggttt	cgccacgttg	ccagggttgg	tcttgaatcc	ctggcctcaa	64500
gcgatccgcc	cgctcagccc	tcccaaaagt	ctgggattac	aggcgtgagc	caccgtgccc	64560
gcccagccta	ggggtacatg	aaactttttt	tttttttttt	ttgagacaga	gtttcactct	64620
gtcctcaggg	tggagtgcag	tggcgtgata	tcggcgtaact	gcaatctccg	cctcccgggt	64680
caagcgattc	tcctgcctca	gcctcccgag	tagctgggat	tgcaggcagc	cgccaccaca	64740
cccagctaat	ttttgtattt	ttagtagaga	cgggctttca	ccatgtggga	caggatgggt	64800
tcgatctcct	gacctctgta	tcggcccgcc	tcagcctccg	aaagtgtctg	gattacaggc	64860
ctgagccacc	gtgcccagcc	atgatgtttt	gatacaggca	tataacgtat	ataaatcaca	64920
tcagggtaaa	tgatgtaacc	atcacatcaa	gcatttatcc	tttgtgttac	aaaaaaaaat	64980
ctaattatac	tttctacttt	attctttttt	tttttttttt	ttgagacgga	gtctccctca	65040
gtcgcgccagg	ctggagtgca	gtggcatgat	ctcagttcac	tgcaggctct	gcctcctagc	65100
tctgcctcct	gggttcatgc	cattctcctg	tctcagcctc	gcgagtagct	gggactacag	65160
gcgcctgcc	ccgtgcccgg	ctaatttttt	tttttgtatt	tttggtagag	acagggtttc	65220
accgtgttag	ccaggatggg	ctcgatctcc	tgacctcata	atccgcccgt	ctcggcctcc	65280
caaagtgtct	ggattacagg	catgagccac	cgccccagc	ctatttatcc	ttaaatgtac	65340
aataaattat	tgttgactcc	agtcacctg	ctgtgctacc	aaatacggat	cttcttcatt	65400
ctatctaact	gtattttctg	acctgttaac	catctctcct	ccacctcacc	ccccaaacc	65460
actacccttc	tcagcctctg	gtaaccatcc	ttctactctc	tatctctatg	agttcaattg	65520
tattaatttt	tagctccccg	gcggggcagc	gtggtctcag	cctgtaatcc	cagcacttca	65580
ggaggctgag	gcagggtgat	cacgaggtea	ggagtgtgag	accagcctgg	ccaacatggg	65640
ggaaccccat	ctctactaaa	aacacaaaaa	ttagctgggc	gtggtgggtg	gcgcttgtag	65700
tcccagctac	ttgggaggct	gaggcaggag	aatcgcttga	aactgggagg	cagagggttc	65760
agtgagccaa	gattgcgcca	ctgcactcca	gtctgggtga	cagagtaaga	ttccatccc	65820
aaaaaaaaaa	agttagctgc	ccacaaataa	gtgagaacac	gtgaagtttc	tctttctgtg	65880
cctcgcttgt	ttcacttaac	ataatgacct	ccagttccat	ccaggttgtt	gctttgttat	65940
aatgacagg	atcttggtca	ggcgagtggt	ctcatgcctg	taatcccagc	actttgggag	66000

10

20

30

gctgaggtgg actgatcatg aggtcaagag atcgagacca tcctgggctaa cacagtgaag 66060
 ccccgctctct actaaaaata caagaaatta gcggggcgtg gtgggtgggca cccattttccg 66120
 ccccttctctg ggacgctgat gcacgacata ttacccatcc ccggaagact aatcctcccc 66180
 cactctatat tgtacctctt cctttctcct ccacgcgatt ccccgagtaa cccgtcttcc 66240
 ctccctctct ggattacgct cacttttccg cttcaatcac gttgctccgt ccccttcccc 66300
 attcgtacca ctccctcactt tcgtcttctt acccccaata tcccttttccg tcctctctat 66360
 tccttactta ctccctcccc ttctcttcat acttcattcc ctccgctctt cccactcgcg 66420
 ctcccaacttt cactagttg ccctcaccta cgttgccatc tcgccccctc ttcagctctc 66480
 ggctctctac ccattctgct tcctctcttac ctctctctct atctcgtctc gacatctctc 66540
 tagactatcc ctcaactttac cttctcagtc gtcttcttcc tatccttccg tctccatgat 66600
 cttcacgtcg ccattctctt tcgccccctt catatgtctc tcttcattgt ctcaactatca 66660
 ttctcatgat cactatcgtt ctcaactact atcaactccc tctttcttca tcaattctct 66720
 tcggtcattc tcgtctctct cttacaaccg ccttccttgt gctatctaac tcaacctatgc 66780
 ctctcctact ctctctctat cgcctctcca tcgcttatgc atctcttctt attgcacacc 66840
 cgcctctcca tcgcttatgc atctcttctt attgcacacc gccctctcat cgttatgca 66900
 tcctcttcta ttgcacatcc tcttctattg cac 66933

<210> 12

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 12

ctgagcggaa ttcgtgagac c

21

<210> 13

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 13

ttggtctcac gtattccgct cga

23

<210> 14

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 14

ctcgagaatt ctggatcctc

20

10

20

30

<210> 15
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 15
ttgaggatcc agaattctcg ag 22

<210> 16
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 16
tgtatgcgaa ttcgctgcgc g 21

<210> 17
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 17
ttcgcgcagc gaattcgcat aca 23

<210> 18
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 18
gtccactgaa ttctcagtga g 21

<210> 19
<211> 23
<212> DNA
<213> Artificial Sequence

<220>

10

20

30

<223> Artificial sequence is a primer.

<400> 19

ttgtcactga gaattcagtg gac

23

<210> 20

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

10

<400> 20

gaatccgaat tcttggtcag c

21

<210> 21

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 21

ttgctgacca ggaattcgga ttc

23

<210> 22

<211> 33

<212> DNA

<213> Artificial Sequence

20

<220>

<223> Artificial sequence is a primer.

<400> 22

cuacuacuac uactgagcgg aattcgtgag acc

33

<210> 23

<211> 32

<212> DNA

<213> Artificial Sequence

30

<220>

<223> Artificial sequence is a primer.

<400> 23

cuacuacuac uactcgagaa ttctggatcc tc

32

<210> 24

<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 24
cuacuacuac uatgtatgcg aattcgctgc gcg 33

<210> 25
<211> 33
<212> DNA
<213> Artificial Sequence

10

<220>
<223> Artificial sequence is a primer.

<400> 25
cuacuacuac uagtcactg aattctcagt gag 33

<210> 26
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

20

<400> 26
cuacuacuac uagaatccga attcctgggc agc 33

<210> 27
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 27
aactggaaga attcgcgcc gcaggaattt tttttttttt ttttt 45

30

<210> 28
<211> 13
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 28 aatcggcac gag	13	
<210> 29 <211> 9 <212> DNA <213> Artificial Sequence		
<220> <223> Artificial sequence is a primer.		
<400> 29 ctcgtgccg	9	10
<210> 30 <211> 14 <212> DNA <213> Artificial Sequence		
<220> <223> Artificial sequence is a primer.		
<400> 30 gtacgacggc cagt	14	
<210> 31 <211> 16 <212> DNA <213> Artificial Sequence		20
<220> <223> Artificial sequence is a primer.		
<400> 31 aacagctatg accatg	16	
<210> 32 <211> 18 <212> DNA <213> Artificial Sequence		
<220> <223> Artificial sequence is a primer.		30
<400> 32 ccaagttctg agaagtcc	18	
<210> 33 <211> 20 <212> DNA		

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 33

aatacctgaa accatacctg

20

<210> 34

<211> 57

<212> DNA

<213> Artificial Sequence

10

<220>

<223> Artificial sequence is a primer.

<400> 34

agctgctcgt agctgtctct cctggatca cgggtacatg tactggacag actgggt

57

<210> 35

<211> 56

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 35

tgagacgcc gccattgagcg ggcaggata gcttattccc tgtgccgcat tacggc

56

20

<210> 36

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 36

agctgctcgt agctgtctct cctgga

27

<210> 37

<211> 27

<212> DNA

<213> Artificial Sequence

30

<220>

<223> Artificial sequence is a primer.

<400> 37

gccgtaatgc ggcacagggg ataagct 27

<210> 38
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Artificial sequence is a primer.

<400> 38
 gagaggctat atccctgggc 20

<210> 39
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Artificial sequence is a primer.

<400> 39
 acagcacgtg tttaaagggg 20

<210> 40
 <211> 163
 <212> DNA
 <213> Homo sapiens 20

<400> 40
 actaaagcgc cgcgcgcgcg ccatggagcc cgagtgcgct cggcgcgggc ccgtccggcc 60
 gccggacaac atggaggcag ctccgcccg ggcgcgctgg ccgtctctgc tgctctctgc 120
 gctgctgctg gcgctgtgcg gctgcccggc cccgcgcgcg gcc 163

<210> 41
 <211> 419
 <212> DNA
 <213> Homo sapiens

<400> 41
 gcccacagc ctgcgcgctc ctgctatttg ccaaccgcgc ggacgtacgg ctggtggacg 60
 ccggcggagt caagctggag tccaccatcg tggtcagcgc cctggaggat gcggccgcag 120
 tggacttcca gttttccaag ggagccgtgt actggacaga cgtgagcgag gaggccatca 180
 agcagacctc cctgaaccag acggggggcg ccgtgcagaa cgtggtcatc tccggcctgg 240
 tctctcccga cggcctcgcc tgcgactggg tgggcaagaa gctgtactgg acggactcag 300
 agaccaaccg catcgagggt gccaacctca atggcacatc ccggaagggt ctcttctggc 360
 aggaccttga ccagccgagg gccatgcct tggacccgcg tcacgggtaa accctgctg 419

<210> 42
 <211> 221

10

20

30

<212> DNA

<213> Homo sapiens

<400> 42

ccccgtcaca	ggtacatgta	ctggacagac	tggggtgaga	cgccccggat	tgagcgggca	60
gggatgggatg	gcagcaccog	gaagatcatt	gtggactcgg	acatttactg	gcccgaatgga	120
ctgaccatcg	acctggaggga	gcagaagctc	tactgggctg	acgccaagct	cagcttcacg	180
caccgtgcca	acctggacgg	ctcgttcogg	taggtaccca	c		221

<210> 43

<211> 221

<212> DNA

<213> Homo sapiens

10

<400> 43

tccctgactg	caggcagaag	gtggtggagg	gcagcctgac	gcaccccttc	gccctgacgc	60
tctccgggga	cactctgtac	tggacagact	ggcagaccog	ctccatccat	gccctgcaaca	120
agcgactgg	ggggaagagg	aaggagatcc	tgagtgcctt	atactcaccg	atggacatcc	180
aggtgctgag	ccaggagcgg	cagccttttt	gtgagtgcgg	g		221

<210> 44

<211> 156

<212> DNA

<213> Homo sapiens

<400> 44

tttctcagtc	cacactcgct	gtgaggagga	caatggcggc	tgggtcccacc	tgtgcctgct	60
gtccccaagc	gagccttttt	acacatgcgc	ctgccccacg	ggtgtgcaga	tgcaggacaa	120
cggcaggacg	tgtaaaggcag	gtgaggcggt	gggacg			156

20

<210> 45

<211> 416

<212> DNA

<213> Homo sapiens

<400> 45

ctccacagga	gccgaggagg	tgctgctgct	ggccccggcg	acggacctac	ggaggatctc	60
gctggacacg	cgggacttca	ccgacatcgt	gctgcagggtg	gacgacatcc	ggcacgccat	120
tgccatcgac	tacgaccgcg	tagagggtca	tgtctactgg	acagatgacg	aggtgcgggc	180
catccgcagg	gcgtacctgg	acgggtctcg	ggcgagacg	ctggtcaaca	ccgagatcaa	240
cgaccccgat	ggcatcgcg	tcgactgggt	ggccccgaaac	ctctactgga	ccgacacggg	300
cacggaccgc	atcgagggtga	cgcgcctcaa	cggcacctcc	cgcaagatcc	tgggtgtcgga	360
ggacctggac	gagccccgag	ccatcgcact	gcaccccgctg	atggggtaag	acgggc	416

30

<210> 46

<211> 198

<212> DNA

<213> Homo sapiens

<400> 46

```

ttctttctcca gctcatgta ctggacagac tggggagaga accctaaaat cgagtgtgcc 60
aacttggatg ggcaggagcg gcgtgtgctg gtcaatgcct cctcgggtg gcccacggc 120
ctggccctgg acctgcagga ggggaagctc tactggggag acgccaagac agacaagatc 180
gaggtgaggg tcctgtgg 198

```

```

<210> 47
<211> 244
<212> DNA
<213> Homo sapiens

```

```

<400> 47
ccgtcctgca ggtgatcaat gttgatggga cgaagaggcg gacctcctg gaggacaagc 60
tcccgacat tttcgggttc acgtgtgctg gggacttcct ctactggact gactggcagc 120
gcgcagcat cgagcgggtg cacaaggtca aggccagccg ggacgtcatc attgaccagc 180
tgcccacct gatggggctc aaagctgtga atgtggccaa ggtcgtcggg gactccgggg 240
ggc 244

```

10

```

<210> 48
<211> 313
<212> DNA
<213> Homo sapiens

```

```

<400> 48
gttcgcttcc aggaaccaac ccgtgtgctg acaggaacgg ggggtgcagc cacctgtgct 60
tctgcacacc ccacgcaacc cgggtgtgct gcccctcctg cctggagctg ctgagtgcac 120
tgaagacctg catcgtgctt gaggcctttt tggctcttcac cagcagagcc gccatccaca 180
ggatctccct cgagaccaat aacaacgacg tggccatccc gctcacgggc gtcaaggagg 240
cctcagccct ggactttgat gtgtccaaca accacatcta ctggacagac gtcagcctga 300
aggtagcgtg ggc 313

```

20

```

<210> 49
<211> 255
<212> DNA
<213> Homo sapiens

```

```

<400> 49
cctgctgcca gaccatcagc cgcgccttca tgaacgggag ctgggtggag cacgtggtgg 60
agtttggcct tgactacccc gagggcatgg ccgttgactg gatgggcaag aacctctact 120
gggcgacac tgggaccaac agaatcgaag tggcgcggct ggacgggcag ttccggcaag 180
tcctcgtgtg gagggacttg gacaacccga ggtcgtctggc cctggatccc accaaggggt 240
aagtgtttgc ctgct 255

```

```

<210> 50
<211> 210
<212> DNA
<213> Homo sapiens

```

30

```

<400> 50
gtgccttcca gctacatcta ctggaccgag tggggcggca agccgaggat cgtgcggggc 60
ttcatggacg ggaccaactg catgacgctg gtggacaagg tgggcggggc caacgacctc 120

```

accattgact acgctgacca gcgcctctac tggaccgacc tggacaccaa catgatcgag 180
tcgtccaaca tgctgggtga gggccgggct 210

<210> 51

<211> 352

<212> DNA

<213> Homo sapiens

<400> 51

gtgttcacgc aggtcaggag cgggtcgtga ttgccgacga tctcccgac ccgttcggtc 60
tgacgcagta cagcgattat atctactgga cagactggaa tctgcacagc attgagcggg 120
ccgacaagac tagcggccgg aaccgcaccc tcctccaggg ccacctggac ttcgtgatgg 180
acatcctggg gttccactcc tcccgccagg atggcctcaa tgactgtatg cacaacaacg 240
ggcagtgtgg gcagctgtgc cttgccatcc ccggcggcca ccgctgcggc tgcgcctcac 300
actacaccct ggacccacgc agccgcaact gcagccgtaa gtgcctcatg gt 352

<210> 52

<211> 225

<212> DNA

<213> Homo sapiens

<400> 52

gcctctctta cgcaccacac cttcttctgtg ttcagccaga aatctgccat cagtcggatg 60
atcccggacg accagcacag ccgggatctc atcctgcccc tgcctggact gaggaacgtc 120
aaagccatcg actatgaccc actggacaag ttcacttact gggtaggatg gccccagaac 180
atcaagcgag ccaaggacga cgggacccag gcaggtgccc tgtgg 225

<210> 53

<211> 235

<212> DNA

<213> Homo sapiens

<400> 53

ctttgtctta cagccctttg ttttgacctc tctgagccaa ggccaaaacc cagacaggca 60
gcccacagac ctcagcatcg acatctacag ccggacactg ttctggacgt gcgaggccac 120
caataccatc aacgtccaca ggctgagcgg ggaagccatg ggggtggtgc tgcgtgggga 180
ccgcagacaag cccagggccca tcgtcgtcaa cgcggagcga gggtaggagg ccaac 235

<210> 54

<211> 218

<212> DNA

<213> Homo sapiens

<400> 54

ccaccctccc gcaggtaacct gtacttcacc aacatgcagg accgggcagc caagatcgaa 60
cgcgacagcc tggacggcac cgagcgcgag gtctcttcca ccaccggcct catccgccct 120
gtggccctgg tggtaggacaa cacactgggc aagctgttct gggtaggacg ggacctgaag 180
cgcattgaga gctgtgacct gtcaggtagc cgcgccgg 218

<210> 55

10

20

30

<211> 234

<212> DNA

<213> Homo sapiens

<400> 55

ggctgcttgc	agggggccaac	cgcttgaccc	tggaggacgc	caacatcgtg	cagcctctgg	60
gcctgaccat	ccttggcaag	catctctact	ggatcgaccg	ccagcagcag	atgatcgagc	120
gtgtggagaa	gaccacgggg	gacaagcgga	ctcgcatcca	gggcggtgtc	gcccacotca	180
ctggcatcca	tgcagtggag	gaagtcagcc	tggaggagtt	ctgtacgtgg	gggc	234

<210> 56

<211> 157

<212> DNA

<213> Homo sapiens

<400> 56

ttgtctttgc	agcagcccaac	ccatgtgccc	gtgacaatgg	tggctgctcc	cacatctgta	60
ttgccaaggg	tgatgggaca	ccacgggtgct	catgcccagt	ccacctcgtg	ctcctgcaga	120
acctgctgac	ctgtggaggt	aggtgtgacc	taggtgc			157

<210> 57

<211> 272

<212> DNA

<213> Homo sapiens

<400> 57

gttctctct	gtccctcccc	cagagccgcc	cacctgctcc	cgggaccagt	ttgcatgtgc	60
cacaggggag	atcgactgta	tccccggggc	ctggcgctgt	gacggctttc	ccgagtgcga	120
tgaccagagc	gacgaggagg	gctgccccgt	gtgctccgcc	gccagttcc	cctgcgcgcg	180
gggtcagtg	gtggacctgc	gcctgcgctg	cgacggcgag	gcagactgtc	aggaccgctc	240
agacgaggtg	gactgtgacg	gtgaggccct	cc			272

<210> 58

<211> 134

<212> DNA

<213> Homo sapiens

<400> 58

tctccttgca	gccatctgcc	tgcccaacca	gttcgggtgt	gcgagcggcc	agtgtgtcct	60
catcaaacag	cagtgcgact	ccttccccga	ctgtatogac	ggctccgacg	agctcatgtg	120
tggtgagcca	gctt					134

<210> 59

<211> 274

<212> DNA

<213> Homo sapiens

<400> 59

gtttgtctct	ggcagaaatc	accaagccgc	cctcagacga	cagcccggcc	cacagcagtg	60
ccatcggggc	cgtcattggc	atcatcctct	ctctctctgt	catgggtggg	gtctatattg	120

10

20

30

tgtgccagcg cgtgggtgtgc cagcgctatg cggggggccaa cggggcccttc ccgcacgagt 180
 atgtcagcgg gaccccgcac gtgcccctca atttcatagc cccggggcggg tcccagcatg 240
 gccccttcac aggtaaggag cctgagatat ggaa 274

<210> 60
 <211> 164
 <212> DNA
 <213> Homo sapiens

<400> 60
 ctcccttgcc aggcacgcga tgcggaaagt ccatgatgag ctccgtgagc ctgatggggg 60
 gccggggcgg ggtgcccctc taagaccgga accacgtcac aggggcctcg tccagcagct 120
 cgtccagcac gaaggccacg ctgtaccgcg cggtgagggg cggg 164

10

<210> 61
 <211> 130
 <212> DNA
 <213> Homo sapiens

<400> 61
 ttggtctctc tcagatcctg aaccgcgcgc cctccccggc caccgacccc tccctgtaca 60
 acatggacat gttctactct tcaaacattc cggccactgc gagaccgtac aggtaggaca 120
 tcccctgcag 130

<210> 62
 <211> 496
 <212> DNA
 <213> Homo sapiens

20

<400> 62
 tcaaacattc cggccactgc gagaccgtac agggccctaca tcattcgagg aatggcgccc 60
 ccgacgacgc cctgcagcac cgacgtgtgt gacagcgact acagcgccag ccgctggaag 120
 gccagcaagt actacctgga tttgaaactg gactcagacc cctatccacc cccaccacg 180
 cccacagcc agtacctgtc ggcgaggac agctgcccgc cctcgcccgc caccgagagg 240
 agctaacttc atctcttccc gcccctcctg tcccctgca cggactcatc ctgacctcgg 300
 ccgggcccact ctggcttctc tgtgcccctg taaatagttt taaatatgaa caaagaaaaa 360
 aatatatttt atgatttaaa aaataaatat aattgggatt ttaaaaacat gagaaatgtg 420
 aactgtgatg ggggtgggcag ggctgggaga actttgtaca gtggagaaat atttataaac 480
 ttaattttgt aaaaca 496

<210> 63
 <211> 22
 <212> DNA
 <213> Artificial Sequence

30

<220>
 <223> Artificial Sequence is a primer.

<400> 63

ttttgggtac acaattcagt cg 22

<210> 64
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence is a primer.

<400> 64
aaaactgtgg gtgcttctgg 20

<210> 65
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence is a primer.

<400> 65
gtgattgagc caatcctgag a 21

<210> 66
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence is a primer.

<400> 66
tgagccaaat aaaccccttc t 21

<210> 67
<211> 20
<212> DNA
<213> Homo sapiens

<400> 67
ctggactacg tggccttctc 20

<210> 68
<211> 19
<212> DNA
<213> Homo sapiens

<400> 68
ttcagaagca cttggctgg 19

10

20

30

<210> 69
<211> 20
<212> DNA
<213> Homo sapiens

<400> 69
ctcagtgcca tgaagatgga 20

<210> 70
<211> 21
<212> DNA
<213> Homo sapiens

10

<400> 70
caagatcact cgatctccag g 21

<210> 71
<211> 20
<212> DNA
<213> Homo sapiens

<400> 71
gtttcaggag actcagagtc 20

<210> 72
<211> 20
<212> DNA
<213> Homo sapiens

20

<400> 72
ttctgcaggt tgctgttgag 20

<210> 73
<211> 20
<212> DNA
<213> Homo sapiens

<400> 73
ttattgtgat ttcccggtggc 20

<210> 74
<211> 21
<212> DNA
<213> Homo sapiens

30

<400> 74
gccctctgtc ctgacttcag g 21

<210> 75
<211> 20

<212> DNA
<213> Homo sapiens

<400> 75
gagaaagaaa taaggggacc 20

<210> 76
<211> 20
<212> DNA
<213> Homo sapiens

<400> 76
tgctttgtaa agcactgaga 20

10

<210> 77
<211> 24
<212> DNA
<213> Homo sapiens

<400> 77
gaagtacggg cagttcagtg gcct 24

<210> 78
<211> 25
<212> DNA
<213> Homo sapiens

<400> 78
atacaccaag gtccatgttc cccgt 25

20

<210> 79
<211> 25
<212> DNA
<213> Homo sapiens

<400> 79
agcctggggc acagcgtgag actac 25

<210> 80
<211> 25
<212> DNA
<213> Homo sapiens

<400> 80
tcccggagct tgcacaccgg ctcca 25

30

<210> 81
<211> 20
<212> DNA
<213> Homo sapiens

<400> 81
catgtgccca cctcattcat

20

<210> 82
<211> 20
<212> DNA
<213> Homo sapiens

<400> 82
caagattctg tagcttctgg

20

<210> 83
<211> 20
<212> DNA
<213> Homo sapiens

10

<400> 83
cagagaagtc aagggacttg

20

<210> 84
<211> 20
<212> DNA
<213> Homo sapiens

<400> 84
atcctctcac atccacact

20

<210> 85
<211> 20
<212> DNA
<213> Homo sapiens

20

<400> 85
caaggctaaa agacgaaaaa

20

<210> 86
<211> 20
<212> DNA
<213> Homo sapiens

<400> 86
tcaggagcat ttcattcttt

20

30

<210> 87
<211> 19
<212> DNA
<213> Homo sapiens

<400> 87
aagtcgaggc tgcaaggag

19

<210> 88
<211> 20
<212> DNA
<213> Homo sapiens

<400> 88
gccctgtgtt cctttcagta

20

<210> 89
<211> 19
<212> DNA
<213> Homo sapiens

10

<400> 89
aaggtgtgag gatcactgg

19

<210> 90
<211> 17
<212> DNA
<213> Homo sapiens

<400> 90
agctcatggg ggctatt

17

<210> 91
<211> 20
<212> DNA
<213> Homo sapiens

20

<400> 91
gcttctccga gtgtatcaac

20

<210> 92
<211> 20
<212> DNA
<213> Homo sapiens

<400> 92
atggcagagg acttagaaca

20

<210> 93
<211> 24
<212> DNA
<213> Homo sapiens

30

<400> 93
gatcagcgaa cttctctctcg gctc

24

<210> 94
<211> 24

<212> DNA
<213> Homo sapiens

<400> 94
tccacattga ggactgtggg aacg 24

<210> 95
<211> 20
<212> DNA
<213> Homo sapiens

<400> 95
gctaatacaca gtctaaccga 20 10

<210> 96
<211> 19
<212> DNA
<213> Homo sapiens

<400> 96
ttgcactgtc ttggatgca 19

<210> 97
<211> 25
<212> DNA
<213> Homo sapiens

<400> 97
gcacagctgt agtgggggttc taggc 25 20

<210> 98
<211> 25
<212> DNA
<213> Homo sapiens

<400> 98
caggcgcaaa ggacatgcac acggc 25

<210> 99
<211> 25
<212> DNA
<213> Homo sapiens 30

<400> 99
caccgatgag tgcacgttca aggag 25

<210> 100
<211> 25
<212> DNA
<213> Homo sapiens

<400> 100
cagacagaga tgctccacgc catac

25

<210> 101
<211> 19
<212> DNA
<213> Homo sapiens

<400> 101
tttctgggtg tgtctgaat

19

<210> 102
<211> 20
<212> DNA
<213> Homo sapiens

10

<400> 102
acacagttgc tctaaagggt

20

<210> 103
<211> 20
<212> DNA
<213> Homo sapiens

<400> 103
catttgggaa atccagaaga

20

<210> 104
<211> 25
<212> DNA
<213> Homo sapiens

20

<400> 104
taggtgtctt attttttgtt gettc

25

<210> 105
<211> 25
<212> DNA
<213> Homo sapiens

<400> 105
gacataccat gaacactata agagg

25

30

<210> 106
<211> 20
<212> DNA
<213> Homo sapiens

<400> 106
caaccatac cagggataag

20

<210> 107
<211> 21
<212> DNA
<213> Homo sapiens

<400> 107
gaacaagagg ggtaagttgg c 21

<210> 108
<211> 22
<212> DNA
<213> Homo sapiens 10

<400> 108
tgaggacaca gatactgatg gg 22

<210> 109
<211> 25
<212> DNA
<213> Homo sapiens

<400> 109
gaagtgttcc ctcttaaatt ctttg 25

<210> 110
<211> 25
<212> DNA
<213> Homo sapiens 20

<400> 110
gaactatatt gtagttagtg aggag 25

<210> 111
<211> 18
<212> DNA
<213> Homo sapiens

<400> 111
cctgtaaccc ccagtccc 18

<210> 112
<211> 22
<212> DNA
<213> Homo sapiens 30

<400> 112
tcttgcttcc taagtttctc gg 22

<210> 113
<211> 21

<212> DNA
<213> Homo sapiens

<400> 113
actccatcca cctcatcact g

21

<210> 114
<211> 20
<212> DNA
<213> Homo sapiens

<400> 114
tgctgtttgc ctcacttgac

20

10

<210> 115
<211> 20
<212> DNA
<213> Homo sapiens

<400> 115
gtggacaggc atagctgagg

20

<210> 116
<211> 21
<212> DNA
<213> Homo sapiens

<400> 116
tgttcactct tctgcctgca g

21

20

<210> 117
<211> 20
<212> DNA
<213> Homo sapiens

<400> 117
agctggactc tcacagaatg

20

<210> 118
<211> 20
<212> DNA
<213> Homo sapiens

<400> 118
caagaggctg gtagaagggtg

20

30

<210> 119
<211> 24
<212> DNA
<213> Homo sapiens

<400> 119
gactccagtc tgggcaataa aagc

24

<210> 120
<211> 22
<212> DNA
<213> Homo sapiens

<400> 120
ggtggcagca tgacctctaa ag

22

<210> 121
<211> 16
<212> DNA
<213> Homo sapiens

10

<400> 121
caggcccagt ctcttg

16

<210> 122
<211> 18
<212> DNA
<213> Homo sapiens

<400> 122
cgtgtccaga tgaaagtg

18

<210> 123
<211> 18
<212> DNA
<213> Homo sapiens

20

<400> 123
acctcacggt gtaatccc

18

<210> 124
<211> 18
<212> DNA
<213> Homo sapiens

<400> 124
cttgaagccc atctttgc

18

30

<210> 125
<211> 23
<212> DNA
<213> Homo sapiens

<400> 125
tatttgcaaa gcttgagact tct

23

<210> 126
<211> 20
<212> DNA
<213> Homo sapiens

<400> 126
aatcactgtg ctttgttgcc

20

<210> 127
<211> 20
<212> DNA
<213> Homo sapiens

10

<400> 127
actttattgt cagcgtgggc

20

<210> 128
<211> 18
<212> DNA
<213> Homo sapiens

<400> 128
actccctega tggcttcc

18

<210> 129
<211> 18
<212> DNA
<213> Homo sapiens

20

<400> 129
gagcagggga gagaaggc

18

<210> 130
<211> 20
<212> DNA
<213> Homo sapiens

<400> 130
cccactggct tgttttattg

20

<210> 131
<211> 25
<212> DNA
<213> Homo sapiens

30

<400> 131
agccacttta ttgttatttt gatgc

25

<210> 132
<211> 25

<212> DNA
<213> Homo sapiens

<400> 132
aagagtgaac aaaagcaaac atacc 25

<210> 133
<211> 18
<212> DNA
<213> Homo sapiens

<400> 133
gtggagtgtg ggattggg 18 10

<210> 134
<211> 25
<212> DNA
<213> Homo sapiens

<400> 134
tactgttctt gataagtatg tcggc 25

<210> 135
<211> 25
<212> DNA
<213> Homo sapiens

<400> 135 20
atgcttttgc atgattctaa ttatt 25

<210> 136
<211> 21
<212> DNA
<213> Homo sapiens

<400> 136
tcccccaaaa gaatgtaaag g 21

<210> 137
<211> 20
<212> DNA
<213> Homo sapiens 30

<400> 137
ctggctcttc ttgtgtgctg 20

<210> 138
<211> 18
<212> DNA
<213> Homo sapiens

<400> 138
atcaccagg ccagggat 18

<210> 139
<211> 24
<212> DNA
<213> Homo sapiens

<400> 139
tcagaagcag aactgttttt aaca 24

<210> 140
<211> 22
<212> DNA
<213> Homo sapiens

<400> 140
cctgcttgaa agttctagag cc 22

<210> 141
<211> 21
<212> DNA
<213> Homo sapiens

<400> 141
caagcccggg ttttattgaa a 21

<210> 142
<211> 18
<212> DNA
<213> Homo sapiens

<400> 142
gatgccagga ccatggac 18

<210> 143
<211> 24
<212> DNA
<213> Homo sapiens

<400> 143
gcatatagaa acaatttatt gccg 24

<210> 144
<211> 20
<212> DNA
<213> Homo sapiens

<400> 144
ctctgaagca gggaccagag 20

10

20

30

<210> 145
<211> 19
<212> DNA
<213> Homo sapiens

<400> 145
ctaccacacc acaccaggc

19

<210> 146
<211> 18
<212> DNA
<213> Homo sapiens

10

<400> 146
caagcgaaag ctgccttc

18

<210> 147
<211> 23
<212> DNA
<213> Homo sapiens

<400> 147
gttgtcttga cttcaggtct gtc

23

<210> 148
<211> 24
<212> DNA
<213> Homo sapiens

20

<400> 148
ttttccttca acaatcacta ctcc

24

<210> 149
<211> 20
<212> DNA
<213> Homo sapiens

<400> 149
gcgtggggat atagaggtca

20

<210> 150
<211> 20
<212> DNA
<213> Homo sapiens

30

<400> 150
tacgtggcca agaagctagg

20

<210> 151
<211> 24

<212> DNA
<213> Homo sapiens

<400> 151
taatatatcc ccagtctaag gcac

24

<210> 152
<211> 18
<212> DNA
<213> Homo sapiens

<400> 152
agcttgcaga tggagccc

18

10

<210> 153
<211> 25
<212> DNA
<213> Homo sapiens

<400> 153
tgggttttaa cctttaatga gaaaa

25

<210> 154
<211> 23
<212> DNA
<213> Homo sapiens

<400> 154
tggtgatcta tacctgttt ccg

23

20

<210> 155
<211> 25
<212> DNA
<213> Homo sapiens

<400> 155
aattatttaa aagagaggaa aggca

25

<210> 156
<211> 20
<212> DNA
<213> Homo sapiens

<400> 156
tggctgtgaa cttcctctga

20

30

<210> 157
<211> 25
<212> DNA
<213> Homo sapiens

<400> 157
ggttacagaa aaacatttga gagat

25

<210> 158
<211> 22
<212> DNA
<213> Homo sapiens

<400> 158
tgagcttttag ttcccttctc tg

22

<210> 159
<211> 23
<212> DNA
<213> Homo sapiens

10

<400> 159
ttgaaaaacc atttatttca ccg

23

<210> 160
<211> 18
<212> DNA
<213> Homo sapiens

<400> 160
tctgcggctg ttggattt

18

<210> 161
<211> 23
<212> DNA
<213> Homo sapiens

20

<400> 161
ttgaaaaacc atttatttca ccg

23

<210> 162
<211> 20
<212> DNA
<213> Homo sapiens

<400> 162
tgttctcttc tcccagcagg

20

30

<210> 163
<211> 23
<212> DNA
<213> Homo sapiens

<400> 163
ctttattgaa aacattgagt gca

23

<210> 164
<211> 20
<212> DNA
<213> Homo sapiens

<400> 164
ttgtcaaatt ccccccaaaa

20

<210> 165
<211> 16
<212> DNA
<213> Homo sapiens

10

<220>
<221> misc_feature
<222> 12
<223> n = A,T,C or G

<400> 165
aaaccacgac cnccaa

16

<210> 166
<211> 20
<212> DNA
<213> Homo sapiens

<400> 166
ccctggaaag gtaagatgct

20

20

<210> 167
<211> 23
<212> DNA
<213> Homo sapiens

<400> 167
cttttggtag agacaaggtc tca

23

<210> 168
<211> 25
<212> DNA
<213> Homo sapiens

<400> 168
tatctgtctg tagtgcttca aatgt

25

30

<210> 169
<211> 19
<212> DNA
<213> Homo sapiens

<400> 169
gacgaagggtg attcagggc

19

<210> 170
<211> 19
<212> DNA
<213> Homo sapiens

<400> 170
actgaagaac tcttgtcct

19

<210> 171
<211> 25
<212> DNA
<213> Homo sapiens

10

<400> 171
cagataaaaag agtcactatg gctca

25

<210> 172
<211> 20
<212> DNA
<213> Homo sapiens

<400> 172
cacttctccc actttgtccc

20

<210> 173
<211> 25
<212> DNA
<213> Homo sapiens

20

<400> 173
ttattgataa gcattagtga acccc

25

<210> 174
<211> 20
<212> DNA
<213> Homo sapiens

<400> 174
tggcaagtta ggcacagtca

20

<210> 175
<211> 21
<212> DNA
<213> Homo sapiens

30

<400> 175
ctatgcccag agatgaacag g

21

<210> 176
<211> 20
<212> DNA
<213> Homo sapiens

<400> 176
tccactaagg gctatgtcgc 20

<210> 177
<211> 24
<212> DNA
<213> Homo sapiens

<400> 177
gccagcttta ttgagtaaac ttcc 24

<210> 178
<211> 22
<212> DNA
<213> Homo sapiens

<400> 178
cactggagac tacaagtggg gg 22

<210> 179
<211> 20
<212> DNA
<213> Homo sapiens

<400> 179
catcccaacc atcactcagt 20

<210> 180
<211> 22
<212> DNA
<213> Homo sapiens

<400> 180
ggggactagc ttacagattt ga 22

<210> 181
<211> 23
<212> DNA
<213> Homo sapiens

<400> 181
agactacatt ttggaaccag tgg 23

<210> 182
<211> 23

10

20

30

<212> DNA
<213> Homo sapiens

<400> 182
tgaaaggata tttatagcct gga 23

<210> 183
<211> 20
<212> DNA
<213> Homo sapiens

<400> 183
gaagggttttg tccctcgatc 20 10

<210> 184
<211> 20
<212> DNA
<213> Homo sapiens

<400> 184
tgaggggttg gaagatcata 20

<210> 185
<211> 18
<212> DNA
<213> Homo sapiens

<400> 185
ccttcatagc cacaccgc 18 20

<210> 186
<211> 21
<212> DNA
<213> Homo sapiens

<400> 186
cagctaactg ttgacatgcc a 21

<210> 187
<211> 25
<212> DNA
<213> Homo sapiens 30

<400> 187
tctttactgt gcttacaact ttcct 25

<210> 188
<211> 20
<212> DNA
<213> Homo sapiens

<400> 188
caacagtgca gtcggtatcg

20

<210> 189
<211> 18
<212> DNA
<213> Homo sapiens

<400> 189
agatcagcaa gcagatag

18

<210> 190
<211> 18
<212> DNA
<213> Homo sapiens

10

<400> 190
cattccacat ggatagac

18

<210> 191
<211> 24
<212> DNA
<213> Homo sapiens

<400> 191
catacctatg aggtgtgcta cagg

24

<210> 192
<211> 21
<212> DNA
<213> Homo sapiens

20

<400> 192
gcattttctc atcatccttg c

21

<210> 193
<211> 20
<212> DNA
<213> Homo sapiens

<400> 193
ttacagccac caaggtttcc

20

30

<210> 194
<211> 20
<212> DNA
<213> Homo sapiens

<400> 194
aggtgtgtgt gccaggttga

20

<210> 195
<211> 25
<212> DNA
<213> Homo sapiens

<400> 195
cactgttatc tcattaactg tgagg

25

<210> 196
<211> 21
<212> DNA
<213> Homo sapiens

10

<400> 196
tttgattttg tgtctcccaa a

21

<210> 197
<211> 20
<212> DNA
<213> Homo sapiens

<400> 197
ccccactccc acttttattt

20

<210> 198
<211> 24
<212> DNA
<213> Homo sapiens

20

<400> 198
ccagtcacct ttactagtcc ttg

24

<210> 199
<211> 20
<212> DNA
<213> Homo sapiens

<400> 199
aggacacagc ctgcatctag

20

<210> 200
<211> 20
<212> DNA
<213> Homo sapiens

30

<400> 200
accaggcatt gcactaaaag

20

<210> 201
<211> 22

<212> DNA
<213> Homo sapiens

<400> 201
gatgggtcac actaacctgt ca 22

<210> 202
<211> 24
<212> DNA
<213> Homo sapiens

<400> 202
acatttatat ttggacatgc aacc 24 10

<210> 203
<211> 22
<212> DNA
<213> Homo sapiens

<400> 203
agcatcttta atgtgtcagg ca 22

<210> 204
<211> 18
<212> DNA
<213> Homo sapiens

<400> 204
atgtgctggg ctggaaag 18 20

<210> 205
<211> 20
<212> DNA
<213> Homo sapiens

<400> 205
tcacattcaa aaatcggcaa 20

<210> 206
<211> 18
<212> DNA
<213> Homo sapiens 30

<400> 206
ctgcctgtgt ggtgtcgc 18

<210> 207
<211> 25
<212> DNA
<213> Homo sapiens

<400> 207
tggttttattt ctcagtacaa agcca

25

<210> 208
<211> 19
<212> DNA
<213> Homo sapiens

<400> 208
gacctcctgt gacaccacg

19

<210> 209
<211> 25
<212> DNA
<213> Homo sapiens

10

<400> 209
ccaccaaatt atttatagtt ctgcg

25

<210> 210
<211> 23
<212> DNA
<213> Homo sapiens

<400> 210
gtaagattct ccaactgttgc acc

23

<210> 211
<211> 20
<212> DNA
<213> Homo sapiens

20

<400> 211
cctataatgg gctggaccaa

20

<210> 212
<211> 21
<212> DNA
<213> Homo sapiens

<400> 212
actcctcatg tgaagtcacc g

21

30

<210> 213
<211> 20
<212> DNA
<213> Homo sapiens

<400> 213
cagtgtgcac gttttcattt

20

<210> 214
 <211> 20
 <212> DNA
 <213> Homo sapiens

<400> 214
 cagcatcttc agcacttacc 20

<210> 215
 <211> 25
 <212> DNA
 <213> Homo sapiens

10

<400> 215
 ctgcatttat tatgagaatc aacag 25

<210> 216
 <211> 20
 <212> DNA
 <213> Homo sapiens

<400> 216
 tgctgctggg agtcagagtc 20

<210> 217
 <211> 23
 <212> DNA
 <213> Homo sapiens

20

<400> 217
 cagggcactg agatacactt acc 23

<210> 218
 <211> 21
 <212> DNA
 <213> Homo sapiens

<400> 218
 aaggatcaag ccaggcattt g 21

<210> 219
 <211> 20
 <212> DNA
 <213> Homo sapiens

30

<400> 219
 acacatctct tctgtgcccc 20

<210> 220
 <211> 18

<212> DNA
<213> Homo sapiens

<400> 220
tgaaccctgg aggacagag 18

<210> 221
<211> 19
<212> DNA
<213> Homo sapiens

<400> 221
cattccccag tttgcagac 19 10

<210> 222
<211> 19
<212> DNA
<213> Homo sapiens

<400> 222
gtgctgggat tacaggtgt 19

<210> 223
<211> 20
<212> DNA
<213> Homo sapiens

<400> 223
gcagagaagt cctgttagcc 20 20

<210> 224
<211> 20
<212> DNA
<213> Homo sapiens

<400> 224
ccatgctaga gaagcacaac 20

<210> 225
<211> 20
<212> DNA
<213> Homo sapiens 30

<400> 225
agtgtggggc aggacctctg 20

<210> 226
<211> 21
<212> DNA
<213> Homo sapiens

<400> 226
cagacagata gccctggggtt c

21

<210> 227
<211> 20
<212> DNA
<213> Homo sapiens

<400> 227
tccctcatcc ccttgtctgt

20

<210> 228
<211> 19
<212> DNA
<213> Homo sapiens

10

<400> 228
agccccctg gggataatc

19

<210> 229
<211> 19
<212> DNA
<213> Homo sapiens

<400> 229
gatgcttacc taccacggc

19

<210> 230
<211> 21
<212> DNA
<213> Homo sapiens

20

<400> 230
aggattccta tctgggctat g

21

<210> 231
<211> 20
<212> DNA
<213> Homo sapiens

<400> 231
tggcagacca tgctccgct

20

30

<210> 232
<211> 20
<212> DNA
<213> Homo sapiens

<400> 232
gagaaggccg ggaggctctg

20

<210> 233
<211> 25
<212> DNA
<213> Homo sapiens

<400> 233
ctccatcaca accagatttg aggct

25

<210> 234
<211> 22
<212> DNA
<213> Homo sapiens

10

<400> 234
gggtgtgagc tgctgctgaa gg

22

<210> 235
<211> 25
<212> DNA
<213> Homo sapiens

<400> 235
agtgggaaac ctcaggtagc tcccg

25

<210> 236
<211> 25
<212> DNA
<213> Homo sapiens

20

<400> 236
cagtttggct cagacatatg ggggc

25

<210> 237
<211> 20
<212> DNA
<213> Homo sapiens

<400> 237
cattagtagt ggggggacag

20

<210> 238
<211> 21
<212> DNA
<213> Homo sapiens

30

<400> 238
caaagcgaca gtgagttagg g

21

<210> 239
<211> 20

<212> DNA
<213> Homo sapiens

<400> 239
ggagtagacc atgattactg 20

<210> 240
<211> 19
<212> DNA
<213> Homo sapiens

<400> 240
catggtctat ttattctcg 19 10

<210> 241
<211> 22
<212> DNA
<213> Homo sapiens

<400> 241
cgccctggat cctcacacta ca 22

<210> 242
<211> 21
<212> DNA
<213> Homo sapiens

<400> 242
gggcatcagg ggatgggtag a 21 20

<210> 243
<211> 23
<212> DNA
<213> Homo sapiens

<400> 243
gctcctatct gtgttttgaa tgg 23

<210> 244
<211> 22
<212> DNA
<213> Homo sapiens 30

<400> 244
ccgtggcata gataagtaaa cg 22

<210> 245
<211> 22
<212> DNA
<213> Homo sapiens

<400> 245
cttggagcgc tatgaggagg gc

22

<210> 246
<211> 23
<212> DNA
<213> Homo sapiens

<400> 246
atgggcaactg accttccgtc ctg

23

<210> 247
<211> 16
<212> DNA
<213> Homo sapiens

10

<400> 247
ttggagtcac aggggc

16

<210> 248
<211> 17
<212> DNA
<213> Homo sapiens

<400> 248
cagcactatc cttgggg

17

<210> 249
<211> 21
<212> DNA
<213> Homo sapiens

20

<400> 249
aacaaagctg cttagcacct g

21

<210> 250
<211> 20
<212> DNA
<213> Homo sapiens

<400> 250
gatgaggacc aactggtgac

20

30

<210> 251
<211> 20
<212> DNA
<213> Homo sapiens

<400> 251
ttttccaata atgtgacttc

20

<210> 252
<211> 20
<212> DNA
<213> Homo sapiens

<400> 252
caatcccaac cgtaacaggc

20

<210> 253
<211> 19
<212> DNA
<213> Homo sapiens

10

<400> 253
cttgatctcg cccaggaac

19

<210> 254
<211> 20
<212> DNA
<213> Homo sapiens

<400> 254
gctcgctgaa ggatgaagac

20

<210> 255
<211> 17
<212> DNA
<213> Homo sapiens

20

<400> 255
gaatcgcttg aacccag

17

<210> 256
<211> 17
<212> DNA
<213> Homo sapiens

<400> 256
ccaggtggtc ttaacgg

17

<210> 257
<211> 21
<212> DNA
<213> Homo sapiens

30

<220>
<221> misc_feature
<222> 8
<223> n = A,T,C or G

<400> 257
gaacgttntt catgtaggcg t 21

<210> 258
<211> 16
<212> DNA
<213> Homo sapiens

<400> 258
taatgggtcgc tgcgcc 16

<210> 259
<211> 22
<212> DNA
<213> Homo sapiens
<400> 259
agggaaaatg gtatgtgggg ag 22

<210> 260
<211> 18
<212> DNA
<213> Homo sapiens

<400> 260
gcagtgtgtg aaggcagg 18

<210> 261
<211> 24
<212> DNA
<213> Homo sapiens

<400> 261
agtggacaaa atgaggaaaa cagg 24

<210> 262
<211> 23
<212> DNA
<213> Homo sapiens

<400> 262
ccaacacagt ttgctcacat gcc 23

<210> 263
<211> 20
<212> DNA
<213> Homo sapiens

<400> 263
tgacatcttt gcattatggc 20

10

20

30

<210> 264
<211> 20
<212> DNA
<213> Homo sapiens

<400> 264
agttatccca cctgataccg 20

<210> 265
<211> 20
<212> DNA
<213> Homo sapiens

<400> 265
agctcttgct tctcagtcca 20

<210> 266
<211> 24
<212> DNA
<213> Homo sapiens

<400> 266
caaaagttgt ttctgtgttc gtcc 24

<210> 267
<211> 22
<212> DNA
<213> Homo sapiens

<400> 267
gcctctcaaa gtagttggaa cc 22

<210> 268
<211> 23
<212> DNA
<213> Homo sapiens

<400> 268
tgtgtatcca tagtgcaaaa cag 23

<210> 269
<211> 19
<212> DNA
<213> Homo sapiens

<400> 269
ctcaaggcca ggcactcact 19

<210> 270
<211> 19

10

20

30

<212> DNA
<213> Homo sapiens

<400> 270
ggactcttcc atgccagtg 19

<210> 271
<211> 19
<212> DNA
<213> Homo sapiens

<400> 271
aatgatgac tcaactctg 19 10

<210> 272
<211> 19
<212> DNA
<213> Homo sapiens

<400> 272
actgaagaac tcttgcct 19

<210> 273
<211> 23
<212> DNA
<213> Homo sapiens

<400> 273
gacatctgtt agtctcataa ttc 23 20

<210> 274
<211> 18
<212> DNA
<213> Homo sapiens

<400> 274
ggtaacagtg tcttgctt 18

<210> 275
<211> 21
<212> DNA
<213> Homo sapiens 30

<400> 275
ctatgtacaa aacaggaaga g 21

<210> 276
<211> 19
<212> DNA
<213> Homo sapiens

<400> 276
atcctagttt cctctcctt

19

<210> 277
<211> 23
<212> DNA
<213> Homo sapiens

<400> 277
gtaaatgaga aacagacaaa tga

23

<210> 278
<211> 23
<212> DNA
<213> Homo sapiens

10

<400> 278
ctattggatg tgatatgtta tgg

23

<210> 279
<211> 22
<212> DNA
<213> Homo sapiens

<400> 279
aagtagaaac aaaatgaggg ac

22

<210> 280
<211> 18
<212> DNA
<213> Homo sapiens

20

<400> 280
cctaccccaa ggtaacag

18

<210> 281
<211> 22
<212> DNA
<213> Homo sapiens

<400> 281
acttcctata aatggaggtg ag

22

30

<210> 282
<211> 18
<212> DNA
<213> Homo sapiens

<400> 282
gaggagcttc aagaggaa

18

<210> 283
<211> 23
<212> DNA
<213> Homo sapiens

<400> 283
cataactccta gactcaagga atc

23

<210> 284
<211> 23
<212> DNA
<213> Homo sapiens

10

<400> 284
gaatgatgta catgaattct ttg

23

<210> 285
<211> 20
<212> DNA
<213> Homo sapiens

<400> 285
gtggttgagga gaaaagcact

20

<210> 286
<211> 19
<212> DNA
<213> Homo sapiens

20

<400> 286
ctcccagtag tcacattcc

19

<210> 287
<211> 23
<212> DNA
<213> Homo sapiens

<400> 287
caagttacaa ataacttaag ccg

23

<210> 288
<211> 23
<212> DNA
<213> Homo sapiens

30

<400> 288
caagacccta tctctacaaa aac

23

<210> 289
<211> 24

<212> DNA		
<213> Homo sapiens		
<400> 289		
tttattagaa gtgactcttg gcc	24	
<210> 290		
<211> 20		
<212> DNA		
<213> Homo sapiens		
<400> 290		
gactacctgc cctcagcttg	20	10
<210> 291		
<211> 20		
<212> DNA		
<213> Homo sapiens		
<400> 291		
ttctcatgta caaagcggtc	20	
<210> 292		
<211> 20		
<212> DNA		
<213> Homo sapiens		
<400> 292		
ccactggctt ctctcttttt	20	20
<210> 293		
<211> 18		
<212> DNA		
<213> Homo sapiens		
<400> 293		
caccagaagg ttgggggtg	18	
<210> 294		
<211> 21		
<212> DNA		
<213> Homo sapiens		30
<400> 294		
actattacga catgaacgcg g	21	
<210> 295		
<211> 18		
<212> DNA		
<213> Homo sapiens		

<400> 295
ctcatgctgg atgacccc 18

<210> 296
<211> 24
<212> DNA
<213> Homo sapiens

<400> 296
ttgcctttct tgaaacttaa ttcc 24

<210> 297
<211> 19
<212> DNA
<213> Homo sapiens

<400> 297
tcacagcctt cagtcaggg 19

<210> 298
<211> 18
<212> DNA
<213> Homo sapiens

<400> 298
acatgctgtg gcaccatg 18

<210> 299
<211> 18
<212> DNA
<213> Homo sapiens

<400> 299
cctgagctac tgccacag 18

<210> 300
<211> 20
<212> DNA
<213> Homo sapiens

<400> 300
ccctgacttg gacagtgtcc 20

<210> 301
<211> 18
<212> DNA
<213> Homo sapiens

<400> 301
tcagagtcac tcctgccc 18

10

20

30

<210> 302
<211> 22
<212> DNA
<213> Homo sapiens

<400> 302
caaattcaag ctcattccaga cc

22

<210> 303
<211> 18
<212> DNA
<213> Homo sapiens

10

<400> 303
cggcatttca tccaggac

18

<210> 304
<211> 20
<212> DNA
<213> Homo sapiens

<400> 304
ggtgtaggag gtgcgacaat

20

<210> 305
<211> 20
<212> DNA
<213> Homo sapiens

20

<400> 305
ttccatttat tgagcacctg

20

<210> 306
<211> 20
<212> DNA
<213> Homo sapiens

<400> 306
cttaagccac tgtgttttgg

20

<210> 307
<211> 20
<212> DNA
<213> Homo sapiens

30

<400> 307
cctcctacac ctgcaaaagc

20

<210> 308
<211> 19

<212> DNA
<213> Homo sapiens

<400> 308
tggaagaacc ccagaggac

19

<210> 309
<211> 23
<212> DNA
<213> Homo sapiens

<400> 309
aaagcacaaa agtaacagca aca

23

10

<210> 310
<211> 20
<212> DNA
<213> Homo sapiens

<400> 310
gtgtgtgggc cacaatattg

20

<210> 311
<211> 20
<212> DNA
<213> Homo sapiens

<400> 311
agagcacctt tcctcagcac

20

20

<210> 312
<211> 20
<212> DNA
<213> Homo sapiens

<400> 312
agaatctcat cacaggggcg

20

<210> 313
<211> 25
<212> DNA
<213> Homo sapiens

<400> 313
aaaaaggaca gtgtctaaaa ttgga

25

30

<210> 314
<211> 25
<212> DNA
<213> Homo sapiens

<400> 314
aattgttttt gtttgttttt tgagt 25

<210> 315
<211> 22
<212> DNA
<213> Homo sapiens

<400> 315
gatttaggga gtacaagtgc gg 22

<210> 316
<211> 25
<212> DNA
<213> Homo sapiens

<400> 316
ggggacaaat tatactttat tcagg 25

<210> 317
<211> 22
<212> DNA
<213> Homo sapiens

<400> 317
ccatcatcat attggtgtga cc 22

<210> 318
<211> 18
<212> DNA
<213> Homo sapiens

<400> 318
tggtgcca agaagaag 18

<210> 319
<211> 24
<212> DNA
<213> Homo sapiens

<400> 319
ttaagatgcc attaaactca tgac 24

<210> 320
<211> 20
<212> DNA
<213> Homo sapiens

<400> 320
ccaaggagat gaccaagtgg 20

10

20

30

<210> 321
<211> 22
<212> DNA
<213> Homo sapiens

<400> 321
ccatctcttt tatcagggtt gg 22

<210> 322
<211> 24
<212> DNA
<213> Homo sapiens

10

<400> 322
ctctgtgcaa gtaagcatct taca 24

<210> 323
<211> 20
<212> DNA
<213> Homo sapiens

<400> 323
cgactgtgta ttttccacag 20

<210> 324
<211> 20
<212> DNA
<213> Homo sapiens

20

<400> 324
agaagcccat atcaatgcac 20

<210> 325
<211> 23
<212> DNA
<213> Homo sapiens

<400> 325
agcttaaagt aggacaacca tgg 23

<210> 326
<211> 20
<212> DNA
<213> Homo sapiens

30

<400> 326
ggatgcttca ctccagaaag 20

<210> 327
<211> 21

<212> DNA
<213> Homo sapiens

<400> 327
tgttggttat ttccacctgc c

21

<210> 328
<211> 18
<212> DNA
<213> Homo sapiens

<400> 328
agagtggctg caggccag

18

10

<210> 329
<211> 25
<212> DNA
<213> Homo sapiens

<400> 329
tttttttttt tacacgaatt tgagg

25

<210> 330
<211> 23
<212> DNA
<213> Homo sapiens

<400> 330
tgaggaagta aaaacaggtc atc

23

20

<210> 331
<211> 23
<212> DNA
<213> Homo sapiens

<400> 331
atgaaatctt aagcagaatc cca

23

<210> 332
<211> 20
<212> DNA
<213> Homo sapiens

<400> 332
cacagagtcc cagggtctgt

20

30

<210> 333
<211> 25
<212> DNA
<213> Homo sapiens

<400> 333
aaaggccttt atttatctct ctctg 25

<210> 334
<211> 18
<212> DNA
<213> Homo sapiens

<400> 334
gcctcagagc tggtaggggt 18

<210> 335
<211> 25 10
<212> DNA
<213> Homo sapiens

<400> 335
gcttctaagt ctttagagtca gctgg 25

<210> 336
<211> 19
<212> DNA
<213> Homo sapiens

<400> 336
agcccacagt cagcctacc 19

<210> 337 20
<211> 20
<212> DNA
<213> Homo sapiens

<400> 337
ttggttaaatt gatgccaga 20

<210> 338
<211> 18
<212> DNA
<213> Homo sapiens

<400> 338
tggtagccact cacatccc 18

<210> 339
<211> 19 30
<212> DNA
<213> Homo sapiens

<400> 339
acacagcatg caggagagag 19

<210> 340
<211> 19
<212> DNA
<213> Homo sapiens

<400> 340
atccctggtg cttaggtgg

19

<210> 341
<211> 21
<212> DNA
<213> Homo sapiens

10

<400> 341
gatggaagta gctcctctcg g

21

<210> 342
<211> 21
<212> DNA
<213> Homo sapiens

<400> 342
ggaaggccag caagtactac c

21

<210> 343
<211> 18
<212> DNA
<213> Homo sapiens

20

<400> 343
ccggtgcttg gaaagatg

18

<210> 344
<211> 20
<212> DNA
<213> Homo sapiens

<400> 344
gaagtgtctc tgttggggga

20

<210> 345
<211> 22
<212> DNA
<213> Homo sapiens

30

<400> 345
ttacaggcat gagtcactac gc

22

<210> 346
<211> 21

<212> DNA
<213> Homo sapiens

<400> 346
accactctca cagcccttac a 21

<210> 347
<211> 18
<212> DNA
<213> Homo sapiens

<400> 347
ccctccctcc acacacac 18 10

<210> 348
<211> 20
<212> DNA
<213> Homo sapiens

<400> 348
gctcactgaa ctttcagggc 20

<210> 349
<211> 20
<212> DNA
<213> Homo sapiens

<400> 349
agatacgggc aaaacactgg 20 20

<210> 350
<211> 21
<212> DNA
<213> Homo sapiens

<400> 350
gttgaatata gagcagggcc c 21

<210> 351
<211> 20
<212> DNA
<213> Homo sapiens 30

<400> 351
ttctgaggtc agggctgtct 20

<210> 352
<211> 21
<212> DNA
<213> Homo sapiens

<400> 352
agcttgga aa atctcgtgtc a

21

<210> 353
<211> 18
<212> DNA
<213> Homo sapiens

<400> 353
actcagtccc tcccaccc

18

<210> 354
<211> 20
<212> DNA
<213> Homo sapiens

10

<400> 354
tcctctcact ccttcccaga

20

<210> 355
<211> 19
<212> DNA
<213> Homo sapiens

<400> 355
gtgatcacgg ctcaacctg

19

<210> 356
<211> 18
<212> DNA
<213> Homo sapiens

20

<400> 356
tggaggactg cttgagcc

18

<210> 357
<211> 19
<212> DNA
<213> Homo sapiens

<400> 357
ctgcagctgc ctcagtttc

19

30

<210> 358
<211> 20
<212> DNA
<213> Homo sapiens

<400> 358
tcaaaagtgc tgggtgacagc

20

<210> 359
<211> 20
<212> DNA
<213> Homo sapiens

<400> 359
atttccagag ccagctcaaa 20

<210> 360
<211> 25
<212> DNA
<213> Homo sapiens 10

<400> 360
ctttaatggt gtgatgacac aaagc 25

<210> 361
<211> 20
<212> DNA
<213> Homo sapiens

<400> 361
gatcatgcac tgttgaccac 20

<210> 362
<211> 25
<212> DNA
<213> Homo sapiens 20

<400> 362
tacatttgaa acatttaaaa cctga 25

<210> 363
<211> 24
<212> DNA
<213> Homo sapiens

<400> 363
aactgagctg taaccagact ggga 24

<210> 364
<211> 22
<212> DNA
<213> Homo sapiens 30

<400> 364
tggaacagtc tggtcctgat gg 22

<210> 365
<211> 25

<212> DNA
<213> Homo sapiens

<400> 365
ttatcccttt attgtttctc ctttg 25

<210> 366
<211> 24
<212> DNA
<213> Homo sapiens

<400> 366
tggtcacctg tatattattgc tagg 24 10

<210> 367
<211> 22
<212> DNA
<213> Homo sapiens

<400> 367
tcttcaaagc ctctgcagta cc 22

<210> 368
<211> 22
<212> DNA
<213> Homo sapiens

<400> 368
ctcatctcca acctgtctaa cc 22 20

<210> 369
<211> 24
<212> DNA
<213> Homo sapiens

<400> 369
gtggctgcag ctaatgtaag acac 24

<210> 370
<211> 24
<212> DNA
<213> Homo sapiens 30

<400> 370
cagcagagac aatggcgtaa gtcc 24

<210> 371
<211> 20
<212> DNA
<213> Homo sapiens

<400> 371
ctgattgaga accagaacag

20

<210> 372
<211> 20
<212> DNA
<213> Homo sapiens

<400> 372
taaagcccta taacctctcc

20

<210> 373
<211> 20
<212> DNA
<213> Homo sapiens

10

<400> 373
tagtaagga ccttcaccag

20

<210> 374
<211> 20
<212> DNA
<213> Homo sapiens

<400> 374
agatgtttgg tatgacttgg

20

<210> 375
<211> 20
<212> DNA
<213> Homo sapiens

20

<400> 375
gatgattaaa ctctcctggc

20

<210> 376
<211> 20
<212> DNA
<213> Homo sapiens

<400> 376
gagacagcta agcactcatg

20

<210> 377
<211> 19
<212> DNA
<213> Homo sapiens

30

<400> 377
gaggtggtgg gcacctgta

19

<210> 378
<211> 20
<212> DNA
<213> Homo sapiens

<400> 378
agaggggagg aacacacctt

20

<210> 379
<211> 20
<212> DNA
<213> Homo sapiens

10

<400> 379
gaccagagtc tgcccagaag

20

<210> 380
<211> 19
<212> DNA
<213> Homo sapiens

<400> 380
tccccagctc tatccaac

19

<210> 381
<211> 20
<212> DNA
<213> Homo sapiens

20

<400> 381
ggaggggatgg acaagtctga

20

<210> 382
<211> 20
<212> DNA
<213> Homo sapiens

<400> 382
gtccagctcg ctgactatcc

20

<210> 383
<211> 20
<212> DNA
<213> Homo sapiens

30

<400> 383
tcaaaacaca gtcattctcca

20

<210> 384
<211> 20

<212> DNA
 <213> Homo sapiens

<400> 384
 gcaaaggctt taccatattg 20

<210> 385
 <211> 16
 <212> DNA
 <213> Homo sapiens

<400> 385
 gctcagcacc cccatt 16 10

<210> 386
 <211> 16
 <212> DNA
 <213> Homo sapiens

<400> 386
 tccctgctcg ggaaac 16

<210> 387
 <211> 20
 <212> DNA
 <213> Homo sapiens

<400> 387
 gttctccaga gagacagcac 20 20

<210> 388
 <211> 19
 <212> DNA
 <213> Homo sapiens

<400> 388
 gagagcaaca ctattgcc 19

<210> 389
 <211> 21
 <212> DNA
 <213> Homo sapiens 30

<400> 389
 tatagacttc agccctgctg c 21

<210> 390
 <211> 20
 <212> DNA
 <213> Homo sapiens

<400> 390 cctctgtagg atgcagttgg	20	
<210> 391 <211> 20 <212> DNA <213> Homo sapiens		
<400> 391 ttgctacgca ctctctact	20	
<210> 392 <211> 20 <212> DNA <213> Homo sapiens		10
<400> 392 gtgaaggcag gaaatgtgac	20	
<210> 393 <211> 20 <212> DNA <213> Homo sapiens		
<400> 393 atcctagacc agaggagccc	20	
<210> 394 <211> 20 <212> DNA <213> Homo sapiens		20
<400> 394 ctccccctgg tccagttatt	20	
<210> 395 <211> 20 <212> DNA <213> Homo sapiens		
<400> 395 aactttcatt tgccaagggga	20	30
<210> 396 <211> 20 <212> DNA <213> Homo sapiens		
<400> 396 agcagatctg ctcttgcgat	20	

<210> 397
<211> 20
<212> DNA
<213> Homo sapiens

<400> 397
acagttgtca tcggtaggca

20

<210> 398
<211> 22
<212> DNA
<213> Homo sapiens

10

<400> 398
aaaagtatga atgggatgga gc

22

<210> 399
<211> 20
<212> DNA
<213> Homo sapiens

<400> 399
gtgcaggtgg cgtttatattt

20

<210> 400
<211> 21
<212> DNA
<213> Homo sapiens

20

<400> 400
ccctatatct ccgtgtgctc c

21

<210> 401
<211> 21
<212> DNA
<213> Homo sapiens

<400> 401
gctctagtgg gaaacctcag g

21

<210> 402
<211> 20
<212> DNA
<213> Homo sapiens

30

<400> 402
gaattccagg ctcttgettg

20

<210> 403
<211> 20

<212> DNA
<213> Homo sapiens

<400> 403
ggtttgggtct caaaggcaaa 20

<210> 404
<211> 20
<212> DNA
<213> Homo sapiens

<400> 404
ccagtacatg gtggtcacca 20

10

<210> 405
<211> 20
<212> DNA
<213> Homo sapiens

<400> 405
gctgccttgg aatttctgtt 20

<210> 406
<211> 18
<212> DNA
<213> Homo sapiens

<400> 406
gtgctgtggt ggggaaag 18

20

<210> 407
<211> 20
<212> DNA
<213> Homo sapiens

<400> 407
attcaagctc atccagaccc 20

<210> 408
<211> 20
<212> DNA
<213> Homo sapiens

30

<400> 408
ggactggccc tttgaaactc 20

<210> 409
<211> 22
<212> DNA
<213> Homo sapiens

<400> 409		
atattgaccg tgcacaaata cg	22	
<210> 410		
<211> 21		
<212> DNA		
<213> Homo sapiens		
<400> 410		
agacctggga aaagtggaga a	21	
<210> 411		
<211> 20		10
<212> DNA		
<213> Homo sapiens		
<400> 411		
attggcagtg gaaaatgctt	20	
<210> 412		
<211> 25		
<212> DNA		
<213> Homo sapiens		
<400> 412		
ttaatctttt gtcaacttcc tgatt	25	
<210> 413		20
<211> 23		
<212> DNA		
<213> Homo sapiens		
<400> 413		
tctgtctctc tttcacgga agc	23	
<210> 414		
<211> 29		
<212> DNA		
<213> Homo sapiens		
<400> 414		
ggataaagaa actccgctct gctggtaga	29	30
<210> 415		
<211> 25		
<212> DNA		
<213> Homo sapiens		
<400> 415		
tcagggcctg tgttgccgca ctctg	25	

<210> 416
<211> 25
<212> DNA
<213> Homo sapiens

<400> 416
agcgatgtaa aggggtaccag tgccg

25

<210> 417
<211> 18
<212> DNA
<213> Homo sapiens

10

<400> 417
aggcatgcaa gcttctta

18

<210> 418
<211> 18
<212> DNA
<213> Homo sapiens

<400> 418
ccgggaggag acatctat

18

<210> 419
<211> 19
<212> DNA
<213> Homo sapiens

20

<400> 419
tggttaagcac agaaaatgc

19

<210> 420
<211> 18
<212> DNA
<213> Homo sapiens

<400> 420
aatggatggg ggattatt

18

<210> 421
<211> 18
<212> DNA
<213> Homo sapiens

30

<400> 421
ctggacgtta tgtctgcc

18

<210> 422
<211> 18

<212> DNA
<213> Homo sapiens

<400> 422
agaggcccag tcacagat 18

<210> 423
<211> 19
<212> DNA
<213> Homo sapiens

<400> 423
atcactctga actgccact 19

10

<210> 424
<211> 20
<212> DNA
<213> Homo sapiens

<400> 424
cccttctgtt tttctgtttt 20

<210> 425
<211> 18
<212> DNA
<213> Homo sapiens

<400> 425
caagctttga aggaagag 18

20

<210> 426
<211> 19
<212> DNA
<213> Homo sapiens

<400> 426
taggacgtta agtgaggac 19

<210> 427
<211> 18
<212> DNA
<213> Homo sapiens

30

<400> 427
gctctgcagt gggtaaaa 18

<210> 428
<211> 18
<212> DNA
<213> Homo sapiens

<400> 428
actctccaag actgtgcg

18

<210> 429
<211> 18
<212> DNA
<213> Homo sapiens

<400> 429
ccctttctga ggcaagat

18

<210> 430
<211> 18
<212> DNA
<213> Homo sapiens

10

<400> 430
gaccacctgg gagagaac

18

<210> 431
<211> 18
<212> DNA
<213> Homo sapiens

<400> 431
cgctatgagt cccatctg

18

<210> 432
<211> 18
<212> DNA
<213> Homo sapiens

20

<400> 432
gatcagctgc aatgaagg

18

<210> 433
<211> 18
<212> DNA
<213> Homo sapiens

<400> 433
ttgagtacac ggggtgac

18

30

<210> 434
<211> 18
<212> DNA
<213> Homo sapiens

<400> 434
cgcaggactg aaagatga

18

<210> 435
<211> 18
<212> DNA
<213> Homo sapiens

<400> 435
acctgtctcc tctcctgg

18

<210> 436
<211> 18
<212> DNA
<213> Homo sapiens

10

<400> 436
tgcttttctt ctgtggga

18

<210> 437
<211> 18
<212> DNA
<213> Homo sapiens

<400> 437
atgaccagca agcattgt

18

<210> 438
<211> 18
<212> DNA
<213> Homo sapiens

20

<400> 438
gtactgggat tacaggcg

18

<210> 439
<211> 18
<212> DNA
<213> Homo sapiens

<400> 439
gcagaagggtc ctttggat

18

<210> 440
<211> 18
<212> DNA
<213> Homo sapiens

30

<400> 440
tttgcaggat tcatgctt

18

<210> 441
<211> 19

<212> DNA
<213> Homo sapiens

<400> 441
cgacattcctt ttctggagg 19

<210> 442
<211> 19
<212> DNA
<213> Homo sapiens

<400> 442
accttttgcac gttggtttt 19 10

<210> 443
<211> 18
<212> DNA
<213> Homo sapiens

<400> 443
gcactttttcc ttccttcc 18

<210> 444
<211> 18
<212> DNA
<213> Homo sapiens

<400> 444
tgcttttgctt tcttctgg 18 20

<210> 445
<211> 19
<212> DNA
<213> Homo sapiens

<400> 445
acagctccag agagaagga 19

<210> 446
<211> 19
<212> DNA
<213> Homo sapiens 30

<400> 446
gcagtcactt gaaaccaga 19

<210> 447
<211> 18
<212> DNA
<213> Homo sapiens

<400> 447 aggcatcaag ctttcctt	18	
<210> 448 <211> 18 <212> DNA <213> Homo sapiens		
<400> 448 ggtttagaga accgagcc	18	
<210> 449 <211> 18 <212> DNA <213> Homo sapiens		10
<400> 449 gtggtgctgc aagttacc	18	
<210> 450 <211> 18 <212> DNA <213> Homo sapiens		
<400> 450 ggaatccctt totttcca	18	
<210> 451 <211> 18 <212> DNA <213> Homo sapiens		20
<400> 451 gaccatttgt tacgcagc	18	
<210> 452 <211> 19 <212> DNA <213> Homo sapiens		
<400> 452 gatgggtgtg aatgaacaa	19	30
<210> 453 <211> 19 <212> DNA <213> Homo sapiens		
<400> 453 ctcaagcttc tgttcatgc	19	

<210> 454
<211> 18
<212> DNA
<213> Homo sapiens

<400> 454
gctgtgagtg tcttggt

18

<210> 455
<211> 18
<212> DNA
<213> Homo sapiens

10

<400> 455
tacagaaaac cgcagctc

18

<210> 456
<211> 18
<212> DNA
<213> Homo sapiens

<400> 456
gccaccaaag gaaagatt

18

<210> 457
<211> 18
<212> DNA
<213> Homo sapiens

20

<400> 457
aaaaggaggg aatcatgg

18

<210> 458
<211> 18
<212> DNA
<213> Homo sapiens

<400> 458
tcacttagca ggaggcag

18

<210> 459
<211> 18
<212> DNA
<213> Homo sapiens

30

<400> 459
ctgagcatcc gatgagac

18

<210> 460
<211> 18

<212> DNA
<213> Homo sapiens

<400> 460
gtgcaaaatg agcagctt

18

<210> 461
<211> 18
<212> DNA
<213> Homo sapiens

<400> 461
tctaaccctt tactgggc

18

10

<210> 462
<211> 18
<212> DNA
<213> Homo sapiens

<400> 462
tcctcaaact gggaatga

18

<210> 463
<211> 18
<212> DNA
<213> Homo sapiens

<400> 463
tttacacagg accagga

18

20

<210> 464
<211> 18
<212> DNA
<213> Homo sapiens

<400> 464
atctcccca ctcagaag

18

<210> 465
<211> 18
<212> DNA
<213> Homo sapiens

<400> 465
gtccacgggc tttattct

18

30

<210> 466
<211> 22
<212> DNA
<213> Homo sapiens

<400> 466
 tgagcataaa tttcattagc tg 22

 <210> 467
 <211> 20
 <212> DNA
 <213> Homo sapiens

 <400> 467
 ggaagagcaa aataaatcca 20

 <210> 468
 <211> 19
 <212> DNA
 <213> Homo sapiens

 <400> 468
 ggtgcacaga attgttcat 19

 <210> 469
 <211> 18
 <212> DNA
 <213> Homo sapiens

 <400> 469
 agcacgctta tttcatgg 18

 <210> 470
 <211> 18
 <212> DNA
 <213> Homo sapiens

 <400> 470
 gtaacaccag cagggaca 18

 <210> 471
 <211> 18
 <212> DNA
 <213> Homo sapiens

 <400> 471
 tctgctgca ttatggat 18

 <210> 472
 <211> 18
 <212> DNA
 <213> Homo sapiens

 <400> 472
 gggggtgaga agtaggaa 18

10

20

30

<210> 473
<211> 18
<212> DNA
<213> Homo sapiens

<400> 473
atgggggatta aatacggg

18

<210> 474
<211> 18
<212> DNA
<213> Homo sapiens

10

<400> 474
agctagcatt gggctctt

18

<210> 475
<211> 18
<212> DNA
<213> Homo sapiens

<400> 475
ctgaggagaa gaggctgg

18

<210> 476
<211> 18
<212> DNA
<213> Homo sapiens

20

<400> 476
cgccttaciaa ggcaagta

18

<210> 477
<211> 18
<212> DNA
<213> Homo sapiens

<400> 477
aggatgcttg ctagggtt

18

<210> 478
<211> 18
<212> DNA
<213> Homo sapiens

30

<400> 478
cacaagtgtc tggaaggc

18

<210> 479
<211> 18

<212> DNA
<213> Homo sapiens

<400> 479
ggtctcagga gcccttta

18

<210> 480
<211> 21
<212> DNA
<213> Homo sapiens

<400> 480
acatgccact cttctcacta a

21

10

<210> 481
<211> 18
<212> DNA
<213> Homo sapiens

<400> 481
acttaaccaa ggatgggg

18

<210> 482
<211> 18
<212> DNA
<213> Homo sapiens

<400> 482
caaccacga gcataaga

18

20

<210> 483
<211> 18
<212> DNA
<213> Homo sapiens

<400> 483
taggctctgc actcttgg

18

<210> 484
<211> 18
<212> DNA
<213> Homo sapiens

<400> 484
accacaggag tctctctc

18

30

<210> 485
<211> 18
<212> DNA
<213> Homo sapiens

<400> 485
taaaggcggg gaagtgag

18

<210> 486
<211> 18
<212> DNA
<213> Homo sapiens

<400> 486
ctaccgctct cctaggct

18

<210> 487
<211> 18
<212> DNA
<213> Homo sapiens

10

<400> 487
tggggccaga taattctt

18

<210> 488
<211> 18
<212> DNA
<213> Homo sapiens

<400> 488
ctgggtgttg gtgggtgtt

18

<210> 489
<211> 18
<212> DNA
<213> Homo sapiens

20

<400> 489
aaggaagagg tcaccagg

18

<210> 490
<211> 18
<212> DNA
<213> Homo sapiens

<400> 490
cacaaattcc atttcca

18

30

<210> 491
<211> 21
<212> DNA
<213> Homo sapiens

<400> 491
tcaatagggtg atccaacatt t

21

<210> 492
 <211> 18
 <212> DNA
 <213> Homo sapiens

<400> 492
 aaagtccac aaagggtc

18

<210> 493
 <211> 18
 <212> DNA
 <213> Homo sapiens

10

<400> 493
 gggtagggg atcttttt

18

<210> 494
 <211> 18
 <212> DNA
 <213> Homo sapiens

<400> 494
 tgtggaacat tcattggc

18

<210> 495
 <211> 18
 <212> DNA
 <213> Homo sapiens

20

<400> 495
 gtcctgggaa agatggaa

18

<210> 496
 <211> 18
 <212> DNA
 <213> Homo sapiens

<400> 496
 tcaaagcgtc tcccataa

18

<210> 497
 <211> 18
 <212> DNA
 <213> Homo sapiens

30

<400> 497
 tctttcgctg tacttggc

18

<210> 498
 <211> 18

<212> DNA
<213> Homo sapiens

<400> 498
tgggaggtca gagtgatg

18

<210> 499
<211> 19
<212> DNA
<213> Homo sapiens

<400> 499
ggacagtgtgta tgtgttggg

19

10

<210> 500
<211> 18
<212> DNA
<213> Homo sapiens

<400> 500
aggcagctgt ttttgtga

18

<210> 501
<211> 18
<212> DNA
<213> Homo sapiens

<400> 501
cttcttgagt cccgtgtg

18

20

<210> 502
<211> 19
<212> DNA
<213> Homo sapiens

<400> 502
caaccgagaa tcctctagc

19

<210> 503
<211> 18
<212> DNA
<213> Homo sapiens

<400> 503
gctgggagag aatcacao

18

30

<210> 504
<211> 18
<212> DNA
<213> Homo sapiens

<400> 504
gctttgcaga agagacca

18

<210> 505
<211> 18
<212> DNA
<213> Homo sapiens

<400> 505
acgctgtcag gtcacact

18

<210> 506
<211> 18
<212> DNA
<213> Homo sapiens

10

<400> 506
ggaggatgct caggtgat

18

<210> 507
<211> 18
<212> DNA
<213> Homo sapiens

<400> 507
tagggggatc tttttcca

18

<210> 508
<211> 18
<212> DNA
<213> Homo sapiens

20

<400> 508
gagcaatttg aaaagcca

18

<210> 509
<211> 18
<212> DNA
<213> Homo sapiens

<400> 509
atgggtccagc tcctctgt

18

<210> 510
<211> 18
<212> DNA
<213> Homo sapiens

30

<400> 510
atagagcacc ccatctcc

18

<210> 511		
<211> 18		
<212> DNA		
<213> Homo sapiens		
<400> 511		
aacattgctg ttagccca	18	
<210> 512		
<211> 18		
<212> DNA		
<213> Homo sapiens		10
<400> 512		
gcaatcgaaa cagcattc	18	
<210> 513		
<211> 18		
<212> DNA		
<213> Homo sapiens		
<400> 513		
atgagttggc agctgaag	18	
<210> 514		
<211> 18		
<212> DNA		
<213> Homo sapiens		20
<400> 514		
aatgaaggtc ttgcctcc	18	
<210> 515		
<211> 20		
<212> DNA		
<213> Homo sapiens		
<400> 515		
gaggagaaga tccacaagcg	20	
<210> 516		
<211> 20		
<212> DNA		
<213> Homo sapiens		30
<400> 516		
tctctggggc atactgaacc	20	
<210> 517		
<211> 18		

<212> DNA
<213> Homo sapiens

<400> 517
ctgagctttt ggcactgt

18

<210> 518
<211> 18
<212> DNA
<213> Homo sapiens

<400> 518
ctgctaggtg acagcagg

18

10

<210> 519
<211> 20
<212> DNA
<213> Homo sapiens

<400> 519
tgtatgagtc tggagggtgt

20

<210> 520
<211> 18
<212> DNA
<213> Homo sapiens

<400> 520
acacctggct gaggaaat

18

20

<210> 521
<211> 18
<212> DNA
<213> Homo sapiens

<400> 521
gcaggggacg tgataata

18

<210> 522
<211> 18
<212> DNA
<213> Homo sapiens

<400> 522
ttttgcttcc taccatgc

18

30

<210> 523
<211> 18
<212> DNA
<213> Homo sapiens

<400> 523 aaaattgtga gcacctcc	18	
<210> 524 <211> 23 <212> DNA <213> Homo sapiens		
<400> 524 tttatattta aagtggcttt gtt	23	
<210> 525 <211> 18 <212> DNA <213> Homo sapiens		10
<400> 525 gtgcaaagcc cacagtat	18	
<210> 526 <211> 18 <212> DNA <213> Homo sapiens		
<400> 526 aggaaaatgc aagagcag	18	
<210> 527 <211> 20 <212> DNA <213> Homo sapiens		20
<400> 527 ccactgaatt gcatactttg	20	
<210> 528 <211> 18 <212> DNA <213> Homo sapiens		
<400> 528 tctgggtcca gtctgcta	18	30
<210> 529 <211> 18 <212> DNA <213> Homo sapiens		
<400> 529 agattttggg gagtcagg	18	

<210> 530
<211> 17
<212> DNA
<213> Homo sapiens

<400> 530
gcgctcaagc aattctc 17

<210> 531
<211> 18
<212> DNA
<213> Homo sapiens

<400> 531
caagccccaa agtagtca 18

<210> 532
<211> 18
<212> DNA
<213> Homo sapiens

<400> 532
gaatcatcca atccacga 18

<210> 533
<211> 18
<212> DNA
<213> Homo sapiens

<400> 533
agcctccagg tgactacc 18

<210> 534
<211> 18
<212> DNA
<213> Homo sapiens

<400> 534
gaaggacatg gtcagcag 18

<210> 535
<211> 18
<212> DNA
<213> Homo sapiens

<400> 535
atgctttcag cattttcg 18

<210> 536
<211> 18

10

20

30

<212> DNA
<213> Homo sapiens

<400> 536
tgatccgtgg tagggtta 18

<210> 537
<211> 18
<212> DNA
<213> Homo sapiens

<400> 537
gtcggattgg tttcacaa 18

10

<210> 538
<211> 19
<212> DNA
<213> Homo sapiens

<400> 538
ttttatggga atttcagcc 19

<210> 539
<211> 21
<212> DNA
<213> Homo sapiens

<400> 539
tttggaag aacagaaatg t 21

20

<210> 540
<211> 19
<212> DNA
<213> Homo sapiens

<400> 540
ggctagtctt tctgaacc 19

<210> 541
<211> 18
<212> DNA
<213> Homo sapiens

30

<400> 541
ccttaatgcc cctgattc 18

<210> 542
<211> 18
<212> DNA
<213> Homo sapiens

<400> 542
gcgttttataa gctgagga

18

<210> 543
<211> 18
<212> DNA
<213> Homo sapiens

<400> 543
tcaagcttgc tttctcaa

18

<210> 544
<211> 18
<212> DNA
<213> Homo sapiens

10

<400> 544
gtagccacgc aagtgtct

18

<210> 545
<211> 18
<212> DNA
<213> Homo sapiens

<400> 545
cctggctgga gataggat

18

<210> 546
<211> 18
<212> DNA
<213> Homo sapiens

20

<400> 546
cttccctctt gcctatgt

18

<210> 547
<211> 18
<212> DNA
<213> Homo sapiens

<400> 547
ggcacgtact tcctacca

18

<210> 548
<211> 18
<212> DNA
<213> Homo sapiens

30

<400> 548
ggtgcttctt acaggcaa

18

<210> 549
<211> 16
<212> DNA
<213> Homo sapiens

<400> 549
acccaggctg gtgtgt 16

<210> 550
<211> 23
<212> DNA
<213> Homo sapiens

<400> 550
actgagttaa ttatcactcc cct 23

<210> 551
<211> 18
<212> DNA
<213> Homo sapiens

<400> 551
gatgcatttt gcttcacc 18

<210> 552
<211> 21
<212> DNA
<213> Homo sapiens

<400> 552
tctgctttta gagctgtag c 21

<210> 553
<211> 19
<212> DNA
<213> Homo sapiens

<400> 553
tcaagcttca aagagcaga 19

<210> 554
<211> 18
<212> DNA
<213> Homo sapiens

<400> 554
ggagtacatc ccaggacc 18

<210> 555
<211> 19

10

20

30

<212> DNA
<213> Homo sapiens

<400> 555
tgggtgctttt aaatccaga

19

<210> 556
<211> 21
<212> DNA
<213> Homo sapiens

<400> 556
ctcccttact tacttgcatt g

21

10

<210> 557
<211> 18
<212> DNA
<213> Homo sapiens

<400> 557
tcttctccca gggaatct

18

<210> 558
<211> 18
<212> DNA
<213> Homo sapiens

<400> 558
tttatgtccc ctgagcac

18

20

<210> 559
<211> 19
<212> DNA
<213> Homo sapiens

<400> 559
tccctggcta tcttgaatc

19

<210> 560
<211> 16
<212> DNA
<213> Homo sapiens

<400> 560
cttgactggg tccacg

16

30

<210> 561
<211> 20
<212> DNA
<213> Homo sapiens

<400> 561
cgagacgccg gtagatacca

20

<210> 562
<211> 20
<212> DNA
<213> Homo sapiens

<400> 562
catcctccat gcctttcagt

20

<210> 563
<211> 20
<212> DNA
<213> Homo sapiens

10

<400> 563
agttccagag aacgagacgc

20

<210> 564
<211> 20
<212> DNA
<213> Homo sapiens

<400> 564
cttgtcatcc tccatgcctt

20

<210> 565
<211> 20
<212> DNA
<213> Homo sapiens

20

<400> 565
gagcgtgaga ggttgaggag

20

<210> 566
<211> 20
<212> DNA
<213> Homo sapiens

<400> 566
aaacaaactc cagacgcacc

20

<210> 567
<211> 24
<212> DNA
<213> Homo sapiens

30

<400> 567
ctgaaccact acctgtatga cctg

24

<210> 568
<211> 25
<212> DNA
<213> Homo sapiens

<400> 568
ctaactactt actcctacag ggccc 25

<210> 569
<211> 23
<212> DNA
<213> Homo sapiens 10

<400> 569
gaagcatttc aatactttaa ctg 23

<210> 570
<211> 20
<212> DNA
<213> Homo sapiens

<400> 570
ccactccagt gcaccaatc 20

<210> 571
<211> 20
<212> DNA
<213> Homo sapiens 20

<400> 571
cttctcctgg ccactctgac 20

<210> 572
<211> 21
<212> DNA
<213> Homo sapiens

<400> 572
ggtttacctt tgaatcccag c 21

<210> 573
<211> 22
<212> DNA
<213> Homo sapiens 30

<400> 573
tgaggatgaa tgagcacata gg 22

<210> 574
<211> 21

<212> DNA
<213> Homo sapiens

<400> 574
tttgtggtcc attgagtagg c

21

<210> 575
<211> 20
<212> DNA
<213> Homo sapiens

<400> 575
aggggaagga atgtgcttgg

20

10

<210> 576
<211> 20
<212> DNA
<213> Homo sapiens

<400> 576
ttcggctgag cgggcagtgt

20

<210> 577
<211> 26
<212> DNA
<213> Homo sapiens

<400> 577
attgaaggctc ctccaaaaga atgctg

26

20

<210> 578
<211> 30
<212> DNA
<213> Homo sapiens

<400> 578
agaacgtcaa catatctttt tgggggacac

30

<210> 579
<211> 22
<212> DNA
<213> Homo sapiens

<400> 579
ttgtatttga ggactttgct cg

22

30

<210> 580
<211> 20
<212> DNA
<213> Homo sapiens

<400> 580
cggtaccatc ctctctctcc 20

<210> 581
<211> 20
<212> DNA
<213> Homo sapiens

<400> 581
tttttgctc atctatgcc 20

<210> 582
<211> 20
<212> DNA
<213> Homo sapiens

<400> 582
gggtgacaga gcaagactcc 20

<210> 583
<211> 18
<212> DNA
<213> Homo sapiens

<400> 583
ttgctcaagt tctcctgg 18

<210> 584
<211> 18
<212> DNA
<213> Homo sapiens 20

<400> 584
accttgtttt gaggggag 18

<210> 585
<211> 18
<212> DNA
<213> Homo sapiens

<400> 585
cttggtatt tggacagc 18

<210> 586
<211> 18
<212> DNA
<213> Homo sapiens 30

<400> 586
gggcatttac tcacttgc 18

<210> 587
<211> 19
<212> DNA
<213> Homo sapiens

<400> 587
cttgtgtcag ttgtcaggg

19

<210> 588
<211> 19
<212> DNA
<213> Homo sapiens

<400> 588
tggaattggt gtgtcttgg

19

<210> 589
<211> 18
<212> DNA
<213> Homo sapiens

<400> 589
ccagttccac tggatggt

18

<210> 590
<211> 18
<212> DNA
<213> Homo sapiens

<400> 590
atgggctgtg tttctcaa

18

<210> 591
<211> 18
<212> DNA
<213> Homo sapiens

<400> 591
ctgcctatcc ctggactt

18

<210> 592
<211> 18
<212> DNA
<213> Homo sapiens

<400> 592
agtttgtccc tagtgccc

18

<210> 593
<211> 19

10

20

30

<212> DNA		
<213> Homo sapiens		
<400> 593		
caacacgtct gacatccat	19	
<210> 594		
<211> 16		
<212> DNA		
<213> Homo sapiens		
<400> 594		
ggatagtgca caccca	16	10
<210> 595		
<211> 22		
<212> DNA		
<213> Homo sapiens		
<400> 595		
tgggtgggtac tattgttccc at	22	
<210> 596		
<211> 20		
<212> DNA		
<213> Homo sapiens		
<400> 596		
agttccagcc cccttaccag	20	20
<210> 597		
<211> 20		
<212> DNA		
<213> Homo sapiens		
<400> 597		
ggccactatc atccctgtgt	20	
<210> 598		
<211> 20		
<212> DNA		
<213> Homo sapiens		
<400> 598		
tttcacatgg gaagaacacg	20	30
<210> 599		
<211> 20		
<212> DNA		
<213> Homo sapiens		

<400> 599		
acagtgcacac tagggacggg	20	
<210> 600		
<211> 20		
<212> DNA		
<213> Homo sapiens		
<400> 600		
tgccaggatg gagataacaa	20	
<210> 601		
<211> 20		10
<212> DNA		
<213> Homo sapiens		
<400> 601		
cctgtggcac acatatcacc	20	
<210> 602		
<211> 20		
<212> DNA		
<213> Homo sapiens		
<400> 602		
acaaccaaga atggagccac	20	
<210> 603		20
<211> 20		
<212> DNA		
<213> Homo sapiens		
<400> 603		
tgctgtgtaa caagtcccca	20	
<210> 604		
<211> 20		
<212> DNA		
<213> Homo sapiens		
<400> 604		
tgaacggagg acctaccaag	20	30
<210> 605		
<211> 20		
<212> DNA		
<213> Homo sapiens		
<400> 605		
gcagggtccg actcactaag	20	

<210> 606
<211> 20
<212> DNA
<213> Homo sapiens

<400> 606
gctgtgagtt ccctttacgc

20

<210> 607
<211> 20
<212> DNA
<213> Homo sapiens

<400> 607
acagtgggga caaagacagg

20

<210> 608
<211> 20
<212> DNA
<213> Homo sapiens

<400> 608
tacagggcac ctcccagtag

20

<210> 609
<211> 21
<212> DNA
<213> Homo sapiens

<400> 609
tcttctgtta aggtttcccc c

21

<210> 610
<211> 20
<212> DNA
<213> Homo sapiens

<400> 610
tgtctcaaac ctccctctgc

20

<210> 611
<211> 20
<212> DNA
<213> Homo sapiens

<400> 611
aacatatttc ctcccagcc

20

<210> 612
<211> 19

10

20

30

<212> DNA
<213> Homo sapiens

<400> 612
cagtcccagc caatgagaa 19

<210> 613
<211> 20
<212> DNA
<213> Homo sapiens

<400> 613
ctcctctgca tgggagaatc 20 10

<210> 614
<211> 20
<212> DNA
<213> Homo sapiens

<400> 614
agacctggga ccagtctgtg 20

<210> 615
<211> 20
<212> DNA
<213> Homo sapiens

<400> 615
gggagacgac gtcacaagat 20 20

<210> 616
<211> 20
<212> DNA
<213> Homo sapiens

<400> 616
tgatgttggg aagatggtga 20

<210> 617
<211> 20
<212> DNA
<213> Homo sapiens 30

<400> 617
caggcatctt ctatgtgcca 20

<210> 618
<211> 20
<212> DNA
<213> Homo sapiens

<400> 618
gggaggcaca agttctttca

20

<210> 619
<211> 20
<212> DNA
<213> Homo sapiens

<400> 619
acttcgtggc actgagtgtg

20

<210> 620
<211> 20
<212> DNA
<213> Homo sapiens

10

<400> 620
cctttctttac ggatgaggca

20

<210> 621
<211> 20
<212> DNA
<213> Homo sapiens

<400> 621
ggctgctgag ctctttctgat

20

<210> 622
<211> 20
<212> DNA
<213> Homo sapiens

20

<400> 622
tgggtctctc tgcttgactt

20

<210> 623
<211> 20
<212> DNA
<213> Homo sapiens

<400> 623
tcacctactt ccagcttccg

20

<210> 624
<211> 20
<212> DNA
<213> Homo sapiens

30

<400> 624
agacctggga ccagtctgtg

20

<210> 625
<211> 20
<212> DNA
<213> Homo sapiens

<400> 625
ctcctctgca tgggagaatc 20

<210> 626
<211> 20
<212> DNA
<213> Homo sapiens

<400> 626
aattcaggag acctgggacc 20

<210> 627
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence is a BstXI-linker adapter.

<400> 627
gtcttcacca cgggg 15

<210> 628
<211> 11
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence is a BstXI-linker adapter.

<400> 628
gtggtgaaga c 11

<210> 629
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence is a primer.

<400> 629
ccaagttctg agaagtcc 18

<210> 630

10

20

30

<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence is a primer.

<400> 630
aatacctgaa accatac 17

<210> 631
<211> 17
<212> DNA
<213> Artificial Sequence

10

<220>
<223> Artificial Sequence is an allele specific oligonucleotide.

<400> 631
agactggggt gagacgc 17

<210> 632
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence is an allele specific oligonucleotide.

20

<400> 632
cagactgggt tgagacgc 19

<210> 633
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence is a primer.

<400> 633
cccgtgtgct ccgccgccca gttc 24

30

<210> 634
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence is a primer.

<400> 634
ggctcagga gctcatcatg gactt

25

<210> 635
<211> 502
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence is a primer.

<400> 635
ccggtgtgct cgcgcgccca gtccccctgc gcgcggggtc agtgtgtgga cctgcgcctg 60
cgctgcgacg gcgaggcaga ctgtcaggac cgtcagacg aggtggactg tgacgccatc 120
tgccctgcccc accagttccg gtgtgcgagc ggccagtgtg tctcatcaa acagcagtgc 180
gactccttcc cgcactgtat cgacggctcc gacgagctca tgtgtgaaat caccaagccg 240
ccctcagacg acagcccggc ccacagcagt gccatcgggc cgtcattgg catcatcctc 300
tctctcttcg tcatgggtgg tgttatattt gtgtgccagc gcgtgggtgtg ccagcgtat 360
gcggggggcca acgggcccctt cccgcacgag tatgtcagcg ggaccccgca cgtgccctc 420
aatttcatag cccggggcgg ttcccagcat ggccccttca caggcatcgc atgcggaaag 480
tccatgatga gtcctgtgag cc 502

10

<210> 636
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence is a primer.

20

<400> 636
agcgaggcca ccatccacag g

21

<210> 637
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence is a primer.

<400> 637
tcgctggtcg gcataatcaa t

21

30

<210> 638
<211> 501
<212> DNA
<213> Artificial Sequence

<220>

<223> Artificial Sequence is a primer.

<400> 638

```

agcagagcca ccatccacag gatctccctg gagactaaca acaacgatgt ggctatccca 60
ctcacgggtg tcaaagaggc ctctgcactg gactttgatg tgtccaacaa tcacatctac 120
tggactgatg ttagcctcaa gacgatcagc cgagccttca tgaatgggag ctcagtggag 180
cacgtgattg agtttggcct cgactaccct gaagggaatgg ctgtggactg gatgggcaag 240
aacctctatt gggcggacac agggaccaac aggattgagg tggcccggtt ggatgggcag 300
ttccggcagg tgcttgtgtg gagagacctt gacaacccca ggtctctggc tctggatcct 360
actaaaggct acatctactg gactgagtgg ggtggcaagc caaggattgt gcgggccttc 420
atggatggga ccaattgtat gacactggta gacaagggtg gccgggcca cgcacctacc 480
attgattatg ccgaccagcg a                                     501

```

10

<210> 639

<211> 26

<212> RNA

<213> Artificial Sequence

<220>

<223> Artificial Sequence is a Zmax1 oligonucleotide.

<400> 639

```

raguacagcu ucuugccaac ccaguc                                     26

```

<210> 640

<211> 26

<212> RNA

<213> Artificial Sequence

20

<220>

<223> Artificial Sequence is a Zmax1 oligonucleotide.

<400> 640

```

ruccuccagg ucgaugguca gcccau                                     26

```

<210> 641

<211> 26

<212> RNA

<213> Artificial Sequence

<220>

<223> Artificial Sequence is a Zmax1 oligonucleotide.

30

<400> 641

```

rgucugaguc cgagucaaaa uccagg                                     26

```

<210> 642

<211> 3081

<212> DNA

<213> Homo sapiens

<400> 642

```

cccgccagcc cagcccagcc caaccctact ccctccccc ggcagggcag cagccgttgc      60
tcagagagaa ggtggaggaa gaaatccaga ccctagcacg cgcgcaccat catggaccat      120
tatgattctc agcaaaccaa cgattacatg cagccagaag aggactggga ccgggacctg      180
ctcctggacc cggcctggga gaagcagcag agaaagacat tcacggcatg gtgtaaactcc      240
cacctccgga aggcggggac acagatcgag aacatcgaag aggacttccg ggtatggcctg      300
aagctcatgc tgctgctgga ggtcatctca ggtgaacgct tggccaagcc agagcgaggc      360
aagatgagag tgcacaagat ctccaacgtc aacaaggccc tggatttcat agccagcaaa      420
ggcgctcaaac tgggtgccat cggagccgaa gaaatcgtgg atgggaatgt gaagatgacc      480
ctgggcatga tctggaccat catcctggcg tttgccatcc aggacatctc cgtggaagag      540
acttcagcca aggaagggct gctcctgtgg tgtcagagaa agacagcccc ttacaaaaat      600
gtcaacatcc agaacttcca cataagctgg aaggatggcc tcggcttctg tgccttgatc      660
caccgacacc ggcccagact gattgactac gggaagctgc ggaaggatga tccactcaca      720
aatctgaata cggcttttga cgtggcagag aagtacctgg acatcccaa gatgctggat      780
gccgaagaca togttggaac tgcccgaacc gatgagaaag ccatcatgac ttacgtgtct      840
agcttctacc acgccttctc tggagcccag aaggcggaga cagcagccaa tcgcatctgc      900
aagggtgttg ccgtcaacca ggagaacgag cagcttatgg aagactacga gaagctggcc      960
agtgatctgt tggagtggat ccgcgcaca atccctggc tggagaaccc ggtgcccgag      1020
aacaccatgc atgccatgca acagaagctg gaggacttcc gggactaccg gcgcctgcac      1080
aagccgcccc aggtgcagga gaagtgccag ctggagatca acttcaacac gctgcagacc      1140
aagctgcggc tcagcaaccg gcctgccttc atgccctctg agggcaggat ggtctcggac      1200
atcaacaatg cctggggctg cctggagcag gtggagaagg gctatgagga gtggttgctg      1260
aatgagatcc ggaggctgga gcgactggac cacctggcag agaagtccg gcagaaggcc      1320
tccatccacc aggcctggac tgacggcaaa gaggccatgc tcgcacagaa ggactatgag      1380
accgcccacc tctcggagat caaggccttg ctcaagaagc atgaggcctt cgagagtgc      1440
ctgggtgccc accaggaccg tgtggagcag attgcgcaca tcgcacagga gctcaatgag      1500
ctggactatt atgactcacc cagtgtcaac gcccgctgcc aaaagatctg tgaccagtgg      1560
gacaatctgg gggccctaac tcagaagcga agggaagctc tggagcggac cgagaaactg      1620
ctggagacca ttgaccagct gtacttgag tatgccaaag gggctgcacc ctccaacaac      1680
tggatggagg gggccatgga ggacctgcag gacaccttca ttgtgcacac cattgaggag      1740
atccaggggc tgaccacagc ccatgagcag ttcaaggcca ccctccctga tgccgacaag      1800
gagcgctcgg ccatcctggg catccacaat gaggtgtcca agattgtcca gacctaccac      1860
gtcaatatgg cgggcaccaa ccctacaca accatcacgc ctcaggagat caatggcaaa      1920
tgggaccacc tgcggcagct ggtgcctcgg agggaccaag ctctgacgga ggagcatgcc      1980
cgacagcagc acaatgagag gctacgcaag cagtttgag ccagggccaa tgtcatcggg      2040
ccctggatcc agaccaagat ggaggagatc gggaggatct ccattgagat gcatgggacc      2100
ctggaggacc agctcagcca cctgcggcag tatgagaaga gcacgtcaa ctacaagcca      2160
aagattgata agctggaggg cgaccaccag ctcacccagg aggcgctcat cttcgacaac      2220
aagcacacca actacaccat ggagcacatc cgtgtgggct gggagcagct gctcaccacc      2280
atcgccagga ccatcaatga ggtagagaac cagatcctga cccgggatgc caagggcac      2340
agccaggagc agatgaatga gttccgggcc tccctcaacc actttgaccg ggatcaactc      2400
ggcacactgg gtcccaggga gttcaaaagg tgcctcatca gcttgggtta tgatatggc      2460
aacgaccccc agggagaagc agaatttgcc cgcacatga gcattgtgga cccaaccgc      2520
ctgggggtag tgacattcca ggcccttcatt gacttcatgt cccgcgagac agccgacaca      2580
gatacagcag accaagtcat ggcttccttc aagatcctgg ctggggacaa gaactacatt      2640
accatggacg agctgcgccg cgagctgcca cccgaccagg ctgagtactg catcgcgcg      2700
atggccccc acaccggccc cgactccgtg ccagggtgctc tggactacat gtccttctcc      2760

```

10

20

30

```

acggcgctgt acggcgagag tgacctetaa tccacccgcg ccggccgccc tegtcttgtg 2820
cgccgtgccc acagatgtga aatgaatgta atctaataga agcctaataca gcccaccatg 2880
ttctccactg aaaaatcctc tttctttggg gttttttttt ctttcttttt tgattttgca 2940
ctggacgggtg acgtcagcct gtacaggctc ccagggttgg cgtcaaattgc tattgaaatt 3000
gcgctgaatc gtagtgccttt tctttttgat aaataaacia tgtaaaaatg tttcaaaaac 3060
ctaataaaat aaataaatac g 3081

```

<210> 643

<211> 1324

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(1324)

<223> n = A,T,C or G

10

<400> 643

```

ggccgccccg cgccccagc agnccgagcc gggggcgaca gncggggcgcc agcccccgcc 60
ccccgccccg attgacatga tgtttccaca aagcaggcat tcgggctcct cgcacctacc 120
ccagcaactc aaattcacca cctcggactc ctgcgaccgc atcaaagacg aatttcagct 180
actgcaagct cagtaccaca gcctcaagct cgaatgtgac aagttggcca gtgagaagtc 240
agagatgcag cgtcactatg tgatgtacta cgagatgtcc tacggcttga acatcgagat 300
gcacaaacag gctgagatcg tcaaaaggct gaacgggatt tgtgcccagg tctgcacctc 360
cctctcccaa gagcaccagc agcaggctct gggagccatt gagaggggcca agcaggctcc 420
cgctccccgag ctgaacteta tcatccgaca gcagctccaa gcccaccagc tgtcccagct 480
gcaggccctg gcctgcccct tgaccccact acccgtgggg ctgcagccgc cttcgtctgc 540
ggcggtcagc gcaggcaccg gcctcctctc gctgtccgcg ctgggttccc aggccaccct 600
ctccaaggaa gacaagaacg ggcacgatgg tgacaccac caggaggatg atggcgagaa 660
gtcggattag cagggggccg ggacaggag gttgggagg ggacagagg ggagacagag 720
gcacggagag aaaggaatgt ttagcacaag acacagcgga gctcgggatt ggctaattct 780
ccatagtatt tatggtggcg ccggcggggc cccagcccag cttgcaggcc acctctagct 840
ttcttcttac cccattccgg cttccctcct cctccctgc agcctgggta ggtggatacc 900
tgccctgaca tgtgaggcaa gctaaggcct ggagggtcag atgggagacc aggtcccaag 960
ggagcaagac ctgcgaagcg cagcagcccc ggcccttccc ccgttttgaa catgtgtaac 1020
cgacagtctg ccttggggcca cagccctctc accctgggtac tgcatgcaag caatgctagc 1080
tgcccttttc ccgtcctggg caccctcagc ctcccccgac cccgggtccc aggtatgctc 1140
ccacctccac ctgccccact caccacctct gctagtcca gacacctcca cggccacctg 1200
gtcctctccc atcgcccaca aaaggggggg cagcaggggc gagcttagct gagctgggag 1260
gagcagggtg aggggtggggc acccaggatt cccctcccc tccccaaata aagatgaggg 1320
tact 1324

```

20

<210> 644

<211> 2377

<212> DNA

<213> Homo sapiens

30

<400> 644

```

ggtgacaaag agccaacaga gacaatagga gacttgctca tttgtcttga tgggctacag 60
ttagagtctg aagttgttac caatggtgaa actacatggt cagaaagtgc ttctcagaat 120

```



```

gatgatggct ccagatccaa ggatgaaaca agagtggagca caaatggatc agatgaccct 180
gaagatgcag gagctggtga aaataggaga gtcagtggga ataattctcc atcactctca 240
aatgggtggtt ttaaaccttc tagacctcca agaccttcac gaccaccacc acccacccca 300
cgtagaccag catctgtcaa tggttcacca tctgccactt ctgaaagtga tgggtctagt 360
acaggctctc tgccgccgac aaatacaaat acaaatacat ctgaaggagc aacatctgga 420
ttaataatc ctcttactat atctggaggc tcaggcccta ggccattaaa tctgttaact 480
caagctccct tggccactgg ttgggagcag agagtggacc agcacgggag agtttactat 540
gtagatcatg ttgagaaaag aacaacatgg gatagaccag aacctctacc tcctggctgg 600
gaacggcggg ttgacaacat gggacgtatt tattatgttg accatttcac aagaacaaca 660
acgtggcaga ggccaacact ggaatccgtc cggaactatg aacaatggca gctacagcgt 720
agtcagcttc aaggagcaat gcagcagttt aaccagagat tcatttatgg gaatcaagat 780
ttatttgcta catcacaag taaagaattt gatcctcttg gtccattgcc acctggatgg 840
gagaagagaa cagacagcaa tggcagagta tatttcgtca accacaacac acgaattaca 900
caatgggaag accccagaag tcaaggtcaa ttaaatgaaa agcccttacc tgaaggttgg 960
gaaatgagat tcacagtggg tggaaattcca tattttgtgg accacaatag aagaactacc 1020
acctatatag atccccgcac aggaaaatct gccctagaca atggaacctca gatagcctat 1080
gttcgggact tcaaaacaaa ggttcagtat ttccggttct ggtgtcagca actggccatg 1140
ccacagcaca taaagattac agtgacaaga aaaacattgt ttgaggattc otttcaacag 1200
ataatgagct tcagtcacca agatctgcga agacgtttgt gggtgatttt tccaggagaa 1260
gaaggtttag attatggagg ttagcaaga gaatggttct ttctttgtc acatgaagtg 1320
ttgaacccaa tgtattgctt gtttgaatat gcagggaagg ataactactg cttgcagata 1380
aaccocgctt cttacatcaa tccagatcac ctgaaatatt ttcgttttat tggcagattt 1440
attgccatgg ctctgttcca tgggaaattc atagacacgg gtttttcttt accattctat 1500
aagcgtatct tgaacaaacc agttggactc aaggatttag aatctattga tccagaattt 1560
tacaattctc tcactctggg taaggaaaac aatattgagg aatgtgattt ggaaatgtac 1620
ttctccgttg acaaaagaat tctaggtgaa attagagtc atgatctgaa acctaatggt 1680
ggcaatattc ttgtaacaga agaaaataaa gaggaatata tcagaatggt agctgagtgg 1740
aggttgtctc gaggtgttga agaacagaca caagctttct ttgaaggctt taatgaaatt 1800
cttccccagc aatatttgca atactttgat gcaaaggaat tagaggtcct tttatgtgga 1860
atgcaagaga ttgatttgaa tgactggcaa agacatgcc tctaccgtca ttatgcaagg 1920
accagcaaac aaatcatgtg gttttggcag tttgttaaag aaattgataa tgagaagaga 1980
atgagacttc tgcagtttgt tactggaacc tgccgattgc cagtaggagg atttgctgat 2040
ctcatgggga gcaatggacc acagaaattc tgcattgaaa aagttgggaa agaaaattgg 2100
ctaccagaa gtcatacctg ttttaatcgc ctggacctgc caccatacaa gagctatgag 2160
caactgaagg aaaagctggt gtttgccata gaagaaacag aaggatttgg acaagagtaa 2220
cttctgagaa cttgcacat gaatgggcaa gaacttattt gcaatgtttg tcttctctg 2280
cctgttgcac atcttgtaaa attggacaat ggctctttag agagttatct gagtgttaagt 2340
aaattaatgt tctcatttaa aaaaaaaaaa aaaaaaa 2377

```

<210> 645

<211> 1295

<212> DNA

<213> Homo sapiens

<400> 645

```

gggcccgcgc ccaccgggac cttgcctcta cctcagtcgt tgcccccca ttttcggctg 60
gagccacacg ccccgccctc cagcccccgc tctagcttcg ccagtagctc ggccagcgac 120
gcgagcaagc cgtccagccc ccggggcagc ctgctgctgg acggggcggg ggtggcgga 180
gctggaggta gccggccctg cagcaatcgc accagcggca tcagcatggg ctacgaccag 240
cgccacggga gccccctgac agcggggcgc tgcctgtttg gccacccct ggccggagca 300

```

10

20

30

```

ccggcaggct attctcccgg aggggtcccg tccgcctacc cggagctcca cgcgcacctg 360
gaccgatgtg acgctcagcg gcccgcgggg ttccgctgcc aggaagcccg ccactcgat 420
ccccggccc tggggcagccc tggagctcta gccggggccc gagggggagc ggcgggggccc 480
ttggagagac ggggggcgca acccgacga cactctgtga ccggctacgg ggaactgcgc 540
gtgggcgccc ggtaccagga cgagctaaca gctttgtctc gcctgacggg gggcacccgt 600
gggcgagaag ccggagcccc cggagaaccc tccgggattg agccgtcggg tctggaggag 660
ccaccaggtc ctttctgttc ggaggccgcc cggggcccga tgcgggagcc agaggccagg 720
gaggactact tccgcacctg tatcaagtgc aacaaaggca tctatgggca gagcaatgcc 780
tgccaggccc tggacagcct ctaccacacc cagtgccttg ttgctgtctc ttgtggcgca 840
actttgcgtt gcaagccttt ctacagtgtc aatggctctg tgtactgtga ggaagattat 900
ctgttttcag ggtttcagga ggcagctgag aaatgctgtg tctgtgggca cttgattttg 960
gagaagatcc tacaagcaat ggggaagtcc tatcatccag gctgtttccg atgcattgtt 1020
tgcaacaagt gcctggatgg catccccttc acagtggact tctccaacca agtatactgt 1080
gtcacccgact accacaaaaa ttatgtcctc aagtgtgcag cctgtggcca acccatcctc 1140
ccctctgagg gctgtgagga catcgtgagg gtgatatoca tggaccggga ttatcacttt 1200
gagtgtctacc actgtgagga ctgccggatg cagctgagtg atgagggaagg ctgctgtctg 1260
ttccctctgg atgggcactt gctctgccat ggttg 1295

```

<210> 646

<211> 3411

<212> DNA

<213> Homo sapiens

<400> 646

```

gggcccgggg tcccgccacc accgcgcgcg ggacagattg attcactttg gagctgtaag 60
tactgatgta ttagggtgca gcgctcattg ttcatlgacg cagagtccca aaatgaatat 120
ccaagagcag ggtttccccc tggacctcgg agcaagtttc accgaagatg ctecccgacc 180
cccagtgcct ggtgaggagg gagaactggt gtccacagac ccgaggcccg ccagctacag 240
tttctgtctc gggaaagggt ttggcattaa aggtgagact tcgacggcca ctccgaggcg 300
ctcgatctg gacctggggg atgagcctga gggcagtgc tccccacccc caccatactt 360
gaagtgggct gagtactgc attccctgct ggatgaccaa gatgggataa gctgttcag 420
gactttcctg aagcaggagg gctgtgccga cttgtggac ttctggtttg cctgactgg 480
cttcaggaag ctggagccct gtgactcgaa cgaggagaag aggctgaagc tggcgagagc 540
catctaccga aagtacattc ttgataacaa tggcatcgtg tcccggcaga ccaagccagc 600
caccaagagc ttcataaagg gctgcatcat gaagcagctg atcgatcctg ccatgtttga 660
ccaggcccag accgaaatcc aggcactat ggaggaaaac acctatccct ccttccttaa 720
gtctgatatt tatttggaat atacgaggac aggtcggag agcccaaaag tctgtagtga 780
ccagagctct gggtcaggga cagggaaggg catatctgga tacctgccga ccttaaatga 840
agatgaggaa tggaaagtgt accaggacat ggacgaggac gatggcagag acgtgtctcc 900
ccccggaaga ctccctcaga agctgtcct ggagacagct gcccgcaggg tctcctccag 960
tagacggtag agcgaaggca gagagttcag gtatggatcc tggcgggagc cagtcaaccc 1020
ctattatgtc aatgcccggc atgcccctgg ccagccacc agtgccaacg acagcgagca 1080
gcagagcctg tccagcgatg cagacacct gtccctcacg gacagcagcg tggatgggat 1140
ccccccatac aggatccgta agcagcaccg caggagagatg caggagagcg cgcagggtcaa 1200
tggcggggtg cccctacctc acattccccg cactacccg gtgcccgaagg aggtcccgct 1260
ggagcctcag aagttcggcg aggagctcat ccaccgcctg gaggctgtgc agcgacgcg 1320
ggaggccgag gagaagctgg aggagcggct gaagcgcgtg cgcattggag aggaaggtga 1380
ggacggcgat ccactgtcag ggccccagg gccgtgtcac aagctgcctc ccgccccgc 1440
ttggcaccac ttcccgcctc gcttgtgttg gacatgggct tgtgccgggc tccgggatgc 1500
acacgaggag aacctgaga gcatcctgga cgagcacgta cagcgtgtgc tgaggacaac 1560

```

10

20

30

```

tggccgccag tgcctggggc ctggccatcg ctccccggac agtgggcacg tggccaagat 1620
gccagtgga cttggggggtg ccgcctcggg gcacgggaag cactatccca agtcaggggc 1680
gaagctggag ggggcccggc tgcaccacca ccgacacgtc caccaccacg tccaccacag 1740
cacagcccgg cccaaggagc aggtggaggg cgaggccacc cgcaggggcc agagcagctt 1800
cgctgggggc ctggaaccac acagccatgg ggcaagggtc cgaggctact cagagagtgt 1860
tggcgctgce cccaacgcca gtgatggcct cgcgccacgt gggaagggtg gcgttgctgt 1920
caaaagaaat gccaaagaag ctgagtcggg gaagagcgcc agcaccgagg tgcaggtg 1980
ctcggaggat ggggagaaga accagaaaat catgcagtgg atcattgagg gggaaaagga 2040
gatcagcagg caccgcaggc ccggccacgg gtcttcgggg acgaggaagc cacagcccca 2100
tgagaactcc agaccyttgt ccttgagca cccctgggcc ggccctcagc tccggacctc 2160
cgtgcagccc tcccacctct tcatccaaga cccaccatg ccacccacc cagctcccaa 2220
ccccctaacc cagctggagg aggcgcgcgc acgtctggag gaggaagaaa agagagccag 2280
ccgagcacc cccaagcaga ggtatgtgca ggaggttatg cggcggggac gcgcctcgt 2340
caggccagcg tgcgcgcgcg tgcctgcagt ggtaccagcc gtgtcggaca tggagctctc 2400
cgagcagag acaagatcgc agaggaaggt gggcgggggg agtgccagc cgtgtgacag 2460
catcgttgtg cgttactact tctgcgggga acccatcccc taccgcaccc tggtaggggg 2520
ccgcgtgtgc accctgggcc agttcaagga gctgctgacc aaaaagggca gctacagata 2580
ctacttcaag aaagtgcgcg acgagtttga ctgtggggtg gtgtttgagg aggttcgaga 2640
ggacgaggcc gtccctgcgcg tctttgagga gaagatcctc ggcaaagtgg agaaggtgga 2700
ctgataggct ggtgggcttg ccgctgtgce aggcgaggcc cttggcgggc accgggtgtc 2760
cggccaggca gatgaacctg tactcaggag ccgatgggg aacagtgttg ggtgtaccac 2820
ccatccctgt ggtctacccg tgtctagagg caggtagggg gtccctccaa gtggtccaca 2880
agcttctgtc ctgcccccaa ggaggcagcc tggaccactc ctcatagcaa tacttgagg 2940
gcccagccca agtgaggcag ccgagggtcc tgcctgcagc ttcagggtgac cccccccat 3000
ccccggcac ctccttggg cactgtgtgt gggatctact ttccctctgg gatttgccca 3060
cgtaccagc tctggctggg gccaggccc ggatgcagag gcctgcaggg cctctgtcaa 3120
ttgtacgcgc cccaagtgcc cttcaacaca gcttgtctct tgcctgccac tgtgtgaatc 3180
ggcgacggag cactgcacct gcctccagcc gccggctgtg cagtctggg tcctccttcc 3240
tgaggggccc tgtaaatatg tacatttctc aggttagggc cagcaggggc tgcccaggtc 3300
tgtttttcat gcgatgcac ttgtacaatt atcttttcaa aggtacttgg ataataatga 3360
aataaaaactg tttttgaacc tgaataaaaa aaaaaaaaaa aaaaaaaaaa a 3411

```

<210> 647

<211> 3140

<212> DNA

<213> Homo sapiens

<400> 647

```

ggctgcgagt aacctcatgg tcccggtggc tgtgacggcg gcagtggcgc ctgtcctgtc 60
cataaacagc gattttctcag atttgcggga aattaaaaag caactgctgc ttattgcggg 120
ccttaccggg gagcggggcc tactacacag tagcaaatgg tcggcggagt tggctttctc 180
tctccctgca ttgcctctgg ccgagctgca accgcctccg cctattacag aggaagatgc 240
ccaggatatg gatgcctata cctggcccaa ggccactttt gacgttaaa agtatgatcg 300
ggcagcacat ttcctgcatg gctgcaatag caagaaagcc tattttctgt atatgtatc 360
cagatatctg tctggagaaa aaaagaagga cgatgaaaca gttgatagct taggccccct 420
ggaaaaagga caagtgaata atgaggcgct tagagaattg agagtggagc tcagcaaaaa 480
acaccaagct cgagaacttg atggatttgg actttatctg tatggtgtgg tgcttcgaaa 540
actggacttg gttaaagagg ccattgatgt gtttgggaa gctactcatg ttttgccctt 600
gcattgggga gcctgggttag aactctgtaa cctgatcaca gacaaagaga tgctgaagtt 660
cctgtctttg ccagacacct ggatgaaaga gttttttctg gctcatatat acacagagtt 720

```

10

20

30

```

gcagttgata gaggaggccc tgcaaaagta tcagaatctc attgatgtgg gcttctctaa 780
gagctcgtat attgtttccc aaattgcagt tgcctatcac aatatcagag atattgacaa 840
agccctctcc atttttaaatg agctaaggaa acaagaccct tacaggattg aaaatatgga 900
cacattctcc aaccttcttt atgtcaggag catgaaatcg gagttgagtt atctggctca 960
taacctctgt gagattgata aataccgtgt agaaaacgtgc tgtgtaattg gcaattatta 1020
cagtttaoct tctcagcatg agaaagcagc cttatatctc cagagagccc tgaaattaaa 1080
tcctcgggat cttgggtgct ggacactaat gggacatgag tacatggaga tgaagaacac 1140
gtctgctgct atccaggctt atagacatgc cattgaggtc aacaaaacggg actacagagc 1200
ttgggtatggc ctgggcagag cctatgaaat ccttaagatg ccattttact gcctttatta 1260
ttatagacgg gccaccagc ttcgacccaa tgattctcgc atgctgggtg ctttaggaga 1320
atgttacgag aaactcaatc aactagtga agccaaaag tgttattgga gagcttacgc 1380
cgtgggagat gtggagaaaa tggctctggt gaaactggca aagcttcacg aacagttgac 1440
tgagtcagaa caggctgccc agtgttacat caaatatctc caagatatct attcctgtgg 1500
ggaaatagta gaacacttgg aggaaagcac tgcctttcgc tatctggccc agtactattt 1560
taagtgcaaa ctgtgggatg aagcttcaac ttgtgcacaa aagtgttggt catttaatga 1620
tacccggaag gaaggttaagg ccttactcgc gcaaatccta cagcttcgga accaaggcga 1680
gactcctacc accgaggtgc ctgctccctt ttctactact gcttactct ctgctaacia 1740
tacccccaca cgcagagttt ctccactcaa cttgtctctc gtcacgccat agttggctac 1800
tctcaagcca gcacattgtt agaccatct taattaagcc ttacctccat gtaaagaaca 1860
gcacgtctgt tccaaggacc tcagctcttc ttgtttctac agatggcaac agctccatag 1920
ggacagcttg tataattacc ttcagaggcc aactgacaga atcctggcag gaacagacat 1980
tatcttgcca gttagaagta cttctgtctc acttatgtcc aaagagtggc tatagatctt 2040
ggccttcttc cctgaatgct tttttttttt ggcccccagg aaagtccctt ttatagcaat 2100
ttagacacag caatgctaca ggaacaaagt ttcaatgctg ctgagagtga aagaaaggag 2160
gaaagtctgc cactctaccc tgagctggca gtagggcact gagtaccctt aggaagaagt 2220
cagagcaatg gatataaatg accttgcctc tggatttgcg gagcatgac cctattctga 2280
tgtcagagat taggtttaaa tggaaatagag ctatccattt gttcttactc tctagggaga 2340
caatcttcca aaacagtttt gggggggtct tctaaagctt tcaaatgga agtaacttta 2400
ttcaactaga gttgaataaa agaagggcaa aaataatctc acagagcttg gaactgctga 2460
tagcccttac tgagggcmaa agatggctat attgttagct atactcctac caaagcaagc 2520
aaggagatag gattatagat aatttcacgg acatttggaa ataacattgg tgattatata 2580
gacaagaata aactcacttc aagctggctc gtttttaata attttcaacg taattgtcta 2640
tttttttccc tccatctgc aacagaatac atttttttca gcctttatct agatgaggta 2700
aagggaatca ttcttatggt gctcttgag agtttcaggc ctgtgcatgt gtgtacagca 2760
ggaggttaata tgcataatg tctgctgtaa tatatttgca cagtagatgc tatggatcat 2820
tctgagctca ggggtccagac tttattctta tcccagaat tttgtgttac gtttttacct 2880
cctaacatat gacacttcat cttatattaa ggaaggttta gaatatctaa tacgacttga 2940
attcatttgt tactaagcct tctcaggcaa gctgtatact agttaactgg ctccactgcc 3000
atgccttttc aaggttccca tggtcagaa tgatgtttga ttcttaattt ttctgtccct 3060
tttataattt gtttttaatga ttttgcata tttggaattc aataaaaaat gtgaacaata 3120
ataaaaaaaa aaaaaaaaaa 3140

```

```

<210> 648
<211> 3513
<212> DNA
<213> Homo sapiens

```

```

<400> 648
ccgtgtacca ggtgctgcta gtgggaagca cgctgctgaa ggaagtgcct tccgggctgc 60
agctggagca gttgccttct cagagcctgc tgaccacat cccaacggcg gggctgcccc 120

```

10

20

30

cttcgctagg	aggaggcctg	ccttactgcc	accaggcctg	gctggatttc	cgaaggcggc	180
tggaaagctct	actacagaac	tgccaggcag	cttgtgccct	gctccagggg	gccatcgaaa	240
gtgtgaaggc	tgtgccccag	cccatggagc	ctggggagggt	cgggtcagctg	ctacagcaga	300
cagaggtcct	gatgcagcag	gtgctagact	cgccatggct	ggcatggcta	caatgccagg	360
ggggccggga	gctgacatgg	ctgaagcaag	agggtccaga	ggtgacctg	agcccagact	420
acaggacggc	aatggacaag	gctgaacgagc	tatatgaccg	ggtggatgga	ttgctgcacc	480
aactgaccct	gcagagcaac	cagcgaatac	aggccctaga	gttggtccaa	acactggagg	540
cccgggaaaag	cggactgcac	cagattgaag	tgtggctgca	gcagggtgggc	tggccagcac	600
tggaggaggc	tggggagccc	tgcctggaca	tgctgctcca	ggcccaaggc	tcttttcagg	660
agctgtacca	ggttgcccag	gagcagggtca	ggcaaggggga	gaagtttctg	cagccgctga	720
ctggctggga	ggcgctgaa	ctggaccccc	ctggggcacg	ctttctggcc	ctgcgagccc	780
agctgactga	attctctagg	gctttggccc	agcggtgcca	gcggctggcg	gatgctgaga	840
ggctgtttca	gctcttcagg	gaggccttga	cgtgggctga	ggaggggcag	cgagtgttgg	900
cagagctgga	gcaggaacgc	ccgggggttg	tgttcagca	gctgcagctg	cactggacca	960
ggcacccctga	cttgctcct	gcccacttcc	gaaagatgtg	ggctctggcc	acggggctgg	1020
gctcagaggc	catccgcag	gagtgccgct	gggcctgggc	gcggtgccag	gacacctggc	1080
tggccctgga	ccaaaagctt	gaggcttcac	tgaagctacc	accggtgggc	agcacagcta	1140
gcctgtgtgt	cagccagggtc	cccgtgcac	ctgccacccc	tcccctgagg	aaggccctaca	1200
gcttcgatcg	gaatctgggg	cagagtctca	gtgaacctgc	ctgccactgc	cacctgctgg	1260
ccactattgc	tgcctgcgcg	agaccagagg	ctggaggagg	tgcctgccc	caggcatccc	1320
ctactgtgca	tccaccaggc	agctctgacc	ccaggagcct	caacaggcta	cagctggctg	1380
tggcagagat	ggtggccacg	gagcgggagt	atgtccgggc	tctagagtac	actatggaga	1440
actatttccc	cgagctggat	cgccccgatg	tgccccaggg	cctccgcggt	cagcgtgccc	1500
acctcttttg	caacctggag	aagctgcggg	acttccactg	ccacttcttc	ctgctgtgag	1560
tggaggcttg	caccgggcac	ccaccaagag	tggcctatgc	cttccctgcg	catagggctg	1620
agtttgggat	gtacgcgctc	tacagcaaga	ataagcctcg	ctccgatgcc	ctgatgtcaa	1680
gctatgggca	caccttcttc	aaggacaagc	agcaagcact	gggggaccac	ctggacctgg	1740
cctcctacct	gctaaagccc	atccagcgca	tgggcaagta	cgcactgctg	ctgcaggagc	1800
tggcacgggc	ctgcgggggc	cccacgcagg	agctcagtg	gctgcgggag	gcccagagcc	1860
ttgtgcactt	ccagctgcgg	cacggaaaag	acctgctggc	catggacgcc	atccaggggc	1920
gtgatgttaa	cctcaaggaa	caggggcagc	tgggtgcgaca	ggatgagttt	gtgggtgcgca	1980
ctgggcgcca	caagtccgtg	cgccgcactc	tcccttttga	ggagctgctg	ctcttcagca	2040
agcctcgcca	tgggcccaca	gggggttgaca	catttgccca	caagcgctcc	ttcaagatgg	2100
cagaccttgg	tctcactgag	tgctgtggga	acagcaacct	gcgcttcgag	atctgggttc	2160
gcgcgcgcaa	ggccagggac	acctttgtgc	tgcaggcctc	cagcctggct	atcaagcagg	2220
cctggacagc	tgacatctcc	cacctgcttt	ggaggcaggc	cgtccacaac	aaggagggtg	2280
gcctggctga	gatgggtgtc	atgggtgttg	ggaacaaggc	cttccgagac	attgctccca	2340
gcgaggaagc	catcaacgac	cgcaccgtca	actatgtcct	gaagtgcga	gaagttcgct	2400
ctcgggcgtc	cattgcgcta	gccccgtttg	acctatgacg	cctctacctg	ggggcctcga	2460
actcccttcc	tggagacctc	gcctcttget	ctgttctggg	gtccctcaac	ctgcacctgt	2520
acagagaccc	agctcttctg	ggtctccgct	gtccctgtga	tcccagcttc	ctagagggaag	2580
cagcactgga	ggetgaggca	gagctgggcg	gccagccctc	tttgactgct	gaggactcag	2640
agatctcgtc	ccaatgcccc	tcagccagtg	gctccagtg	ctctgacagc	agctgtgtgt	2700
cagggcaggc	cctgggtagg	ggcctggagg	acttaccctg	tgtctgagcc	cgggactgga	2760
cgagcagtag	atccagcagc	ctgcagctcc	aaggaaacatt	gcctctctgg	atctgctgtg	2820
accagggtgt	ggctgacacc	tgggctacct	ccaacctaca	tgtgcaacgc	tgttgactac	2880
cctttctgat	gtgtgtggcc	attggactaa	ctggcacggg	gcctctctag	ggaagtctgg	2940
ttgtagagcc	tgaataggct	cctggcccca	tgacccttcc	tccctgtccc	agctccctac	3000
ccagtgttgg	gttaagaata	ggctagagca	gacattgggt	gtttccatgc	tgtaggctgg	3060
tgggggacca	tgtgcctcta	ggcagtgact	aggggtgccc	cacccctcag	gaagaacaca	3120

10

20

30

```

gggtgggctcc tagcagctga tccccaatgc ctggccttaa agccgagctc agttaccata 3180
gggacaggtc caccctctact gggccctcat gcttgccctt cctggccccc aggccccagcc 3240
cctttttact ggggcagttt cgttatcttg acttgatgcc ttttgaataa ctttcaatag 3300
aattgtctaa aattatctta ctggttgtta ggcctttggt gtctcagaga aggagtctag 3360
gtcctttgatg tgtgatttaa tcttttattt gtttataata aaaaatagac tgatttgtaa 3420
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3480
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa 3513

```

<210> 649

<211> 3597

<212> DNA

<213> Homo sapiens

<400> 649

```

catgccagtt acttcctcag gaaaatattt tcttgccctt ttctttcagt atgggttttaa 60
atgtgggaac agtggataac ccaagtgtcc cacaggccaa ggtatatcc aatggcagca 120
tgatccctgc acccaaagcc agccctctaa gcctacccct tgtgcacccg cagcctggta 180
agtgaagcttg gctgcttggt aggagctaca agtgaagag aagttatttt aaataaatcc 240
caaagtttga ggcagactgt ccaggactgt tcccaggaag aagcaggagt taccacacagg 300
aaaagtctct gacctgggtcc cctcaggccc agctacctgc gccaccagc agtgaaggtt 360
gagtgtactgg cccagcatct ccacctcccc catgcaacca ggtccctggt accgtgtctc 420
ccgttgcatg tctggcttct gctgtgtct ctcctgccac gagcatctc cctgtccctc 480
ctcattccac cgtgtctctc ctgcacacat agcctctgtc ccaggggcat ttatccactt 540
gagtacagga gctgtctcaga cctctcagcc cagccctctg tgaactgccc agccccatcc 600
taccacaccc aaagtgcctt tctgtgctgt aggagctccc tctgtctagc aaggccctat 660
gggtccccc cagaggatcc acaagcaatg acttcccaaa tgacctccac tgcaagaaga 720
atcctttacca ctgtttccag agccgtgaac gatgctgtga tggcccagggt ctcagacca 780
ccctctgtga cctaaaaaga aaagctcaat ttccatctgt cttctttccc aggaccaagg 840
ggacacagta atgtgaagtc aaatacttaa ccgagcaaa ggcagtggt gttatcagtc 900
aaggacaaac ctcccacctc acagacagcc aagcagtgag ggaaagacag acagacatag 960
gtaggaaggt gctctgcagg cacaaggccc agagaagccc ctctccggga acttcccctg 1020
ctccttccag gaacagtgag ccagtgagc agtcccagcc agctcttcaa ggccttcaag 1080
gggtctttcc atgactgagt cacctccagg agctcacctg acccccagag aagacctacc 1140
ccaggcagct ccgtgccctg gcttctcccc atgccccaaa tcccccccag ccatccctcc 1200
tggtcctcgt ctacatcaag ggcctcttcc cctcttctc ccagctctca ggacaggtga 1260
ctgggagacc ttgaaccctc agcctcttcc tttaaaaaaa acaaaacaaa acaaaactgt 1320
gggccattta tttgggattt tggagttgtt tgggttttgt ttttatctt taatagtctg 1380
aaagtaagaa gggagccctg ctatggatgt taagtccaaa ttactcgggt agtgggagca 1440
aaacctatga cttccaaggg gatgaggaga gggtcagagg acaggaggag cctcccccat 1500
tgaaaaaaa aaaatgggtc aggacattcc ctggatgagg acaatgctag ggggtggcatc 1560
tcacatggct gctgtctatt ctggtgcttc occacacttt tgacagatgg agtccctctc 1620
ctaccgcctc ctgccacctc accctacagg cattctctat gtaggaacaa agagccttat 1680
cttatagagt ggggagctga gacacagcct caggtaacac tgacacagct cccgaatgag 1740
gctgggacac tctgcaaacc tctcctcatg gtgctaaggg tggcatgctc ttgacaggaa 1800
acctaaatga ccactcctct catttggaag gtaatccact gcagtaaaag tttcagacat 1860
gcaagagaga gttttttttt ttttttacta caaatttttg ctccccata aaattatttt 1920
tttattagag ggagtatcca agttttaaaa gtatatagaa ttttttggtt gtaagagaaa 1980
tacatactca ttaggatccc gattaaattc cttgagtaga ctgggtgcta ccagaaagca 2040
aagcaaagtt aaacaaaacg aaacaaaate cttcatatac aaaaagaact ttctgtttgt 2100
attggcagag gtatgtagggt gattcaggta ggctgaaaat cctgggtgtc gggagcctca 2160

```

10

20

30

```

ctttattcca ttcccacccg ctttgatgtc tatgcttggc tctctgggct gcccttggtg 2220
ctgccgaatc ctacacatct cttatcagct ttcctcaaac tttaaggagg ctctgtgagg 2280
gatgggtcat ggggaagaccc aagccttccc tccgccagga ttgcaaaagc aagtagactt 2340
gggtctatgca gctcttcttc caacaatttc tttatttga attagaactt cttttgttag 2400
tatctttgat cttttgactc aagcacattt tgggaagggt cccttacaaa agtagaattt 2460
aaaacagagg atacagttaa agagcaaccc aaaggacgct taagaaaccc agaccacttc 2520
accgaacagg actaagggaac actttcgtgc acagaagtca gccgcaatcc aggcacagga 2580
cgaagatggg atacacgtgc tcatctgtct gtctctcttc cctctccctc ccgacgttc 2640
tagttagctt gttgacttgt taaaccttct gttcttaaaa tgaaaagcta gcttacctca 2700
aagaatcttg tttccattcg gaaaccaacg attttgtgtt ttagaatgga cagccctccc 2760
ctcaccactc cctaccttgg cctggtgtcc ttgagacata cggctcttgc ttagtcgtgt 2820
gttggtgctt ttgagcagga acgaggcctc caggccctga ggtgggaagg aaggattgga 2880
tgccactgcc ctctccccc ctttagcatg taggggccag ccctctctct ccagcagggt 2940
cctgctgagt taccatagca accagcaact ccagggtacc acaacagaca atggctcagc 3000
gagccgacgt gtggggatga tgcaggggtt ttggcccgag cagaggaccc agagttgagc 3060
ttcaaatgct agagaagggg agaaacagga tgggaagggt gtttaaggaa ctggcagggg 3120
tctttgagtc acatagagaa gccgttgaag gaggtagggt aggttatctc tgttccagtc 3180
accccttccc agcccatcc cacttctgtt tcaaaactaa gctcccacct cgaacattga 3240
ccctttgtta gaacaaagca aagcatatct ttaaacaaac gtgttaaaat gagcctcaa 3300
tgtatgtgga tgagatctct aagaagaggg tcttctggtt ttgattttta aagaagagta 3360
tcctagttaa atattaaaa aaaattaaaa agtttttaaa aaggaaacct atgctattta 3420
aattggagcc cagttgtaac ttggtaaagg caagcttctg tacctttgtt ataattaatt 3480
gtataoctgt gtatgtaaat ataaggcatt cctattttgc agttcagaac aaaaaaaact 3540
tatttgaat atagaataaa gtttattaaa aaataagaaa aaaaaaaaaa aaaaaaa 3597

```

```

<210> 650
<211> 855
<212> DNA
<213> Homo sapiens

```

10

20

```

<400> 650
cgctcaatta tctactcgag tctagactcg aggcggccgc ccattgtgca ctaaagcagg 60
ggatagcaac ggcgtccctc ctcccgcctc agctgcagcc cgcagtcctc acagtggtaa 120
catgccacgt ggtagtctct gtccatggac accacacgga tggttgtctc gcagccctgt 180
gcaggagga taggacgggc acaggaggcg ctttttgggt caaaaaccgt gtgatagtct 240
cgcacgcagt agatgttgtt ctccacgtcc acggtgaagg gaaccccgct caggcactca 300
ttgcacacgg agcaccggaa gcagcctggg ttgtaggact tgcccagggc ctgcaggatc 360
atttccatga tgagatgtcc acacacgtcg ctttgtcgcc ccgtctgctg gaacccggag 420
tacaggaagt cctcctggca gtacactttc tcacccacgt tgtagaacgc cttcccacgg 480
agtctgtctc cacacgagtc gcaggtgaag cagtcagtgt gataaagact cccattgcc 540
tgccacgcct gctgggctcc gtagatgcca agcccacact tgatgcaaat gccgaagtag 600
tcccgcgcgc tgcgcgcctc gagcgcgcgc tccagctccc gggtagagcg ctcagccgcg 660
cgctcggccg cgttggggcc gccctcccgg ccagggggca gcgggagtg aggcagcggg 720
aagggggccg gccccgcagg ctccggggag cgagcggggc cggggcaggc gccgggcggg 780
aggaagtcag cgtaggcttt agtgcacaat gggaattcgg atccggtcga cactagtctc 840
aggatctctt tttttt
855

```

```

<210> 651
<211> 3791
<212> DNA

```

30

<213> Homo sapiens

<400> 651

```

acagacggcg ggtgaacatg gcgtcctcga cttgggtctga gacgtgatag gcctgccttc      60
tggttgaaga tgtggcgagt gaaaaaactg agcctcagcc tgtegccttc gcccagacg      120
ggaaaaacat ctatgagaac tcctctccgt gaacttaacc tgcagcccg gcccctcacc      180
acctctggaa aaagatcccc cgcttctctc tcgctgaacc catcactgtg caagctgggg      240
ctgcaggaag gcagcaacaa ctctctcca gtggattttg taaataacaa gaggacagac      300
ttatcttcag aacatttcag tcattcctca aagtggctag aaacttgtca gcatgaatca      360
gatgagcagc ctctagatcc aattcccaa attagctota ctctaaaaac gtctgaggaa      420
gcagtagacc cactgggcaa ttatatggtt aaaaccatcg tccttgtacc atctccactg      480
gggcagcaac aagacatgat atttgaggcc cgttttagata ccatggcaga gacaaacagc      540
atatctttaa atggaccttt gagaacagac gatctggtga gagaggaggt ggcacctgc      600
atgggagaca ggttttcaga agttgtctgt gtatctgaga aacctatctt tcaggaatct      660
ccgtcccatc tcttagagga gtctccacca aatccctggt ctgaacaact acattgtctc      720
aaggaaagcc tgagcagtag aactgaggct gtgcgtgagg acttagtacc ttctgaaagt      780
aacgccttct tgccttctct tgttctctgt ctttccctt caactgcctt ggcagcagat      840
ttcctgttca atcatgtgga ccagaggag gaaattgtag agcatggagc tatggaggaa      900
agagaaatga ggtttccac acatcctaag gagtctgaaa cagaagatca agcacttgtc      960
tcaagtgtgg aagatattct gtccacatgc ctgacaccaa atctagtaga aatggaatcc     1020
caagaagctc caggccagc agtagaagat gttggttaga ttcttggctc tgatacacag      1080
tcttggatgt cccactggc ctggctggaa aaagggtgtaa atacctccgt catgctggaa     1140
aatctccgcc aaagcttctc ccttccctcg atgcttcggg atgctgcaat tggcactacc     1200
cctttctcta cttgtcgggt ggggacttgg tttactcctt cagcaccaca ggaaaagagt     1260
acaaacacat ccagacagc cctggttggc accaagcaca gtactctctga gacagagcag     1320
ctcctgtgtg gcgggcctcc agatctgact gccttgtctc gacatgactt ggaagataac     1380
ctgctgagct ctctgtcat tgtggagttt ctctccgcc agcttcggga ctggaagagc     1440
cagctggctg tcctcaccac agaaacccag gacagtagca cacagactga cacatctcac     1500
agtgggataa ctaataaact tcagcatctt aaggagagcc atgagatggg acaggcccta     1560
cagcaggcca gaaatgtcat gcaatcatgg gtgcttatct ctaaagagct gatatccttg     1620
cttcacctat cctgttgca tttagaagaa gataagacta ctgtgaatca ggagtctcgg     1680
cgtgcagaaa cattggtctg ttgctgtttt gatattgtga agaaattgag ggcaaagctc     1740
cagagcctca aagcagaaag ggaggaggca aggcacagag aggaaatggc tctcagaggc     1800
aaggatgctg cagagatagt gttggaggct ttctgtgcac acgccagcca gcgcacagc     1860
cagctggaac aggacctagc atccatgctg gaattcagag gccttctgaa ggatgccagc     1920
acccaactgg tagggcttca tgccaagcaa gaagagctgg ttcagcagac agtgagtctt     1980
acttctacct tgcaacaaga ctggaggctc atgcaactgg attatacaac atggacagct     2040
ttgctgagtc ggtcccgaca actcacagag aaactcacag tcaagagcca gcaagccctg     2100
caggaaactg atgtggcaat tgaggaaaag caggagggtt ctagggtgct ggaacaagtc     2160
tctgccagct tagaggagtg caaaggccaa acagaacaac tggagtggga aaacattcgt     2220
ctagcaacag atctccgggc tcagttgcag attctggcca acatggacag ccagctaaaa     2280
gagctacaga gtcagcatac ccattgtgcc caggacctgg ctatgaagga tgagttactc     2340
tgccagctta ccagagcaa tgaggagcag gctgctcaat gcgtaaagga agagatggca     2400
ctaaaaacaca tgcaggcaga actgcagcag caacaagctg tccctggcaa agagggtcgg     2460
gacctgaaag agaccttggg gtttgcagac caggagaatc aggttgtctc cctggagctg     2520
ggtcagggtt agtgtcaatt gaaaaccaca ctggaagtgc tccgggagcg cagcttgagc     2580
tgttagaacc tcaaggacac tgtagagaac ctaacggcta aactggccag caccatagca     2640
gataaaccag agcaagatct ggagaaaaca cggcagta ctcaaaagct agggctgctg     2700
actgagcaac tacagagcct gactctctt ctacagacaa aactaaagga gaagactgaa     2760
caagagaccc ttctgctgag tacagctgt cctccacccc aggaacaccc tctgcctaatt     2820

```

10

20

30


```

gacaggacct tcttgggaag catcttgaca gcagtggcag atgaagagcc agaatacaact 2880
cctgtgccct tgccttgggaag tgacaagagt gctttcaccc gagtagcatc aatgggtttcc 2940
cttcagcccc cagagacccc aggcattggag gagagcctgg cagaaatgag tattatgact 3000
actgagcttc agagtctttt ttccctgcta caagagtcta aagaagaagc catcaggact 3060
ctgcagcgaa aaatttgtga gctgcaagct aggcctgcagg ccaggaaga acagcatcag 3120
gaagtccaga aggcataaaga agcagacata gagaagctga accaggcctt gtgcttgccg 3180
tacaagaatg aaaaggagct ccagggaagt atacagcaga atgagaagat cctagaacag 3240
atagacaaga gtggcgagct cataagcctt agagaggagg tgaccacct taccgctca 3300
cttcggcgtg cggagacaga gaccaaagt ctccaggagg cctggcaggc cagctggact 3360
ccaactgcc aacctatggc accaattgga tccaggagaa agtgtggctc tctcaggagg 3420
tggacaaact gagagtgatg ttccctggaga tgaaaaatga gaaggaaaac tctgatcaa 3480
gttccagagc ccatagaaat atcctagagg agaacccttc gcgctctgac aaggagttag 3540
aaaaactaga tgacattgtt cagcatattt ataagacct gctctctatt ccagggtgg 3600
tgaggggatg caaagaacta cagggattgc tggaaattct gagctaagaa actgaaagcc 3660
agaatttgtt tcacctcttt ttacctgcaa tccccctta cccaatacc aagaccaact 3720
ggcatagagc caactgagat aaatgctatt taaataaagt gtatttaatg aaaaaaaaaa 3780
aaaaaaaaa a 3791

```

<210> 652

<211> 1683

<212> DNA

<213> Homo sapiens

<400> 652

```

ctctgagtgt ccagtggtea gttgccccag gatggggacc acagccagag cagccttgg 60
cttgacctat ttggctgttg cttctgctgc ctctgagga ggcttcacgg ctacaggaca 120
gaggcagctg aggccagagc actttcaaga agttggctac gcagctccc cctccccacc 180
cctatcccga agcctcccca tggatcacc tgactcctct cagcatggcc ctcccttga 240
gggacagagt caagtgcagc cccctccctc tcaggaggcc acccctctcc aacaggaaaa 300
gctgtacct gcccaactcc ctgctgaaaa ggaagtgggt cccctctcc ctccaggagc 360
tgtccccctc caaaaagagc tgccctctct ccagcaccac aatgaacaga aggaaggaa 420
gccagctcca ttggtggacc agagccatcc agaaccctgag tcttggaatg cagcccagca 480
ctgccaacag gaccggtccc aagggggctg gggccaccgg ctggatggct tccccctgg 540
gcgcccttct ccagacaatc tgaaccaa atgccttctt aaccgtcagc atgtggtata 600
tggctccctg aacctaccac agtccagcta ctccacctc actcgccagg gtgagacct 660
caatttctct gagattggat attcccgctg ctgccactgc cgcagccaca caaacgcct 720
agagtgtgct aaacttgtgt gggaggaagc aatgagccga ttctgtgagg ccgagttctc 780
ggtcaagacc cgaccccaact ggtgtgcac gcggcagggg gaggctcgg tctctgctt 840
ccaggaggaa gctccccagc cactactcca gctccgggct tgccccagcc atcagcctga 900
tatttctctg ggtcttgagc tgcccttccc tcttgggggt cccacattgg acaatatcaa 960
gaacatctgc cactgaggc gcttccgctc tgtgccacgc aacctgccag ctactgacc 1020
cctacaaagg gagctgctgg cactgatcca gctggagagg gagttccagc gctgctgccg 1080
ccagggggaa aatcacacct gtacatggaa ggcctgggag gatacccttg acaaatactg 1140
tgaccgggag tatgtgtga agacccacca ccacttgtgt tgccggccacc ctccagccc 1200
tactcgggat gagtgccttg cccgtcgggc tccttaccac aactatgacc gggacatctt 1260
gaccattgac atcagtcgag tcaccccaa cctcatgggc cactctgtgt gaaaccaaag 1320
agttctcacc aagcataaac atattcctgg gctgatccac aacatgactg ccgctgctg 1380
tgacctgcca ttccagaac aggcctgctg tgcagaggag gagaaattaa ccttcatcaa 1440
tgatctgtgt ggtcccgac gtaacatctg gcgagacct gccctctgct gttacctgag 1500
tcttggggat gaacaggtca actgcttcaa catcaattat ctgaggaaag tggctctagt 1560

```

10

20

30

```

gtctggagac actgagaacg ccaagggcca gggggagcag ggctcaactg gaggaacaaa 1620
tatcagctcc acctctgagc ccaaggaaga atgagtcacc ccagagccct agagggtcag 1680
atg 1683

```

<210> 653

<211> 1696

<212> DNA

<213> Homo sapiens

<400> 653

```

cacctaaaaa ccaaaatggg aaaggaaaaa actcatatca acattgtcgt cattggacac 60
gtagattcgg gcaagtccac cactactggc catctgatct ataaatgcgg tggcatcgac 120
aaaagaacca ttgaaaaatt tgagaaggag gctgctgaga tgggaaaggg ctccctcaag 180
tatgcctggg tcttgataaa actgaaagct gagcgtgaac gtggtatcac cattgatate 240
tccttgtgga aatttgagac cagcaagtac tatgtgacta tcattgatgc cccaggacac 300
agagacttta tcaaaaacat gattacaggg acatctcagg ctgactgtgc tgcctgatt 360
gttctgctg gtgttggtga atttgaagct ggtatctcca agaattgggca gacccgagag 420
catgcccttc tggcttacac actgggtgtg aaacaactaa ttgtcgggtg taacaaaatg 480
gattccactg agccacccta cagccagaag agatatgagg aaattgttaa ggaagtcagc 540
acttacatta agaaaattgg ctacaacccc gacacagtag catttgtgcc aatttctggt 600
tggaatggtg acaacatgct ggagccaagt gctaacatgc cttggttcaa gggatggaaa 660
gtcaccgcta aggatggcaa tgccagtggg accacgctgc ttgaggctgt ggactgcac 720
ctaccaccaa ctgctccaac tgacaagccc ttgcgcctgc ctctccagga tgtctacaaa 780
attggttgta ttggtactgt tcctgttggc cgagtggaga ctggtgttct caaacccggt 840
atggttggtc ccttctgctc agtcaacggt acaacggaag taaaatctgt cgaaatgcac 900
catgaagctt tgagtgaagc tcttcctggg gacaatgtgg gcttcaatgt caagaatgtg 960
tctgtcaagg atgttcgtcg tggcaacggt gctggtgaca gcaaaaatga cccaccaatg 1020
gaagcagctg gcttcactgc tcaggtgatt atcctgaacc atccaggcca aataagcgcc 1080
ggctatgcc ctgtattgga ttgccacacg gctcacattg catgcaagtt tgctgagctg 1140
aaggaaaaga ttgatcgccg ttctggtaaa aagctggaag atggccctaa attcttgaag 1200
tctggtgatg ctgccattgt tgatatggtt cctggcaagc ccatgtgtgt tgagagcttc 1260
tcagactatc cacctttggg tcgctttgct gttcgtgata tgagacagac agttgcgggtg 1320
ggtgtcatca aagcagtggg caagaaggct gctggagctg gcaaggtcac caagtctgcc 1380
cagaaagctc agaaggctaa atgaatatta tccttaatac ctgccacccc actcttaatc 1440
agtgggtggg gaacgggtctc agaactgttt gtttcaattg gccatttaag tttagtagta 1500
aaagactggt taatgataac aatgcacgtg aaaaccttca gaaggaaagg agaattgttt 1560
gtggaccact ttggttttct tttttgcgtg tggcagtttt aagttattag tttttaaaat 1620
cagtactttt taatggaaac aacttgacca aaaatttgtc acagaatttt gagaccatt 1680
aaaaaagtta atgag 1696

```

<210> 654

<211> 7680

<212> DNA

<213> Homo sapiens

<400> 654

```

gaagagcaag aggcaggctc agcaaatggt tcagccccag tccccggtgg ctgtcagtca 60
aagcaagccc ggttggtatg acaatggaaa acactatcag ataaatcaac agtgggagcg 120
gacctaccta ggtaatgtgt tggttgttac ttgttatgga ggaagccgag gttttaaactg 180
cgaaagtaaa cctgaagctg aagagacttg ctttgacaag tacactggga acacttaccg 240

```

10

20

30

agtgggtgac	acttatgagc	gtcctaaaga	ctccatgato	tgggactgta	cctgcatcgg	300
ggctgggcca	gggagataaa	gctgtaccat	cgcaaacccg	tgccatgaag	ggggtcagtc	360
ctacaagatt	ggtgacacct	ggaggagacc	acatgagact	ggtgggtaca	tgtagagtg	420
tgtgtgtctt	ggtaatggaa	aaggagaatg	gacctgcaag	cccatagctg	agaagtgttt	480
tgatcatgct	gctgggactt	cctatgtggt	cggagaaacg	tgggagaagc	cctaccaagg	540
ctggatgatg	gtagattgta	cttgctctgg	agaaggcagc	ggacgcatca	cttgcaactc	600
tagaaataga	tgcaacgato	aggacacaag	gacatcctat	agaattggag	acacctggag	660
caagaaggat	aatcgaggaa	acctgctcca	gtgcatctgc	acaggcaacg	gccgaggaga	720
gtggaagtgt	gagaggcaca	cctctgtgca	gaccacatcg	agcggatctg	gccccctcac	780
cgatgttcgt	gcagctgttt	accaaccgca	gcctcacccc	cagcctcctc	cctatggcca	840
ctgtgtcaca	gacagtgggt	tggtctactc	tgtggggatg	cagtgggtga	agacacaagg	900
aaataagcaa	atgctttgca	cgtgctctgg	caacggagtc	agctgccaag	agacagctgt	960
aaccagact	tacgttgcca	acttaaatgg	agagccatgt	gtcttaccat	tcacctacaa	1020
tggcaggacg	ttctactcct	gcaccacgga	agggcgacag	gacggacatc	tttgggtcag	1080
cacaacttcg	aattatgagc	aggaccagaa	atactcttct	tgacacagac	acactgtttt	1140
ggttcagact	caaggaggaa	attccaatgg	tgccttctgc	cacttccccct	tcctatacaa	1200
caaccacaat	tacactgatt	gcacttctga	gggcagaaga	gacaacatga	agtgggtgtg	1260
gaccacacag	aactatgatg	cgcaccagaa	gtttgggttc	tgccccatgg	ctgccccaga	1320
ggaaatctgc	acaaccaatg	aagggtcat	gtaccgcatt	ggagatcagt	gggataagca	1380
gcatgacatg	ggtcacatga	tgaggtgcac	gtgtgttggg	aatggctcgt	gggaatggac	1440
atgcatttgc	tactcgcaac	ttcgagatca	gtgcattgtt	gatgacatca	cttacaatgt	1500
gaacgacaca	ttccacaagc	gtcatgaaga	ggggcacatg	ctgaactgta	catgcttcgg	1560
tcagggctcg	ggcaggtgga	agtgtgatcc	cgtcgaccaa	tgccaggatt	cagagactgg	1620
gacgttttat	caaattggag	attcatggga	gaagtatgtg	catggtgtca	gataccagtg	1680
ctactgctat	ggcgttgcca	ttggggagtg	gcattgcca	cctttacaga	cctatccaag	1740
ctcaagtggg	cctgtcgaag	tatttatcac	tgagactcgg	agtcagccca	actcccaccc	1800
catccagtgg	aatgcaccac	agccatctca	catttccaag	tacattctca	ggtggagacc	1860
taaaaaattct	gtaggccgtt	ggaagggaagc	taccatacca	ggccacttaa	actcctacac	1920
catcaaaggc	ctgaagcctg	gtgtgtgata	cgaggggcag	ctcatcagca	tccagcagta	1980
cggccacca	gaagtgaact	gctttgactt	caccaccacc	agcaccagca	cacctgtgac	2040
cagcaacacc	gtgacaggag	agacgactcc	cttttctcct	cttgtggcca	cttctgaatc	2100
tgtgaccgaa	atcacagcca	gtagctttgt	ggtctcctgg	gtctcagctt	cgcacaocgt	2160
gtcgggattc	cgggtggaat	atgagctgag	tgaggaggga	gatgagccac	agtacctgga	2220
tcttccaagc	acagccactt	ctgtgaacat	ccctgacctg	cttctctggc	gaaaatacat	2280
tgtaaatgtc	tatcagatat	ctgaggatgg	ggagcagagt	ttgatcctgt	ctacttcaca	2340
aacaacagcg	cctgatgccc	ctcctgaccc	gactgtggac	caagttgatg	acacctcaat	2400
tgttgttcgc	tggagcagac	cccaggctcc	catcacaggg	tacagaatag	tctattcgcc	2460
atcagtagaa	ggtagcagca	cagaactcaa	ccttcctgaa	actgcaaaact	ccgtcacccct	2520
cagtgaactt	caacctgggt	ttcagtataa	catcactatc	tatgctgtgg	aagaaaatca	2580
agaaagtaca	cctgtttgtca	ttcaacaaga	aaccactggc	accccacgct	cagatacagt	2640
gccccctccc	agggacctgc	agtttgtgga	agtgacagac	gtgaagggtca	ccatcatgtg	2700
gacaccgcct	gagagtgcag	tgaccggcta	ccgtgtggat	gtgatccccg	tcaacctgoc	2760
tggcgagcac	gggcagaggc	tgcccatcag	caggaacacc	tttgcaagaag	tcaccgggct	2820
gtccccctgg	gtcacctatt	acttcaaagt	ctttgcagtg	agccatggga	gggagagcaa	2880
gcctctgact	gotcaacaga	caaccaaact	ggatgtctcc	actaacctcc	agtttgtcaa	2940
tgaaactgat	tctactgtcc	tggtagatg	gactccacct	cgggcccaga	taacaggata	3000
cgcactgacc	gtgggacctt	cccgaagagg	ccagcccagg	cagtacaatg	tgggtccctc	3060
tgtctccaag	taccccttga	ggaatctgca	gcctgcctct	gagtacaccg	tatccctcgt	3120
ggccataaag	ggcaaccaag	agagccccaa	agccactgga	gtctttacca	cactgcagcc	3180
tgggagctct	attccacctt	acaacaccga	ggtgactgag	accaccatcg	tgatcacatg	3240

10

20

30

gacgcctgct	ccaagaattg	gttttaagct	gggtgtacga	ccaagccagg	gaggagagggc	3300
accacgagaa	gtgacttccg	actcaggaag	catcggttg	tccggcttga	ctccaggagt	3360
agaatbacgtc	tacaccatcc	aagtcctgag	agatggacag	gaaagagatg	cgccaattgt	3420
aaacaaagt	gtgacaccat	tgtctccacc	aacaaacttg	catctggagg	caaaccctga	3480
cactggagtg	ctcacagtct	cctgggagag	gagcaccacc	ccagacatta	ctggttatag	3540
aattaccaca	acccctacaa	acggccagca	gggaaattct	ttggaagaag	tgggtccatgc	3600
tgatcagagc	tcctgcactt	ttgataacct	gagtcccggc	ctggagtaca	atgtcagtg	3660
ttacactgtc	aaggatgaca	aggaaaagt	ccctatctct	gataccatca	tcccagctgt	3720
tcctcctccc	actgacctgc	gattcaccaa	cattgggtcca	gacaccoatgc	gtgtcacctg	3780
gggtccaccc	ccatccattg	atttaaccaa	cttcctgggtg	cgttactcac	ctgtgaaaaa	3840
tgaggaagat	gttgacagag	tgtcaatttc	tccttcagac	aatgcagtgg	tcttaacaaa	3900
tctcctgcct	ggtacagaat	atgtagttag	tgtctccagt	gtctacgaac	aacatgagag	3960
cacacctctt	agaggaagac	agaaaacagg	tcttgattcc	ccaactggca	ttgacttttc	4020
tgatattact	gccaaactctt	ttactgtgca	ctggattgct	cctcgagcca	ccatcactgg	4080
ctacaggatc	cgccatcatc	cagagcactt	cagtgggaga	cctcgagaag	atcgggtgcc	4140
ccactctcgg	aattccatca	ccctcaccaa	cctcactcca	ggcacagagt	atgtggtcag	4200
catcggtgct	cttaattggca	gagaggaaag	tccttattg	attggccaac	aatcaacagt	4260
ttctgatgtt	ccgagggacc	tggaagtgtg	tgtcgcgacc	cccaccagcc	tactgatcag	4320
ctgggatgct	cctgctgtca	cagtggagata	ttacaggatc	acttacggag	aaacaggagg	4380
aaatagccct	gtccagggat	tcactgtgcc	tgggagcaag	tctacagcta	ccatcagcgg	4440
ccttaaacct	ggagttgatt	ataccatcac	tgtgtatgct	gtcactggcc	gtggagacag	4500
ccccgcaagc	agcaagccaa	tttccattaa	ttaccgaaca	gaaattgaca	aacctccca	4560
gatgcaagt	accgatgttc	aggacaacag	cattagtgtc	aagtggctgc	cttcaagtcc	4620
cctgttact	ggttacagag	taaccaccac	tcccaaaaat	ggaccaggac	caacaaaaac	4680
taaaactgca	ggtccagatc	aaacagaaat	gactattgaa	ggcttgacgc	ccacagtggg	4740
gtatgtgggt	agtgtctatg	ctcagaatcc	aagcggagag	agtcagcctc	tggttcagac	4800
tgcagtaacc	aacattgatc	gccctaaagg	actggcattc	actgatgtgg	atgtcgattc	4860
catcaaaatt	gcttgggaaa	gcccacaggg	gcaagtcttc	aggtacaggg	tgacctactc	4920
gagccctgag	gatggaaatc	atgagctatt	ccctgcacct	gatggtgaag	aagacactgc	4980
agagctgcaa	ggcctcagac	cgggttctga	gtacacagtc	agtgtggttg	ccttgcaacga	5040
tgatatggag	agccagcccc	tgattggaac	ccagtccaca	gctattcctg	caccaactga	5100
cctgaagtcc	actcagggtc	caccacaag	cctgagcgcc	cagtggacac	cacccaatgt	5160
tcagctcact	ggatatcgag	tgcgggtgac	ccccaaaggag	aagaccggac	caatgaaaga	5220
aatcaacctt	gctcctgaca	gctcatccgt	ggttgtatca	ggacttatgg	tggccaccaa	5280
atatgaagt	agtgtctatg	ctcttaagga	cactttgaca	agcagaccag	ctcagggtgt	5340
tgtcaccact	ctggagaatg	tcagccccc	aagaagggtc	cgtgtgacag	atgctactga	5400
gaccaccatc	accattagct	ggagaaccaa	gactgagacg	atcactggct	tccaagttag	5460
tgcctgtcca	gccaatggcc	agactccaat	ccagagaacc	atcaagccag	atgtcagaag	5520
ctacaccatc	acaggtttac	aaccaggcac	tgactacaag	atctacctgt	acaccttgaa	5580
tgacaatgct	cggagctccc	ctgtggtcat	cgacgcctcc	actgccattg	atgcaccatc	5640
caacctgctg	ttcttggtcca	ccacacccaa	ttccttgctg	gtatcatggc	agccgccacg	5700
tgccaggatt	accggtacac	tcataagta	tgagaagcct	gggtctcctc	ccagagaagt	5760
ggtccctcgg	ccccgccctg	gtgtcacaga	ggctactatt	actggcctgg	aaccgggaac	5820
cgaatatata	atttatgtca	ttgccctgaa	gaataatcag	aagagcgagc	ccctgattgg	5880
aaggaaaaag	acagacgagc	ttccccaa	ggtaaccctt	ccacacccca	atcttcatgg	5940
accagagatc	ttggatgttc	cttccacagt	tcaaaagacc	cctttcgtca	cccaccctgg	6000
gtatgacact	ggaaatggta	ttcagcttcc	tggcacttct	ggtcagcaac	ccagtgttgg	6060
gcaacaaaatg	atctttgagg	aacatgggtt	taggcggacc	acaccgcca	caacggccac	6120
ccccataagg	cataggccaa	gaccataccc	gccgaatgta	ggacaagaag	ctctctctca	6180
gacaaccatc	tcattggggcc	cattccagga	cacttctgag	tacatcattt	catgtcatcc	6240

10

20

30

tggtggcact	gatgaagaac	ccttacagtt	caggggttct	ggaacttcta	ccagtgccac	6300
tctgacagc	ctcaccagag	gtgccaccta	caacatcata	gtggaggcac	tgaagacca	6360
gcagaggcat	aaggttggg	aagaggttgt	taccgtgggc	aactctgtca	acgaaggctt	6420
gaaccaacct	acggatgact	cgtgctttga	ccctacaca	gtttccatt	atgccgttgg	6480
agatgagtgg	gaacgaatgt	ctgaatcagg	ctttaaactg	ttgtgccagt	gcttaggctt	6540
tggaagtgg	catttcagat	gtgattcatc	tagatgggtg	catgacaatg	gtgtgaacta	6600
caagattgga	gagaagtggg	accgtcagg	agaaaatggc	cagatgatga	gctgcacatg	6660
tcttgggaac	ggaaaaggag	aattcaagt	tgacctcat	gaggcaacgt	gttacgatga	6720
tgggaagaca	taccacgtag	gagaacagt	gcagaaggaa	tatctoggtg	ccatttgctc	6780
ctgcacatgc	tttggaggcc	agcggggctg	gcgctgtgac	aactgccgca	gacctggggg	6840
tgaaccag	ccgaaggca	ctactggcca	gtcctacaac	cagtattctc	agagatacca	6900
tcagagaaca	aactaataatg	ttaattgccc	aattgagtgc	ttcatgcctt	tagatgtaca	6960
ggctgacaga	gaagattccc	gagagtaa	catctttcca	atccagagga	acaagcatgt	7020
ctctctgcca	agatccatct	aaactggagt	gatgttagca	gaccagctt	agatttcttc	7080
ttctcttctt	aagcccttgg	ctctggagga	agttctccag	cttcagctca	actcacagct	7140
ttccaagca	tcacctggg	agtttctga	gggttttctc	ataaatgagg	gctgcacatt	7200
gctgttctg	cttcgaagta	ttcaataccg	ctcagtattt	ttaatgaagt	gattctaaga	7260
tttggtttgg	gatcaaatagg	aaagcatatg	cagccaacca	agatgcaaat	gttttgaaat	7320
gatatgacca	aaatttttaag	taggaaagtc	acccaaacac	ttctgctttc	acttaagtgt	7380
ctggcccgca	atactgtagg	aacaagcatg	atcttggtac	tgtgatattt	taaatatcca	7440
cagtactcac	tttttccaaa	tgatcctagt	aattgcctag	aaatatcttt	ctcttacctg	7500
ttatttatca	atttttccca	gtatttttat	acggaaaaaa	ttgtattgaa	aacacttagt	7560
atgcagttga	taagaggaat	ttggtataat	tatggtgggt	gattattttt	tatactgtat	7620
gtgcaaaagc	tttactactg	tggaaagaca	actgttttaa	taaaagattt	acattccaca	7680

<210> 655

<211> 1316

<212> DNA

<213> Homo sapiens

<400> 655

tcctaatac	actcactata	gggtctgagc	ggccgcccgg	gcaggtcgaa	tgcaggcgac	60
ttgcgagctg	ggagcgattt	aaaaacgttt	ggattccccc	ggcctgggtg	gggagagcga	120
gctgggtg	ccctagatcc	cccgcccccg	cacctcatga	gccgacctc	ggctccatgg	180
agcccgcaa	ttatgccacc	ttggatggag	ccaaggatat	cgaaggcttg	ctgggagcgg	240
gagggggcg	gaatctgtg	gcccactccc	ctctgaccag	ccaccagcg	gcgcctacgc	300
tgatgcctgc	tgtaactat	gcccccttgg	atctgcccag	ctcggcgagg	ccgccaagc	360
aatgccacc	atgcctggg	gtgcccagg	ggacgtcccc	agctcccg	ccttatgggt	420
actttggagg	cgggtactac	tcctgcccag	tgtcccggag	ctcgtgaaa	ccctgtgccc	480
aggcagccac	cctggccg	taccccgcg	agactccac	ggccggggaa	gagtacccca	540
gtcgcgccac	tgagtttgcc	ttctatccgg	gatatccggg	aacctaccac	gctatggcca	600
gttaacctga	cgtgtctgtg	gtgcagactc	tgggtgctcc	tggagaaccg	cgacatgact	660
ccctgttgcc	tgtggacagt	taccagtctt	gggtctctgc	tgggtggctg	aacagccaga	720
tgtgttgcca	gggagaacag	aaccaccag	gtcccttttg	gaaggcagca	tttgcagact	780
ccagcgggca	gcacctcct	gacgcctg	cctttcgtg	cggccgcaag	aaacgcattc	840
cgtacagcaa	ggggcagttg	cgggagctgg	agcgggagta	tgcggctaac	aagttcatca	900
ccaaggacaa	gaggcgcaag	atctcggcag	ccaccagcct	ctcggagcgc	cagattacca	960
tctggtttca	gaaccgccc	gtcaaagaga	agaaggttct	cggcaagggtg	aagaacagcg	1020
ctacccctta	agagatctcc	ttgcctgggt	gggaggagcg	aaagtggggg	tgtcctgggg	1080

10

20

30

```

agaccagaaa cctgccaaagc ccaggctggg gccagggact ctgctgagag gccctagag 1140
acaacaccct tcccaggcca ctggctgctg gactgttcct caggagcggc ctgggtaccc 1200
agtatgtgca gggagacgga accccatgtg acaggcccac tccaccaggg tcccaaaga 1260
acctggccca gtcataatca ttcattcctca cagtggcaat aatcacgata accagt 1316

```

<210> 656

<211> 566

<212> DNA

<213> Homo sapiens

<400> 656

```

cccaccaaac ccataaagag ggtgggtcga cccacgcgctc cgcggacgag tgggaaatta 60
ttgaattgga aacagaaata gaaaagttaa aagctgagaa cgcattctta gctaaacttc 120
gcattgaacg agaaagtgc ttggaaaac tcaggaaaga aattgcagac ttcgaacaac 180
agaaagcaaa agaattagct cgaatagaag agtttaaaaa ggaggagatg aggaagctac 240
aaaagggaacg taaagttttt gaaaagtata ctacagctgc aagaactttt ccagataaaa 300
aggaacgtga agaaatacag actttaaaac agcaaatagc agatttacgg gaagatttga 360
aaagaaagga gaccaaattg tcaagtacac acagccgtct cagaagccag atacaaatgt 420
tagtcagaga gaacacagac ctccgggaag aaataaaagt gatggaaaga ttccgactgg 480
atgcctggaa gagagcagaa gccatagaga gcagcctcga ggtggagaag aaggacaagc 540
ttgogaacac atctgttcga tttcaa 566

```

<210> 657

<211> 5067

<212> DNA

<213> Homo sapiens

<400> 657

```

gcccgacac ctgtctgcag catggataag tatgacgacc tgggcctgga ggccagtaaa 60
ttcatcgagg acctgaacat gtatgaggcc tctaaggatg ggctcttcag agtggacaag 120
ggtgcaggca acaaccccg gtttgaggaa actcgcaggg tgttcgccac caagatggcc 180
aaaatccacc tccagcagca gcagcagcag ctccctgcagg aggagactct gccaggggg 240
agtagaggcc ctgtcaatgg agggggccgc ctggggccac agggccgttg ggaagtgtg 300
ggcagcaagc tgactgtgga tgggtctgcc aagcctcctc ttgctgcctc gacaggggca 360
cctggggcag tcaccaccct cgctgctggg cagcccccgt acccaccgca ggagcagaga 420
tccaggccat acctgcatgg cagcaggcat ggcagccagg actgtgggtc caggagagagc 480
ctggcgactt ctgagatgtc tgctttccac cagccaggcc cctgtgagga tccctcctgc 540
ctcactcatg gagactatta tgacaacctc tccttgga gcccaggatg gggtgacaaa 600
ccaggagtgt cccccagcat cggcctgagt gtagggagtg ggtggcctag ctcccgggg 660
agtgaccac cactgcccac accctgcggg gaccatccc taaatcacag acagctctcc 720
ctgagctcca gcaggctctc tgagggtagc ctcggtggtc agaatagtg cattggtggc 780
cgcagcagcg agaagccaac aggccttttg tccactgcct cctccagcgg ggtgagccct 840
ggcctgcctt ccccaaactt ggagaacgga gcaccagctg tggggcctgt tcagccagg 900
actcctctg tgtcagcacc cttggccctg agctgcccc ggcaggagg tcttccaaga 960
tcaaactcgg ggtgggggg tgagggttca ggtgtgatgt ccaaaccaca tgtggacccc 1020
caaccctgg tccaggatgg gcccaaactc tacctttcca gttctgcccc gtcattctcg 1080
ccagctggcc tggacgggtc acagcagggt gcggtccctg ggctggggcc gaagcctggc 1140
tgacacagac ttggcactgg tcccaagctc agccccacca gtcttgtcca tccagtgatg 1200
tccaccctgc ctgagttatc ttgtaaagag ggccccctg gctggtcttc tgatggtagc 1260
ctgggatctg tgctcctgga cagccccagc tcccctaggg taaggctgcc ctgccagccc 1320

```

10

20

30

ctcgtcccag	gtcctgagct	gagaccctct	gctgctgagt	tgaaattaga	agccctcacc	1380
caacgctctg	agcgagagat	ggatgctcac	cgaaggctg	attactttgg	agcctgtgtg	1440
aaatgcagca	aaggggtggt	tggggctggc	caggcctgtc	aggccatggg	gaacctctac	1500
catgacacat	gcttcacctg	tgcagcttgc	agccggaagc	tgagaggaaa	agccttttat	1560
tttgtcaacg	gcaaatgtgt	ttgtgaagaa	gacttcctgt	actctgggtt	ccagcagtct	1620
gctgacaggt	gttttctttg	tggacatctg	atcatggaca	tgatcctgca	agccctgggg	1680
aagtccctacc	accccggtct	tttccgctgt	gtcatctgta	atgagtgttt	ggatgggggt	1740
cccttcaccg	tggactcaga	gaacaagatc	tactgtgtcc	gagattacca	caagggtgct	1800
gcccccaagt	gtgcagcctg	tgggcttccc	atccttccac	ctgagggtcc	agatgagacc	1860
atccgtgtcg	tgtccatgga	cagagactac	cacgtggagt	gttaccactg	cgaggactgt	1920
ggctctggagc	tcaatgatga	agatggccac	cgctgttatc	cgctggagga	ccacctgttc	1980
tgctactcct	gccacgtgaa	gaggctggag	aagagaccct	catctacagc	ccttcaccag	2040
caccactttct	agccagagcc	acttgccagac	atcacggcag	gggatgagga	gccgggggtg	2100
ctgctgctgc	ttccgggtggc	ccctgggggt	gaagtggggg	aggggaagag	gaggggcagg	2160
agggagaggt	cctgtgagca	tgtggggggg	gcctttcctt	taaccaggga	ggtgaacaat	2220
acctgcctcc	tgcgtgtatt	ttccaagtgc	ttttctctgt	tgccacattt	tcctcaggtt	2280
actcaggaaa	atgctccagc	atgtgcgagc	acatgacctg	aggttgcata	atagcaccaa	2340
aggaatcctc	ctgtcccttc	tgggaacatt	tcatgcttca	gagggagagg	tttttattga	2400
gcttgtttca	caatatcccc	ttgaagggac	agctcagctg	ccaatacatt	caacctcttc	2460
tcttctctca	ggaaaatacc	tatacccaaa	tgttccctcc	cccgcataat	atcatggcat	2520
gacttaagga	ttcttttccac	ctgagagctt	cagttcttct	gcagaatggc	tgcaaattta	2580
attgcattaa	ggcaagaagg	aagctctaata	gtgtgctttg	tatcctaaga	taaatttgcct	2640
tagaaaacca	gagtcaagat	ttgaaatagg	tgaggcaggg	tttctccttt	agacactgac	2700
agcattctcc	gtaccctctc	aaatccttac	tctcctaaag	gcagctgagt	ccgcgacaga	2760
aatttgcctc	atgggagtaa	aacatacttt	gggagaagaa	cttgggtcag	gcaccaggat	2820
tttttttttt	gcccacgtgt	ttgcgctgtt	tttctctgga	gttctcaaga	gttgggtgact	2880
tgggaaggccg	cttctgcaag	gcaagtctca	ggaacccatg	caggtaacac	gcttgacact	2940
gttttttagct	tatttaatat	cgggcctttg	ggaagagctg	ccgcataact	gagagacagc	3000
ttcttataaa	caaggagagt	ttttgtgtgt	gcgagatctc	taagccagcg	tgggagggag	3060
cgccctagga	taagtattta	tattcatttc	gttggtttct	ctcctgcccc	attcttggca	3120
caggcattat	gtttgaagaa	accaggataa	ggtacactgc	ttttgtctgt	ttaatTTTTT	3180
tagttgtttt	ccttcaacttt	cagtcttcca	cacacaaaaa	atacctcaca	gagcttcacc	3240
aaatcacaga	ttcaggagga	atttggcttt	cacactggac	tcagatacct	tcttcagtgt	3300
gttggaaaac	actggcttca	cacaggccca	actccaactg	gtcaggggcag	agtgatcgta	3360
actaaaggtc	agtggggaat	agatccgatt	cagtgccttt	gccttatgca	tttcagcatc	3420
ctggctcccc	agggtggcag	gagctgagga	agggccacac	actggccaaga	tttcaagacc	3480
actctctgca	ctgaagaggt	aaaatttgca	ctgcaagtca	catccctgag	gccagaggtc	3540
agtacccttt	ggtatttctga	ttagaagaag	ctgcaaaaaga	aaggcagccc	attttaccat	3600
tgccagccag	gccgggggaca	caggagccgg	tgtgtgcact	ctgcctccta	acattgcacc	3660
cagagcaaga	ggactgggtg	ctgggctgca	gaggccgggt	agtggagccc	ctagcacgtg	3720
tgaactcagg	cttttcattg	ggcccggctc	cacttctagg	ccatgttttg	actcatttgg	3780
taaccattgc	ctgtaagcag	cacagaattg	gtgccatgga	ttatcttttc	catgttgatg	3840
gaattcattc	tgttggaaat	ctttggccag	atgtcaactc	agccagggtg	tgcacatca	3900
ttgggtcttt	ttcaccaggct	gagcctcctg	aaaacccatg	aacgctgggg	ctgggggaagt	3960
gaacctgag	gtgggggacc	tctcttccca	tcaaatacat	cagctcagtg	tggggcgtgg	4020
caggggggta	aatgaagcca	gccaatgtgt	taacctgtct	ctgtcaacct	aagaatgttg	4080
gccttactga	cacacccttg	ctccatgttc	aagaccagaa	gtagctggga	tttgtttgca	4140
aattgggtaa	ttagttttaa	aatctgtgat	tacattttta	aatgaaattt	tcaaagtggc	4200
ctagattgag	gtgattcaga	taggtttgcg	aatataccat	tttatattgt	tgagaaagaa	4260
caaaaaggga	atttccagat	gtcctagaaa	tcctagcaac	agatttctct	ggttgtcagt	4320

10

20

30

```

ttccctggag aaggcgccag ataggaatct ccaatcagtt gtttttctct tcgcttcagg 4380
cccttacaca aaagccatga agagatgttc acctaccggg tatttttaa gttctgtaaa 4440
ttattagcca aatagaactg taatgggggt gtattttatgg ggcctagaa agaaaacaca 4500
aggacttggg aggccaggaa gaaaagattt taaaatttag aatgaatagc ccttctgggt 4560
tttctttttg acaattcttg gacttgaggt aaaacaagga ggattgtggc cggatttcag 4620
atcccaaagc cagcctccat cttaggcctt tgcctcattg tgccttttag gttttcttac 4680
ccacogtctc ctgtttttgtc ttttttttct tttctcctac cctatcttg ggacattcag 4740
aaactgcctg ggtggtttga gaagagacaa ccagtttga tctgcaatac aaggatccat 4800
tcgtaatctc tctctcactg atgttattcc cccatctgcc gtcttggttc atctcaccac 4860
agaagggcat ttagtcctac ccagccatcg gctgcggtat gacagcagga tggcacttcc 4920
catttctctg tggttagtgc tcgagtgaac acctctttca gctgagtcct ctgaggttct 4980
gctgttgagt cctgggtggc tgatggaatg attgaggagg tctggtcacc ctcaagcgcc 5040
gtcatcgctc tgtttccatg ggcttct 5067

```

10

```

<210> 658
<211> 950
<212> DNA
<213> Homo sapiens

```

```

<400> 658
tcgaccggat ccgaattccc attgtgcact aaagcgtctc cctgctccgc ggcccgggct 60
ggcgggcggg cgctcggtg gggctgcag cagcagaggg agaccgcggg caaccgccgg 120
aaccaggggc tcggcgctgc tgccaccatg acgggaagca atatgtcgga cgccttggcc 180
aacgcggtgt gccagcgctg ccaggccccg ttctcccccg ccgagcgcat tgtcaacage 240
aatggggagc tgtaccatga gcactgcttc gtgtgtgcc agtgcttcgg gcccttccc 300
gaggggctct tctatgagtt tgaaggccgg aagtactgcg aacacgactt ccaaatgctg 360
tttgcctcgt gctgtggatc ctgcggtgag ttcatcattg gccgcgtcat caaggccatg 420
aacaacaact ggcacccggg ctgcttccgc tgcgagctgt gtgatgtgga gctggctgac 480
ctgggctttg tgaagaatgc cggcaggcat ctctgccggc cttgccacaa cgtgagaag 540
gccaaaggcc tgggcaagta catctgccag cgggtgccacc tggtcacga cagcagccc 600
ctcatgttca ggagcgacgc ctaccacct gaccacttoa actgcaccca ctgtgggaag 660
gagctgacag ccgaggcccc cgagctgaag ggtgagctct actgcctgcc ctgccatgac 720
aagatgggag ttcccatctg cggggcctgc cgccggccca tcgaggggcg agtgggtcac 780
gcgctgggca agcagtggca cgtggagcac tttgtctgtg ccaagtgtga gaagccattc 840
ctggggcacc ggcactatga gaagaaggc ctggcctact gcgagcttta gtgcacaatg 900
ggcgcccgcc tcgagtcctag actcgagtag ataattgagc ggaatttctt 950

```

20

```

<210> 659
<211> 2346
<212> DNA
<213> Homo sapiens

```

30

```

<400> 659
ccgcgctgc cccgcctcc cctgcctca ggggctgcc ccgccagcg gcccgcgct 60
ccccggggc ttgcagcggg ccccgcccg gctggagggg ccccgacccc agctctggtg 120
gcgggcagca gcgcgcggg ccccttccct caccgggact cggccctgaa cagcaggag 180
aaggagtgtc agcggcggt gaagcgtctc taccgggccc tggacgaaca agagacggc 240
ctgctcggg cctggagccc gaaggacaa ttcagctaca tggcctctc tcagaacaac 300
ctgcgggtgc actacaaaag tcatggcaaa accccaaaag atgcgcgctc agttcgagcc 360

```



```

acgcatccaa taccagcagc ctgtgggatt tattattttg aagtaaaaat tgtcagtaag 420
ggaagagatg gttacatggg aatttgtctt tctgctcaag gtgtgaacat gaatagacta 480
ccaggttggg ataagcattc atatggttac catggggatg atggacattc gttttgttct 540
tctggaactg gacaacctta tggaccaact ttcactactg gtgatgtcat tggctgttgt 600
gttaattctta tcaacaatac ctgcttttac accaagaatg gacatagtgt aggtattgct 660
ttcactgacc taccgcaaaa tttgtatcct actgtggggc ttcaaaccac aggagaagtg 720
gtcgatgcca attttgggca acatcctttc gtgtttgata tagaagacta tatgcgggag 780
tggagaacca aaatccaggc acagatagat cgatttccta tcggagatcg agaaggagaa 840
tggcagacca tgatacaaaa aatgggttca tcttatttag tccaccatgg gtactgtgcc 900
acagcagagg cctttgccag atctacagac cagaccgttc tagaagaatt agcttccatt 960
aagaatagac aaagaattca gaaattggta ttagcaggaa gaatgggaga agccattgaa 1020
acaacacaac agttataccc aagtttactt gaaagaaatc ctaatctcct ttccacatta 1080
aaagtgcgtc agtttataga aatgtgtgaat ggtacagata gtgaagtacg atgtttggga 1140
ggcgaagtc caaagtctca agacagttat cctgttagtc ctgcaccttt tagtagtcca 1200
agtatgagcc ccagccatgg aatgaatata cacaatttag catcaggcaa aggaagcacc 1260
gcacattttt caggttttga aagttgtagt aatggtgtaa tatcaaataa agcacatcaa 1320
tcataattgcc atagtaataa acaccagtca tccaacttga atgtaccaga actaaacagt 1380
ataaatatgt caagatcaca gcaagttaat aacttcacca gtaatgatgt agacatggaa 1440
acagatcact actccaatgg agttggagaa acttcaccca atgggttccct aaatggtagc 1500
tctaaccatg accacgaaat ggaagattgt gacaccgaaa tgggaagtga ttcaagtcag 1560
ttgagacgcc agttgtgtgg aggaagtcag gccgccatag aaagaatgat ccactttgga 1620
cgagagctgc aagcaatgag tgaacagcta aggagagact gtggcaagaa cactgcaaac 1680
aaaaaaatgt tgaaggatgc attcagtcct ctagcatatt cagatccctg gaacagccca 1740
gttggaaatc agcttgaccc gattcagaga gaacctgtgt gctcagctct taacagtga 1800
atattagaaa cccacaatct gccaaagcaa cctccacttg ccctagcaat gggacaggcc 1860
acacaatgtc taggactgat ggctcgatca ggaattggat cctgcgcatt tgccacagt 1920
gaagactacc tacatttagc atgcatttca agagctcaca cttatatgtt ggcataatgt 1980
caacatggaa gtagaccagc tctgctgatt tgaaatttag atttttttaa ttatgtactg 2040
gggacagggt tttgtcgctt tacattgctt cctagtttac agcatgatgc aaatgatttt 2100
ctaacttagt gtttagagaa attattttcc atctttaacc tcttagttgt ctaagagtta 2160
aatattactg aatttcagac gttcaaattg atcatcacia atcctttaa acaattacct 2220
aaaagaaacc aaaaatcctg ccttctttgt gggggagggg ggagagaggg gaaggaaatg 2280
gaacaagttg tgtttgtgtt agcatgtggg tgatgtaaac ttcaaattgg gagatgttcc 2340
gacccc 2346

```

<210> 660

<211> 2512

<212> DNA

<213> Homo sapiens

<400> 660

```

caatgcactg acggatatga gtgggatoct gtgagacagc aatgcaaaga tattgatgaa 60
tgtgacattg tcccagacgc ttgtaaaggt ggaatgaagt gtgtcaacca ctatggagga 120
tacctctgcc ttccgaaaac agcccagatt attgtcaata atgaacagcc tcagcaggaa 180
acacaaccag cagaagggaac ctccaggggca accacggggg ttgtagctgc cagcagcatg 240
gcaaccagtg gagtgttgcc cgggggtggg tttgtggcca gtgctgctgc agtcgcaggc 300
cctgaaatgc agactggcgg aaataacttt gtcacccggc ggaaccagc tgaccctcag 360
cgcattccct ccaacccttc ccaccgtatc cagtgtgcag caggctacga gcaaagtga 420
cacaacgtgt gccaaagacat agacgagtg cactgcaggga cgcacaactg tagagcagac 480
caagtgtgca tcaatttacg gggatccttt gcatgtcagt gccctcctgg atatcagaag 540

```

10

20

30

```

cgaggggagc agtgcgtaga catagatgaa tgtaccatcc ctccatattg ccaccaaaga 600
tgcgtgaata caccaggctc atttttattgc cagtgcagtc ctgggtttca attggcagca 660
aacaactata cctgcgtaga tataaatgaa tgtgatgcca gcaatcaatg tgctcagcag 720
tgctacaaca ttcttgggtc attcatctgt cagtgcacac aaggatatga gctaagcagt 780
gacaggctca actgtgaaga cattgatgaa tgcagaacct caagctacct gtgtcaatat 840
caatgtgtca atgaacctgg gaaattctca tgtatgtgcc ccaggggata ccaagtgggtg 900
agaagtagaa catgtcaaga tataaatgag tgtgagacca caaatgaatg ccgggaggat 960
gaaatgtgtt ggaattatca tggcggcttc cgttgtttat caggaatcc ttgtcaagat 1020
ccctacattc taacaccaga gaaccgatgt gtttgcccag tctcaaatgc catgtgccga 1080
gaactgcccc agtcaatagt ctacaaatac atgagcatcc gatctgatag gtctgtgccca 1140
tcagacatct tccagataca ggccacaact atttatgcca acaccatcaa tacttttcggg 1200
attaaatctg gaaatgaaaa tgggaggttc tacctacgac aaacaagtcc tgtaagtgca 1260
atgcttgtgc tegtgaagtc attatcagga ccaagagaac atatcgtgga cctggagatg 1320
ctgacagtca gcagtatagg gaccttcgcc acaagctctg tgttaagatt gacaataata 1380
gtggggccat ttctatttta gtcttttcta agagtcaacc acaggcattt aagtcagcca 1440
aagaatattg ttaccttaaa gcactatttt atttatagat atatctagtg catctacatc 1500
tctatactgt acactcaccg ataacaaaca attacaccat ggtataaagt gggcatttaa 1560
tatgtaaaga ttcaaaagtt gtctttatta ctatatgtaa attagacatt aatccactaa 1620
actggtcttc ttcaagagag ctaagtatac actatctggt gaaacttgga ttctttccta 1680
taaaagtggg accaagcaat gatgatcttc tgtggtgctt aaggaaactt actagagctc 1740
cactaacagt tccataagga ggcagccatc ataaccattg aatagcatgc aagggttaaga 1800
atgagttttt aactgctttg taagaaaatg gaaaaggcca ataaagatat atttctttag 1860
aaaaagggga tctgccatat ttgtgttggt ttttattttc atatccagcc taaagggtggt 1920
tgtttattat atagtaataa atcattgctg tacaaatgct tggtttctgt aggggtatttt 1980
taattttgtc agaaatttta gattgtgaat attttgtaaa aaacagtaag caaaattttc 2040
cagaattccc aaaatgaacc agataccccc tagaaaatta tactattgag aaatctatgg 2100
ggaggatatt agaaaataaa ttcttcttaa accacattgg aactgacctg aagaagcaaa 2160
ctcggaaaat ataataacat cctgaattc aggcattcac aagatgcaga acaaaatgga 2220
taaaagggtat ttactggag aagttttaat ttctaagtaa aatttaaatc ctaacacttc 2280
actaatttat aactaaaatt tctcatcttc gtacttgatg ctccacagagg aagaaaatga 2340
tgatgggtttt tattcctggc atccagagtg acagtgaact taagcaaat accctctac 2400
ccaattctat ggaatatatt atacgtctcc ttgtttaaaa tctgactgct ttactttgat 2460
gtatcatatt tttaataaaa aataaatatt cctttagaag atcactctaa aa 2512

```

<210> 661

<211> 2306

<212> DNA

<213> Homo sapiens

<400> 661

```

gggcgggagc tgcacgcgcc gtggctccgg atctctctgt ctttgcagcg tacgcccag 60
tcggtcagcg ccggaggacc tcagcagcca tgtcgagcc ccatagtga gcccggactg 120
ccttcattca gaccagcag ctgcacgcag ccattggtga cacattcctg gagcacatgt 180
gccgcctgga cattgattca ccacccatca cagcccggaa cactggcatc atctgtacca 240
ttggcccagc ttcccgatca gtggagacgt tgaaggagat gattaagtct ggaatgaatg 300
tggctcgtct gaactctct catggaactc atgagtacca tgcggagacc atcaagaatg 360
tgcgcacagc cagggaaagc tttgcttctg accctacct ctaccggccc gttgctgtgg 420
ctctagacac taaaggacct gagatccgaa ctgggctcat caagggcagc ggcaactgcag 480
agctggagct gaagaaggga gccactctca aaatcaogct ggataacgca tacatggaaa 540
agtgtgacga gaacatcctg tggctggact acaagaacat ctgcaagggt gtggaagtgg 600

```

10

20

30

```

gcagcaagat ctacgtggat gatgggctta tttctctcca ggtgaagcag aaaggtgccg 660
acttctctggg gacggaggtg gaaaatggtg gctccttggtg cagcaagaag ggtgtgaacc 720
ttctctggggc tgotgtggac ttgctctctg tgtcggagaa ggacatccag gatctgaagt 780
ttgggggtcga gcaggatgtt gatattggtgt ttgcgtcatt catccgcaag gcattctgatg 840
tccatgaagt taggaaggtc ctgggagaga agggaaagaa catcaagatt atcagcaaaa 900
tcgagaatca tgagggggtt cggaggtttg atgaaatcct ggaggccagt gatgggatca 960
tgggtggtctg tgggtgatcta ggcattgaga ttctctgcaga gaaggtcttc cttgctcaga 1020
agatgatgat tggacggtgc aaccgagctg ggaagcctgt catctgtgct actcagatgc 1080
tggagagcat gatcaagaag ccccgcccca ctccggctga aggcagtgtg gtggccaatg 1140
cagtcctgga tggagccgac tgcattcatgc tgcctggaga aacagccaaa ggggactatc 1200
ctctggaggc tgtgcgcatg cagcacctga ttgcccgtga ggcagaggct gccatctacc 1260
acttgcaatt atttgaggaa ctccgcccgc tggcgcccat taccagcgac cccacagaag 1320
ccaccgccgt ggggtgccgtg gaggcctcct tcaagtgtct cagtggggcc ataactgtcc 1380
tcaccaagtc tggcaggtct gctcaccagg tggccagata ccgccacgt gccccatca 1440
ttgctgtgac ggggaatccc cagacagctc gtcaggccca cctgtaccgt ggcattcttc 1500
ctgtgtctgt caaggaccca gtccaggagg cctgggctga ggacgtggac ctccgggtga 1560
actttgccat gaattgttggc aaggcccgag gcttcttcaa gaaggagat gtggtcattg 1620
tgetgaccgg atggcgccct ggtcctggct tcaccaacac catgctgtgt gtctctgtgc 1680
cgtgatggac cccagagccc ctctccagc ccctgtccca ccccttccc ccagccatc 1740
cattaggcca gcaacgcttg tagaactcac tctgggctgt aacgtggcac tggtaggttg 1800
ggacaccagg gaagaagatc aacgcctcac tgaaacatgg ctgtgtttgc agcctgtct 1860
agtgggacag cccagagcct ggctgcccac catgtggccc caccacatca agggaagaag 1920
gaggaatgct ggactggagg cccctggagc cagatggcaa gaggtgaca gcttcttttc 1980
ctgtgtgtac tctgtccagt tcttttagaa aaaatggatg cccagaggac tcccaacct 2040
ggcttggggg caagaaacag ccagcaagag ttaggggcct tagggcactg ggctgtgtgt 2100
ccattgaagc cgactctggc cctggccctt acttgcttct ctactctctt aggcctctcc 2160
agtttgcacc tgteccacc ctccactcag ctgtcctgca gcaaacctc caccctccac 2220
cttccatttt cccccactac tgcagcacct ccaggcctgt tgctatagag cctacctgta 2280
tgtcaataaa caacagctga agcacc 2306

```

<210> 662

<211> 2656

<212> DNA

<213> Homo sapiens

<400> 662

```

gaattcgcgg ccgcagagtc cccgggccaa gatggctgct cgggtgctcca cagctggtt 60
gctggtggtt gtggggaccc cgcggctgcc ggctatatcg ggtagagggg cccggccgcc 120
cagggagggc gtggtggggg catggctgag ccgcaagctg agcgtcccg cctttgcgtc 180
ttcctgacc tcttgcgcc cccgagcgt gctgacattg agacctggtg tcagccttac 240
aggaacaaaa cataaccctt tcatltgtac tgctccttc cacacgagt ccccttggc 300
caaagaagat tattatcaga tattaggagt gcctcgaaat gccagccaga aagagatcaa 360
gaaagcctat tatcagcttg ccaagaagta tcacctgac acaaataagg atgatccaa 420
agccaaggag aagttctccc agctggcaga agcctatgag gttttgagt atgaggtgaa 480
gaggaagcag tacgatgctt acggctctgc aggettcgat cctggggcca gggctccca 540
gcatagctac tggaaaggag gccccactgt ggaccccgag gagctgttca ggaagatctt 600
tggcgagttc tcatcctctt catttgagga ttccagacc gtgtttgatc agcctcagga 660
atacttcatg gagttgacat tcaatcaagc tgcaaaaggg gtcaacaagg agttcaccgt 720
gaacatcatg gacacgtgtg agcgtgcaa cggcaagggg aacgagcccg gcaccaaggt 780
gcagcattgc cactactgtg ggggtccgg catggaaacc atcaacacag gcccttttgt 840

```

10

20

30

```

gatgcgttcc acgtgtagga gatgtggtgg ccgcggtccc atcatcatat cgccctgtgt 900
ggtctgcagg ggagcaggac aagccaagca gaaaaagcga gtgatgatcc ctgtgcctgc 960
aggagtcgag gatggccaga ccgtgaggat gcctgtggga aaaagggaaa ttttcattac 1020
gttcagggtg cagaaaaagcc ctgtgttccg gagggaacggc gcagacatcc actccgacct 1080
ctttatctct atagctcagg ctcttcttgg gggaacagcc agagcccgag gcctgtacga 1140
gacgatcaac gtgacgatcc cccctgggac tcagacagac cagaagattc ggatgggtgg 1200
gaaaggcatc ccccggtatta acagctacgg ctacggagac cactacatcc acatcaagat 1260
acgagtcca aagaggctaa cgagccggca gcagagcctg atcctgagct acgcccaggga 1320
cgagacagat gtggaggggga cggtgaaagg cgtcacccctc accagctctg gtggcagcac 1380
catggatagc tccgcaggaa gcaaggctag gcgtgaggct ggggaggagc aggagggatt 1440
cctttccaaa cttaaagaaa tgtttacctc atgatatccc agccgaggaa aaagatccac 1500
tggaacttag gccgggaagc agcagccctc ccaaggggcca gggcacctgg gagacgggag 1560
gattccagaa cagcagcact gagctccac ccgcagagcc tctggacggc cttggcaaca 1620
gcaaatcatc gggacaacac ctctctccac ggaaaggcca cagtggacag cccgggcagt 1680
aggatgcagc ccagagggtt ggtggcagtt tccgttccat tggtaggtga cggccccctg 1740
gtcagcagag gagagggttag atcttgacgg ctaaaactct aatttggaat tgaatatgt 1800
ggatatctta gttaaaggcc atgcttacag cttagaaatg aagccttaag ctgcatcaag 1860
ttacgaagtg attaatcttc ttctcagcaa acctccggga ggttcagaa tgagtctctc 1920
ctgacagggt gtcttcactg ggagcgtggg gcccccaggc cccaccagca ccgtccccc 1980
ctaattgagg gccctgccga ggcatcagct gctctgctca gttagttttt attcccggg 2040
taccgaagcag ctgcacagtc ggtgcctggg aagcacgtta aaggcccgag gagatccctg 2100
gggttctgct ctgaccgtgt ggggtggtgat ccttgtcagg atgtacagtc cttgtcccca 2160
ccccatccgg gatggccgcc tgtccctgac tattgagtcc tgttgttcta agccaggcat 2220
ggagggtccc tgcctctctg ctgagccaca gcccatgtga gcactgtgct ggccagactt 2280
cagctgcctt ggggaactga gccctgccac tgttgctagt caggggcttg gttctccac 2340
ttacactgtt gacatctatt ttctgaagtg tgtttaaatt attcagtgt aatcattgtt 2400
ttttctcttg taaatgttga ttccagaaaag gaaagcacag gctaagcagt tgaaggttcc 2460
ccaccattca gtgagagcag aacccccatt cccagcctc tgcctgtagc atgtgcagat 2520
ttccatgtgt ttccaggatc tcgggctgtc gtttagacag ttaatgaaga acacttctca 2580
acagtcttct ttttgttttc ctttataatt cactaaaata aagcatctat tagtgtctga 2640
aaaaaaaaa aaaaaa 2656

```

<210> 663

<211> 2217

<212> DNA

<213> Homo sapiens

<400> 663

```

gcggacccgg ccgcgaggcg gccacccgag acgcggcgcg cacgctccgg cctgcgcagc 60
ccggcccggc catggcggcc ccccgcccggt ctcccgcgat ctccgtttcg gtctcggctc 120
cggcttttta cgcgcccgag aagaagtctg gccctgtggt ggccccaaag cccaaagtga 180
atcccttccg gcccggggac agcagacctc ccccggcacc cggggcccag cgcgcacaga 240
tgggccgggt gggcgagatt ccccgccgcg ccccggaaga ctttccctg cctccacctc 300
cccttgctgg ggtggcgac gatgcagagg gtgctctggg aggtgccttc ccgcccgc 360
ctccccgat cgaggaatca tttccctctg ccctctgga ggaggagatc ttcccttccc 420
cgccgcctcc tccggaggag gagggagggc ctgaggcccc catacggccc ccaccacagc 480
ccaggagagaa ggtgagcagt attgatttgg agatcgactc tctgtcctca ctgctggatg 540
acatgaccaa gaatgatcct ttcaaagccc ggggtgtatc tggatatgtg cccccaccag 600
tgcccactcc attcagttcc aagtccagta ccaagcctgc agccgggggc acagcacc 660
tgccctcctg gaagtccctc tccagctccc agcctctgcc ccagggttcg gctccggctc 720

```

10

20

30

```

agagccagac acagttccat gttcagcccc agccccagcc caagcctcag gtccaactcc 780
atgtccagtc ccagaccagc cctgtgtctt tggctaaccac ccagccccga gggccccccag 840
cctcatctcc ggctccagcc cctaagtttt ctccagtga cctaagttt actcctgtgg 900
cttcccaagtt cagtccctgga gccccagggtg gatctgggtc acaaccaa atcaaatgtg 960
ggcaccocga agctctttct gctggcacag gctccccca acctcccagc ttccactatg 1020
cccagcagag ggagaagccc cgagtgcagg agaagcagca ccccggtccc ccacgggctc 1080
agaacccaaa ccagggtgcgc tcccctgggg cccagggcc cctgactctg aaggagggtg 1140
aggagctgga gcagctgacc cagcagctaa tgcaggacat ggagcatcct cagaggcaga 1200
atgtggctgt caacgaactc tggggccgat gccatcaacc cctggcccgg gcgcagccag 1260
ccgtccgcgc tctagggcag ctgttccaca tgcctgtctt cactgccac cagtgtgcgc 1320
agcagctcca gggccagcag ttctacagtc tggagggggc gccgtactgc gagggtgtgt 1380
acactgacac cctggagaag tgtaaacact gcggggagcc catcactgac cgcattgtga 1440
gggcccacgg caaggcctat caccgcact gcttcacctg tgtggctctg gccgcgcccc 1500
tggagggcac ctcttctatc gtggaccagg ccaaccggcc cactgtgtgc ccgactacc 1560
acaagcagta cgcgccaggg tgctccgtct gctctgagcc catcatgcct gagcctggcc 1620
gagatgagac tgtgcgagtg gtgcgctgg acaagaactt ccacatgaag tgttacaagt 1680
gtgaggactg cgggaagccc ctgtcgattg aggcagatga caatggctgc ttccccctgg 1740
acgggtcacgt gctctgtcgg aagtgcacac ctgctagagc ccagacctga gtgaggacag 1800
gccctcttca gaccgcagtc catgcacct tgtggaccac ccacactgag accacctgcc 1860
cccacctcag ttattgtttt gatgtctagc cctcccatt tccaaccct ccctagcatc 1920
ccaggtgccc tgaccagga cccaacatgg tctaggatg caggatcccc gcctgggggt 1980
ctggctctcg cccatcctgc agggattgcc caccgtcttc cagacacccc acctgagggg 2040
ggcaccaggt ttagtgctgc tgccttctact gctgcacccg cgcctcgggc cggccccccg 2100
agcagccttt gtactctgct tgcggagggc tgggagaccc tccaggacat tcccacctc 2160
ccccatgctg ccaagttgta gctatagcta caaataaaaa aaaccttgt ttccag 2217

```

<210> 664

<211> 8906

<212> DNA

<213> Homo sapiens

<400> 664

```

gaggcggcca aggacctggc cgacatcgcg gccttcttcc gatccgggtt tcgaaaaaac 60
gatgaaatga aagctatgga tgttttaacca attttgaagg aaaaagttgc atacctttca 120
ggtgggagag ataaacgtgg aggtccatt ttaacgttcc cggcccgag caatcatgac 180
agaatacgac agggagatct caggagactc atttccatc tagcctgtat tccagcgag 240
gagggtctga agogtggctt caggtgtatc gtggacatgc gtgggtccaa gtgggactcc 300
atcaagcccc ttctgaagat cctgcaggag tcttccccct gctgcacca tgtggccctg 360
atcatcaagc cagacaactt ctggcagaaa caggaggacta attttggcag ttctaaattt 420
gaatttgaga caaatatggc ctcttttaga ggctttacca aagtagttga tccttctcag 480
ctaactcctg agtttgatgg ctgcctggaa tacaaccacg aagaatggat tgaatcaga 540
gttgcttttg aagactacat tagcaatgcc accacatgc tgtctcggct ggaggaaactt 600
caggacatcc tagctaagaa ggagctgctt caggatttag agggggctcg gaatatgatc 660
gaggaacatt ctacgtgaa gaagaagggtg attaaggccc ccacgagga cctggatttg 720
gaggggacaga agctgttcca gaggatacag agcagtga aa gctttccaa aaagaactca 780
ggctcaggca atgcggacct gcagaacctc ttgcccagg tgtccaccat gctggaccgg 840
ctgcactcga caccgcagca tctgcaccag atgtggcatg tgagggaagt gaagctggac 900
cagtgtcttc agctgaggtt gtttgaacag gatgtgaga agatgtttga ctggatcaca 960
cacaacaaag gctgtttct aaacagctac acagagattg ggaccagcca cctcatgac 1020
atggagcttc agacgcagca caatcacttt gccatgaact gtatgaacgt gtatgtaaat 1080

```

10

20

30

ataaacgcga	tcattgtcgg	ggccaatcgt	ctgggtggagt	ctggccacta	tgccctgcag	1140
cagatcaggc	agatcgcgag	tcagctggag	caggagtggg	aggcgtttgc	ggcagccctg	1200
gatgagcgga	gcaccttgct	ggacatgtcc	tccattttcc	accagaaggc	cgaaaagtat	1260
atgagcaacg	tggattcatg	gtgtaaaagc	tgccgtgagg	tagaccttcc	ctcagagctg	1320
caggacctag	aagatgccat	tcattaccac	cagggaatat	atgaacatat	cactcttgct	1380
tattctgagg	tcagccaaga	tgggaagtgc	ctccttgaca	agctccagcg	gcccttgact	1440
cccggcagct	cggattccct	gacagcctct	gccaactact	ccaaggccgt	gcaccatgtc	1500
ctggatgtca	tccacgaggt	gctgcaccac	cagcggcacg	tgagaacaat	ctggcaaac	1560
cgcaaggctc	ggctgcatca	gaggctgcag	ctgtgtgttt	tccagcagga	agttcagcag	1620
gtgctagact	ggatcgagaa	ccacggagaa	gcattttctg	gcaaacatac	aggtgtgggg	1680
aaatctcttc	atcggggccag	agcattgcag	aaacgtcatg	aagattttga	agaagtggca	1740
cagaacacat	acaccaatgc	ggataaatta	ctggaagcag	cagaacagct	ggctcagact	1800
ggggaatgtg	accccggaaga	gatttatcag	gctgcccatc	agctggaaga	ccggattcaa	1860
gatttcgttc	ggcgtgttga	gcagcgaag	atcctactgg	acatgtcagt	gtcctttcac	1920
acccatgtga	aagagctgtg	gacgtggctg	gaggagctgc	agaaggagct	gctggacgac	1980
gtgtatgccc	agtccgttga	ggcgtgcag	gacctcatca	agcgttttgg	ccagcagcag	2040
cagaccaccc	tgccaggtgac	tgtcaacgtg	atcaaggag	gggaggacct	catccagcag	2100
ctcagggaact	ctgccatctc	cagtaacaag	acccccaca	acagctccat	caaccacatt	2160
gagacgggtg	tgccagcagct	ggacgaggcg	cagtgcgaga	tggaggagct	cttccaggag	2220
cgcaagatca	agctggagct	cttctgcac	gtgcgcacat	tcgagaggga	cgccatcgac	2280
attatctcag	acctcgagtc	ttggaatgat	gagctttctc	agcaaatgaa	tgacttcgac	2340
acagaagatc	tcacgattgc	agagcagcgc	ctccagcacc	atgcagacaa	agccttgacc	2400
atgaacaact	tgacttttga	cgtcatccac	caagggcaag	atcttctgca	gtatgtcaat	2460
gaggtccagg	cctctggtgt	ggagctgctg	tgtgatagag	atgtagacat	ggcaactcgg	2520
gtccaggacc	tgctggagtt	tcttcatgaa	aaacagcagg	aattggattt	agccgcagag	2580
cagcatcgga	aacacctgga	gcagtgcgtg	cagctgcgcc	acctgcaggc	agaagtgaag	2640
cagggtgctgg	gttggatccg	caacggagag	tccatgttaa	atgccggact	tatcacagcc	2700
agctcgttac	aagaggcaga	gcagctccag	cgagagcag	agcagttcca	gcattgccatt	2760
gagaaaacac	atcagagcgc	gctgcagggt	cagcagaagg	cagaagccat	gtacaggccc	2820
aaccactacg	acatggacat	gatccgggac	tgccgcgaga	aggtggcgct	tcaactggca	2880
cagctcatgc	tcaagatgga	agatcgcttc	aagctcgtca	acgcctctgt	cgctttctac	2940
aaaacctcag	agcaggtctg	cagcgtcctc	gagagcctgg	aacaggagta	caagagagaa	3000
gaagactggg	gtggcggggc	ggataagctg	ggcccaact	ctgagacgga	ccacgtgacg	3060
cccatgatca	gcaagcacct	ggagcagaag	gaggcattcc	tgaaggcttg	cacctttgct	3120
cggagggaatg	cagacgtctt	cctgaaatac	ctgcacagga	acagcgtgaa	catgccaggga	3180
atggtgagcg	acatcaaagc	tcttgaacag	caagtgaata	atatcttgaa	tgaactcttc	3240
caacggggaga	acagggtatt	gcattactgg	acctgagga	agagacggct	ggaccagtgt	3300
cagcagtagc	tggctctttg	gaggagtgc	aagcaggctt	tggaaatggat	ccatgacaa	3360
ggcgagttct	acctttccac	acacacctcc	acgggctcca	gtatacagca	caccaggag	3420
ctcctgaaag	agcacgagga	gttccagata	actgcaaagc	aaaccaaaga	gagagtgaag	3480
ctattgatac	agctggctga	tggcttttgt	gaaaaggggc	atgcccattgc	ggcagagata	3540
aaaaaatgtg	ttactgctgt	ggataagagg	tacagagatt	tctctctgcg	gatggagaag	3600
tacaggacct	ctttggaaaa	agccctgggg	atttcttcag	attccaacaa	atcgagtaaa	3660
agtctccagc	tagatatcat	tccagccagt	atccctggct	cagaggtgaa	acttcgagat	3720
gctgctcatg	aacttaatat	agagaagcgg	aaatctgccc	gcaggaaaga	gttcataatg	3780
gctgagctca	ttcaaaactga	aaaggcttat	gtaagagacc	tccgggaatg	tatggatacg	3840
tacctgtggg	aatgaccag	tggcgtggaa	gagattccac	ctggcattgt	aaacaaagaa	3900
ctcatcatct	tcggaaacat	gcaagaaatc	tacgaatttc	ataataacat	attcctaaag	3960
gagctggaaa	aatatgaaca	gttgccagag	gatgttgag	attgttttgt	tacttgggca	4020
gacaagtttc	agatgtatgt	cacatattgc	aaaaataagc	ctgattctac	tcagctgata	4080

10

20

30

ttggaacatg	caggggtccta	ttttgacgag	atacagcagc	gacatggatt	agccaattcc	4140
atcttctcct	accttattaa	accagttcag	cgaataacga	aatatcagct	ccttttataa	4200
gagctgctga	cgtgctgtga	ggaaggaaag	ggagagatta	aagatggcct	ggagggtgatg	4260
ctcagcgtgc	cgaagcgagc	caatgacgcc	atgcacctca	gcatgctgga	aggggttgat	4320
gaaaacattg	agtctcaggg	agaactcatt	ctacaggaat	ccttccaagt	gtgggaccca	4380
aaaaccttaa	ttcgaaaggg	tcgagaacgg	catctcttcc	tttttgaaat	gtccttagta	4440
tttagttaaag	aagtgaagaa	ttccagtggg	agaagcaagt	acctttataa	aagcaaattg	4500
tttacctcag	agttgggtgt	cacagaacat	gttgaaggag	acctttgcaa	atgtgactg	4560
tgggtgggga	gaacaccaac	ttcagataat	aaaattgtcc	ttaaggcttc	cagcatagag	4620
aacaagcagg	actggataaa	gcataatccg	gaagtcatcc	aggagcggac	gatccacctg	4680
aaggaggccc	tgaaggagcc	cattcacatc	cctaagaccg	ctcccggcac	aagacagaag	4740
ggaaggaggg	atggagagga	tctggacagc	caaggagacg	gcagcagcca	gcctgatacg	4800
atctccatcg	cctcagggac	gtctcagaac	acgctggaca	gcgataagct	ctctggtggc	4860
tgtgagctga	cagtggtgat	ccatgacttc	accgcttgca	acagcaacga	gctgaccatc	4920
cgacgggggc	agaccgtgga	agttctggag	cggcccgcatg	acaagcctga	ctgggtgtctg	4980
gtgcggacca	ctgaccgctc	cccagcggca	gaaggcctgg	tcccctgtgg	ttcactgtgc	5040
atcgccact	ccagaagtag	catggaaatg	gagggcatct	tcaaccacaa	agactcgctc	5100
tccgtctcca	gcaatgacgc	cagtcacccc	gcatccgtgg	cttccctcca	gccccacatg	5160
atcgggggcc	agagctcgcc	gggcccgaag	cggccgggca	acacctcgcg	caagtggctc	5220
accagccccg	tgcggcggtc	cagcagcggc	aaggccgacg	ggcacgtgaa	gaagctggcg	5280
cacaagcaca	agaagagccg	cgaggtccgc	aagagccgcg	acgcccggctc	gcagaaggac	5340
tccgacgaca	gtgcggccac	cccgacggac	gagacggtcg	aggagagagg	ccggaacgag	5400
ggcctgagca	goggtactct	ctccaaatcc	tcctcctcgg	ggatgcagag	ctgtggagaa	5460
gagggaaggcg	aggagggggc	cgacgcctgt	cccctgcgcg	cacccatggc	catccagcag	5520
cacagcctcc	tccagccaga	ctcacaggat	gacaaggcct	cttctcggtt	attagtcccg	5580
cccaccagct	cggaaacacc	gagtgacggc	gagctcgtca	gtgcaattga	ggaactcgtg	5640
aaaagcaaga	tggcactgga	ggatcgcccc	agctcactcc	ttggtgacca	gggagatagt	5700
agcagccctt	ccttcaaccc	ttcggataat	tcccttctct	cttctcctcc	gcccattgat	5760
gagatggaag	aaaggaaatc	cagctcttta	aagagaagac	actacgtttt	gcaagaacta	5820
gtggagacag	agcgtgacta	tgtgcgggac	cttggctatg	tgggtgaggg	ctacatggca	5880
cttatgaaag	aagatgggtg	tcctgatgac	atgaaaggaa	aagacaaaat	tgtgttcggc	5940
aacatccatc	agatttacga	ctggcacaga	gacttttttt	taggagagtt	agagaagtgc	6000
cttgaagatc	cagaaaaact	aggatccctt	tttgttaaac	acgagagaag	gttgccatg	6060
tacatagctt	attgtcaaaa	taaaccacaa	tctgagcaca	ttgtctcaga	atacattgat	6120
accttttttg	aggacttaaa	gcagcgtctt	ggccacaggt	tacagctcac	agatctgttg	6180
atcaaaaccag	tgcagagaat	catgaagtat	cagctgttac	tgaaggactt	cctcaagtat	6240
tccaaaaagg	ccagcctgga	tacatcagaa	ttagagagag	ctgtggaagt	catgtgcata	6300
gtacccaggc	ggtgcaacga	catgatgaac	gtggggcggc	tgcaggattt	cgaaggga	6360
atcgttgccc	agggtaaaat	gctcttgacg	gacacattct	tggtcacaga	ccaagatgca	6420
ggacttctgc	ctcgctgcag	agagaggcgc	atcttctctt	ttgagcagat	cgtcatatcc	6480
agcgaaccac	ttgataaaaa	gaagggtctt	tccatgcggg	gattcctgtt	taagaacagt	6540
atcaagggtg	gttgcccttg	cctggaggaa	aatgtggaaa	atgatccctg	taaatttgct	6600
ctgacatcga	ggacgggtga	cgtggtagag	accttcattt	tgcattcatt	tagtccaagt	6660
gtccggcaaa	cttggatcca	tgaatcaaac	caaattttag	aaaaccagcg	caatttttta	6720
aatgccttga	catcgccaat	cgagtaaccg	aggaaacaca	gcggggggcg	cggcgggcgg	6780
ggcagcgggg	cagcggcggg	ggtgggggca	gcggcgcgcg	cggggccccc	agtggcggca	6840
gcggccacag	tggcgggccc	agcagctgcg	gcggcgcccc	cagcagcagc	aggagccggc	6900
cctcccggtg	ccccagcctt	gtccgacacc	acccccccgt	gtggtctctc	tctgcagcct	6960
cgagccaggc	agaggcagac	agatgtcagc	agtgaagca	gcagcagtag	caacatctcc	7020
accatgtgtg	tgacacacga	ttacacggca	gtgaaggagg	atgagatcaa	cgtctaccaa	7080

10

20

30

```

ggagagggtcg ttcaaattct ggccagcaac cagcagaaca tgtttctggt gttccgagcc 7140
gccactgacc agtgcctccg agctgagggc tggattccag gctttgtcct gggccacacc 7200
agtgcagtc tegtggagaa cccggacggg actctcaaga agtcaacatc ttggcacaca 7260
gcactccgtt taaggaaaaa atctgagaaa aaagataaag acggcaaaag ggaaggcaag 7320
ttagagaacg gttatcggaa gtcacgggaa ggactcagca acaagggtatc tgtgaagctt 7380
ctcaatccca actacattta tgacgttccc ccagaattcg tcattccatt gagtgaaggtc 7440
acgtgtgaga caggggagac cgttcttctt agatgtcgag tctgtggccg ccccaaagcc 7500
tcaattacct ggaagggccc tgaacacaac accttgaaca acgatggtea ctacagcatc 7560
tcctacagtg acctgggaga ggccacgctg aagattgtgg gcgtgaccac ggaagatgac 7620
ggcatctaca cgtgcacgcg tgtcaatgac atgggttcag cctcatcatc ggccagcctg 7680
agggctcctag gtccagggat ggatgggac atgggtgacct ggaaagacaa ctttgactcc 7740
ttctacagtg aagtggctga gcttggcagg ggcagattct ctgtcgtaa gaaatgtgat 7800
cagaaggaa ccaagcgagc agtggccact aagtttgtga acaagaagtt gatgaagcgc 7860
gaccaggtca cccatgagct tggcatcctg cagagcctcc agcaccocct gcttgcggc 7920
ctcctcgaca cctttgagac cccaccagc tacatcctgg tcttagaat ggctgaccag 7980
ggctgcctcc tggactgcgt ggtgcgatgg ggaagcctca ctgaaggaa gatcaggcg 8040
cacctggggg aggttctgga agctgtccgg tacctgcaca actgcaggat agcacacctg 8100
gacctaaagc ctgagaatat cctgttggtg gagagttag ccaagccaac catcaaactg 8160
gctgactttg gagatgctgt tcagctcaac acgacctact acatccacca gttactgggg 8220
aaccctgaat tcgcagcccc tgaaatcatc ctcggaacc ctgtctccct gacctcggt 8280
acgtggagtg ttggagtgc cactacgta cttcttagtg gcgtgtccc ctctctggat 8340
gacagtgtgg aagagacctg cctgaacatt tgccgcttag actttagctt cccagatgac 8400
tactttaaag gagtgaagca gaaggccaag gagttcgtgt gcttctcct gcaggaggac 8460
cccgccaagc gtccctcggc tgcgtggcc ctccaggagc agtggctgca ggccggcaac 8520
ggcagaagca cgggcgtcct cgacacgtcc agactgactt ccttcattga gggcgcaaa 8580
caccagaatg atgttcgacc tatccgtagc attaaaaact ttctgcagag caggcttctg 8640
cctagagttt gacctatcca gaagttctt ctcatctct ttcacctgcc aatcagctgt 8700
taatctgaat tttcaagaga aaacaagcaa acataactga tcagctgccc gtatgttcat 8760
cgtgtgaaat tgcattccaa gtgagctgtg ctccagagtg cttggacaca gagctgcaag 8820
ctgcgtggg gtggaggacc gtcacttaca ctctgccaa gacggaggtc gcattgctgt 8880
atcacagtat tttttacgga tttctg 8906

```

```

<210> 665
<211> 1204
<212> DNA
<213> Homo sapiens

```

```

<400> 665
tcggcggcgg tggatcggc ggccagctgtg aggggggtcc gggaagatgg tgctgatcaa 60
ggaattccgt gtggttttgc catgttctgt tcaggagtat cagggtgggc agctttactc 120
tggtgcagaa gctagtaaga atgagactgg tgggtggagaa ggaattgaag tcttaagaa 180
tgaaccttat gagaaggatg gaaaaaggg acagtatacg caaaaattt atcacctaaa 240
gagcaaagtg cctgcattcg tgaggatgat tgctcccgag ggctccttgg tgtttcatga 300
gaaagcctgg aatgcgtacc cctactgtag aacaattgta acgaatgaat atatgaaaga 360
tgatttcttc attaaaatcg aaacatggca caaccagac ttgggaacat tagaaaatgt 420
acatggttta gatccaaaca catggaaaac tgttgaatt gtccatatag atattgcaga 480
tagaagtcaa gttgaaccag cagactacaa agctgatgaa gaccagcat tattccagtc 540
agtcaagacc aagagaggcc ctttgggacc caactggaag aaggagctgg caaacagccc 600
tgactgtccc cagatgtgtg cctataagct ggtgaccatc aaattcaagt ggtggggact 660

```

10

20

30


```

gcaaagcaaa gtagaaaact tcattcaaaa gcaagaaaaa cggatattta caaacttcca 720
tcgccagctt tttgttggg ttgacaagtg gatcgatctc acgatggaag acattaggag 780
aatggaagac gagactcaga aagaactaga aacaatgcgt aagaggggtt ccgttcgagg 840
cacgtcggct gctgatgtct agatgagtc cctgtagggt cagagacaat gtcaaactgt 900
ttacgtaatc aaggtaagt gaggggaaca agcgcagcca gtgatgagtg aacaacaatc 960
tgaccagtat cttgcagtgt tgacgtttcc cagatgtgtg cttgtgatga tacacacaca 1020
tgcacaggtt ctcaaccacg tgtgtatata tgtatgtgtg catatgtctg tagctgtata 1080
taaagcgcat gtagagctac agatccagat acacacactt gtgtatatat gtacatacag 1140
acatactgaa gggattagta caatttctcc aaagtactgt acctatcttc agcaagaatg 1200
caaa 1204

```

<210> 666

<211> 892

<212> PRT

<213> Homo sapiens

10

<400> 666

```

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

```

20

30

30

```

        660                665                670
Gln Tyr Glu Lys Ser Ile Val Asn Tyr Lys Pro Lys Ile Asp Gln Leu
    675                680                685
Glu Gly Asp His Gln Leu Ile Gln Glu Ala Leu Ile Phe Asp Asn Lys
    690                695                700
His Thr Asn Tyr Thr Met Glu His Ile Arg Val Gly Trp Glu Gln Leu
    705                710                715                720
Leu Thr Thr Ile Ala Arg Thr Ile Asn Glu Val Glu Asn Gln Ile Leu
    725                730                735
Thr Arg Asp Ala Lys Gly Ile Ser Gln Glu Gln Met Asn Glu Phe Arg
    740                745                750
Ala Ser Phe Asn His Phe Asp Arg Asp His Ser Gly Thr Leu Gly Pro
    755                760                765
Glu Glu Phe Lys Ala Cys Leu Ile Ser Leu Gly Tyr Asp Ile Gly Asn
    770                775                780
Asp Pro Gln Gly Glu Ala Glu Phe Ala Arg Ile Met Ser Ile Val Asp
    785                790                795                800
Pro Asn Arg Leu Gly Val Val Thr Phe Gln Ala Phe Ile Asp Phe Met
    805                810                815
Ser Arg Glu Thr Ala Asp Thr Asp Thr Ala Asp Gln Val Met Ala Ser
    820                825                830
Phe Lys Ile Leu Ala Gly Asp Lys Asn Tyr Ile Thr Met Asp Glu Leu
    835                840                845
Arg Arg Glu Leu Pro Pro Asp Gln Ala Glu Tyr Cys Ile Ala Arg Met
    850                855                860
Ala Pro Tyr Thr Gly Pro Asp Ser Val Pro Gly Ala Leu Asp Tyr Met
    865                870                875                880
Ser Phe Ser Thr Ala Leu Tyr Gly Glu Ser Asp Leu
    885                890

```

```

<210> 667
<211> 197
<212> PRT
<213> Homo sapiens

```

```

<400> 667
Met Met Phe Pro Gln Ser Arg His Ser Gly Ser Ser His Leu Pro Gln
  1                5                10                15
Gln Leu Lys Phe Thr Thr Ser Asp Ser Cys Asp Arg Ile Lys Asp Glu
  20                25                30
Phe Gln Leu Leu Gln Ala Gln Tyr His Ser Leu Lys Leu Glu Cys Asp
  35                40                45
Lys Leu Ala Ser Glu Lys Ser Glu Met Gln Arg His Tyr Val Met Tyr
  50                55                60
Tyr Glu Met Ser Tyr Gly Leu Asn Ile Glu Met His Lys Gln Ala Glu
  65                70                75                80
Ile Val Lys Arg Leu Asn Gly Ile Cys Ala Gln Val Leu Pro Tyr Leu
  85                90                95
Ser Gln Glu His Gln Gln Gln Val Leu Gly Ala Ile Glu Arg Ala Lys
  100                105                110

```

10

20

30

Gln Val Thr Ala Pro Glu Leu Asn Ser Ile Ile Arg Gln Gln Leu Gln
 115 120 125
 Ala His Gln Leu Ser Gln Leu Gln Ala Leu Ala Leu Pro Leu Thr Pro
 130 135 140
 Leu Pro Val Gly Leu Gln Pro Pro Ser Leu Pro Ala Val Ser Ala Gly
 145 150 155 160
 Thr Gly Leu Leu Ser Leu Ser Ala Leu Gly Ser Gln Ala His Leu Ser
 165 170 175
 Lys Glu Asp Lys Asn Gly His Asp Gly Asp Thr His Gln Glu Asp Asp
 180 185 190
 Gly Glu Lys Ser Asp
 195

<210> 668
 <211> 739
 <212> PRT
 <213> Homo sapiens

10

<400> 668
 Gly Asp Lys Glu Pro Thr Glu Thr Ile Gly Asp Leu Ser Ile Cys Leu
 1 5 10 15
 Asp Gly Leu Gln Leu Glu Ser Glu Val Val Thr Asn Gly Glu Thr Thr
 20 25 30
 Cys Ser Glu Ser Ala Ser Gln Asn Asp Asp Gly Ser Arg Ser Lys Asp
 35 40 45
 Glu Thr Arg Val Ser Thr Asn Gly Ser Asp Asp Pro Glu Asp Ala Gly
 50 55 60
 Ala Gly Glu Asn Arg Arg Val Ser Gly Asn Asn Ser Pro Ser Leu Ser
 65 70 75 80
 Asn Gly Gly Phe Lys Pro Ser Arg Pro Pro Arg Pro Ser Arg Pro Pro
 85 90 95
 Pro Pro Thr Pro Arg Arg Pro Ala Ser Val Asn Gly Ser Pro Ser Ala
 100 105 110
 Thr Ser Glu Ser Asp Gly Ser Ser Thr Gly Ser Leu Pro Pro Thr Asn
 115 120 125
 Thr Asn Thr Asn Thr Ser Glu Gly Ala Thr Ser Gly Leu Ile Ile Pro
 130 135 140
 Leu Thr Ile Ser Gly Gly Ser Gly Pro Arg Pro Leu Asn Pro Val Thr
 145 150 155 160
 Gln Ala Pro Leu Pro Pro Gly Trp Glu Gln Arg Val Asp Gln His Gly
 165 170 175
 Arg Val Tyr Tyr Val Asp His Val Glu Lys Arg Thr Thr Trp Asp Arg
 180 185 190
 Pro Glu Pro Leu Pro Pro Gly Trp Glu Arg Arg Val Asp Asn Met Gly
 195 200 205
 Arg Ile Tyr Tyr Val Asp His Phe Thr Arg Thr Thr Thr Trp Gln Arg
 210 215 220
 Pro Thr Leu Glu Ser Val Arg Asn Tyr Glu Gln Trp Gln Leu Gln Arg
 225 230 235 240
 Ser Gln Leu Gln Gly Ala Met Gln Gln Phe Asn Gln Arg Phe Ile Tyr

20

30

30

```

        645                650                655
Asn Glu Lys Arg Met Arg Leu Leu Gln Phe Val Thr Gly Thr Cys Arg
        660                665                670
Leu Pro Val Gly Gly Phe Ala Asp Leu Met Gly Ser Asn Gly Pro Gln
        675                680                685
Lys Phe Cys Ile Glu Lys Val Gly Lys Glu Asn Trp Leu Pro Arg Ser
        690                695                700
His Thr Cys Phe Asn Arg Leu Asp Leu Pro Pro Tyr Lys Ser Tyr Glu
705                710                715                720
Gln Leu Lys Glu Lys Leu Leu Phe Ala Ile Glu Glu Thr Glu Gly Phe
        725                730                735
Gly Gln Glu

```

```

<210> 669
<211> 431
<212> PRT
<213> Homo sapiens

```

```

<400> 669
Gly Pro Pro Pro Thr Arg Ala Leu Pro Leu Pro Gln Ser Leu Pro Pro
 1                5                10                15
Asp Phe Arg Leu Glu Pro Thr Ala Pro Ala Leu Ser Pro Arg Ser Ser
 20                25                30
Phe Ala Ser Ser Ser Ala Ser Asp Ala Ser Lys Pro Ser Ser Pro Arg
 35                40                45
Gly Ser Leu Leu Leu Asp Gly Ala Gly Ala Gly Ala Gly Gly Ser
 50                55                60
Arg Pro Cys Ser Asn Arg Thr Ser Gly Ile Ser Met Gly Tyr Asp Gln
 65                70                75                80
Arg His Gly Ser Pro Leu Pro Ala Gly Pro Cys Leu Phe Gly Pro Pro
 85                90                95
Leu Ala Gly Ala Pro Ala Gly Tyr Ser Pro Gly Gly Val Pro Ser Ala
100                105                110
Tyr Pro Glu Leu His Ala Ala Leu Asp Arg Leu Tyr Ala Gln Arg Pro
115                120                125
Ala Gly Phe Gly Cys Gln Glu Ser Arg His Ser Tyr Pro Pro Ala Leu
130                135                140
Gly Ser Pro Gly Ala Leu Ala Gly Ala Arg Val Gly Ala Ala Gly Pro
145                150                155                160
Leu Glu Arg Arg Gly Ala Gln Pro Gly Arg His Ser Val Thr Gly Tyr
165                170                175
Gly Asp Cys Ala Val Gly Ala Arg Tyr Gln Asp Glu Leu Thr Ala Leu
180                185                190
Leu Arg Leu Thr Val Gly Thr Gly Gly Arg Glu Ala Gly Ala Arg Gly
195                200                205
Glu Pro Ser Gly Ile Glu Pro Ser Gly Leu Glu Glu Pro Pro Gly Pro
210                215                220
Phe Val Pro Glu Ala Ala Arg Ala Arg Met Arg Glu Pro Glu Ala Arg
225                230                235                240

```

10

20

30

10

20

30

```

145          150          155          160
Phe Arg Lys Leu Glu Pro Cys Asp Ser Asn Glu Glu Lys Arg Leu Lys
165          170          175
Leu Ala Arg Ala Ile Tyr Arg Lys Tyr Ile Leu Asp Asn Asn Gly Ile
180          185          190
Val Ser Arg Gln Thr Lys Pro Ala Thr Lys Ser Phe Ile Lys Gly Cys
195          200          205
Ile Met Lys Gln Leu Ile Asp Pro Ala Met Phe Asp Gln Ala Gln Thr
210          215          220
Glu Ile Gln Ala Thr Met Glu Glu Asn Thr Tyr Pro Ser Phe Leu Lys
225          230          235          240
Ser Asp Ile Tyr Leu Glu Tyr Thr Arg Thr Gly Ser Glu Ser Pro Lys
245          250          255
Val Cys Ser Asp Gln Ser Ser Gly Ser Gly Thr Gly Lys Gly Ile Ser
260          265          270
Gly Tyr Leu Pro Thr Leu Asn Glu Asp Glu Glu Trp Lys Cys Asp Gln
275          280          285
Asp Met Asp Glu Asp Asp Gly Arg Asp Ala Ala Pro Pro Gly Arg Leu
290          295          300
Pro Gln Lys Leu Leu Leu Glu Thr Ala Ala Pro Arg Val Ser Ser Ser
305          310          315          320
Arg Arg Tyr Ser Glu Gly Arg Glu Phe Arg Tyr Gly Ser Trp Arg Glu
325          330          335
Pro Val Asn Pro Tyr Tyr Val Asn Ala Gly Tyr Ala Leu Ala Pro Ala
340          345          350
Thr Ser Ala Asn Asp Ser Glu Gln Gln Ser Leu Ser Ser Asp Ala Asp
355          360          365
Thr Leu Ser Leu Thr Asp Ser Ser Val Asp Gly Ile Pro Pro Tyr Arg
370          375          380
Ile Arg Lys Gln His Arg Arg Glu Met Gln Glu Ser Ala Gln Val Asn
385          390          395          400
Gly Arg Val Pro Leu Pro His Ile Pro Arg Thr Tyr Arg Val Pro Lys
405          410          415
Glu Val Arg Val Glu Pro Gln Lys Phe Ala Glu Glu Leu Ile His Arg
420          425          430
Leu Glu Ala Val Gln Arg Thr Arg Glu Ala Glu Glu Lys Leu Glu Glu
435          440          445
Arg Leu Lys Arg Val Arg Met Glu Glu Glu Gly Glu Asp Gly Asp Pro
450          455          460
Ser Ser Gly Pro Pro Gly Pro Cys His Lys Leu Pro Pro Ala Pro Ala
465          470          475          480
Trp His His Phe Pro Pro Arg Leu Cys Trp Thr Trp Ala Cys Ala Gly
485          490          495
Leu Arg Asp Ala His Glu Glu Asn Pro Glu Ser Ile Leu Asp Glu His
500          505          510
Val Gln Arg Val Leu Arg Thr Thr Gly Arg Gln Ser Pro Gly Pro Gly
515          520          525
His Arg Ser Pro Asp Ser Gly His Val Ala Lys Met Pro Val Ala Leu
530          535          540
Gly Gly Ala Ala Ser Gly His Gly Lys His Val Pro Lys Ser Gly Ala

```

10

20

30

545 550 555 560
 Lys Leu Asp Ala Ala Gly Leu His His His Arg His Val His His His
 565 570 575
 Val His His Ser Thr Ala Arg Pro Lys Glu Gln Val Glu Ala Glu Ala
 580 585 590
 Thr Arg Arg Ala Gln Ser Ser Phe Ala Trp Gly Leu Glu Pro His Ser
 595 600 605
 His Gly Ala Arg Ser Arg Gly Tyr Ser Glu Ser Val Gly Ala Ala Pro
 610 615 620
 Asn Ala Ser Asp Gly Leu Ala His Ser Gly Lys Val Gly Val Ala Cys
 625 630 635 640
 Lys Arg Asn Ala Lys Lys Ala Glu Ser Gly Lys Ser Ala Ser Thr Glu
 645 650 655
 Val Pro Gly Ala Ser Glu Asp Ala Glu Lys Asn Gln Lys Ile Met Gln
 660 665 670
 Trp Ile Ile Glu Gly Glu Lys Glu Ile Ser Arg His Arg Arg Thr Gly
 675 680 685
 His Gly Ser Ser Gly Thr Arg Lys Pro Gln Pro His Glu Asn Ser Arg
 690 695 700
 Pro Leu Ser Leu Glu His Pro Trp Ala Gly Pro Gln Leu Arg Thr Ser
 705 710 715 720
 Val Gln Pro Ser His Leu Phe Ile Gln Asp Pro Thr Met Pro Pro His
 725 730 735
 Pro Ala Pro Asn Pro Leu Thr Gln Leu Glu Glu Ala Arg Arg Arg Leu
 740 745 750
 Glu Glu Glu Glu Lys Arg Ala Ser Arg Ala Pro Ser Lys Gln Arg Tyr
 755 760 765
 Val Gln Glu Val Met Arg Arg Gly Arg Ala Cys Val Arg Pro Ala Cys
 770 775 780
 Ala Pro Val Leu His Val Val Pro Ala Val Ser Asp Met Glu Leu Ser
 785 790 795 800
 Glu Thr Glu Thr Arg Ser Gln Arg Lys Val Gly Gly Gly Ser Ala Gln
 805 810 815
 Pro Cys Asp Ser Ile Val Val Ala Tyr Tyr Phe Cys Gly Glu Pro Ile
 820 825 830
 Pro Tyr Arg Thr Leu Val Arg Gly Arg Ala Val Thr Leu Gly Gln Phe
 835 840 845
 Lys Glu Leu Leu Thr Lys Lys Gly Ser Tyr Arg Tyr Tyr Phe Lys Lys
 850 855 860
 Val Ser Asp Glu Phe Asp Cys Gly Val Val Phe Glu Glu Val Arg Glu
 865 870 875 880
 Asp Glu Ala Val Leu Pro Val Phe Glu Glu Lys Ile Ile Gly Lys Val
 885 890 895
 Glu Lys Val Asp
 900

10

20

30

<210> 671
 <211> 591
 <212> PRT
 <213> Homo sapiens

<400> 671

```

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

```

10

20

30

```

385          390          395          400
Gln Thr Tyr Glu Ile Leu Lys Met Pro Phe Tyr Cys Leu Tyr Tyr Tyr
          405          410          415
Arg Arg Ala His Gln Leu Arg Pro Asn Asp Ser Arg Met Leu Val Ala
          420          425          430
Leu Gly Glu Cys Tyr Glu Lys Leu Asn Gln Leu Val Glu Ala Lys Lys
          435          440          445
Cys Tyr Trp Arg Ala Tyr Ala Val Gly Asp Val Glu Lys Met Ala Leu
          450          455          460
Val Lys Leu Ala Lys Leu His Glu Gln Leu Thr Glu Ser Glu Gln Ala
          465          470          475          480
Ala Gln Cys Tyr Ile Lys Tyr Ile Gln Asp Ile Tyr Ser Cys Gly Glu
          485          490          495
Ile Val Glu His Leu Glu Glu Ser Thr Ala Phe Arg Tyr Leu Ala Gln
          500          505          510
Tyr Tyr Phe Lys Cys Lys Leu Trp Asp Glu Ala Ser Thr Cys Ala Gln
          515          520          525
Lys Cys Cys Ala Phe Asn Asp Thr Arg Glu Glu Gly Lys Ala Leu Leu
          530          535          540
Arg Gln Ile Leu Gln Leu Arg Asn Gln Gly Glu Thr Pro Thr Thr Glu
          545          550          555          560
Val Pro Ala Pro Phe Phe Leu Pro Ala Ser Leu Ser Ala Asn Asn Thr
          565          570          575
Pro Thr Arg Arg Val Ser Pro Leu Asn Leu Ser Ser Val Thr Pro
          580          585          590

```

<210> 672

<211> 914

<212> PRT

<213> Homo sapiens

<400> 672

```

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

```

10

20

30

Asp Lys Ala Asp Glu Leu Tyr Asp Arg Val Asp Gly Leu Leu His Gln
 145 150 155 160
 Leu Thr Leu Gln Ser Asn Gln Arg Ile Gln Ala Leu Glu Leu Val Gln
 165 170 175
 Thr Leu Glu Ala Arg Glu Ser Gly Leu His Gln Ile Glu Val Trp Leu
 180 185 190
 Gln Gln Val Gly Trp Pro Ala Leu Glu Glu Ala Gly Glu Pro Ser Leu
 195 200 205
 Asp Met Leu Leu Gln Ala Gln Gly Ser Phe Gln Glu Leu Tyr Gln Val
 210 215 220
 Ala Gln Glu Gln Val Arg Gln Gly Glu Lys Phe Leu Gln Pro Leu Thr
 225 230 235 240
 Gly Trp Glu Ala Ala Glu Leu Asp Pro Pro Gly Ala Arg Phe Leu Ala
 245 250 255
 Leu Arg Ala Gln Leu Thr Glu Phe Ser Arg Ala Leu Ala Gln Arg Cys
 260 265 270
 Gln Arg Leu Ala Asp Ala Glu Arg Leu Phe Gln Leu Phe Arg Glu Ala
 275 280 285
 Leu Thr Trp Ala Glu Glu Gly Gln Arg Val Leu Ala Glu Leu Glu Gln
 290 295 300
 Glu Arg Pro Gly Val Val Leu Gln Gln Leu Gln Leu His Trp Thr Arg
 305 310 315 320
 His Pro Asp Leu Pro Pro Ala His Phe Arg Lys Met Trp Ala Leu Ala
 325 330 335
 Thr Gly Leu Gly Ser Glu Ala Ile Arg Gln Glu Cys Arg Trp Ala Trp
 340 345 350
 Ala Arg Cys Gln Asp Thr Trp Leu Ala Leu Asp Gln Lys Leu Glu Ala
 355 360 365
 Ser Leu Lys Leu Pro Pro Val Gly Ser Thr Ala Ser Leu Cys Val Ser
 370 375 380
 Gln Val Pro Ala Ala Pro Ala His Pro Pro Leu Arg Lys Ala Tyr Ser
 385 390 395 400
 Phe Asp Arg Asn Leu Gly Gln Ser Leu Ser Glu Pro Ala Cys His Cys
 405 410 415
 His His Ala Ala Thr Ile Ala Ala Cys Arg Arg Pro Glu Ala Gly Gly
 420 425 430
 Gly Ala Leu Pro Gln Ala Ser Pro Thr Val Pro Pro Pro Gly Ser Ser
 435 440 445
 Asp Pro Arg Ser Leu Asn Arg Leu Gln Leu Val Leu Ala Glu Met Val
 450 455 460
 Ala Thr Glu Arg Glu Tyr Val Arg Ala Leu Glu Tyr Thr Met Glu Asn
 465 470 475 480
 Tyr Phe Pro Glu Leu Asp Arg Pro Asp Val Pro Gln Gly Leu Arg Gly
 485 490 495
 Gln Arg Ala His Leu Phe Gly Asn Leu Glu Lys Leu Arg Asp Phe His
 500 505 510
 Cys His Phe Phe Leu Arg Glu Leu Glu Ala Cys Thr Arg His Pro Pro
 515 520 525
 Arg Val Ala Tyr Ala Phe Leu Arg His Arg Val Gln Phe Gly Met Tyr
 530 535 540

10

20

30

Ala Leu Tyr Ser Lys Asn Lys Pro Arg Ser Asp Ala Leu Met Ser Ser
 545 550 555 560
 Tyr Gly His Thr Phe Lys Asp Lys Gln Gln Ala Leu Gly Asp His
 565 570 575
 Leu Asp Leu Ala Ser Tyr Leu Leu Lys Pro Ile Gln Arg Met Gly Lys
 580 585 590
 Tyr Ala Leu Leu Leu Gln Glu Leu Ala Arg Ala Cys Gly Gly Pro Thr
 595 600 605
 Gln Glu Leu Ser Ala Leu Arg Glu Ala Gln Ser Leu Val His Phe Gln
 610 615 620
 Leu Arg His Gly Asn Asp Leu Leu Ala Met Asp Ala Ile Gln Gly Cys
 625 630 635 640
 Asp Val Asn Leu Lys Glu Gln Gly Gln Leu Val Arg Gln Asp Glu Phe
 645 650 655
 Val Val Arg Thr Gly Arg His Lys Ser Val Arg Arg Ile Phe Leu Phe
 660 665 670
 Glu Glu Leu Leu Leu Phe Ser Lys Pro Arg His Gly Pro Thr Gly Val
 675 680 685
 Asp Thr Phe Ala Tyr Lys Arg Ser Phe Lys Met Ala Asp Leu Gly Leu
 690 695 700
 Thr Glu Cys Cys Gly Asn Ser Asn Leu Arg Phe Glu Ile Trp Phe Arg
 705 710 715 720
 Arg Arg Lys Ala Arg Asp Thr Phe Val Leu Gln Ala Ser Ser Leu Ala
 725 730 735
 Ile Lys Gln Ala Trp Thr Ala Asp Ile Ser His Leu Leu Trp Arg Gln
 740 745 750
 Ala Val His Asn Lys Glu Val Arg Met Ala Glu Met Val Ser Met Gly
 755 760 765
 Val Gly Asn Lys Ala Phe Arg Asp Ile Ala Pro Ser Glu Glu Ala Ile
 770 775 780
 Asn Asp Arg Thr Val Asn Tyr Val Leu Lys Cys Arg Glu Val Arg Ser
 785 790 795 800
 Arg Ala Ser Ile Ala Val Ala Pro Phe Asp His Asp Ser Leu Tyr Leu
 805 810 815
 Gly Ala Ser Asn Ser Leu Pro Gly Asp Pro Ala Ser Cys Ser Val Leu
 820 825 830
 Gly Ser Leu Asn Leu His Leu Tyr Arg Asp Pro Ala Leu Leu Gly Leu
 835 840 845
 Arg Cys Pro Leu Tyr Pro Ser Phe Leu Glu Glu Ala Ala Leu Glu Ala
 850 855 860
 Glu Ala Glu Leu Gly Gly Gln Pro Ser Leu Thr Ala Glu Asp Ser Glu
 865 870 875 880
 Ile Ser Ser Gln Cys Pro Ser Ala Ser Gly Ser Ser Gly Ser Asp Ser
 885 890 895
 Ser Cys Val Ser Gly Gln Ala Leu Gly Arg Gly Leu Glu Asp Leu Pro
 900 905 910
 Cys Val

10

20

30

<211> 277
 <212> PRT
 <213> Homo sapiens

<400> 673

```

Leu Asn Tyr Leu Leu Glu Ser Arg Leu Glu Ala Ala Ala His Cys Ala
 1          5          10          15
Leu Lys Gln Gly Ile Ala Thr Ala Ser Leu Leu Pro Ala Gln Leu Gln
 20          25          30
Pro Ala Val Leu Thr Val Val Thr Cys His Val Val Val Ser Val His
 35          40          45
Gly His His Thr Asp Gly Cys Leu Ala Ala Leu Cys Arg Glu Asp Arg
 50          55          60
Thr Gly Thr Gly Gly Ala Phe Trp Cys Lys Asn Arg Val Ile Val Ser
 65          70          75          80
His Ala Val Asp Val Val Leu His Val His Gly Glu Gly Asn Pro Val
 85          90          95
Gln Ala Leu Ile Ala His Gly Ala Pro Glu Ala Ala Trp Val Val Gly
 100         105         110
Leu Ala Gln Gly Leu Gln Asp His Phe His Asp Glu Met Ser Thr His
 115         120         125
Ala Ala Phe Val Gly Arg Leu Leu Glu Pro Gly Val Gln Glu Val Leu
 130         135         140
Leu Ala Val His Phe Leu Thr His Val Val Glu Arg Leu Pro Thr Glu
 145         150         155         160
Ser Ser Pro Thr Arg Val Ala Gly Glu Ala Val Ser Val Ile Lys Thr
 165         170         175
Pro His Cys Leu Ala Arg Leu Leu Gly Ser Val Asp Ala Lys Pro Thr
 180         185         190
Leu Asp Ala Asn Ala Glu Val Val Pro Arg Arg Ala Arg Leu Glu Arg
 195         200         205
Pro Leu Gln Leu Pro Gly Glu Arg Leu Gln Pro Pro Leu Gly Arg Ala
 210         215         220
Trp Ala Ala Leu Pro Ala Arg Gly Gln Arg Glu Cys Arg Gln Arg Glu
 225         230         235         240
Gly Gly Arg Pro Arg Arg Leu Arg Gly Ala Ser Gly Arg Gly Ala Gly
 245         250         255
Ala Gly Arg Glu Glu Val Ser Val Gly Phe Ser Ala Gln Trp Glu Phe
 260         265         270
Gly Ser Gly Arg His
 275

```

<210> 674
 <211> 1120
 <212> PRT
 <213> Homo sapiens

<400> 674

```

Met Trp Arg Val Lys Lys Leu Ser Leu Ser Leu Ser Pro Ser Pro Gln
 1          5          10          15

```

10

20

30

Thr Gly Lys Pro Ser Met Arg Thr Pro Leu Arg Glu Leu Thr Leu Gln
 20 25 30
 Pro Gly Ala Leu Thr Thr Ser Gly Lys Arg Ser Pro Ala Cys Ser Ser
 35 40 45
 Leu Thr Pro Ser Leu Cys Lys Leu Gly Leu Gln Glu Gly Ser Asn Asn
 50 55 60
 Ser Ser Pro Val Asp Phe Val Asn Asn Lys Arg Thr Asp Leu Ser Ser
 65 70 75 80
 Glu His Phe Ser His Ser Ser Lys Trp Leu Glu Thr Cys Gln His Glu
 85 90 95
 Ser Asp Glu Gln Pro Leu Asp Pro Ile Pro Gln Ile Ser Ser Thr Pro
 100 105 110
 Lys Thr Ser Glu Glu Ala Val Asp Pro Leu Gly Asn Tyr Met Val Lys
 115 120 125
 Thr Ile Val Leu Val Pro Ser Pro Leu Gly Gln Gln Gln Asp Met Ile
 130 135 140
 Phe Glu Ala Arg Leu Asp Thr Met Ala Glu Thr Asn Ser Ile Ser Leu
 145 150 155 160
 Asn Gly Pro Leu Arg Thr Asp Asp Leu Val Arg Glu Glu Val Ala Pro
 165 170 175
 Cys Met Gly Asp Arg Phe Ser Glu Val Ala Ala Val Ser Glu Lys Pro
 180 185 190
 Ile Phe Gln Glu Ser Pro Ser His Leu Leu Glu Glu Ser Pro Pro Asn
 195 200 205
 Pro Cys Ser Glu Gln Leu His Cys Ser Lys Glu Ser Leu Ser Ser Arg
 210 215 220
 Thr Glu Ala Val Arg Glu Asp Leu Val Pro Ser Glu Ser Asn Ala Phe
 225 230 235 240
 Leu Pro Ser Ser Val Leu Trp Leu Ser Pro Ser Thr Ala Leu Ala Ala
 245 250 255
 Asp Phe Arg Val Asn His Val Asp Pro Glu Glu Glu Ile Val Glu His
 260 265 270
 Gly Ala Met Glu Glu Arg Glu Met Arg Phe Pro Thr His Pro Lys Glu
 275 280 285
 Ser Glu Thr Glu Asp Gln Ala Leu Val Ser Ser Val Glu Asp Ile Leu
 290 295 300
 Ser Thr Cys Leu Thr Pro Asn Leu Val Glu Met Glu Ser Gln Glu Ala
 305 310 315 320
 Pro Gly Pro Ala Val Glu Asp Val Gly Arg Ile Leu Gly Ser Asp Thr
 325 330 335
 Glu Ser Trp Met Ser Pro Leu Ala Trp Leu Glu Lys Gly Val Asn Thr
 340 345 350
 Ser Val Met Leu Glu Asn Leu Arg Gln Ser Leu Ser Leu Pro Ser Met
 355 360 365
 Leu Arg Asp Ala Ala Ile Gly Thr Thr Pro Phe Ser Thr Cys Ser Val
 370 375 380
 Gly Thr Trp Phe Thr Pro Ser Ala Pro Gln Glu Lys Ser Thr Asn Thr
 385 390 395 400
 Ser Gln Thr Gly Leu Val Gly Thr Lys His Ser Thr Ser Glu Thr Glu
 405 410 415

10

20

30

Gln Leu Leu Cys Gly Arg Pro Pro Asp Leu Thr Ala Leu Ser Arg His
 420 425 430
 Asp Leu Glu Asp Asn Leu Leu Ser Ser Leu Val Ile Val Glu Phe Leu
 435 440 445
 Ser Arg Gln Leu Arg Asp Trp Lys Ser Gln Leu Ala Val Pro His Pro
 450 455 460
 Glu Thr Gln Asp Ser Ser Thr Gln Thr Asp Thr Ser His Ser Gly Ile
 465 470 475 480
 Thr Asn Lys Leu Gln His Leu Lys Glu Ser His Glu Met Gly Gln Ala
 485 490 495
 Leu Gln Gln Ala Arg Asn Val Met Gln Ser Trp Val Leu Ile Ser Lys
 500 505 510
 Glu Leu Ile Ser Leu Leu His Leu Ser Leu Leu His Leu Glu Glu Asp
 515 520 525
 Lys Thr Thr Val Asn Gln Glu Ser Arg Arg Ala Glu Thr Leu Val Cys
 530 535 540
 Cys Cys Phe Asp Leu Leu Lys Lys Leu Arg Ala Lys Leu Gln Ser Leu
 545 550 555 560
 Lys Ala Glu Arg Glu Glu Ala Arg His Arg Glu Glu Met Ala Leu Arg
 565 570 575
 Gly Lys Asp Ala Ala Glu Ile Val Leu Glu Ala Phe Cys Ala His Ala
 580 585 590
 Ser Gln Arg Ile Ser Gln Leu Glu Gln Asp Leu Ala Ser Met Arg Glu
 595 600 605
 Phe Arg Gly Leu Leu Lys Asp Ala Gln Thr Gln Leu Val Gly Leu His
 610 615 620
 Ala Lys Gln Glu Glu Leu Val Gln Gln Thr Val Ser Leu Thr Ser Thr
 625 630 635 640
 Leu Gln Gln Asp Trp Arg Ser Met Gln Leu Asp Tyr Thr Thr Trp Thr
 645 650 655
 Ala Leu Leu Ser Arg Ser Arg Gln Leu Thr Glu Lys Leu Thr Val Lys
 660 665 670
 Ser Gln Gln Ala Leu Gln Glu Arg Asp Val Ala Ile Glu Glu Lys Gln
 675 680 685
 Glu Val Ser Arg Val Leu Glu Gln Val Ser Ala Gln Leu Glu Glu Cys
 690 695 700
 Lys Gly Gln Thr Glu Gln Leu Glu Leu Glu Asn Ile Arg Leu Ala Thr
 705 710 715 720
 Asp Leu Arg Ala Gln Leu Gln Ile Leu Ala Asn Met Asp Ser Gln Leu
 725 730 735
 Lys Glu Leu Gln Ser Gln His Thr His Cys Ala Gln Asp Leu Ala Met
 740 745 750
 Lys Asp Glu Leu Leu Cys Gln Leu Thr Gln Ser Asn Glu Glu Gln Ala
 755 760 765
 Ala Gln Cys Val Lys Glu Glu Met Ala Leu Lys His Met Gln Ala Glu
 770 775 780
 Leu Gln Gln Gln Gln Ala Val Leu Ala Lys Glu Val Arg Asp Leu Lys
 785 790 795 800
 Glu Thr Leu Glu Phe Ala Asp Gln Glu Asn Gln Val Ala His Leu Glu
 805 810 815

10

20

30

Leu Gly Gln Val Glu Cys Gln Leu Lys Thr Thr Leu Glu Val Leu Arg
 820 825 830
 Glu Arg Ser Leu Gln Cys Glu Asn Leu Lys Asp Thr Val Glu Asn Leu
 835 840 845
 Thr Ala Lys Leu Ala Ser Thr Ile Ala Asp Asn Gln Glu Gln Asp Leu
 850 855 860
 Glu Lys Thr Arg Gln Tyr Ser Gln Lys Leu Gly Leu Leu Thr Glu Gln
 865 870 875 880
 Leu Gln Ser Leu Thr Leu Phe Leu Gln Thr Lys Leu Lys Glu Lys Thr
 885 890 895
 Glu Gln Glu Thr Leu Leu Leu Ser Thr Ala Cys Pro Pro Thr Gln Glu
 900 905 910
 His Pro Leu Pro Asn Asp Arg Thr Phe Leu Gly Ser Ile Leu Thr Ala
 915 920 925
 Val Ala Asp Glu Glu Pro Glu Ser Thr Pro Val Pro Leu Leu Gly Ser
 930 935 940
 Asp Lys Ser Ala Phe Thr Arg Val Ala Ser Met Val Ser Leu Gln Pro
 945 950 955 960
 Ala Glu Thr Pro Gly Met Glu Glu Ser Leu Ala Glu Met Ser Ile Met
 965 970 975
 Thr Thr Glu Leu Gln Ser Leu Cys Ser Leu Leu Gln Glu Ser Lys Glu
 980 985 990
 Glu Ala Ile Arg Thr Leu Gln Arg Lys Ile Cys Glu Leu Gln Ala Arg
 995 1000 1005
 Leu Gln Ala Gln Glu Glu Gln His Gln Glu Val Gln Lys Ala Lys Glu
 1010 1015 1020
 Ala Asp Ile Glu Lys Leu Asn Gln Ala Leu Cys Leu Arg Tyr Lys Asn
 1025 1030 1035 1040
 Glu Lys Glu Leu Gln Glu Val Ile Gln Gln Asn Glu Lys Ile Leu Glu
 1045 1050 1055
 Gln Ile Asp Lys Ser Gly Glu Leu Ile Ser Leu Arg Glu Glu Val Thr
 1060 1065 1070
 His Leu Thr Arg Ser Leu Arg Arg Ala Glu Thr Glu Thr Lys Val Leu
 1075 1080 1085
 Gln Glu Ala Trp Gln Ala Ser Trp Thr Pro Thr Ala Ser Leu Trp Pro
 1090 1095 1100
 Pro Ile Gly Ser Arg Arg Lys Cys Gly Ser Leu Arg Arg Trp Thr Asn
 1105 1110 1115 1120

<210> 675

<211> 540

<212> PRT

<213> Homo sapiens

<400> 675

Met Gly Thr Thr Ala Arg Ala Ala Leu Val Leu Thr Tyr Leu Ala Val
 1 5 10 15
 Ala Ser Ala Ala Ser Glu Gly Gly Phe Thr Ala Thr Gly Gln Arg Gln
 20 25 30
 Leu Arg Pro Glu His Phe Gln Glu Val Gly Tyr Ala Ala Pro Pro Ser

10

20

30

35 40 45
 Pro Pro Leu Ser Arg Ser Leu Pro Met Asp His Pro Asp Ser Ser Gln
 50 55 60
 His Gly Pro Pro Phe Glu Gly Gln Ser Gln Val Gln Pro Pro Pro Ser
 65 70 75 80
 Gln Glu Ala Thr Pro Leu Gln Gln Glu Lys Leu Leu Pro Ala Gln Leu
 85 90 95
 Pro Ala Glu Lys Glu Val Gly Pro Pro Leu Pro Gln Glu Ala Val Pro
 100 105 110
 Leu Gln Lys Glu Leu Pro Ser Leu Gln His Pro Asn Glu Gln Lys Glu
 115 120 125
 Gly Thr Pro Ala Pro Phe Gly Asp Gln Ser His Pro Glu Pro Glu Ser
 130 135 140
 Trp Asn Ala Ala Gln His Cys Gln Gln Asp Arg Ser Gln Gly Gly Trp
 145 150 155 160
 Gly His Arg Leu Asp Gly Phe Pro Pro Gly Arg Pro Ser Pro Asp Asn
 165 170 175
 Leu Asn Gln Ile Cys Leu Pro Asn Arg Gln His Val Val Tyr Gly Pro
 180 185 190
 Trp Asn Leu Pro Gln Ser Ser Tyr Ser His Leu Thr Arg Gln Gly Glu
 195 200 205
 Thr Leu Asn Phe Leu Glu Ile Gly Tyr Ser Arg Cys Cys His Cys Arg
 210 215 220
 Ser His Thr Asn Arg Leu Glu Cys Ala Lys Leu Val Trp Glu Glu Ala
 225 230 235 240
 Met Ser Arg Phe Cys Glu Ala Glu Phe Ser Val Lys Thr Arg Pro His
 245 250 255
 Trp Cys Cys Thr Arg Gln Gly Glu Ala Arg Phe Ser Cys Phe Gln Glu
 260 265 270
 Glu Ala Pro Gln Pro His Tyr Gln Leu Arg Ala Cys Pro Ser His Gln
 275 280 285
 Pro Asp Ile Ser Ser Gly Leu Glu Leu Pro Phe Pro Pro Gly Val Pro
 290 295 300
 Thr Leu Asp Asn Ile Lys Asn Ile Cys His Leu Arg Arg Phe Arg Ser
 305 310 315 320
 Val Pro Arg Asn Leu Pro Ala Thr Asp Pro Leu Gln Arg Glu Leu Leu
 325 330 335
 Ala Leu Ile Gln Leu Glu Arg Glu Phe Gln Arg Cys Cys Arg Gln Gly
 340 345 350
 Asn Asn His Thr Cys Thr Trp Lys Ala Trp Glu Asp Thr Leu Asp Lys
 355 360 365
 Tyr Cys Asp Arg Glu Tyr Ala Val Lys Thr His His His Leu Cys Cys
 370 375 380
 Arg His Pro Pro Ser Pro Thr Arg Asp Glu Cys Phe Ala Arg Arg Ala
 385 390 395 400
 Pro Tyr Pro Asn Tyr Asp Arg Asp Ile Leu Thr Ile Asp Ile Ser Arg
 405 410 415
 Val Thr Pro Asn Leu Met Gly His Leu Cys Gly Asn Gln Arg Val Leu
 420 425 430
 Thr Lys His Lys His Ile Pro Gly Leu Ile His Asn Met Thr Ala Arg

10

20

30

```

      435              440              445
Cys Cys Asp Leu Pro Phe Pro Glu Gln Ala Cys Cys Ala Glu Glu Glu
 450              455              460
Lys Leu Thr Phe Ile Asn Asp Leu Cys Gly Pro Arg Arg Asn Ile Trp
 465              470              475              480
Arg Asp Pro Ala Leu Cys Cys Tyr Leu Ser Pro Gly Asp Glu Gln Val
      485              490              495
Asn Cys Phe Asn Ile Asn Tyr Leu Arg Asn Val Ala Leu Val Ser Gly
      500              505              510
Asp Thr Glu Asn Ala Lys Gly Gln Gly Glu Gln Gly Ser Thr Gly Gly
      515              520              525
Thr Asn Ile Ser Ser Thr Ser Glu Pro Lys Glu Glu
      530              535              540

```

<210> 676

<211> 462

<212> PRT

<213> Homo sapiens

<400> 676

```

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

```

10

20

30

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

10

<210> 677
 <211> 2328
 <212> PRT
 <213> Homo sapiens

20

<400> 677
 Lys Ser Lys Arg Gln Ala Gln Gln Met Val Gln Pro Gln Ser Pro Val
 1 5 10 15
 Ala Val Ser Gln Ser Lys Pro Gly Cys Tyr Asp Asn Gly Lys His Tyr
 20 25 30
 Gln Ile Asn Gln Gln Trp Glu Arg Thr Tyr Leu Gly Asn Val Leu Val
 35 40 45
 Cys Thr Cys Tyr Gly Gly Ser Arg Gly Phe Asn Cys Glu Ser Lys Pro
 50 55 60
 Glu Ala Glu Glu Thr Cys Phe Asp Lys Tyr Thr Gly Asn Thr Tyr Arg
 65 70 75 80
 Val Gly Asp Thr Tyr Glu Arg Pro Lys Asp Ser Met Ile Trp Asp Cys
 85 90 95
 Thr Cys Ile Gly Ala Gly Arg Gly Arg Ile Ser Cys Thr Ile Ala Asn
 100 105 110
 Arg Cys His Glu Gly Gly Gln Ser Tyr Lys Ile Gly Asp Thr Trp Arg

30

115	120	125
Arg Pro His Glu Thr Gly Gly Tyr Met Leu Glu Cys Val Cys Leu Gly		
130	135	140
Asn Gly Lys Gly Glu Trp Thr Cys Lys Pro Ile Ala Glu Lys Cys Phe		
145	150	155
Asp His Ala Ala Gly Thr Ser Tyr Val Val Gly Glu Thr Trp Glu Lys		160
	165	170
		175
Pro Tyr Gln Gly Trp Met Met Val Asp Cys Thr Cys Leu Gly Glu Gly		
	180	185
		190
Ser Gly Arg Ile Thr Cys Thr Ser Arg Asn Arg Cys Asn Asp Gln Asp		
	195	200
		205
Thr Arg Thr Ser Tyr Arg Ile Gly Asp Thr Trp Ser Lys Lys Asp Asn		
	210	215
		220
Arg Gly Asn Leu Leu Gln Cys Ile Cys Thr Gly Asn Gly Arg Gly Glu		
	225	230
		235
Trp Lys Cys Glu Arg His Thr Ser Val Gln Thr Thr Ser Ser Gly Ser		240
	245	250
		255
Gly Pro Phe Thr Asp Val Arg Ala Ala Val Tyr Gln Pro Gln Pro His		
	260	265
		270
Pro Gln Pro Pro Pro Tyr Gly His Cys Val Thr Asp Ser Gly Val Val		
	275	280
		285
Tyr Ser Val Gly Met Gln Trp Leu Lys Thr Gln Gly Asn Lys Gln Met		
	290	295
		300
Leu Cys Thr Cys Leu Gly Asn Gly Val Ser Cys Gln Glu Thr Ala Val		
	305	310
		315
Thr Gln Thr Tyr Gly Gly Asn Leu Asn Gly Glu Pro Cys Val Leu Pro		
	320	325
		330
		335
Phe Thr Tyr Asn Gly Arg Thr Phe Tyr Ser Cys Thr Thr Glu Gly Arg		
	340	345
		350
Gln Asp Gly His Leu Trp Cys Ser Thr Thr Ser Asn Tyr Glu Gln Asp		
	355	360
		365
Gln Lys Tyr Ser Phe Cys Thr Asp His Thr Val Leu Val Gln Thr Gln		
	370	375
		380
Gly Gly Asn Ser Asn Gly Ala Leu Cys His Phe Pro Phe Leu Tyr Asn		
	385	390
		395
Asn His Asn Tyr Thr Asp Cys Thr Ser Glu Gly Arg Arg Asp Asn Met		
	400	405
		410
		415
Lys Trp Cys Gly Thr Thr Gln Asn Tyr Asp Ala Asp Gln Lys Phe Gly		
	420	425
		430
Phe Cys Pro Met Ala Ala His Glu Glu Ile Cys Thr Thr Asn Glu Gly		
	435	440
		445
Val Met Tyr Arg Ile Gly Asp Gln Trp Asp Lys Gln His Asp Met Gly		
	450	455
		460
His Met Met Arg Cys Thr Cys Val Gly Asn Gly Arg Gly Glu Trp Thr		
	465	470
		475
Cys Ile Ala Tyr Ser Gln Leu Arg Asp Gln Cys Ile Val Asp Asp Ile		
	480	485
		490
		495
Thr Tyr Asn Val Asn Asp Thr Phe His Lys Arg His Glu Glu Gly His		
	500	505
		510
Met Leu Asn Cys Thr Cys Phe Gly Gln Gly Arg Gly Arg Trp Lys Cys		

10

20

30

```

515          520          525
Asp Pro Val Asp Gln Cys Gln Asp Ser Glu Thr Gly Thr Phe Tyr Gln
530          535          540
Ile Gly Asp Ser Trp Glu Lys Tyr Val His Gly Val Arg Tyr Gln Cys
545          550          555          560
Tyr Cys Tyr Gly Arg Gly Ile Gly Glu Trp His Cys Gln Pro Leu Gln
565          570          575
Thr Tyr Pro Ser Ser Gly Pro Val Glu Val Phe Ile Thr Glu Thr
580          585          590
Pro Ser Gln Pro Asn Ser His Pro Ile Gln Trp Asn Ala Pro Gln Pro
595          600          605
Ser His Ile Ser Lys Tyr Ile Leu Arg Trp Arg Pro Lys Asn Ser Val
610          615          620
Gly Arg Trp Lys Glu Ala Thr Ile Pro Gly His Leu Asn Ser Tyr Thr
625          630          635          640
Ile Lys Gly Leu Lys Pro Gly Val Val Tyr Glu Gly Gln Leu Ile Ser
645          650          655
Ile Gln Gln Tyr Gly His Gln Glu Val Thr Arg Phe Asp Phe Thr Thr
660          665          670
Thr Ser Thr Ser Thr Pro Val Thr Ser Asn Thr Val Thr Gly Glu Thr
675          680          685
Thr Pro Phe Ser Pro Leu Val Ala Thr Ser Glu Ser Val Thr Glu Ile
690          695          700
Thr Ala Ser Ser Phe Val Val Ser Trp Val Ser Ala Ser Asp Thr Val
705          710          715          720
Ser Gly Phe Arg Val Glu Tyr Glu Leu Ser Glu Glu Gly Asp Glu Pro
725          730          735
Gln Tyr Leu Asp Leu Pro Ser Thr Ala Thr Ser Val Asn Ile Pro Asp
740          745          750
Leu Leu Pro Gly Arg Lys Tyr Ile Val Asn Val Tyr Gln Ile Ser Glu
755          760          765
Asp Gly Glu Gln Ser Leu Ile Leu Ser Thr Ser Gln Thr Thr Ala Pro
770          775          780
Asp Ala Pro Pro Asp Pro Thr Val Asp Gln Val Asp Asp Thr Ser Ile
785          790          795          800
Val Val Arg Trp Ser Arg Pro Gln Ala Pro Ile Thr Gly Tyr Arg Ile
805          810          815
Val Tyr Ser Pro Ser Val Glu Gly Ser Ser Thr Glu Leu Asn Leu Pro
820          825          830
Glu Thr Ala Asn Ser Val Thr Leu Ser Asp Leu Gln Pro Gly Val Gln
835          840          845
Tyr Asn Ile Thr Ile Tyr Ala Val Glu Glu Asn Gln Glu Ser Thr Pro
850          855          860
Val Val Ile Gln Gln Glu Thr Thr Gly Thr Pro Arg Ser Asp Thr Val
865          870          875          880
Pro Ser Pro Arg Asp Leu Gln Phe Val Glu Val Thr Asp Val Lys Val
885          890          895
Thr Ile Met Trp Thr Pro Pro Glu Ser Ala Val Thr Gly Tyr Arg Val
900          905          910
Asp Val Ile Pro Val Asn Leu Pro Gly Glu His Gly Gln Arg Leu Pro

```

10

20

30

915 920 925
 Ile Ser Arg Asn Thr Phe Ala Glu Val Thr Gly Leu Ser Pro Gly Val
 930 935 940
 Thr Tyr Tyr Phe Lys Val Phe Ala Val Ser His Gly Arg Glu Ser Lys
 945 950 955 960
 Pro Leu Thr Ala Gln Gln Thr Thr Lys Leu Asp Ala Pro Thr Asn Leu
 965 970 975
 Gln Phe Val Asn Glu Thr Asp Ser Thr Val Leu Val Arg Trp Thr Pro
 980 985 990
 Pro Arg Ala Gln Ile Thr Gly Tyr Arg Leu Thr Val Gly Leu Thr Arg
 995 1000 1005
 Arg Gly Gln Pro Arg Gln Tyr Asn Val Gly Pro Ser Val Ser Lys Tyr
 1010 1015 1020
 Pro Leu Arg Asn Leu Gln Pro Ala Ser Glu Tyr Thr Val Ser Leu Val
 1025 1030 1035 1040
 Ala Ile Lys Gly Asn Gln Glu Ser Pro Lys Ala Thr Gly Val Phe Thr
 1045 1050 1055
 Thr Leu Gln Pro Gly Ser Ser Ile Pro Pro Tyr Asn Thr Glu Val Thr
 1060 1065 1070
 Glu Thr Thr Ile Val Ile Thr Trp Thr Pro Ala Pro Arg Ile Gly Phe
 1075 1080 1085
 Lys Leu Gly Val Arg Pro Ser Gln Gly Gly Glu Ala Pro Arg Glu Val
 1090 1095 1100
 Thr Ser Asp Ser Gly Ser Ile Val Val Ser Gly Leu Thr Pro Gly Val
 1105 1110 1115 1120
 Glu Tyr Val Tyr Thr Ile Gln Val Leu Arg Asp Gly Gln Glu Arg Asp
 1125 1130 1135
 Ala Pro Ile Val Asn Lys Val Val Thr Pro Leu Ser Pro Pro Thr Asn
 1140 1145 1150
 Leu His Leu Glu Ala Asn Pro Asp Thr Gly Val Leu Thr Val Ser Trp
 1155 1160 1165
 Glu Arg Ser Thr Thr Pro Asp Ile Thr Gly Tyr Arg Ile Thr Thr Thr
 1170 1175 1180
 Pro Thr Asn Gly Gln Gln Gly Asn Ser Leu Glu Glu Val Val His Ala
 1185 1190 1195 1200
 Asp Gln Ser Ser Cys Thr Phe Asp Asn Leu Ser Pro Gly Leu Glu Tyr
 1205 1210 1215
 Asn Val Ser Val Tyr Thr Val Lys Asp Asp Lys Glu Ser Val Pro Ile
 1220 1225 1230
 Ser Asp Thr Ile Ile Pro Ala Val Pro Pro Pro Thr Asp Leu Arg Phe
 1235 1240 1245
 Thr Asn Ile Gly Pro Asp Thr Met Arg Val Thr Trp Ala Pro Pro Pro
 1250 1255 1260
 Ser Ile Asp Leu Thr Asn Phe Leu Val Arg Tyr Ser Pro Val Lys Asn
 1265 1270 1275 1280
 Glu Glu Asp Val Ala Glu Leu Ser Ile Ser Pro Ser Asp Asn Ala Val
 1285 1290 1295
 Val Leu Thr Asn Leu Leu Pro Gly Thr Glu Tyr Val Val Ser Val Ser
 1300 1305 1310
 Ser Val Tyr Glu Gln His Glu Ser Thr Pro Leu Arg Gly Arg Gln Lys

10

20

30

```

      1315      1320      1325
Thr Gly Leu Asp Ser Pro Thr Gly Ile Asp Phe Ser Asp Ile Thr Ala
1330      1335      1340
Asn Ser Phe Thr Val His Trp Ile Ala Pro Arg Ala Thr Ile Thr Gly
1345      1350      1355      1360
Tyr Arg Ile Arg His His Pro Glu His Phe Ser Gly Arg Pro Arg Glu
      1365      1370      1375
Asp Arg Val Pro His Ser Arg Asn Ser Ile Thr Leu Thr Asn Leu Thr
      1380      1385      1390
Pro Gly Thr Glu Tyr Val Val Ser Ile Val Ala Leu Asn Gly Arg Glu
      1395      1400      1405
Glu Ser Pro Leu Leu Ile Gly Gln Gln Ser Thr Val Ser Asp Val Pro
      1410      1415      1420
Arg Asp Leu Glu Val Val Ala Ala Thr Pro Thr Ser Leu Leu Ile Ser
1425      1430      1435      1440
Trp Asp Ala Pro Ala Val Thr Val Arg Tyr Tyr Arg Ile Thr Tyr Gly
      1445      1450      1455
Glu Thr Gly Gly Asn Ser Pro Val Gln Glu Phe Thr Val Pro Gly Ser
      1460      1465      1470
Lys Ser Thr Ala Thr Ile Ser Gly Leu Lys Pro Gly Val Asp Tyr Thr
      1475      1480      1485
Ile Thr Val Tyr Ala Val Thr Gly Arg Gly Asp Ser Pro Ala Ser Ser
      1490      1495      1500
Lys Pro Ile Ser Ile Asn Tyr Arg Thr Glu Ile Asp Lys Pro Ser Gln
1505      1510      1515      1520
Met Gln Val Thr Asp Val Gln Asp Asn Ser Ile Ser Val Lys Trp Leu
      1525      1530      1535
Pro Ser Ser Ser Pro Val Thr Gly Tyr Arg Val Thr Thr Thr Pro Lys
      1540      1545      1550
Asn Gly Pro Gly Pro Thr Lys Thr Lys Thr Ala Gly Pro Asp Gln Thr
      1555      1560      1565
Glu Met Thr Ile Glu Gly Leu Gln Pro Thr Val Glu Tyr Val Val Ser
      1570      1575      1580
Val Tyr Ala Gln Asn Pro Ser Gly Glu Ser Gln Pro Leu Val Gln Thr
1585      1590      1595      1600
Ala Val Thr Asn Ile Asp Arg Pro Lys Gly Leu Ala Phe Thr Asp Val
      1605      1610      1615
Asp Val Asp Ser Ile Lys Ile Ala Trp Glu Ser Pro Gln Gly Gln Val
      1620      1625      1630
Ser Arg Tyr Arg Val Thr Tyr Ser Ser Pro Glu Asp Gly Ile His Glu
      1635      1640      1645
Leu Phe Pro Ala Pro Asp Gly Glu Glu Asp Thr Ala Glu Leu Gln Gly
      1650      1655      1660
Leu Arg Pro Gly Ser Glu Tyr Thr Val Ser Val Val Ala Leu His Asp
1665      1670      1675      1680
Asp Met Glu Ser Gln Pro Leu Ile Gly Thr Gln Ser Thr Ala Ile Pro
      1685      1690      1695
Ala Pro Thr Asp Leu Lys Phe Thr Gln Val Thr Pro Thr Ser Leu Ser
      1700      1705      1710
Ala Gln Trp Thr Pro Pro Asn Val Gln Leu Thr Gly Tyr Arg Val Arg

```

10

20

30

1715 1720 1725
 Val Thr Pro Lys Glu Lys Thr Gly Pro Met Lys Glu Ile Asn Leu Ala
 1730 1735 1740
 Pro Asp Ser Ser Ser Val Val Val Ser Gly Leu Met Val Ala Thr Lys
 1745 1750 1755 1760
 Tyr Glu Val Ser Val Tyr Ala Leu Lys Asp Thr Leu Thr Ser Arg Pro
 1765 1770 1775
 Ala Gln Gly Val Val Thr Thr Leu Glu Asn Val Ser Pro Pro Arg Arg
 1780 1785 1790
 Ala Arg Val Thr Asp Ala Thr Glu Thr Thr Ile Thr Ile Ser Trp Arg
 1795 1800 1805
 Thr Lys Thr Glu Thr Ile Thr Gly Phe Gln Val Asp Ala Val Pro Ala
 1810 1815 1820
 Asn Gly Gln Thr Pro Ile Gln Arg Thr Ile Lys Pro Asp Val Arg Ser
 1825 1830 1835 1840
 Tyr Thr Ile Thr Gly Leu Gln Pro Gly Thr Asp Tyr Lys Ile Tyr Leu
 1845 1850 1855
 Tyr Thr Leu Asn Asp Asn Ala Arg Ser Ser Pro Val Val Ile Asp Ala
 1860 1865 1870
 Ser Thr Ala Ile Asp Ala Pro Ser Asn Leu Arg Phe Leu Ala Thr Thr
 1875 1880 1885
 Pro Asn Ser Leu Leu Val Ser Trp Gln Pro Pro Arg Ala Arg Ile Thr
 1890 1895 1900
 Gly Tyr Ile Ile Lys Tyr Glu Lys Pro Gly Ser Pro Pro Arg Glu Val
 1905 1910 1915 1920
 Val Pro Arg Pro Arg Pro Gly Val Thr Glu Ala Thr Ile Thr Gly Leu
 1925 1930 1935
 Glu Pro Gly Thr Glu Tyr Thr Ile Tyr Val Ile Ala Leu Lys Asn Asn
 1940 1945 1950
 Gln Lys Ser Glu Pro Leu Ile Gly Arg Lys Lys Thr Asp Glu Leu Pro
 1955 1960 1965
 Gln Leu Val Thr Leu Pro His Pro Asn Leu His Gly Pro Glu Ile Leu
 1970 1975 1980
 Asp Val Pro Ser Thr Val Gln Lys Thr Pro Phe Val Thr His Pro Gly
 1985 1990 1995 2000
 Tyr Asp Thr Gly Asn Gly Ile Gln Leu Pro Gly Thr Ser Gly Gln Gln
 2005 2010 2015
 Pro Ser Val Gly Gln Gln Met Ile Phe Glu Glu His Gly Phe Arg Arg
 2020 2025 2030
 Thr Thr Pro Pro Thr Thr Ala Thr Pro Ile Arg His Arg Pro Arg Pro
 2035 2040 2045
 Tyr Pro Pro Asn Val Gly Gln Glu Ala Leu Ser Gln Thr Thr Ile Ser
 2050 2055 2060
 Trp Ala Pro Phe Gln Asp Thr Ser Glu Tyr Ile Ile Ser Cys His Pro
 2065 2070 2075 2080
 Val Gly Thr Asp Glu Glu Pro Leu Gln Phe Arg Val Pro Gly Thr Ser
 2085 2090 2095
 Thr Ser Ala Thr Leu Thr Gly Leu Thr Arg Gly Ala Thr Tyr Asn Ile
 2100 2105 2110
 Ile Val Glu Ala Leu Lys Asp Gln Gln Arg His Lys Val Arg Glu Glu

10

20

30

```

2115          2120          2125
Val Val Thr Val Gly Asn Ser Val Asn Glu Gly Leu Asn Gln Pro Thr
2130          2135          2140
Asp Asp Ser Cys Phe Asp Pro Tyr Thr Val Ser His Tyr Ala Val Gly
2145          2150          2155          2160
Asp Glu Trp Glu Arg Met Ser Glu Ser Gly Phe Lys Leu Leu Cys Gln
2165          2170          2175
Cys Leu Gly Phe Gly Ser Gly His Phe Arg Cys Asp Ser Ser Arg Trp
2180          2185          2190
Cys His Asp Asn Gly Val Asn Tyr Lys Ile Gly Glu Lys Trp Asp Arg
2195          2200          2205
Gln Gly Glu Asn Gly Gln Met Met Ser Cys Thr Cys Leu Gly Asn Gly
2210          2215          2220
Lys Gly Glu Phe Lys Cys Asp Pro His Glu Ala Thr Cys Tyr Asp Asp
2225          2230          2235          2240
Gly Lys Thr Tyr His Val Gly Glu Gln Trp Gln Lys Glu Tyr Leu Gly
2245          2250          2255
Ala Ile Cys Ser Cys Thr Cys Phe Gly Gly Gln Arg Gly Trp Arg Cys
2260          2265          2270
Asp Asn Cys Arg Arg Pro Gly Gly Glu Pro Ser Pro Glu Gly Thr Thr
2275          2280          2285
Gly Gln Ser Tyr Asn Gln Tyr Ser Gln Arg Tyr His Gln Arg Thr Asn
2290          2295          2300
Thr Asn Val Asn Cys Pro Ile Glu Cys Phe Met Pro Leu Asp Val Gln
2305          2310          2315          2320
Ala Asp Arg Glu Asp Ser Arg Glu
2325

```

```

<210> 678
<211> 188
<212> PRT
<213> Homo sapiens

```

```

<400> 678
His Gln Thr His Lys Glu Gly Gly Ser Thr His Ala Ser Ala Asp Ala
1      5      10      15
Trp Glu Ile Ile Glu Leu Glu Thr Glu Ile Glu Lys Phe Lys Ala Glu
20     25     30
Asn Ala Ser Leu Ala Lys Leu Arg Ile Glu Arg Glu Ser Ala Leu Glu
35     40     45
Lys Leu Arg Lys Glu Ile Ala Asp Phe Glu Gln Gln Lys Ala Lys Glu
50     55     60
Leu Ala Arg Ile Glu Glu Phe Lys Lys Glu Glu Met Arg Lys Leu Gln
65     70     75     80
Lys Glu Arg Lys Val Phe Glu Lys Tyr Thr Thr Ala Ala Arg Thr Phe
85     90     95
Pro Asp Lys Lys Glu Arg Glu Glu Ile Gln Thr Leu Lys Gln Gln Ile
100    105    110
Ala Asp Leu Arg Glu Asp Leu Lys Arg Lys Glu Thr Lys Trp Ser Ser
115    120    125

```

10

20

30

Thr His Ser Arg Leu Arg Ser Gln Ile Gln Met Leu Val Arg Glu Asn
 130 135 140
 Thr Asp Leu Arg Glu Glu Ile Lys Val Met Glu Arg Phe Arg Leu Asp
 145 150 155 160
 Ala Trp Lys Arg Ala Glu Ala Ile Glu Ser Ser Leu Glu Val Glu Lys
 165 170 175
 Lys Asp Lys Leu Ala Asn Thr Ser Val Arg Phe Gln
 180 185

<210> 679

<211> 284

<212> PRT

<213> Homo sapiens

10

<400> 679

Met Glu Pro Gly Asn Tyr Ala Thr Leu Asp Gly Ala Lys Asp Ile Glu
 1 5 10 15
 Gly Leu Leu Gly Ala Gly Gly Gly Arg Asn Leu Val Ala His Ser Pro
 20 25 30
 Leu Thr Ser His Pro Ala Ala Pro Thr Leu Met Pro Ala Val Asn Tyr
 35 40 45
 Ala Pro Leu Asp Leu Pro Gly Ser Ala Glu Pro Pro Lys Gln Cys His
 50 55 60
 Pro Cys Pro Gly Val Pro Gln Gly Thr Ser Pro Ala Pro Val Pro Tyr
 65 70 75 80
 Gly Tyr Phe Gly Gly Tyr Tyr Ser Cys Arg Val Ser Arg Ser Ser
 85 90 95
 Leu Lys Pro Cys Ala Gln Ala Ala Thr Leu Ala Ala Tyr Pro Ala Glu
 100 105 110
 Thr Pro Thr Ala Gly Glu Glu Tyr Pro Ser Arg Pro Thr Glu Phe Ala
 115 120 125
 Phe Tyr Pro Gly Tyr Pro Gly Thr Tyr His Ala Met Ala Ser Tyr Leu
 130 135 140
 Asp Val Ser Val Val Gln Thr Leu Gly Ala Pro Gly Glu Pro Arg His
 145 150 155 160
 Asp Ser Leu Leu Pro Val Asp Ser Tyr Gln Ser Trp Ala Leu Ala Gly
 165 170 175
 Gly Trp Asn Ser Gln Met Cys Cys Gln Gly Glu Gln Asn Pro Pro Gly
 180 185 190
 Pro Phe Trp Lys Ala Ala Phe Ala Asp Ser Ser Gly Gln His Pro Pro
 195 200 205
 Asp Ala Cys Ala Phe Arg Arg Gly Arg Lys Lys Arg Ile Pro Tyr Ser
 210 215 220
 Lys Gly Gln Leu Arg Glu Leu Glu Arg Glu Tyr Ala Ala Asn Lys Phe
 225 230 235 240
 Ile Thr Lys Asp Lys Arg Arg Lys Ile Ser Ala Ala Thr Ser Leu Ser
 245 250 255
 Glu Arg Gln Ile Thr Ile Trp Phe Gln Asn Arg Arg Val Lys Glu Lys
 260 265 270
 Lys Val Leu Ala Lys Val Lys Asn Ser Ala Thr Pro

20

30

275

280

<210> 680

<211> 676

<212> PRT

<213> Homo sapiens

<400> 680

```

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

```

10

20

30

Phe Gln Asp Gly Pro Lys Ser Tyr Leu Ser Ser Ser Ala Pro Ser Ser
 340 345 350
 Ser Pro Ala Gly Leu Asp Gly Ser Gln Gln Gly Ala Val Pro Gly Leu
 355 360 365
 Gly Pro Lys Pro Gly Cys Thr Asp Leu Gly Thr Gly Pro Lys Leu Ser
 370 375 380
 Pro Thr Ser Leu Val His Pro Val Met Ser Thr Leu Pro Glu Leu Ser
 385 390 395 400
 Cys Lys Glu Gly Pro Leu Gly Trp Ser Ser Asp Gly Ser Leu Gly Ser
 405 410 415
 Val Leu Leu Asp Ser Pro Ser Ser Pro Arg Val Arg Leu Pro Cys Gln
 420 425 430
 Pro Leu Val Pro Gly Pro Glu Leu Arg Pro Ser Ala Ala Glu Leu Lys
 435 440 445
 Leu Glu Ala Leu Thr Gln Arg Leu Glu Arg Glu Met Asp Ala His Pro
 450 455 460
 Lys Ala Asp Tyr Phe Gly Ala Cys Val Lys Cys Ser Lys Gly Val Phe
 465 470 475 480
 Gly Ala Gly Gln Ala Cys Gln Ala Met Gly Asn Leu Tyr His Asp Thr
 485 490 495
 Cys Phe Thr Cys Ala Ala Cys Ser Arg Lys Leu Arg Gly Lys Ala Phe
 500 505 510
 Tyr Phe Val Asn Gly Lys Val Phe Cys Glu Glu Asp Phe Leu Tyr Ser
 515 520 525
 Gly Phe Gln Gln Ser Ala Asp Arg Cys Phe Leu Cys Gly His Leu Ile
 530 535 540
 Met Asp Met Ile Leu Gln Ala Leu Gly Lys Ser Tyr His Pro Gly Cys
 545 550 555 560
 Phe Arg Cys Val Ile Cys Asn Glu Cys Leu Asp Gly Val Pro Phe Thr
 565 570 575
 Val Asp Ser Glu Asn Lys Ile Tyr Cys Val Arg Asp Tyr His Lys Val
 580 585 590
 Leu Ala Pro Lys Cys Ala Ala Cys Gly Leu Pro Ile Leu Pro Pro Glu
 595 600 605
 Gly Ser Asp Glu Thr Ile Arg Val Val Ser Met Asp Arg Asp Tyr His
 610 615 620
 Val Glu Cys Tyr His Cys Glu Asp Cys Gly Leu Glu Leu Asn Asp Glu
 625 630 635 640
 Asp Gly His Arg Cys Tyr Pro Leu Glu Asp His Leu Phe Cys His Ser
 645 650 655
 Cys His Val Lys Arg Leu Glu Lys Arg Pro Ser Ser Thr Ala Leu His
 660 665 670
 Gln His His Phe
 675

10

20

30

<210> 681

<211> 296

<212> PRT

<213> Homo sapiens

<400> 681

```

Ser Thr Gly Ser Glu Phe Pro Leu Cys Thr Lys Ala Ser Pro Cys Ser
 1          5          10          15
Ala Ala Arg Ala Gly Gly Arg Ala Leu Gly Trp Arg Leu Gln Gln Gln
 20          25          30
Arg Glu Thr Arg Gly Asn Pro Gly Asn Pro Gly Leu Gly Val Ala Ala
 35          40          45
Thr Met Thr Gly Ser Asn Met Ser Asp Ala Leu Ala Asn Ala Val Cys
 50          55          60
Gln Arg Cys Gln Ala Arg Phe Ser Pro Ala Glu Arg Ile Val Asn Ser
 65          70          75          80
Asn Gly Glu Leu Tyr His Glu His Cys Phe Val Cys Ala Gln Cys Phe
 85          90          95
Arg Pro Phe Pro Glu Gly Leu Phe Tyr Glu Phe Glu Gly Arg Lys Tyr
100          105          110
Cys Glu His Asp Phe Gln Met Leu Phe Ala Pro Cys Cys Gly Ser Cys
115          120          125
Gly Glu Phe Ile Ile Gly Arg Val Ile Lys Ala Met Asn Asn Asn Trp
130          135          140
His Pro Gly Cys Phe Arg Cys Glu Leu Cys Asp Val Glu Leu Ala Asp
145          150          155          160
Leu Gly Phe Val Lys Asn Ala Gly Arg His Leu Cys Arg Pro Cys His
165          170          175
Asn Arg Glu Lys Ala Lys Gly Leu Gly Lys Tyr Ile Cys Gln Arg Cys
180          185          190
His Leu Val Ile Asp Glu Gln Pro Leu Met Phe Arg Ser Asp Ala Tyr
195          200          205
His Pro Asp His Phe Asn Cys Thr His Cys Gly Lys Glu Leu Thr Ala
210          215          220
Glu Ala Arg Glu Leu Lys Gly Glu Leu Tyr Cys Leu Pro Cys His Asp
225          230          235          240
Lys Met Gly Val Pro Ile Cys Gly Ala Cys Arg Arg Pro Ile Glu Gly
245          250          255
Arg Val Val Asn Ala Leu Gly Lys Gln Trp His Val Glu His Phe Val
260          265          270
Cys Ala Lys Cys Glu Lys Pro Phe Leu Gly His Arg His Tyr Glu Lys
275          280          285
Lys Gly Leu Ala Tyr Cys Glu Leu
290          295

```

<210> 682

<211> 500

<212> PRT

<213> Homo sapiens

<400> 682

```

Met Gly Ile Gly Leu Ser Ala Gln Gly Val Asn Met Asn Arg Leu Pro
 1          5          10          15
Gly Trp Asp Lys His Ser Tyr Gly Tyr His Gly Asp Asp Gly His Ser
 20          25          30

```

10

20

30

Phe Cys Ser Ser Gly Thr Gly Gln Pro Tyr Gly Pro Thr Phe Thr Thr
 35 40 45
 Gly Asp Val Ile Gly Cys Cys Val Asn Leu Ile Asn Asn Thr Cys Phe
 50 55 60
 Tyr Thr Lys Asn Gly His Ser Leu Gly Ile Ala Phe Thr Asp Leu Pro
 65 70 75 80
 Pro Asn Leu Tyr Pro Thr Val Gly Leu Gln Thr Pro Gly Glu Val Val
 85 90 95
 Asp Ala Asn Phe Gly Gln His Pro Phe Val Phe Asp Ile Glu Asp Tyr
 100 105 110
 Met Arg Glu Trp Arg Thr Lys Ile Gln Ala Gln Ile Asp Arg Phe Pro
 115 120 125
 Ile Gly Asp Arg Glu Gly Glu Trp Gln Thr Met Ile Gln Lys Met Val
 130 135 140
 Ser Ser Tyr Leu Val His His Gly Tyr Cys Ala Thr Ala Glu Ala Phe
 145 150 155 160
 Ala Arg Ser Thr Asp Gln Thr Val Leu Glu Glu Leu Ala Ser Ile Lys
 165 170 175
 Asn Arg Gln Arg Ile Gln Lys Leu Val Leu Ala Gly Arg Met Gly Glu
 180 185 190
 Ala Ile Glu Thr Thr Gln Gln Leu Tyr Pro Ser Leu Leu Glu Arg Asn
 195 200 205
 Pro Asn Leu Leu Phe Thr Leu Lys Val Arg Gln Phe Ile Glu Met Val
 210 215 220
 Asn Gly Thr Asp Ser Glu Val Arg Cys Leu Gly Gly Arg Ser Pro Lys
 225 230 235 240
 Ser Gln Asp Ser Tyr Pro Val Ser Pro Arg Pro Phe Ser Ser Pro Ser
 245 250 255
 Met Ser Pro Ser His Gly Met Asn Ile His Asn Leu Ala Ser Gly Lys
 260 265 270
 Gly Ser Thr Ala His Phe Ser Gly Phe Glu Ser Cys Ser Asn Gly Val
 275 280 285
 Ile Ser Asn Lys Ala His Gln Ser Tyr Cys His Ser Asn Lys His Gln
 290 295 300
 Ser Ser Asn Leu Asn Val Pro Glu Leu Asn Ser Ile Asn Met Ser Arg
 305 310 315 320
 Ser Gln Gln Val Asn Asn Phe Thr Ser Asn Asp Val Asp Met Glu Thr
 325 330 335
 Asp His Tyr Ser Asn Gly Val Gly Glu Thr Ser Ser Asn Gly Phe Leu
 340 345 350
 Asn Gly Ser Ser Lys His Asp His Glu Met Glu Asp Cys Asp Thr Glu
 355 360 365
 Met Glu Val Asp Ser Ser Gln Leu Arg Arg Gln Leu Cys Gly Gly Ser
 370 375 380
 Gln Ala Ala Ile Glu Arg Met Ile His Phe Gly Arg Glu Leu Gln Ala
 385 390 395 400
 Met Ser Glu Gln Leu Arg Arg Asp Cys Gly Lys Asn Thr Ala Asn Lys
 405 410 415
 Lys Met Leu Lys Asp Ala Phe Ser Leu Leu Ala Tyr Ser Asp Pro Trp
 420 425 430

10

20

30

Asn Ser Pro Val Gly Asn Gln Leu Asp Pro Ile Gln Arg Glu Pro Val
 435 440 445
 Cys Ser Ala Leu Asn Ser Ala Ile Leu Glu Thr His Asn Leu Pro Lys
 450 455 460
 Gln Pro Pro Leu Ala Leu Ala Met Gly Gln Ala Thr Gln Cys Leu Gly
 465 470 475 480
 Leu Met Ala Arg Ser Gly Ile Gly Ser Cys Ala Phe Ala Thr Val Glu
 485 490 495
 Asp Tyr Leu His
 500

<210> 683

<211> 387

<212> PRT

<213> Homo sapiens

10

<400> 683

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

20

30

260 265 270
 Cys Pro Val Ser Asn Ala Met Cys Arg Glu Leu Pro Gln Ser Ile Val
 275 280 285
 Tyr Lys Tyr Met Ser Ile Arg Ser Asp Arg Ser Val Pro Ser Asp Ile
 290 295 300
 Phe Gln Ile Gln Ala Thr Thr Ile Tyr Ala Asn Thr Ile Asn Thr Phe
 305 310 315 320
 Arg Ile Lys Ser Gly Asn Glu Asn Gly Glu Phe Tyr Leu Arg Gln Thr
 325 330 335
 Ser Pro Val Ser Ala Met Leu Val Leu Val Lys Ser Leu Ser Gly Pro
 340 345 350
 Arg Glu His Ile Val Asp Leu Glu Met Leu Thr Val Ser Ser Ile Gly
 355 360 365
 Thr Phe Arg Thr Ser Ser Val Leu Arg Leu Thr Ile Ile Val Gly Pro
 370 375 380
 Phe Ser Phe
 385

<210> 684

<211> 531

<212> PRT

<213> Homo sapiens

<400> 684

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

10

20

30

Val Asn Leu Pro Gly Ala Ala Val Asp Leu Pro Ala Val Ser Glu Lys
 210 215 220
 Asp Ile Gln Asp Leu Lys Phe Gly Val Glu Gln Asp Val Asp Met Val
 225 230 235 240
 Phe Ala Ser Phe Ile Arg Lys Ala Ser Asp Val His Glu Val Arg Lys
 245 250 255
 Val Leu Gly Glu Lys Gly Lys Asn Ile Lys Ile Ile Ser Lys Ile Glu
 260 265 270
 Asn His Glu Gly Val Arg Arg Phe Asp Glu Ile Leu Glu Ala Ser Asp
 275 280 285
 Gly Ile Met Val Ala Arg Gly Asp Leu Gly Ile Glu Ile Pro Ala Glu
 290 295 300
 Lys Val Phe Leu Ala Gln Lys Met Met Ile Gly Arg Cys Asn Arg Ala
 305 310 315 320
 Gly Lys Pro Val Ile Cys Ala Thr Gln Met Leu Glu Ser Met Ile Lys
 325 330 335
 Lys Pro Arg Pro Thr Arg Ala Glu Gly Ser Asp Val Ala Asn Ala Val
 340 345 350
 Leu Asp Gly Ala Asp Cys Ile Met Leu Ser Gly Glu Thr Ala Lys Gly
 355 360 365
 Asp Tyr Pro Leu Glu Ala Val Arg Met Gln His Leu Ile Ala Arg Glu
 370 375 380
 Ala Glu Ala Ala Ile Tyr His Leu Gln Leu Phe Glu Glu Leu Arg Arg
 385 390 395 400
 Leu Ala Pro Ile Thr Ser Asp Pro Thr Glu Ala Thr Ala Val Gly Ala
 405 410 415
 Val Glu Ala Ser Phe Lys Cys Cys Ser Gly Ala Ile Ile Val Leu Thr
 420 425 430
 Lys Ser Gly Arg Ser Ala His Gln Val Ala Arg Tyr Arg Pro Arg Ala
 435 440 445
 Pro Ile Ile Ala Val Thr Arg Asn Pro Gln Thr Ala Arg Gln Ala His
 450 455 460
 Leu Tyr Arg Gly Ile Phe Pro Val Leu Cys Lys Asp Pro Val Gln Glu
 465 470 475 480
 Ala Trp Ala Glu Asp Val Asp Leu Arg Val Asn Phe Ala Met Asn Val
 485 490 495
 Gly Lys Ala Arg Gly Phe Phe Lys Lys Gly Asp Val Val Ile Val Leu
 500 505 510
 Thr Gly Trp Arg Pro Gly Ser Gly Phe Thr Asn Thr Met Arg Val Val
 515 520 525
 Pro Val Pro
 530

<210> 685
 <211> 480
 <212> PRT
 <213> Homo sapiens

<400> 685
 Met Ala Ala Arg Cys Ser Thr Arg Trp Leu Leu Val Val Val Gly Thr

10

20

30

```

1           5           10           15
Pro Arg Leu Pro Ala Ile Ser Gly Arg Gly Ala Arg Pro Pro Arg Glu
20
Gly Val Val Gly Ala Trp Leu Ser Arg Lys Leu Ser Val Pro Ala Phe
35
Ala Ser Ser Leu Thr Ser Cys Gly Pro Arg Ala Leu Leu Thr Leu Arg
50
Pro Gly Val Ser Leu Thr Gly Thr Lys His Asn Pro Phe Ile Cys Thr
65
Ala Ser Phe His Thr Ser Ala Pro Leu Ala Lys Glu Asp Tyr Tyr Gln
85
Ile Leu Gly Val Pro Arg Asn Ala Ser Gln Lys Glu Ile Lys Lys Ala
100
Tyr Tyr Gln Leu Ala Lys Lys Tyr His Pro Asp Thr Asn Lys Asp Asp
115
Pro Lys Ala Lys Glu Lys Phe Ser Gln Leu Ala Glu Ala Tyr Glu Val
130
Leu Ser Asp Glu Val Lys Arg Lys Gln Tyr Asp Ala Tyr Gly Ser Ala
145
Gly Phe Asp Pro Gly Ala Ser Gly Ser Gln His Ser Tyr Trp Lys Gly
165
Gly Pro Thr Val Asp Pro Glu Glu Leu Phe Arg Lys Ile Phe Gly Glu
180
Phe Ser Ser Ser Ser Phe Gly Asp Phe Gln Thr Val Phe Asp Gln Pro
195
Gln Glu Tyr Phe Met Glu Leu Thr Phe Asn Gln Ala Ala Lys Gly Val
210
Asn Lys Glu Phe Thr Val Asn Ile Met Asp Thr Cys Glu Arg Cys Asn
225
Gly Lys Gly Asn Glu Pro Gly Thr Lys Val Gln His Cys His Tyr Cys
245
Gly Gly Ser Gly Met Glu Thr Ile Asn Thr Gly Pro Phe Val Met Arg
260
Ser Thr Cys Arg Arg Cys Gly Gly Arg Gly Ser Ile Ile Ile Ser Pro
275
Cys Val Val Cys Arg Gly Ala Gly Gln Ala Lys Gln Lys Lys Arg Val
290
Met Ile Pro Val Pro Ala Gly Val Glu Asp Gly Gln Thr Val Arg Met
305
Pro Val Gly Lys Arg Glu Ile Phe Ile Thr Phe Arg Val Gln Lys Ser
325
Pro Val Phe Arg Arg Asp Gly Ala Asp Ile His Ser Asp Leu Phe Ile
340
Ser Ile Ala Gln Ala Leu Leu Gly Gly Thr Ala Arg Ala Gln Gly Leu
355
Tyr Glu Thr Ile Asn Val Thr Ile Pro Pro Gly Thr Gln Thr Asp Gln
370
Lys Ile Arg Met Gly Gly Lys Gly Ile Pro Arg Ile Asn Ser Tyr Gly
385
Tyr Gly Asp His Tyr Ile His Ile Lys Ile Arg Val Pro Lys Arg Leu
400

```

10

20

30

```

          405          410          415
Thr Ser Arg Gln Gln Ser Leu Ile Leu Ser Tyr Ala Glu Asp Glu Thr
          420          425          430
Asp Val Glu Gly Thr Val Asn Gly Val Thr Leu Thr Ser Ser Gly Gly
          435          440          445
Ser Thr Met Asp Ser Ser Ala Gly Ser Lys Ala Arg Arg Glu Ala Gly
          450          455          460
Glu Asp Glu Glu Gly Phe Leu Ser Lys Leu Lys Lys Met Phe Thr Ser
465          470          475          480

```

<210> 686

<211> 572

<212> PRT

<213> Homo sapiens

10

<400> 686

```

Met Ala Ala Pro Arg Pro Ser Pro Ala Ile Ser Val Ser Val Ser Ala
 1          5          10          15
Pro Ala Phe Tyr Ala Pro Gln Lys Lys Phe Gly Pro Val Val Ala Pro
          20          25          30
Lys Pro Lys Val Asn Pro Phe Arg Pro Gly Asp Ser Glu Pro Pro Pro
          35          40          45
Ala Pro Gly Ala Gln Arg Ala Gln Met Gly Arg Val Gly Glu Ile Pro
          50          55          60
Pro Pro Pro Pro Glu Asp Phe Pro Leu Pro Pro Pro Leu Ala Gly
65          70          75          80
Asp Gly Asp Asp Ala Glu Gly Ala Leu Gly Gly Ala Phe Pro Pro Pro
          85          90          95
Pro Pro Pro Ile Glu Glu Ser Phe Pro Pro Ala Pro Leu Glu Glu Glu
          100          105          110
Ile Phe Pro Ser Pro Pro Pro Pro Glu Glu Glu Gly Gly Pro Glu
          115          120          125
Ala Pro Ile Pro Pro Pro Pro Gln Pro Arg Glu Lys Val Ser Ser Ile
          130          135          140
Asp Leu Glu Ile Asp Ser Leu Ser Ser Leu Leu Asp Asp Met Thr Lys
145          150          155          160
Asn Asp Pro Phe Lys Ala Arg Val Ser Ser Gly Tyr Val Pro Pro Pro
          165          170          175
Val Ala Thr Pro Phe Ser Ser Lys Ser Ser Thr Lys Pro Ala Ala Gly
          180          185          190
Gly Thr Ala Pro Leu Pro Pro Trp Lys Ser Pro Ser Ser Ser Gln Pro
          195          200          205
Leu Pro Gln Val Pro Ala Pro Ala Gln Ser Gln Thr Gln Phe His Val
          210          215          220
Gln Pro Gln Pro Gln Pro Lys Pro Gln Val Gln Leu His Val Gln Ser
225          230          235          240
Gln Thr Gln Pro Val Ser Leu Ala Asn Thr Gln Pro Arg Gly Pro Pro
          245          250          255
Ala Ser Ser Pro Ala Pro Ala Pro Lys Phe Ser Pro Val Thr Pro Lys
          260          265          270

```

20

30

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

<210> 687

<211> 2861

<212> PRT

<213> Homo sapiens

<400> 687

Met Lys Ala Met Asp Val Leu Pro Ile Leu Lys Glu Lys Val Ala Tyr
 1 5 10 15
 Leu Ser Gly Gly Arg Asp Lys Arg Gly Gly Pro Ile Leu Thr Phe Pro
 20 25 30
 Ala Arg Ser Asn His Asp Arg Ile Arg Gln Glu Asp Leu Arg Arg Leu

10

20

30

30

30

835 840 845
 Gln Leu Arg His Leu Gln Ala Glu Val Lys Gln Val Leu Gly Trp Ile
 850 855 860
 Arg Asn Gly Glu Ser Met Leu Asn Ala Gly Leu Ile Thr Ala Ser Ser
 865 870 875 880
 Leu Gln Glu Ala Glu Gln Leu Gln Arg Glu His Glu Gln Phe Gln His
 885 890 895
 Ala Ile Glu Lys Thr His Gln Ser Ala Leu Gln Val Gln Gln Lys Ala
 900 905 910
 Glu Ala Met Leu Gln Ala Asn His Tyr Asp Met Asp Met Ile Arg Asp
 915 920 925
 Cys Ala Glu Lys Val Ala Ser His Trp Gln Gln Leu Met Leu Lys Met
 930 935 940
 Glu Asp Arg Leu Lys Leu Val Asn Ala Ser Val Ala Phe Tyr Lys Thr
 945 950 955 960
 Ser Glu Gln Val Cys Ser Val Leu Glu Ser Leu Glu Gln Glu Tyr Lys
 965 970 975
 Arg Glu Glu Asp Trp Cys Gly Gly Ala Asp Lys Leu Gly Pro Asn Ser
 980 985 990
 Glu Thr Asp His Val Thr Pro Met Ile Ser Lys His Leu Glu Gln Lys
 995 1000 1005
 Glu Ala Phe Leu Lys Ala Cys Thr Leu Ala Arg Arg Asn Ala Asp Val
 1010 1015 1020
 Phe Leu Lys Tyr Leu His Arg Asn Ser Val Asn Met Pro Gly Met Val
 1025 1030 1035 1040
 Thr His Ile Lys Ala Pro Glu Gln Gln Val Lys Asn Ile Leu Asn Glu
 1045 1050 1055
 Leu Phe Gln Arg Glu Asn Arg Val Leu His Tyr Trp Thr Met Arg Lys
 1060 1065 1070
 Arg Arg Leu Asp Gln Cys Gln Gln Tyr Val Val Phe Glu Arg Ser Ala
 1075 1080 1085
 Lys Gln Ala Leu Glu Trp Ile His Asp Asn Gly Glu Phe Tyr Leu Ser
 1090 1095 1100
 Thr His Thr Ser Thr Gly Ser Ser Ile Gln His Thr Gln Glu Leu Leu
 1105 1110 1115 1120
 Lys Glu His Glu Glu Phe Gln Ile Thr Ala Lys Gln Thr Lys Glu Arg
 1125 1130 1135
 Val Lys Leu Leu Ile Gln Leu Ala Asp Gly Phe Cys Glu Lys Gly His
 1140 1145 1150
 Ala His Ala Ala Glu Ile Lys Lys Cys Val Thr Ala Val Asp Lys Arg
 1155 1160 1165
 Tyr Arg Asp Phe Ser Leu Arg Met Glu Lys Tyr Arg Thr Ser Leu Glu
 1170 1175 1180
 Lys Ala Leu Gly Ile Ser Ser Asp Ser Asn Lys Ser Ser Lys Ser Leu
 1185 1190 1195 1200
 Gln Leu Asp Ile Ile Pro Ala Ser Ile Pro Gly Ser Glu Val Lys Leu
 1205 1210 1215
 Arg Asp Ala Ala His Glu Leu Asn Glu Glu Lys Arg Lys Ser Ala Arg
 1220 1225 1230
 Arg Lys Glu Phe Ile Met Ala Glu Leu Ile Gln Thr Glu Lys Ala Tyr

10

20

30

1235	1240	1245
Val Arg Asp Leu Arg Glu Cys Met Asp Thr Tyr Leu Trp Glu Met Thr		
1250	1255	1260
Ser Gly Val Glu Glu Ile Pro Pro Gly Ile Val Asn Lys Glu Leu Ile		
1265	1270	1275
Ile Phe Gly Asn Met Gln Glu Ile Tyr Glu Phe His Asn Asn Ile Phe		1280
1285	1290	1295
Leu Lys Glu Leu Glu Lys Tyr Glu Gln Leu Pro Glu Asp Val Gly His		
1300	1305	1310
Cys Phe Val Thr Trp Ala Asp Lys Phe Gln Met Tyr Val Thr Tyr Cys		
1315	1320	1325
Lys Asn Lys Pro Asp Ser Thr Gln Leu Ile Leu Glu His Ala Gly Ser		
1330	1335	1340
Tyr Phe Asp Glu Ile Gln Gln Arg His Gly Leu Ala Asn Ser Ile Ser		
1345	1350	1355
Ser Tyr Leu Ile Lys Pro Val Gln Arg Ile Thr Lys Tyr Gln Leu Leu		1360
1365	1370	1375
Leu Lys Glu Leu Leu Thr Cys Cys Glu Glu Gly Lys Gly Glu Ile Lys		
1380	1385	1390
Asp Gly Leu Glu Val Met Leu Ser Val Pro Lys Arg Ala Asn Asp Ala		
1395	1400	1405
Met His Leu Ser Met Leu Glu Gly Phe Asp Glu Asn Ile Glu Ser Gln		
1410	1415	1420
Gly Glu Leu Ile Leu Gln Glu Ser Phe Gln Val Trp Asp Pro Lys Thr		
1425	1430	1435
Leu Ile Arg Lys Gly Arg Glu Arg His Leu Phe Leu Phe Glu Met Ser		
1445	1450	1455
Leu Val Phe Ser Lys Glu Val Lys Asp Ser Ser Gly Arg Ser Lys Tyr		
1460	1465	1470
Leu Tyr Lys Ser Lys Leu Phe Thr Ser Glu Leu Gly Val Thr Glu His		
1475	1480	1485
Val Glu Gly Asp Pro Cys Lys Phe Ala Leu Trp Val Gly Arg Thr Pro		
1490	1495	1500
Thr Ser Asp Asn Lys Ile Val Leu Lys Ala Ser Ser Ile Glu Asn Lys		
1505	1510	1515
Gln Asp Trp Ile Lys His Ile Arg Glu Val Ile Gln Glu Arg Thr Ile		
1525	1530	1535
His Leu Lys Gly Ala Leu Lys Glu Pro Ile His Ile Pro Lys Thr Ala		
1540	1545	1550
Pro Ala Thr Arg Gln Lys Gly Arg Arg Asp Gly Glu Asp Leu Asp Ser		
1555	1560	1565
Gln Gly Asp Gly Ser Ser Gln Pro Asp Thr Ile Ser Ile Ala Ser Arg		
1570	1575	1580
Thr Ser Gln Asn Thr Leu Asp Ser Asp Lys Leu Ser Gly Gly Cys Glu		
1585	1590	1595
Leu Thr Val Val Ile His Asp Phe Thr Ala Cys Asn Ser Asn Glu Leu		1600
1605	1610	1615
Thr Ile Arg Arg Gly Gln Thr Val Glu Val Leu Glu Arg Pro His Asp		
1620	1625	1630
Lys Pro Asp Trp Cys Leu Val Arg Thr Thr Asp Arg Ser Pro Ala Ala		

10

20

30

1635 1640 1645
 Glu Gly Leu Val Pro Cys Gly Ser Leu Cys Ile Ala His Ser Arg Ser
 1650 1655 1660
 Ser Met Glu Met Glu Gly Ile Phe Asn His Lys Asp Ser Leu Ser Val
 1665 1670 1675 1680
 Ser Ser Asn Asp Ala Ser Pro Pro Ala Ser Val Ala Ser Leu Gln Pro
 1685 1690 1695
 His Met Ile Gly Ala Gln Ser Ser Pro Gly Pro Lys Arg Pro Gly Asn
 1700 1705 1710
 Thr Leu Arg Lys Trp Leu Thr Ser Pro Val Arg Arg Leu Ser Ser Gly
 1715 1720 1725
 Lys Ala Asp Gly His Val Lys Lys Leu Ala His Lys His Lys Lys Ser
 1730 1735 1740
 Arg Glu Val Arg Lys Ser Ala Asp Ala Gly Ser Gln Lys Asp Ser Asp
 1745 1750 1755 1760
 Asp Ser Ala Ala Thr Pro Gln Asp Glu Thr Val Glu Glu Arg Gly Arg
 1765 1770 1775
 Asn Glu Gly Leu Ser Ser Gly Thr Leu Ser Lys Ser Ser Ser Ser Gly
 1780 1785 1790
 Met Gln Ser Cys Gly Glu Glu Glu Gly Glu Glu Gly Ala Asp Ala Val
 1795 1800 1805
 Pro Leu Pro Pro Pro Met Ala Ile Gln Gln His Ser Leu Leu Gln Pro
 1810 1815 1820
 Asp Ser Gln Asp Asp Lys Ala Ser Ser Arg Leu Leu Val Arg Pro Thr
 1825 1830 1835 1840
 Ser Ser Glu Thr Pro Ser Ala Ala Glu Leu Val Ser Ala Ile Glu Glu
 1845 1850 1855
 Leu Val Lys Ser Lys Met Ala Leu Glu Asp Arg Pro Ser Ser Leu Leu
 1860 1865 1870
 Val Asp Gln Gly Asp Ser Ser Ser Pro Ser Phe Asn Pro Ser Asp Asn
 1875 1880 1885
 Ser Leu Leu Ser Ser Ser Ser Pro Ile Asp Glu Met Glu Glu Arg Lys
 1890 1895 1900
 Ser Ser Ser Leu Lys Arg Arg His Tyr Val Leu Gln Glu Leu Val Glu
 1905 1910 1915 1920
 Thr Glu Arg Asp Tyr Val Arg Asp Leu Gly Tyr Val Val Glu Gly Tyr
 1925 1930 1935
 Met Ala Leu Met Lys Glu Asp Gly Val Pro Asp Asp Met Lys Gly Lys
 1940 1945 1950
 Asp Lys Ile Val Phe Gly Asn Ile His Gln Ile Tyr Asp Trp His Arg
 1955 1960 1965
 Asp Phe Phe Leu Gly Glu Leu Glu Lys Cys Leu Glu Asp Pro Glu Lys
 1970 1975 1980
 Leu Gly Ser Leu Phe Val Lys His Glu Arg Arg Leu His Met Tyr Ile
 1985 1990 1995 2000
 Ala Tyr Cys Gln Asn Lys Pro Lys Ser Glu His Ile Val Ser Glu Tyr
 2005 2010 2015
 Ile Asp Thr Phe Phe Glu Asp Leu Lys Gln Arg Leu Gly His Arg Leu
 2020 2025 2030
 Gln Leu Thr Asp Leu Leu Ile Lys Pro Val Gln Arg Ile Met Lys Tyr

10

20

30

2035	2040	2045
Gln Leu Leu Leu Lys Asp Phe Leu Lys Tyr Ser Lys Lys Ala Ser Leu		
2050	2055	2060
Asp Thr Ser Glu Leu Glu Arg Ala Val Glu Val Met Cys Ile Val Pro		
2065	2070	2075
Arg Arg Cys Asn Asp Met Met Asn Val Gly Arg Leu Gln Gly Phe Asp		2080
2085	2090	2095
Gly Lys Ile Val Ala Gln Gly Lys Leu Leu Leu Gln Asp Thr Phe Leu		
2100	2105	2110
Val Thr Asp Gln Asp Ala Gly Leu Leu Pro Arg Cys Arg Glu Arg Arg		
2115	2120	2125
Ile Phe Leu Phe Glu Gln Ile Val Ile Phe Ser Glu Pro Leu Asp Lys		
2130	2135	2140
Lys Lys Gly Phe Ser Met Pro Gly Phe Leu Phe Lys Asn Ser Ile Lys		
2145	2150	2155
Val Ser Cys Leu Cys Leu Glu Glu Asn Val Glu Asn Asp Pro Cys Lys		2160
2165	2170	2175
Phe Ala Leu Thr Ser Arg Thr Gly Asp Val Val Glu Thr Phe Ile Leu		
2180	2185	2190
His Ser Ser Ser Pro Ser Val Arg Gln Thr Trp Ile His Glu Ile Asn		
2195	2200	2205
Gln Ile Leu Glu Asn Gln Arg Asn Phe Leu Asn Ala Leu Thr Ser Pro		
2210	2215	2220
Ile Glu Tyr Gln Arg Asn His Ser Gly Gly Gly Gly Gly Gly Ser		
2225	2230	2235
Gly Ala Ala Ala Gly Val Gly Ala Ala Ala Ala Gly Pro Pro Val		
2245	2250	2255
Ala Ala Ala Ala Thr Val Ala Ala Pro Ala Ala Ala Ala Ala Pro Pro		
2260	2265	2270
Ala Arg Ala Gly Ala Gly Pro Pro Gly Ser Pro Ser Leu Ser Asp Thr		
2275	2280	2285
Thr Pro Pro Cys Trp Ser Pro Leu Gln Pro Arg Ala Arg Gln Arg Gln		
2290	2295	2300
Thr Arg Cys Gln Ser Glu Ser Ser Ser Ser Ser Asn Ile Ser Thr Met		
2305	2310	2315
Leu Val Thr His Asp Tyr Thr Ala Val Lys Glu Asp Glu Ile Asn Val		
2325	2330	2335
Tyr Gln Gly Glu Val Val Gln Ile Leu Ala Ser Asn Gln Gln Asn Met		
2340	2345	2350
Phe Leu Val Phe Arg Ala Ala Thr Asp Gln Cys Pro Ala Ala Glu Gly		
2355	2360	2365
Trp Ile Pro Gly Phe Val Leu Gly His Thr Ser Ala Val Ile Val Glu		
2370	2375	2380
Asn Pro Asp Gly Thr Leu Lys Lys Ser Thr Ser Trp His Thr Ala Leu		
2385	2390	2395
Arg Leu Arg Lys Lys Ser Glu Lys Lys Asp Lys Asp Gly Lys Arg Glu		2400
2405	2410	2415
Gly Lys Leu Glu Asn Gly Tyr Arg Lys Ser Arg Glu Gly Leu Ser Asn		
2420	2425	2430
Lys Val Ser Val Lys Leu Leu Asn Pro Asn Tyr Ile Tyr Asp Val Pro		

10

20

30

2435 2440 2445
 Pro Glu Phe Val Ile Pro Leu Ser Glu Val Thr Cys Glu Thr Gly Glu
 2450 2455 2460
 Thr Val Val Leu Arg Cys Arg Val Cys Gly Arg Pro Lys Ala Ser Ile
 2465 2470 2475 2480
 Thr Trp Lys Gly Pro Glu His Asn Thr Leu Asn Asn Asp Gly His Tyr
 2485 2490 2495
 Ser Ile Ser Tyr Ser Asp Leu Gly Glu Ala Thr Leu Lys Ile Val Gly
 2500 2505 2510
 Val Thr Thr Glu Asp Asp Gly Ile Tyr Thr Cys Ile Ala Val Asn Asp
 2515 2520 2525
 Met Gly Ser Ala Ser Ser Ser Ala Ser Leu Arg Val Leu Gly Pro Gly
 2530 2535 2540
 Met Asp Gly Ile Met Val Thr Trp Lys Asp Asn Phe Asp Ser Phe Tyr
 2545 2550 2555 2560
 Ser Glu Val Ala Glu Leu Gly Arg Gly Arg Phe Ser Val Val Lys Lys
 2565 2570 2575
 Cys Asp Gln Lys Gly Thr Lys Arg Ala Val Ala Thr Lys Phe Val Asn
 2580 2585 2590
 Lys Lys Leu Met Lys Arg Asp Gln Val Thr His Glu Leu Gly Ile Leu
 2595 2600 2605
 Gln Ser Leu Gln His Pro Leu Leu Val Gly Leu Leu Asp Thr Phe Glu
 2610 2615 2620
 Thr Pro Thr Ser Tyr Ile Leu Val Leu Glu Met Ala Asp Gln Gly Arg
 2625 2630 2635 2640
 Leu Leu Asp Cys Val Val Arg Trp Gly Ser Leu Thr Glu Gly Lys Ile
 2645 2650 2655
 Arg Ala His Leu Gly Glu Val Leu Glu Ala Val Arg Tyr Leu His Asn
 2660 2665 2670
 Cys Arg Ile Ala His Leu Asp Leu Lys Pro Glu Asn Ile Leu Val Asp
 2675 2680 2685
 Glu Ser Leu Ala Lys Pro Thr Ile Lys Leu Ala Asp Phe Gly Asp Ala
 2690 2695 2700
 Val Gln Leu Asn Thr Thr Tyr Tyr Ile His Gln Leu Leu Gly Asn Pro
 2705 2710 2715 2720
 Glu Phe Ala Ala Pro Glu Ile Ile Leu Gly Asn Pro Val Ser Leu Thr
 2725 2730 2735
 Ser Asp Thr Trp Ser Val Gly Val Leu Thr Tyr Val Leu Leu Ser Gly
 2740 2745 2750
 Val Ser Pro Phe Leu Asp Asp Ser Val Glu Glu Thr Cys Leu Asn Ile
 2755 2760 2765
 Cys Arg Leu Asp Phe Ser Phe Pro Asp Asp Tyr Phe Lys Gly Val Ser
 2770 2775 2780
 Gln Lys Ala Lys Glu Phe Val Cys Phe Leu Leu Gln Glu Asp Pro Ala
 2785 2790 2795 2800
 Lys Arg Pro Ser Ala Ala Leu Ala Leu Gln Glu Gln Trp Leu Gln Ala
 2805 2810 2815
 Gly Asn Gly Arg Ser Thr Gly Val Leu Asp Thr Ser Arg Leu Thr Ser
 2820 2825 2830
 Phe Ile Glu Arg Arg Lys His Gln Asn Asp Val Arg Pro Ile Arg Ser

10

20

30

2835 2840 2845
 Ile Lys Asn Phe Leu Gln Ser Arg Leu Leu Pro Arg Val
 2850 2855 2860

<210> 688
 <211> 271
 <212> PRT
 <213> Homo sapiens

<400> 688
 Met Val Leu Ile Lys Glu Phe Arg Val Val Leu Pro Cys Ser Val Gln
 1 5 10 15
 Glu Tyr Gln Val Gly Gln Leu Tyr Ser Val Ala Glu Ala Ser Lys Asn
 20 25 30
 Glu Thr Gly Gly Gly Glu Gly Ile Glu Val Leu Lys Asn Glu Pro Tyr
 35 40 45
 Glu Lys Asp Gly Glu Lys Gly Gln Tyr Thr His Lys Ile Tyr His Leu
 50 55 60
 Lys Ser Lys Val Pro Ala Phe Val Arg Met Ile Ala Pro Glu Gly Ser
 65 70 75 80
 Leu Val Phe His Glu Lys Ala Trp Asn Ala Tyr Pro Tyr Cys Arg Thr
 85 90 95
 Ile Val Thr Asn Glu Tyr Met Lys Asp Asp Phe Phe Ile Lys Ile Glu
 100 105 110
 Thr Trp His Lys Pro Asp Leu Gly Thr Leu Glu Asn Val His Gly Leu
 115 120 125
 Asp Pro Asn Thr Trp Lys Thr Val Glu Ile Val His Ile Asp Ile Ala
 130 135 140
 Asp Arg Ser Gln Val Glu Pro Ala Asp Tyr Lys Ala Asp Glu Asp Pro
 145 150 155 160
 Ala Leu Phe Gln Ser Val Lys Thr Lys Arg Gly Pro Leu Gly Pro Asn
 165 170 175
 Trp Lys Lys Glu Leu Ala Asn Ser Pro Asp Cys Pro Gln Met Cys Ala
 180 185 190
 Tyr Lys Leu Val Thr Ile Lys Phe Lys Trp Trp Gly Leu Gln Ser Lys
 195 200 205
 Val Glu Asn Phe Ile Gln Lys Gln Glu Lys Arg Ile Phe Thr Asn Phe
 210 215 220
 His Arg Gln Leu Phe Cys Trp Ile Asp Lys Trp Ile Asp Leu Thr Met
 225 230 235 240
 Glu Asp Ile Arg Arg Met Glu Asp Glu Thr Gln Lys Glu Leu Glu Thr
 245 250 255
 Met Arg Lys Arg Gly Ser Val Arg Gly Thr Ser Ala Ala Asp Val
 260 265 270

<210> 642
 <211> 11
 <212> DNA
 <213> Homo sapiens

10

20

30

<400> 642
gatgatgatg a 11

<210> 643
<211> 11
<212> DNA
<213> Homo sapiens

<400> 643
gatgatgatg a 11

<210> 644
<211> 11
<212> DNA
<213> Homo sapiens

10

<400> 644
gatgatgatg a 11

<210> 645
<211> 11
<212> DNA
<213> Homo sapiens

<400> 645
gatgatgatg a 11

<210> 646
<211> 11
<212> DNA
<213> Homo sapiens

20

<400> 646
gatgatgatg a 11

<210> 647
<211> 11
<212> DNA
<213> Homo sapiens

<400> 647
gatgatgatg a 11

30

<210> 648
<211> 11
<212> DNA
<213> Homo sapiens

<400> 648
gatgatgatg a 11

<210> 649
<211> 11
<212> DNA
<213> Homo sapiens

<400> 649
gatgatgatg a 11

<210> 650
<211> 11
<212> DNA
<213> Homo sapiens

<400> 650
gatgatgatg a 11

10

<210> 651
<211> 11
<212> DNA
<213> Homo sapiens

<400> 651
gatgatgatg a 11

<210> 652
<211> 11
<212> DNA
<213> Homo sapiens

20

<400> 652
gatgatgatg a 11

<210> 653
<211> 11
<212> DNA
<213> Homo sapiens

<400> 653
gatgatgatg a 11

<210> 654
<211> 11
<212> DNA
<213> Homo sapiens

30

<400> 654
gatgatgatg a 11

<210> 655
<211> 11

<212> DNA
<213> Homo sapiens

<400> 655
gatgatgatg a 11

<210> 656
<211> 11
<212> DNA
<213> Homo sapiens

<400> 656
gatgatgatg a 11

10

<210> 657
<211> 11
<212> DNA
<213> Homo sapiens

<400> 657
gatgatgatg a 11

<210> 658
<211> 11
<212> DNA
<213> Homo sapiens

<400> 658
gatgatgatg a 11

20

<210> 659
<211> 11
<212> DNA
<213> Homo sapiens

<400> 659
gatgatgatg a 11

<210> 660
<211> 11
<212> DNA
<213> Homo sapiens

<400> 660
gatgatgatg a 11

30

<210> 661
<211> 11
<212> DNA
<213> Homo sapiens

<400> 661 gatgatgatg a	11	
<210> 662 <211> 11 <212> DNA <213> Homo sapiens		
<400> 662 gatgatgatg a	11	
<210> 663 <211> 11 <212> DNA <213> Homo sapiens		10
<400> 663 gatgatgatg a	11	
<210> 664 <211> 11 <212> DNA <213> Homo sapiens		
<400> 664 gatgatgatg a	11	
<210> 665 <211> 11 <212> DNA <213> Homo sapiens		20
<400> 665 gatgatgatg a	11	
<210> 666 <211> 11 <212> DNA <213> Homo sapiens		
<400> 666 gatgatgatg a	11	30
<210> 667 <211> 11 <212> DNA <213> Homo sapiens		
<400> 667 gatgatgatg a	11	

<210> 668
<211> 11
<212> DNA
<213> Homo sapiens

<400> 668
gatgatgatg a 11

<210> 669
<211> 11
<212> DNA
<213> Homo sapiens

<400> 669
gatgatgatg a 11

<210> 670
<211> 11
<212> DNA
<213> Homo sapiens

<400> 670
gatgatgatg a 11

<210> 671
<211> 11
<212> DNA
<213> Homo sapiens

<400> 671
gatgatgatg a 11

<210> 672
<211> 11
<212> DNA
<213> Homo sapiens

<400> 672
gatgatgatg a 11

<210> 673
<211> 11
<212> DNA
<213> Homo sapiens

<400> 673
gatgatgatg a 11

<210> 674
<211> 11

10

20

30

<212> DNA
<213> Homo sapiens

<400> 674
gatgatgatg a 11

<210> 675
<211> 11
<212> DNA
<213> Homo sapiens

<400> 675
gatgatgatg a 11

10

<210> 676
<211> 11
<212> DNA
<213> Homo sapiens

<400> 676
gatgatgatg a 11

<210> 677
<211> 11
<212> DNA
<213> Homo sapiens

<400> 677
gatgatgatg a 11

20

<210> 678
<211> 11
<212> DNA
<213> Homo sapiens

<400> 678
gatgatgatg a 11

<210> 679
<211> 11
<212> DNA
<213> Homo sapiens

<400> 679
gatgatgatg a 11

30

<210> 680
<211> 11
<212> DNA
<213> Homo sapiens

<400> 680 gatgatgatg a	11	
<210> 681 <211> 11 <212> DNA <213> Homo sapiens		
<400> 681 gatgatgatg a	11	
<210> 682 <211> 11 <212> DNA <213> Homo sapiens		10
<400> 682 gatgatgatg a	11	
<210> 683 <211> 11 <212> DNA <213> Homo sapiens		
<400> 683 gatgatgatg a	11	
<210> 684 <211> 11 <212> DNA <213> Homo sapiens		20
<400> 684 gatgatgatg a	11	
<210> 685 <211> 11 <212> DNA <213> Homo sapiens		
<400> 685 gatgatgatg a	11	30
<210> 686 <211> 11 <212> DNA <213> Homo sapiens		
<400> 686 gatgatgatg a	11	

<210> 687
 <211> 11
 <212> DNA
 <213> Homo sapiens

<400> 687
 gatgatgatg a 11

<210> 688
 <211> 11
 <212> DNA
 <213> Homo sapiens

<400> 688
 gatgatgatg a 11

<210> 689
 <211> 21
 <212> DNA
 <213> Homo sapiens

<400> 689
 gtcagcctgg aggagttctc a 21

<210> 690
 <211> 22
 <212> DNA
 <213> Homo sapiens

<400> 690
 tcacccttgg caatacagat gt 22

<210> 691
 <211> 21
 <212> DNA
 <213> Homo sapiens

<400> 691
 cccacccatg tgcccgtgac a 21

<210> 692
 <211> 19
 <212> DNA
 <213> Homo sapiens

<400> 692
 cgtgattgcc gacgatctc 19

<210> 693
 <211> 19

10

20

30

<212> DNA
<213> Homo sapiens

<400> 693
ttccgggcgc tagtcttgt 19

<210> 694
<211> 26
<212> DNA
<213> Homo sapiens

<400> 694
cgcacccgtt cggctctgacg cagtac 26 10

<210> 695
<211> 22
<212> DNA
<213> Homo sapiens

<400> 695
ctttccccac gagtatgttg gt 22

<210> 696
<211> 19
<212> DNA
<213> Homo sapiens

<400> 696
aagggaccgt gctgtgagc 19 20

<210> 697
<211> 29
<212> DNA
<213> Homo sapiens

<400> 697
agccccctcat gtgcctctca acttcatag 29

<210> 698
<211> 6
<212> PRT
<213> Artificial Sequence 30

<220>
<223> partial signal sequence of HBM and Zmax1

<400> 698
Ser Pro Ala His Ser Ser
1 5

<210> 699
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic duplex

<400> 699
tcgaccacca tggaggcagc gccgc

25

<210> 700
<211> 25
<212> DNA
<213> Artificial Sequence

10

<220>
<223> synthetic duplex

<400> 700
ggtggtacct cegtcgcggc gggcc

25

<210> 701
<211> 55
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic duplex

20

<400> 701
catgtgtgaa atcaccaagc cgccttcaga cgacagcccg gccacagca gtggc

55

<210> 702
<211> 63
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic duplex

<400> 702
tcgagtacac actttagtgg ttccggcggga gtctgctgtc gggccgggtg tcgtcacgc 60
cgg 63

30

<210> 703
<211> 17
<212> PRT
<213> Artificial Sequence

<220>
<223> peptide from Zmax1 sequence (SEQ ID NO: 3)

<400> 703
Val Glu Thr Pro Arg Ile Glu Arg Ala Gly Met Asp Gly Ser Thr Arg
1 5 10 15
Lys

<210> 704
<211> 15
<212> PRT
<213> Artificial Sequence

10

<220>
<223> peptide from Zmax1 sequence (SEQ ID NO: 3)

<400> 704
Asn Lys Arg Thr Gly Gly Lys Arg Lys Glu Ile Leu Ser Ala Leu
1 5 10 15

<210> 705
<211> 12
<212> PRT
<213> Artificial Sequence

20

<220>
<223> peptide from Zmax1 sequence (SEQ ID NO: 3)

<400> 705
Glu Arg Gln Pro Phe Phe His Thr Arg Cys Glu Glu
1 5 10

<210> 706
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> peptide from Zmax1 sequence (SEQ ID NO: 3)

<400> 706
Val Asp Gly Thr Lys Arg Arg Thr Leu Leu Glu Asp Lys Leu Pro
1 5 10 15

30

<210> 707

<211> 15
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> peptide from Zmax1 sequence (SEQ ID NO: 3)

<400> 707
 Asp Gly Leu Asn Asp Cys Met His Asn Asn Gly Gln Cys Gly Gln
 1 5 10 15

<210> 708
 <211> 12
 <212> PRT
 <213> Artificial Sequence

10

<220>
 <223> peptide from Zmax1 sequence (SEQ ID NO: 3)

<400> 708
 Pro Phe Val Leu Thr Ser Leu Ser Gln Gly Gln Asn
 1 5 10

<210> 709
 <211> 15
 <212> PRT
 <213> Artificial Sequence

20

<220>
 <223> peptide from Zmax1 sequence (SEQ ID NO: 3)

<400> 709
 Tyr Ala Gly Ala Asn Gly Pro Phe Pro His Glu Tyr Val Ser Gly
 1 5 10 15

<210> 710
 <211> 13
 <212> PRT
 <213> Artificial Sequence

30

<220>
 <223> peptide from Zmax1 sequence (SEQ ID NO: 3)

<400> 710
 Ala Cys Gly Lys Ser Met Met Ser Ser Val Ser Leu Met
 1 5 10

<210> 711
 <211> 18
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> peptide from Zmax1 sequence (SEQ ID NO: 3)

<400> 711
 Arg Trp Lys Ala Ser Lys Tyr Tyr Leu Asp Leu Asn Ser Asp Ser Asp
 1 5 10 15
 Pro Tyr

10

<210> 712
 <211> 15
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> peptide from Zmax1 sequence (SEQ ID NO: 3)

<400> 712
 Ser Gly Trp Asn Glu Cys Ala Ser Ser Asn Gly His Cys Ser His
 1 5 10 15

20

<210> 713
 <211> 14
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> peptide from Zmax1 sequence (SEQ ID NO: 3)

<400> 713
 Asn Gly Asp Ala Asn Cys Gln Asp Lys Ser Asp Glu Lys Asn
 1 5 10

<210> 714
 <211> 22
 <212> PRT
 <213> Homo sapiens

30

<400> 714
 Ile Lys Arg Ala Lys Asp Asp Gly Thr Gln Pro Phe Val Leu Thr Ser
 1 5 10 15
 Leu Ser Gln Gly Gln Asn

20

<210> 715
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 715
cgtaacgtaaa gcggccgctt ggcaatacag atgtggga

38

10

<210> 716
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 716
cgtaacgtaaa gcggccgcag tagctcctct cggtaggc

37

<210> 717
<211> 37
<212> DNA
<213> Artificial Sequence

20

<220>
<223> primer

<400> 717
cgtaacgtaaa gcggccgcgc tcacatcatgga ctttcgc

37

<210> 718
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

30

<400> 718
cgtaacgtaaa gcggccgcgc actgctgttt gatgagg

37

<210> 719
<211> 38
<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 719

cgtacgtaaa gcggccgcga gtgtggaaga aaggctgc 38

<210> 720

<211> 38

<212> DNA

<213> Artificial Sequence

10

<220>

<223> primer

<400> 720

cgtacgtaaa gcggccgcag tagagcttcc cctcctgc 38

<210> 721

<211> 38

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 721

cgtacgtaaa gcggccgcgt ccatcacgaa gtccaggt 38

20

<210> 722

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 722

cagcggcctg gaggatgc 18

<210> 723

<211> 19

<212> DNA

<213> Artificial Sequence

30

<220>

<223> primer

<400> 723

cggtccagta gaggtttcg

19

<210> 724

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 724

catcagccgc gccttcatg

19

<210> 725

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 725

cctgcatgtt ggtgaagtac

20

<210> 726

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 726

gcccgaacc tctactggac cgac

24

<210> 727

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 727

gcccaccca tcacagttca catt

24

<210> 728

<211> 18

<212> DNA

<213> Artificial Sequence

10

20

30

<220>
<223> primer

<400> 728
cagcggcctg gaggatgc 18

<210> 729
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> primer 10

<400> 729
cgggtacatg tactggacag ctgattagc 29

<210> 730
<211> 21
<212> DNA
<213> Homo sapiens

<400> 730
gtcagcctgg aggagttctc a 21

<210> 731
<211> 22
<212> DNA 20
<213> Homo sapiens

<400> 731
tcacccttgg caatacagat gt 22

<210> 732
<211> 21
<212> DNA
<213> Homo sapiens

<400> 732
cccacccatg tgcccgtagc a 21

<210> 733
<211> 19
<212> DNA
<213> Homo sapiens 30

<400> 733
cgtgattgcc gacgatctc 19

<210> 734

<211> 19
<212> DNA
<213> Homo sapiens

<400> 734
ttccggccgc tagtcttgt 19

<210> 735
<211> 26
<212> DNA
<213> Homo sapiens

<400> 735
cgcacccggt cggctctgacg cagtac 26

<210> 736
<211> 22
<212> DNA
<213> Mouse

<400> 736
ctttccccac gagtatgttg gt 22

<210> 737
<211> 19
<212> DNA
<213> Mouse

<400> 737
aagggaccgt gctgtgagc 19

<210> 738
<211> 29
<212> DNA
<213> Mouse

<400> 738
agccctcat gtgcctctca acttcacag 29

<210> 739
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 739
gagcgggcag ggatggatgg 20

10

20

30

<210> 740
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 740
aggttggcac ggtggatgaa gc 22

<210> 741
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 741
aagcttggtt aaactgggca tgggtggcaca tggttgtaat 40

<210> 742
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 742
gggcttccac ccagtcagtc cagtacatgt acct 34

<210> 743
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 743
ctgactgggt ggaagcacc cggatcgagc 30

<210> 744
<211> 33
<212> DNA
<213> Artificial Sequence

<220>

10

20

30

<223> primer

<400> 744

gaattcatcg gtacctgtgc ggccgttca ttg

33

<210> 745

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

10

<400> 745

tgagatgtcc tgtctgtggc

20

<210> 746

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 746

tccttccttc cctacagttg

20

<210> 747

<211> 25

<212> DNA

<213> Artificial Sequence

20

<220>

<223> primer

<400> 747

cctaaggatc tccttgtgtc tgtgg

25

<210> 748

<211> 22

<212> DNA

<213> Artificial Sequence

30

<220>

<223> primer

<400> 748

ctgcagcagg tcagtagcct gc

22

<210> 749

<211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> primer

<400> 749
 tgagatgtcc tgtctgtggc 20

<210> 750
 <211> 20
 <212> DNA
 <213> Artificial Sequence

10

<220>
 <223> primer

<400> 750
 tccttccttc cctacagttg 20

<210> 751
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> primer

20

<400> 751
 gcttctggcg tgtgaccggc g 21

<210> 752
 <211> 22
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> primer

<400> 752
 gccgcacagc gccagcagca gc 22

30

<210> 753
 <211> 22
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> primer

<400> 753 caccacgccc ccacagccag ta	22	
<210> 754 <211> 24 <212> DNA <213> Artificial Sequence		
<220> <223> primer		
<400> 754 atttgccctc ccatatgtcc ttcc	24	10
<210> 755 <211> 24 <212> DNA <213> Artificial Sequence		
<220> <223> primer		
<400> 755 ttcctcccag ccctcctcca tcag	24	
<210> 756 <211> 22 <212> DNA <213> Artificial Sequence		20
<220> <223> primer		
<400> 756 gccgcacagc gccagcagca gc	22	
<210> 757 <211> 19 <212> DNA <213> Artificial Sequence		
<220> <223> primer		30
<400> 757 gaatggcgcc cccgacgac	19	
<210> 758 <211> 24 <212> DNA		

<213> Artificial Sequence

<220>

<223> primer

<400> 758

gctcccatc atcagttcca tagg

24

<210> 759

<211> 5278

<212> DNA

<213> Mouse

<400> 759

tctagactcg aggcggcgc ccattgtgca etaaagcgcc gccgcgcgc catggagccc 60
 gagtgcgcgc ggcgcgggccc cgtccggccc ccggacaaca tggaggcagc gccgcgcggg 120
 ccgcgcgtggc cgtctgtgct gctgctgctg ctgctgctgg cgtctgtgcg ctgcccgccc 180
 cccgcgcgcgc cctcgcgcgt cctgtctatt gccaaacggc gggacgtacg gctggtggac 240
 gccgcgcggag tcaagctgga gtccaccatc gtggtcagcg gcctggagga tgcggccgca 300
 gtggacttcc agttttccaa gggagccgtg tactggacag acgtgagcga ggaggccatc 360
 aagcagacct acctgaacca gacgggggccc gccgtgcaga acgtggtcat ctccggcctg 420
 gtctctcccg acggcctcgc ctgcgactgg gtgggcaaga agctgtactg gacggactca 480
 gagaccaacc gcatcgaggt ggccaacctc aatggcacat cccggaaggt gctcttctgg 540
 caggaccttg accagccgag gcccatcgcc ttggaccccg ctcaagggtg catgtactgg 600
 acagactggg ttgagacgcc ccggtattgag cgggcaggga tggatggcag caccgggaag 660
 atcattgttg actcggacat ttactggccc aatggactga ccatcgacct ggaggagcag 720
 aagctctact gggctgacgc caagctcagc ttcatccacc gtgccaacct ggacggctcg 780
 ttccggcaga aggtggtgga gggcagcctg acgcacccct tcgcctgac gctctccggg 840
 gacactctgt actggacaga ctggcagacc cgtccatcc atgctgcaa caagcgact 900
 ggggggaaga ggaaggagat cctgagtgc ctctactcac ccatggacat ccaggtgctg 960
 agccaggagc ggcagccttt ctccacact cgtctgtagg aggacaatgg cggctgctcc 1020
 cacctgtgcc tgctgtcccc aagcagacct ttctacacat gcgcctgccc caccgggtgtg 1080
 cagctgcagg acaacggcag gacgtgtaag gcaggagccg aggaggtgct gctgctggcc 1140
 cggcggacgg acctacggag gatctcgtg gacacgcggc acttcaccga catcgtgctg 1200
 cagggtggacg acatccggca cgcattgcc atcgactacg acccgctaga gggctatgtc 1260
 tactggacag atgacgaggt gcgggccatc cgcaggcgct acctgacgg gtctggggcg 1320
 cagacgctgg tcaacaccga gatcaacgac ccgatggca tgcggtcga ctgggtggcc 1380
 cgaaacctct actggaccga caccggcacg gaccgcacg aggtgacgg cctcaacggc 1440
 acctcccgca agatcctggt gtccgaggac ctggacgagc cccgagccat cgcactgcac 1500
 cccgtgatgg gcctcatgta ctggacagac tggggagaga accctaaaat cgagtgtgcc 1560
 aacttgatg ggcaggagcg gcgtgtgctg gtcaatgcct cctcgggtg gcccaacggc 1620
 ctggccctgg acctgcagga ggggaagctc tactggggag acgccaagac agacaagatc 1680
 gagggtgatca atgttgatgg gacgaagagg cggaccctcc tggaggacaa gctcccgcac 1740
 attttcgggt tcacgctgct gggggacttc atctactgga ctgactggca gcgcgcagc 1800
 atcgagcggg tgcacaaggt caaggccagc cgggacgtca tcattgacca gctgcccagc 1860
 ctgatggggc tcaaagctgt gaatgtggcc aaggtcgtcg gaaccaaccc gtgtgcggac 1920
 aggaacgggg ggtgcagcca cctgtgcttc ttcacacccc acgcaacccc gtgtggctgc 1980
 cccatcgccc tggagctgct gagtgcacatg aagacctgca tcgtgcctga ggccttcttg 2040
 gtcttcacca gcagagccgc catccacagg atctccctcg agaccaataa caacgacgtg 2100
 cccatccgcg tcacgggcgt caaggaggcc tcagccctgg actttgatgt gtccaacaac 2160

10

20

30

cacatctact ggacagacgt cagcctgaag accatcagcc gcgccttcat gaacgggagc 2220
tcggtgggagc acgtgggtgga gtttggcctt gactaccccg agggcatggc cgttgactgg 2280
atgggcaaga acctctactg ggccgacact gggaccaaca gaatcgaagt ggcgcggctg 2340
gacgggcagt tccggcaagt cctcgtgtgg agggacttgg acaaccogag gtcgctggcc 2400
ctggatccca ccaagggcta catctactgg accgagtggg gcggcaagcc gaggatcgtg 2460
cgggccttca tggacgggac caactgcatg acgctgggtg acaagggtggg cggggccaac 2520
gacctcacca ttgactacgc tgaccagcgc ctctactgga ccgacctgga caccaacatg 2580
atcgagtcgt ccaacatgct gggtcaggag cgggtcgtga ttgcccagca tctcccgac 2640
ccgttcggtc tgacgcagta cagcgattat atctactgga cagactggaa tctgcacagc 2700
attgagcggg ccgacaagac tagcggcccg aaccgcaccc tcatccaggg ccacctggac 2760
ttcgtgatgg tcatcctggg gttccactcc tcccgcagg atggcctcaa tgactgtatg 2820
cacaacaacg ggcagtgtgg gcagctgtgc cttgccatcc ccggcggcca ccgctgcggc 2880
tgccctcac actacaccct ggaccccagc agccgcaact gcagcccgc caccaccttc 2940
ttgtgttca gccagaaatc tgccatcagt cggatgatcc cggacgacca gcacagcccg 3000
gatctcatcc tgcccctgca tggactgagg aacgtcaaag ccactgacta tgaccactg 3060
gacaagtcca tctactgggt ggatgggccc cagaacatca agcgagccaa ggacgacggg 3120
accagccct ttgttttgac ctctctgagc caaggccaaa accagacag gcagccccc 3180
gacctcagca tgcacatcta cagccggaca ctgttctgga cgtgcgaggc caccaatacc 3240
atcaacgtcc acaggctgag cggggaagcc atgggggtgg tgctgcgtgg ggaccgcgac 3300
aagcccaggg ccategtcgt caacgcggag caggggtacc tgtacttcac caaatgcag 3360
gacggggcag ccaagatcga acgcgcagcc ctggacggca ccgagcgca ggtcctcttc 3420
accacgggcc tcatccgccc tgtggccctg gtggtagaca acacactggg caagctgttc 3480
tgggtggacg cggacctgaa gcgcattgag agctgtgacc tgtcaggggc caaccgcctg 3540
acctggagg acgccaacat cgtgcagcct ctgggcctga ccactcttg caagcatctc 3600
tactggatcg accgccagca gcagatgatc gagcgtgtgg agaagaccac cggggacaag 3660
cggactcgca tccagggccg tgtcggccac ctactggca tccatgcagt ggaggaagtc 3720
agctggagg agttctcagc ccacccatgt gcccgtagca atgggtggctg ctcccacac 3780
tgtattgcca aggtgtgatg gacaccacgg tgcctatgcc cagtccacct cgtgctcctg 3840
cagaacctgc tgacctgtgg agagccgccc acctgctccc cggaccagtt tgcagtgtgc 3900
acaggggaga tgcactgtat ccccggggcc tggcgtctgt acggctttcc cagtgctgat 3960
gaccagagcg acgaggaggg ctgcccctgt tgcctccgag cccagttccc ctgcgcggg 4020
ggtcagtgtg tggacctgag cctgcgtgac gacggcgagg cagactgtca ggaccgctca 4080
gacgaggcgg actgtgagc catctgcctg cccaaccagt tccggtgtgc gagcggccag 4140
tgtgtcctca tcaaacagca gtgcgactcc ttcccgcact gtatcgacgg ctccgacgag 4200
ctcatgtgtg aaatcaccaa gccgcctcca gacgacagcc cggcccacag cagtgcctac 4260
gggcocgtca ttggcatcat cctctctctc ttctcatagg gtgggtgtcta ttttgtgtgc 4320
cagcgcgtgg tgtgccagcg ctatgcgggg gccaaacgggc ccttcccga cagtatgtc 4380
agcgggaccc cgcacgtgcc cctcaatttc atagcccgg gcggttccca gcatggcccc 4440
ttcacaggca tgcctgagg aaagtccatg atgagctccg tgagcctgat ggggggcccg 4500
ggcgggggtg cctctacga cgggaaccac gtcacagggg cctcgtccag cagctcgtcc 4560
agcaggaagg ccacgctgta ccgcgcgac ctgaacccgc cgcctcccc ggccacggac 4620
ccctccctgt acaacatgga catgttctac tcttcaaaca ttccggccac tgcgagaccg 4680
tacaggccct acatcattcg aggaatggcg ccccgacga cgcctgcag caccgacgtg 4740
tgtgacagcg actacagcgc cagccgctgg aaggccagca agtactacct ggatttgaac 4800
tcggactcag acccctatcc acccccaccc acgcccacac gccagtaact gtcggcggag 4860
gacagctgcc cgcctcgccc cgcacccag aggagctact tccatctctt ccgcgccct 4920
ccgtccccct gcacggactc atcctgacct cggccgggccc actctggctt ctctgtgccc 4980
ctgtaaatag ttttaaatat gaacaaagaa aaaaatatat tttatgatit aaaaaataa 5040
tataattggg attttaaaaa catgagaaat gtgaactgtg atgggggtgg cagggctggg 5100
agaactttgt acagtggaga aatatattata aacttaattt tgtaaaacag aaaaaaaaa 5160

10

20

30

aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 5220
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gcggccgc 5278

<210> 760

<211> 5278

<212> DNA

<213> Homo sapiens

<400> 760

tctagactcg aggcggccgc ccattgtgca ctaaagcgcc gccggccgcg catggagccc 60
 gagtgagcgc ggcgcggggc cgtccggccg ccggacaaca tggaggcagc gccggccggg 120
 ccgcgctggc cgtcgtcgtc gctcgtcgtc ctgctcgtgg cgctgtgcgg ctgcccggcc 180
 ccgcgccggc cctcgcgcgt cctcgtatct gccaacgcgc gggacgtacg gctgggtggc 240
 gccggcggag tcaagctgga gtccaccatc gtggtcagcg gcctggagga tgcggccgca 300
 gtggacttcc agttttccaa gggagccgtg tactggacag acgtgagcga ggaggccatc 360
 aagcagacct acctgaacca gacggggggc gcctgcaga acgtggtcat ctccggcctg 420
 gtctctcccg acggcctcgc ctgcgactgg gtgggcaaga agctgtactg gacggactca 480
 gagaccaacc gcctcgaggt ggccaacctc aatggcacat cccggaaggc gctcttctgg 540
 caggaccttg accagccgag ggccatcgcc ttggaccccg ctacacggta catgtactgg 600
 acagactggg gtgagacgcc ccggtatgag cgggcaggga tggatggcag caccgggaag 660
 atcattgttg actcggacat ttactggccc aatggactga ccacgacct ggaggagcag 720
 aagctctact gggctgacgc caagctcagc ttcatccacc gtgccaaact ggacggctcg 780
 ttccggcaga aggtggtgga gggcagcctg acgcacccct tcgcccctgac gctctccggg 840
 gacactctgt actggacaga ctggcagacc cgctccatcc atgcctgcaa caagcgcact 900
 ggggggaaga ggaaggagat cctgagtgcg ctctactcac ccattggacat ccagggtgctg 960
 agccaggagc ggcagccttt ctccacact cgctgtgagg aggacaatgg cggtgctcc 1020
 cacctgtgcc tgcgtgcccc aagcgagcct ttctacacat gcgcctgccc cagggtgctg 1080
 cagctgcagg acaacggcag gacgtgtaag gcaggagccg agggagtgct gctgctggcc 1140
 cggcgagcgg acctacggag gatctcgtg gacacgcggc acctcaccga catcgtgctg 1200
 caggctggag acatccggca cgccattgcc atcgactacg acccgctaga gggctatgct 1260
 tactggacag atgacgaggt gcgggcccac cgcaggcgct acctggacgg gtctggggcg 1320
 cagacgctgg tcaacaocga gatcaacgac ccgatggca tcgctgctga ctgggtggcc 1380
 cgaaacctct actggaccga caccggcagc gaccgcatcg aggtgacgcg cctcaacggc 1440
 acctcccgca agatcctggt gtgggaggac ctggacgagc cccgagccat cgcactgcac 1500
 cccgtgatgg gcctcatgta ctggacagac tggggagaga accctaaaat cgagtgtgac 1560
 aacttggatg ggcaggagcg gcgtgtgctg gtcaatgcct ccctcggtg gcccaacggc 1620
 ctggccctgg acctgcagga ggggaagctc tactggggag acgccaagac agacaagatc 1680
 gaggtgatca atgttgatgg gacgaagagg cggacctcc tggaggacaa gctcccgcac 1740
 attttcgggt tcacgtgct gggggacttc atctactgga ctgactggca gcgcgcgacg 1800
 atcgagcggg tgcacaaggt caaggccagc cgggacgtca tcattgacca gctgcccgcg 1860
 ctgatggggc tcaaaagctg gaatgtggcc aaggtcgtcg gaaccaaccc gtgtgcggac 1920
 aggaacgggg ggtgcagcca cctgtgcttc ttcacacccc acgcaacccc gtgtggctgc 1980
 cccatcgccc tggagctgct gagtgcacat aagacctgca tcgtgcctga ggccctcttg 2040
 gtcttcacca gcagagccgc catccacagg atctccctcg agaccaataa caacgacgtg 2100
 gccatccccc tcacgggcgt caaggaggcc tcagccctgg accttgatgt gtccaacaac 2160
 cacatctact ggacagacgt cagcctgaag accatcagcc gcgccttcat gaacgggagc 2220
 tcggtggagc acgtggtgga gtttgccott gactacccc agggcatggc cgttgactgg 2280
 atgggcaaga acctctactg ggccgacact gggaccaaca gaatcgaagt ggcgcggctg 2340
 gacgggcagt tccggcaagt cctcgtgtgg agggacttgg acaaccggag gtccgtggcc 2400
 ctggatccca ccaagggcta catctactgg accgagtggg gcggcaagcc gaggatcgtg 2460

10

20

30

cgggacctca tggacgggac caactgcatg acgctgggtgg acaagggtggg ccggggccaac 2520
 gacctcacca ttgactacgc tgaccagcgc ctctactgga ccgacctgga caccaacatg 2580
 atcagtcgtt ccaacatgct gggtcaggag cgggtcgtga ttgcccagca tctcccgcac 2640
 ccgttcgggtc tgacgcagta cagcgattat atctactgga cagactggaa tctgcacagc 2700
 attgagcggg ccgacaagac tagcggccgg aaccgcaccc tcatccaggg ccacctggac 2760
 ttctgtgatgg acatcctggt gtccactccc tcccgcagg atggcctcaa tgactgtatg 2820
 cacaacaacg ggcagtgtgg gcagctgtgc cttgccatcc ccggcggcca ccgctgcggc 2880
 tgcgcctcac actacacct ggaccccagc agccgcgaact gcagcccgc caccaccttc 2940
 ttgtctgttca gccagaaatc tgccatcagt cggatgatcc cggacgacca gcacagcccg 3000
 gatctcatcc tgcacctgca tggactgagg aacgtcaaag ccactgacta tgacctactg 3060
 gacaagttca tctactgggt ggatgggcgc cagaacatca agcagaccaa ggacgacggg 3120
 acccagccct ttgttttgac ctctctgagc caaggccaaa acccagacag gcagcccccac 3180
 gacctcagca tcgacatcta cagccggaca ctgttctgga cgtgcgaggc caccataacc 3240
 atcaacgtcc acaggctgag cggggaagcc atgggggtgg tgctgcgtgg ggaccgcgac 3300
 aagcccaggg ccactcgtct caacgcggag cgagggtacc tgtacttcac caacatgcag 3360
 gaccgggcag ccaagatcga acgcgcagcc ctggacggca ccgagcgcga ggtcctcttc 3420
 accacgggcc tcatccggcc tgtggccctg tgggtagaca acacactggg caagctgttc 3480
 tgggtggacg cggacctgaa gcgcattgag agctgtgacc tgtcaggggc caaccgcctg 3540
 accctggagg acgccaacat cgtgcagcct ctgggcctga ccactcctgg caagcatctc 3600
 tactggatcg accgccagca gcagatgatc gagcgtgtgg agaagaccac cggggacaag 3660
 cggactcgca tccaggcccg tgtcggccac ctactggca tccatgcagt ggaggaaagt 3720
 agcctggagg agttctcagc ccacccatgt gccgtgaca atgggtgctg ctcccacatc 3780
 tgtattgcca agggtgatgg gacaccacgg tgcctatgcc cagtcacact cgtgctcctg 3840
 cagaacctgc tgacctgtgg agagccgccc acctgctccc cggaccagtt tgcattgtgc 3900
 acaggggaga tgcactgtat ccccggggccc tggcgctgtg acggcttccc cagtgcgatg 3960
 gaccagagcg acgaggaggg ctgcccctgt tgcctcgccc ccagttccc ctgcccggc 4020
 ggtcagtggt tggacctgcg cctgcgtgac gacggcgagg cagactgtca ggaccgtcca 4080
 gacgaggcgg actgtgacgc catctgcctg cccaaccagt tccggtgtgc gagcggccag 4140
 tgtgtcctca tcaaacagca gtgcgactcc tccccgact gtatcgacgg ctccgacgag 4200
 ctcatgtgtg aaatcaccaa gccgcctcca gacgacagcc cggcccacag cagtgccttc 4260
 gggcccgctc ttggcatcat cctctctctc ttogtcatgg tgggtgtcta ttttgtgtgc 4320
 cagcgcgtgg tgtgcccagc ctatgcgggg gccaacgggc ccttcccga cagtatgtc 4380
 agcgggaccc cgcacgtgcc cctcaatttc atagcccgg gcggttccc gcatggcccc 4440
 ttcacaggca tggcatgccc aaagtccatg atgagctccg tgagcctgat ggggggccc 4500
 ggggggggtg ccctctacga ccggaaccac gtcacagggg cctcgtccag cagctcgtcc 4560
 agcaggaagg ccacgtgtga cccgcgcgac ctgaaccgcc cgcctcccc ggccacggac 4620
 ccctccctgt acaacatgga catgtttctc tcttcaaaac ttcgggccac tgcgagaccg 4680
 tacaggccct acatcattcg aggaatggcg ccccgcagca cgcctgcag caccgacgtg 4740
 tgtgacagcg actacagcgc cagcgcctgg aaggccagca agtactacct ggatttgaac 4800
 tccgactcag acccctatcc acccccaccc acgcccacac gccagtacct gtcggcggag 4860
 gacagctgcc cgcctctgcc cgcaccagag agggactact tccatctctt cccgccccct 4920
 ccgtccccct gcacggactc atcctgacct cggccgggccc actctggttt ctctgtgccc 4980
 ctgtaaatag ttttaaatat gaacaaagaa aaaaatatat tttatgatit aaaaaataaa 5040
 tataattggg attttaaaaa catgagaaat gtgaactgtg atgggggtggg cagggctggg 5100
 agaactttgt acagtggaga aatatttata aacttaattt tgtaaaacag aaaaaaaaaa 5160
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 5220
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gcggccgc 5278

<210> 761

<211> 1615

10

20

30

<212> PRT

<213> Homo sapiens

<400> 761

```

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

```

10

20

30

Val Asp Asp Ile Arg His Ala Ile Ala Ile Asp Tyr Asp Pro Leu Glu
 370 375 380
 Gly Tyr Val Tyr Trp Thr Asp Asp Glu Val Arg Ala Ile Arg Arg Ala
 385 390 395 400
 Tyr Leu Asp Gly Ser Gly Ala Gln Thr Leu Val Asn Thr Glu Ile Asn
 405 410 415
 Asp Pro Asp Gly Ile Ala Val Asp Trp Val Ala Arg Asn Leu Tyr Trp
 420 425 430
 Thr Asp Thr Gly Thr Asp Arg Ile Glu Val Thr Arg Leu Asn Gly Thr
 435 440 445
 Ser Arg Lys Ile Leu Val Ser Glu Asp Leu Asp Glu Pro Arg Ala Ile
 450 455 460
 Ala Leu His Pro Val Met Gly Leu Met Tyr Trp Thr Asp Trp Gly Glu
 465 470 475 480
 Asn Pro Lys Ile Glu Cys Ala Asn Leu Asp Gly Gln Glu Arg Arg Val
 485 490 495
 Leu Val Asn Ala Ser Leu Gly Trp Pro Asn Gly Leu Ala Leu Asp Leu
 500 505 510
 Gln Glu Gly Lys Leu Tyr Trp Gly Asp Ala Lys Thr Asp Lys Ile Glu
 515 520 525
 Val Ile Asn Val Asp Gly Thr Lys Arg Arg Thr Leu Leu Glu Asp Lys
 530 535 540
 Leu Pro His Ile Phe Gly Phe Thr Leu Leu Gly Asp Phe Ile Tyr Trp
 545 550 555 560
 Thr Asp Trp Gln Arg Arg Ser Ile Glu Arg Val His Lys Val Lys Ala
 565 570 575
 Ser Arg Asp Val Ile Ile Asp Gln Leu Pro Asp Leu Met Gly Leu Lys
 580 585 590
 Ala Val Asn Val Ala Lys Val Val Gly Thr Asn Pro Cys Ala Asp Arg
 595 600 605
 Asn Gly Gly Cys Ser His Leu Cys Phe Phe Thr Pro His Ala Thr Arg
 610 615 620
 Cys Gly Cys Pro Ile Gly Leu Glu Leu Leu Ser Asp Met Lys Thr Cys
 625 630 635 640
 Ile Val Pro Glu Ala Phe Leu Val Phe Thr Ser Arg Ala Ala Ile His
 645 650 655
 Arg Ile Ser Leu Glu Thr Asn Asn Asn Asp Val Ala Ile Pro Leu Thr
 660 665 670
 Gly Val Lys Glu Ala Ser Ala Leu Asp Phe Asp Val Ser Asn Asn His
 675 680 685
 Ile Tyr Trp Thr Asp Val Ser Leu Lys Thr Ile Ser Arg Ala Phe Met
 690 695 700
 Asn Gly Ser Ser Val Glu His Val Val Glu Phe Gly Leu Asp Tyr Pro
 705 710 715 720
 Glu Gly Met Ala Val Asp Trp Met Gly Lys Asn Leu Tyr Trp Ala Asp
 725 730 735
 Thr Gly Thr Asn Arg Ile Glu Val Ala Arg Leu Asp Gly Gln Phe Arg
 740 745 750
 Gln Val Leu Val Trp Arg Asp Leu Asp Asn Pro Arg Ser Leu Ala Leu
 755 760 765

10

20

30

Asp Pro Thr Lys Gly Tyr Ile Tyr Trp Thr Glu Trp Gly Gly Lys Pro
 770 775 780
 Arg Ile Val Arg Ala Phe Met Asp Gly Thr Asn Cys Met Thr Leu Val
 785 790 795 800
 Asp Lys Val Gly Arg Ala Asn Asp Leu Thr Ile Asp Tyr Ala Asp Gln
 805 810 815
 Arg Leu Tyr Trp Thr Asp Leu Asp Thr Asn Met Ile Glu Ser Ser Asn
 820 825 830
 Met Leu Gly Gln Glu Arg Val Val Ile Ala Asp Asp Leu Pro His Pro
 835 840 845
 Phe Gly Leu Thr Gln Tyr Ser Asp Tyr Ile Tyr Trp Thr Asp Trp Asn
 850 855 860
 Leu His Ser Ile Glu Arg Ala Asp Lys Thr Ser Gly Arg Asn Arg Thr
 865 870 875 880
 Leu Ile Gln Gly His Leu Asp Phe Val Met Asp Ile Leu Val Phe His
 885 890 895
 Ser Ser Arg Gln Asp Gly Leu Asn Asp Cys Met His Asn Asn Gly Gln
 900 905 910
 Cys Gly Gln Leu Cys Leu Ala Ile Pro Gly Gly His Arg Cys Gly Cys
 915 920 925
 Ala Ser His Tyr Thr Leu Asp Pro Ser Ser Arg Asn Cys Ser Pro Pro
 930 935 940
 Thr Thr Phe Leu Leu Phe Ser Gln Lys Ser Ala Ile Ser Arg Met Ile
 945 950 955 960
 Pro Asp Asp Gln His Ser Pro Asp Leu Ile Leu Pro Leu His Gly Leu
 965 970 975
 Arg Asn Val Lys Ala Ile Asp Tyr Asp Pro Leu Asp Lys Phe Ile Tyr
 980 985 990
 Trp Val Asp Gly Arg Gln Asn Ile Lys Arg Ala Lys Asp Asp Gly Thr
 995 1000 1005
 Gln Pro Phe Val Leu Thr Ser Leu Ser Gln Gly Gln Asn Pro Asp Arg
 1010 1015 1020
 Gln Pro His Asp Leu Ser Ile Asp Ile Tyr Ser Arg Thr Leu Phe Trp
 1025 1030 1035 1040
 Thr Cys Glu Ala Thr Asn Thr Ile Asn Val His Arg Leu Ser Gly Glu
 1045 1050 1055
 Ala Met Gly Val Val Leu Arg Gly Asp Arg Asp Lys Pro Arg Ala Ile
 1060 1065 1070
 Val Val Asn Ala Glu Arg Gly Tyr Leu Tyr Phe Thr Asn Met Gln Asp
 1075 1080 1085
 Arg Ala Ala Lys Ile Glu Arg Ala Ala Leu Asp Gly Thr Glu Arg Glu
 1090 1095 1100
 Val Leu Phe Thr Thr Gly Leu Ile Arg Pro Val Ala Leu Val Val Asp
 1105 1110 1115 1120
 Asn Thr Leu Gly Lys Leu Phe Trp Val Asp Ala Asp Leu Lys Arg Ile
 1125 1130 1135
 Glu Ser Cys Asp Leu Ser Gly Ala Asn Arg Leu Thr Leu Glu Asp Ala
 1140 1145 1150
 Asn Ile Val Gln Pro Leu Gly Leu Thr Ile Leu Gly Lys His Leu Tyr
 1155 1160 1165

10

20

30

Trp Ile Asp Arg Gln Gln Gln Met Ile Glu Arg Val Glu Lys Thr Thr
 1170 1175 1180
 Gly Asp Lys Arg Thr Arg Ile Gln Gly Arg Val Ala His Leu Thr Gly
 1185 1190 1195 1200
 Ile His Ala Val Glu Glu Val Ser Leu Glu Glu Phe Ser Ala His Pro
 1205 1210 1215
 Cys Ala Arg Asp Asn Gly Gly Cys Ser His Ile Cys Ile Ala Lys Gly
 1220 1225 1230
 Asp Gly Thr Pro Arg Cys Ser Cys Pro Val His Leu Val Leu Leu Gln
 1235 1240 1245
 Asn Leu Leu Thr Cys Gly Glu Pro Pro Thr Cys Ser Pro Asp Gln Phe
 1250 1255 1260
 Ala Cys Ala Thr Gly Glu Ile Asp Cys Ile Pro Gly Ala Trp Arg Cys
 1265 1270 1275 1280
 Asp Gly Phe Pro Glu Cys Asp Asp Gln Ser Asp Glu Glu Gly Cys Pro
 1285 1290 1295
 Val Cys Ser Ala Ala Gln Phe Pro Cys Ala Arg Gly Gln Cys Val Asp
 1300 1305 1310
 Leu Arg Leu Arg Cys Asp Gly Glu Ala Asp Cys Gln Asp Arg Ser Asp
 1315 1320 1325
 Glu Ala Asp Cys Asp Ala Ile Cys Leu Pro Asn Gln Phe Arg Cys Ala
 1330 1335 1340
 Ser Gly Gln Cys Val Leu Ile Lys Gln Gln Cys Asp Ser Phe Pro Asp
 1345 1350 1355 1360
 Cys Ile Asp Gly Ser Asp Glu Leu Met Cys Glu Ile Thr Lys Pro Pro
 1365 1370 1375
 Ser Asp Asp Ser Pro Ala His Ser Ser Ala Ile Gly Pro Val Ile Gly
 1380 1385 1390
 Ile Ile Leu Ser Leu Phe Val Met Gly Gly Val Tyr Phe Val Cys Gln
 1395 1400 1405
 Arg Val Val Cys Gln Arg Tyr Ala Gly Ala Asn Gly Pro Phe Pro His
 1410 1415 1420
 Glu Tyr Val Ser Gly Thr Pro His Val Pro Leu Asn Phe Ile Ala Pro
 1425 1430 1435 1440
 Gly Gly Ser Gln His Gly Pro Phe Thr Gly Ile Ala Cys Gly Lys Ser
 1445 1450 1455
 Met Met Ser Ser Val Ser Leu Met Gly Gly Arg Gly Gly Val Pro Leu
 1460 1465 1470
 Tyr Asp Arg Asn His Val Thr Gly Ala Ser Ser Ser Ser Ser Ser
 1475 1480 1485
 Thr Lys Ala Thr Leu Tyr Pro Pro Ile Leu Asn Pro Pro Pro Ser Pro
 1490 1495 1500
 Ala Thr Asp Pro Ser Leu Tyr Asn Met Asp Met Phe Tyr Ser Ser Asn
 1505 1510 1515 1520
 Ile Pro Ala Thr Ala Arg Pro Tyr Arg Pro Tyr Ile Ile Arg Gly Met
 1525 1530 1535
 Ala Pro Pro Thr Thr Pro Cys Ser Thr Asp Val Cys Asp Ser Asp Tyr
 1540 1545 1550
 Ser Ala Ser Arg Trp Lys Ala Ser Lys Tyr Tyr Leu Asp Leu Asn Ser
 1555 1560 1565

10

20

30

Asp Ser Asp Pro Tyr Pro Pro Pro Pro Thr Pro His Ser Gln Tyr Leu
 1570 1575 1580
 Ser Ala Glu Asp Ser Cys Pro Pro Ser Pro Ala Thr Glu Arg Ser Tyr
 1585 1590 1595 1600
 Phe His Leu Phe Pro Pro Pro Ser Pro Cys Thr Asp Ser Ser
 1605 1610 1615

<210> 762
 <211> 1564
 <212> PRT
 <213> Mouse

<400> 762
 Met Glu Thr Ala Pro Thr Arg Ala Pro Pro Pro Pro Pro Pro Leu
 1 5 10 15
 Leu Leu Leu Val Leu Tyr Cys Ser Leu Val Pro Ala Ala Ala Ser Pro
 20 25 30
 Leu Leu Leu Phe Ala Asn Arg Arg Asp Val Arg Leu Val Asp Ala Gly
 35 40 45
 Gly Val Lys Leu Glu Ser Thr Ile Val Ala Ser Gly Leu Glu Asp Ala
 50 55 60
 Ala Ala Val Asp Phe Gln Phe Ser Lys Gly Ala Val Tyr Trp Thr Asp
 65 70 75 80
 Val Ser Glu Glu Ala Ile Lys Gln Thr Tyr Leu Asn Gln Thr Gly Ala
 85 90 95
 Ala Ala Gln Asn Ile Val Ile Ser Gly Leu Val Ser Pro Asp Gly Leu
 100 105 110
 Ala Cys Asp Trp Val Gly Lys Lys Leu Tyr Trp Thr Asp Ser Glu Thr
 115 120 125
 Asn Arg Ile Glu Val Ala Asn Leu Asn Gly Thr Ser Arg Lys Val Leu
 130 135 140
 Phe Trp Gln Asp Leu Asp Gln Pro Arg Ala Ile Ala Leu Asp Pro Ala
 145 150 155 160
 His Gly Tyr Met Tyr Trp Thr Asp Trp Gly Glu Ala Pro Arg Ile Glu
 165 170 175
 Arg Ala Gly Met Asp Gly Ser Thr Arg Lys Ile Ile Val Asp Ser Asp
 180 185 190
 Ile Tyr Trp Pro Asn Gly Leu Thr Ile Asp Leu Glu Glu Gln Lys Leu
 195 200 205
 Tyr Trp Ala Asp Ala Lys Leu Ser Phe Ile His Arg Ala Asn Leu Asp
 210 215 220
 Gly Ser Phe Arg Gln Lys Val Val Glu Gly Ser Leu Thr His Pro Phe
 225 230 235 240
 Ala Leu Thr Leu Ser Gly Asp Thr Leu Tyr Trp Thr Asp Trp Gln Thr
 245 250 255
 Arg Ser Ile His Ala Cys Asn Lys Trp Thr Gly Glu Gln Arg Lys Glu
 260 265 270
 Ile Leu Ser Ala Leu Tyr Ser Pro Met Asp Ile Gln Val Leu Ser Gln
 275 280 285

10

20

30

Glu Arg Gln Pro Pro Phe His Thr Pro Cys Glu Glu Asp Asn Gly Gly
 290 295 300
 Cys Ser His Leu Cys Leu Leu Ser Pro Arg Glu Pro Phe Tyr Ser Cys
 305 310 315 320
 Ala Cys Pro Thr Gly Val Gln Leu Gln Asp Asn Gly Lys Thr Cys Lys
 325 330 335
 Thr Gly Ala Glu Glu Val Leu Leu Leu Ala Arg Arg Thr Asp Leu Arg
 340 345 350
 Arg Ile Ser Leu Asp Thr Pro Asp Phe Thr Asp Ile Val Leu Gln Val
 355 360 365
 Gly Asp Ile Arg His Ala Ile Ala Ile Asp Tyr Asp Pro Leu Glu Gly
 370 375 380
 Tyr Val Tyr Trp Thr Asp Asp Glu Val Arg Ala Ile Arg Arg Ala Tyr
 385 390 395 400
 Leu Asp Gly Ser Gly Ala Gln Thr Leu Val Asn Thr Glu Ile Asn Asp
 405 410 415
 Pro Asp Gly Ile Ala Val Asp Trp Val Ala Arg Asn Leu Tyr Trp Thr
 420 425 430
 Asp Thr Gly Thr Asp Arg Ile Glu Val Thr Arg Leu Asn Gly Thr Ser
 435 440 445
 Arg Lys Ile Leu Val Ser Glu Asp Leu Asp Glu Pro Arg Ala Ile Val
 450 455 460
 Leu His Pro Val Met Gly Leu Met Tyr Trp Thr Asp Trp Gly Glu Asn
 465 470 475 480
 Pro Lys Ile Glu Cys Ala Asn Leu Asp Gly Arg Asp Arg His Val Leu
 485 490 495
 Val Asn Thr Ser Leu Gly Trp Pro Asn Gly Leu Ala Leu Asp Leu Gln
 500 505 510
 Glu Gly Lys Leu Tyr Trp Gly Asp Ala Lys Thr Asp Lys Ile Glu Val
 515 520 525
 Ile Asn Ile Asp Gly Thr Lys Arg Lys Thr Leu Leu Glu Asp Lys Leu
 530 535 540
 Pro His Ile Phe Gly Phe Thr Leu Leu Gly Asp Phe Ile Tyr Trp Thr
 545 550 555 560
 Asp Trp Gln Arg Arg Ser Ile Glu Arg Val His Lys Val Lys Ala Ser
 565 570 575
 Arg Asp Val Ile Ile Asp Gln Leu Pro Asp Leu Met Gly Leu Lys Ala
 580 585 590
 Val Asn Val Ala Lys Val Val Gly Thr Asn Pro Cys Ala Asp Gly Asn
 595 600 605
 Gly Gly Cys Ser His Leu Cys Phe Phe Thr Pro Arg Ala Thr Lys Cys
 610 615 620
 Gly Cys Pro Ile Gly Leu Glu Leu Leu Ser Asp Met Lys Thr Cys Ile
 625 630 635 640
 Ile Pro Glu Ala Phe Leu Val Phe Thr Ser Arg Ala Thr Ile His Arg
 645 650 655
 Ile Ser Leu Glu Thr Asn Asn Asn Asp Val Ala Ile Pro Leu Thr Gly
 660 665 670
 Val Lys Glu Ala Ser Ala Leu Asp Phe Asp Val Ser Asn Asn His Ile
 675 680 685

10

20

30

Tyr Trp Thr Asp Val Ser Leu Lys Thr Ile Ser Arg Ala Phe Met Asn
 690 695 700
 Gly Ser Ser Val Glu His Val Ile Glu Phe Gly Leu Asp Tyr Pro Glu
 705 710 715 720
 Gly Met Ala Val Asp Trp Met Gly Lys Asn Leu Tyr Trp Ala Asp Thr
 725 730 735
 Gly Thr Asn Arg Ile Glu Val Ala Arg Leu Asp Gly Gln Phe Arg Gln
 740 745 750
 Val Leu Val Trp Arg Asp Leu Asp Asn Pro Arg Ser Leu Ala Leu Asp
 755 760 765
 Pro Thr Lys Gly Tyr Ile Tyr Trp Thr Glu Trp Gly Gly Lys Pro Arg
 770 775 780
 Ile Val Arg Ala Phe Met Asp Gly Thr Asn Cys Met Thr Leu Val Asp
 785 790 795 800
 Lys Val Gly Arg Ala Asn Asp Leu Thr Ile Asp Tyr Ala Asp Gln Arg
 805 810 815
 Leu Tyr Trp Thr Asp Leu Asp Thr Asn Met Ile Glu Ser Ser Asn Met
 820 825 830
 Leu Gly Gln Glu Arg Met Val Ile Ala Asp Asp Leu Pro Tyr Pro Phe
 835 840 845
 Gly Leu Thr Gln Tyr Ser Asp Tyr Ile Tyr Trp Thr Asp Trp Asn Leu
 850 855 860
 His Ser Ile Glu Arg Ala Asp Lys Thr Ser Gly Arg Asn Arg Thr Leu
 865 870 875 880
 Ile Gln Gly His Leu Asp Phe Val Met Asp Ile Leu Val Phe His Ser
 885 890 895
 Ser Arg Gln Asp Gly Leu Asn Asp Cys Val His Ser Asn Gly Gln Cys
 900 905 910
 Gly Gln Leu Cys Leu Ala Ile Pro Gly Gly His Arg Cys Gly Cys Ala
 915 920 925
 Ser His Tyr Thr Leu Asp Pro Ser Ser Arg Asn Cys Ser Pro Pro Ser
 930 935 940
 Thr Phe Leu Leu Phe Ser Gln Lys Phe Ala Ile Ser Arg Met Ile Pro
 945 950 955 960
 Asp Asp Gln Leu Ser Pro Asp Leu Val Leu Pro Leu His Gly Leu Arg
 965 970 975
 Asn Val Lys Ala Ile Asn Tyr Asp Pro Leu Asp Lys Phe Ile Tyr Trp
 980 985 990
 Val Asp Gly Arg Gln Asn Ile Lys Arg Ala Lys Asp Asp Gly Thr Gln
 995 1000 1005
 Pro Ser Met Leu Thr Ser Pro Ser Gln Ser Leu Ser Pro Asp Arg Gln
 1010 1015 1020
 Pro His Asp Leu Ser Ile Asp Ile Tyr Ser Arg Thr Leu Phe Trp Thr
 1025 1030 1035 1040
 Cys Glu Ala Thr Asn Thr Ile Asn Val His Arg Leu Asp Gly Asp Ala
 1045 1050 1055
 Met Gly Val Val Leu Arg Gly Asp Arg Asp Lys Pro Arg Ala Ile Ala
 1060 1065 1070
 Val Asn Ala Glu Arg Gly Tyr Met Tyr Phe Thr Asn Met Gln Asp His
 1075 1080 1085

10

20

30

Ala Ala Lys Ile Glu Arg Ala Ser Leu Asp Gly Thr Glu Arg Glu Val
 1090 1095 1100
 Leu Phe Thr Thr Gly Leu Ile Arg Pro Val Ala Leu Val Val Asp Asn
 1105 1110 1115 1120
 Ala Leu Gly Lys Leu Phe Trp Val Asp Ala Asp Leu Lys Arg Ile Glu
 1125 1130 1135
 Ser Cys Asp Leu Ser Gly Ala Asn Arg Leu Thr Leu Glu Asp Ala Asn
 1140 1145 1150
 Ile Val Gln Pro Val Gly Leu Thr Val Leu Gly Arg His Leu Tyr Trp
 1155 1160 1165
 Ile Asp Arg Gln Gln Gln Met Ile Glu Arg Val Glu Lys Thr Thr Gly
 1170 1175 1180
 Asp Lys Arg Thr Arg Val Gln Gly Arg Val Thr His Leu Thr Gly Ile
 1185 1190 1195 1200
 His Ala Val Glu Glu Val Ser Leu Glu Glu Phe Ser Ala His Pro Cys
 1205 1210 1215
 Ala Arg Asp Asn Gly Gly Cys Ser His Ile Cys Ile Ala Lys Gly Asp
 1220 1225 1230
 Gly Thr Pro Arg Cys Ser Cys Pro Val His Leu Val Leu Leu Gln Asn
 1235 1240 1245
 Leu Leu Thr Cys Gly Glu Pro Pro Thr Cys Ser Pro Asp Gln Phe Ala
 1250 1255 1260
 Cys Thr Thr Gly Glu Ile Asp Cys Ile Pro Gly Ala Trp Arg Cys Asp
 1265 1270 1275 1280
 Gly Phe Pro Glu Cys Ala Asp Gln Ser Asp Glu Glu Gly Cys Pro Val
 1285 1290 1295
 Cys Ser Ala Ser Gln Phe Pro Cys Ala Arg Gly Gln Cys Val Asp Leu
 1300 1305 1310
 Arg Leu Arg Cys Asp Gly Glu Ala Asp Cys Gln Asp Arg Ser Asp Glu
 1315 1320 1325
 Ala Asn Cys Asp Ala Val Cys Leu Pro Asn Gln Phe Arg Cys Thr Ser
 1330 1335 1340
 Gly Leu Phe Val Met Gly Gly Val Tyr Phe Val Cys Gln Arg Val Met
 1345 1350 1355 1360
 Cys Gln Arg Tyr Thr Gly Ala Ser Gly Pro Phe Pro His Glu Tyr Val
 1365 1370 1375
 Gly Gly Ala Pro His Val Pro Leu Asn Phe Ile Ala Pro Gly Gly Ser
 1380 1385 1390
 Gln His Gly Pro Phe Pro Gly Ile Pro Cys Ser Lys Ser Val Met Ser
 1395 1400 1405
 Ser Met Ser Leu Val Gly Gly Arg Gly Ser Val Pro Leu Tyr Asp Arg
 1410 1415 1420
 Asn His Val Thr Gly Ala Ser Ser Ser Ser Ser Ser Thr Lys Ala
 1425 1430 1435 1440
 Thr Leu Tyr Pro Pro Ile Leu Asn Pro Pro Pro Ser Pro Ala Thr Asp
 1445 1450 1455
 Pro Ser Leu Tyr Asn Val Asp Val Phe Tyr Ser Ser Gly Ile Pro Ala
 1460 1465 1470
 Thr Ala Arg Pro Tyr Arg Pro Tyr Val Ile Arg Gly Met Ala Pro Pro
 1475 1480 1485

10

20

30

Thr Thr Pro Cys Ser Thr Asp Val Cys Asp Ser Asp Tyr Ser Ile Ser
 1490 1495 1500
 Arg Trp Lys Ser Ser Lys Tyr Tyr Leu Asp Leu Asn Ser Asp Ser Asp
 1505 1510 1515 1520
 Pro Tyr Pro Pro Pro Pro Thr Pro His Ser Gln Tyr Leu Ser Ala Glu
 1525 1530 1535
 Asp Ser Cys Pro Pro Ser Pro Gly Thr Glu Arg Ser Tyr Cys His Leu
 1540 1545 1550
 Phe Pro Pro Pro Pro Ser Pro Cys Thr Asp Ser Ser
 1555 1560

<210> 763
 <211> 1665
 <212> PRT
 <213> Homo sapiens

10

<400> 763
 Met Glu Ala Ala Pro Gly Pro Pro Trp Pro Leu Leu Leu Leu Leu
 1 5 10 15
 Leu Leu Leu Leu Ala Leu Cys Gly Cys Pro Ala Pro Ala Ala Ser
 20 25 30
 Pro Leu Leu Leu Phe Ala Asn Arg Arg Asp Val Arg Leu Val Asp Ala
 35 40 45
 Gly Gly Val Lys Leu Glu Ser Thr Ile Val Val Ser Gly Leu Glu Asp
 50 55 60
 Ala Ala Ala Val Asp Phe Gln Phe Ser Lys Gly Ala Val Tyr Trp Thr
 65 70 75 80
 Asp Val Ser Glu Glu Ala Ile Lys Gln Thr Tyr Leu Asn Gln Thr Gly
 85 90 95
 Ala Ala Val Gln Asn Val Val Ile Ser Gly Leu Val Ser Pro Asp Gly
 100 105 110
 Leu Ala Cys Asp Trp Val Gly Lys Lys Leu Tyr Trp Thr Asp Ser Glu
 115 120 125
 Thr Asn Arg Ile Glu Val Ala Asn Leu Asn Gly Thr Ser Arg Lys Val
 130 135 140
 Leu Phe Trp Gln Asp Leu Asp Gln Pro Arg Ala Ile Ala Leu Asp Pro
 145 150 155 160
 Ala His Gly Tyr Met Tyr Trp Thr Asp Trp Gly Glu Thr Pro Arg Ile
 165 170 175
 Glu Arg Ala Gly Met Asp Gly Ser Thr Arg Lys Ile Ile Val Asp Ser
 180 185 190
 Asp Ile Tyr Trp Pro Asn Gly Leu Thr Ile Asp Leu Glu Glu Gln Lys
 195 200 205
 Leu Tyr Trp Ala Asp Ala Lys Leu Ser Phe Ile His Arg Ala Asn Leu
 210 215 220
 Asp Gly Ser Phe Arg Gln Lys Val Val Glu Gly Ser Leu Thr His Pro
 225 230 235 240
 Phe Ala Leu Thr Leu Ser Gly Asp Thr Leu Tyr Trp Thr Asp Trp Gln
 245 250 255

20

30

Thr Arg Ser Ile His Ala Cys Asn Lys Arg Thr Gly Gly Lys Arg Lys
 260 265 270
 Glu Ile Leu Ser Ala Leu Tyr Ser Pro Met Asp Ile Gln Val Leu Ser
 275 280 285
 Gln Glu Arg Gln Pro Phe Phe His Thr Arg Cys Glu Glu Asp Asn Gly
 290 295 300
 Gly Cys Ser His Leu Cys Leu Leu Ser Pro Ser Glu Pro Phe Tyr Thr
 305 310 315 320
 Cys Ala Cys Pro Thr Gly Val Gln Leu Gln Asp Asn Gly Arg Thr Cys
 325 330 335
 Lys Ala Gly Ala Glu Glu Val Leu Leu Leu Ala Arg Arg Thr Asp Leu
 340 345 350
 Arg Arg Ile Ser Leu Asp Thr Pro Asp Phe Thr Asp Ile Val Leu Gln
 355 360 365
 Val Asp Asp Ile Arg His Ala Ile Ala Ile Asp Tyr Asp Pro Leu Glu
 370 375 380
 Gly Tyr Val Tyr Trp Thr Asp Asp Glu Val Arg Ala Ile Arg Arg Ala
 385 390 395 400
 Tyr Leu Asp Gly Ser Gly Ala Gln Thr Leu Val Asn Thr Glu Ile Asn
 405 410 415
 Asp Pro Asp Gly Ile Ala Val Asp Trp Val Ala Arg Asn Leu Tyr Trp
 420 425 430
 Thr Asp Thr Gly Thr Asp Arg Ile Glu Val Thr Arg Leu Asn Gly Thr
 435 440 445
 Ser Arg Lys Ile Leu Val Ser Glu Asp Leu Asp Glu Pro Arg Ala Ile
 450 455 460
 Ala Leu His Pro Val Met Gly Leu Met Tyr Trp Thr Asp Trp Gly Glu
 465 470 475 480
 Asn Pro Lys Ile Glu Cys Ala Asn Leu Asp Gly Gln Glu Arg Arg Val
 485 490 495
 Leu Val Asn Ala Ser Leu Gly Trp Pro Asn Gly Leu Ala Leu Asp Leu
 500 505 510
 Gln Glu Gly Lys Leu Tyr Trp Gly Asp Ala Lys Thr Asp Lys Ile Glu
 515 520 525
 Val Ile Asn Val Asp Gly Thr Lys Arg Arg Thr Leu Leu Glu Asp Lys
 530 535 540
 Leu Pro His Ile Phe Gly Phe Thr Leu Leu Gly Asp Phe Ile Tyr Trp
 545 550 555 560
 Thr Asp Trp Gln Arg Arg Ser Ile Glu Arg Val His Lys Val Lys Ala
 565 570 575
 Ser Arg Asp Val Ile Ile Asp Gln Leu Pro Asp Leu Met Gly Leu Lys
 580 585 590
 Ala Val Asn Val Ala Lys Val Val Gly Thr Asn Pro Cys Ala Asp Arg
 595 600 605
 Asn Gly Gly Cys Ser His Leu Cys Phe Phe Thr Pro His Ala Thr Arg
 610 615 620
 Cys Gly Cys Pro Ile Gly Leu Glu Leu Leu Ser Asp Met Lys Thr Cys
 625 630 635 640
 Ile Val Pro Glu Ala Phe Leu Val Phe Thr Ser Arg Ala Ala Ile His
 645 650 655

10

20

30

Arg Ile Ser Leu Glu Thr Asn Asn Asn Asp Val Ala Ile Pro Leu Thr
 660 665 670
 Gly Val Lys Glu Ala Ser Ala Leu Asp Phe Asp Val Ser Asn Asn His
 675 680 685
 Ile Tyr Trp Thr Asp Val Ser Leu Lys Thr Ile Ser Arg Ala Phe Met
 690 695 700
 Asn Gly Ser Ser Val Glu His Val Val Glu Phe Gly Leu Asp Tyr Pro
 705 710 715 720
 Glu Gly Met Ala Val Asp Trp Met Gly Lys Asn Leu Tyr Trp Ala Asp
 725 730 735
 Thr Gly Thr Asn Arg Ile Glu Val Ala Arg Leu Asp Gly Gln Phe Arg
 740 745 750
 Gln Val Leu Val Trp Arg Asp Leu Asp Asn Pro Arg Ser Leu Ala Leu
 755 760 765
 Asp Pro Thr Lys Gly Tyr Ile Tyr Trp Thr Glu Trp Gly Gly Lys Pro
 770 775 780
 Arg Ile Val Arg Ala Phe Met Asp Gly Thr Asn Cys Met Thr Leu Val
 785 790 795 800
 Asp Lys Val Gly Arg Ala Asn Asp Leu Thr Ile Asp Tyr Ala Asp Gln
 805 810 815
 Arg Leu Tyr Trp Thr Asp Leu Asp Thr Asn Met Ile Glu Ser Ser Asn
 820 825 830
 Met Leu Gly Gln Glu Arg Val Val Ile Ala Asp Asp Leu Pro His Pro
 835 840 845
 Phe Gly Leu Thr Gln Tyr Ser Asp Tyr Ile Tyr Trp Thr Asp Trp Asn
 850 855 860
 Leu His Ser Ile Glu Arg Ala Asp Lys Thr Ser Gly Arg Asn Arg Thr
 865 870 875 880
 Leu Ile Gln Gly His Leu Asp Phe Val Met Asp Ile Leu Val Phe His
 885 890 895
 Ser Ser Arg Gln Asp Gly Leu Asn Asp Cys Met His Asn Asn Gly Gln
 900 905 910
 Cys Gly Gln Leu Cys Leu Ala Ile Pro Gly Gly His Arg Cys Gly Cys
 915 920 925
 Ala Ser His Tyr Thr Leu Asp Pro Ser Ser Arg Asn Cys Ser Pro Pro
 930 935 940
 Thr Thr Phe Leu Leu Phe Ser Gln Lys Ser Ala Ile Ser Arg Met Ile
 945 950 955 960
 Pro Asp Asp Gln His Ser Pro Asp Leu Ile Leu Pro Leu His Gly Leu
 965 970 975
 Arg Asn Val Lys Ala Ile Asp Tyr Asp Pro Leu Asp Lys Phe Ile Tyr
 980 985 990
 Trp Val Asp Gly Arg Gln Asn Phe Ser Gln Lys Ser Ala Ile Asn Arg
 995 1000 1005
 Met Val Ile Asp Glu Gln Gln Ser Pro Asp Ile Ile Leu Pro Ile His
 1010 1015 1020
 Ser Leu Arg Asn Val Arg Ala Ile Asp Tyr Asp Pro Leu Asp Lys Gln
 1025 1030 1035 1040
 Leu Tyr Trp Ile Asp Ser Arg Gln Asn Ile Lys Arg Ala Lys Asp Asp
 1045 1050 1055

10

20

30

Gly Thr Gln Pro Phe Val Leu Thr Ser Leu Ser Gln Gly Gln Asn Pro
 1060 1065 1070
 Asp Arg Gln Pro His Asp Leu Ser Ile Asp Ile Tyr Ser Arg Thr Leu
 1075 1080 1085
 Phe Trp Thr Cys Glu Ala Thr Asn Thr Ile Asn Val His Arg Leu Ser
 1090 1095 1100
 Gly Glu Ala Met Gly Val Val Leu Arg Gly Asp Arg Asp Lys Pro Arg
 1105 1110 1115 1120
 Ala Ile Val Val Asn Ala Glu Arg Gly Tyr Leu Tyr Phe Thr Asn Met
 1125 1130 1135
 Gln Asp Arg Ala Ala Lys Ile Glu Arg Ala Ala Leu Asp Gly Thr Glu
 1140 1145 1150
 Arg Glu Val Leu Phe Thr Thr Gly Leu Ile Arg Pro Val Ala Leu Val
 1155 1160 1165
 Val Asp Asn Thr Leu Gly Lys Leu Phe Trp Val Asp Ala Asp Leu Lys
 1170 1175 1180
 Arg Ile Glu Ser Cys Asp Leu Ser Gly Ala Asn Arg Leu Thr Leu Glu
 1185 1190 1195 1200
 Asp Ala Asn Ile Val Gln Pro Leu Gly Leu Thr Ile Leu Gly Lys His
 1205 1210 1215
 Leu Tyr Trp Ile Asp Arg Gln Gln Gln Met Ile Glu Arg Val Glu Lys
 1220 1225 1230
 Thr Thr Gly Asp Lys Arg Thr Arg Ile Gln Gly Arg Val Ala His Leu
 1235 1240 1245
 Thr Gly Ile His Ala Val Glu Glu Val Ser Leu Glu Glu Phe Ser Ala
 1250 1255 1260
 His Pro Cys Ala Arg Asp Asn Gly Gly Cys Ser His Ile Cys Ile Ala
 1265 1270 1275 1280
 Lys Gly Asp Gly Thr Pro Arg Cys Ser Cys Pro Val His Leu Val Leu
 1285 1290 1295
 Leu Gln Asn Leu Leu Thr Cys Gly Glu Pro Pro Thr Cys Ser Pro Asp
 1300 1305 1310
 Gln Phe Ala Cys Ala Thr Gly Glu Ile Asp Cys Ile Pro Gly Ala Trp
 1315 1320 1325
 Arg Cys Asp Gly Phe Pro Glu Cys Asp Asp Gln Ser Asp Glu Glu Gly
 1330 1335 1340
 Cys Pro Val Cys Ser Ala Ala Gln Phe Pro Cys Ala Arg Gly Gln Cys
 1345 1350 1355 1360
 Val Asp Leu Arg Leu Arg Cys Asp Gly Glu Ala Asp Cys Gln Asp Arg
 1365 1370 1375
 Ser Asp Glu Ala Asp Cys Asp Ala Ile Cys Leu Pro Asn Gln Phe Arg
 1380 1385 1390
 Cys Ala Ser Gly Gln Cys Val Leu Ile Lys Gln Gln Cys Asp Ser Phe
 1395 1400 1405
 Pro Asp Cys Ile Asp Gly Ser Asp Glu Leu Met Cys Glu Ile Thr Lys
 1410 1415 1420
 Pro Pro Ser Asp Asp Ser Pro Ala His Ser Ser Ala Ile Gly Pro Val
 1425 1430 1435 1440
 Ile Gly Ile Ile Leu Ser Leu Phe Val Met Gly Gly Val Tyr Phe Val
 1445 1450 1455

10

20

30

Cys Gln Arg Val Val Cys Gln Arg Tyr Ala Gly Ala Asn Gly Pro Phe
 1460 1465 1470
 Pro His Glu Tyr Val Ser Gly Thr Pro His Val Pro Leu Asn Phe Ile
 1475 1480 1485
 Ala Pro Gly Gly Ser Gln His Gly Pro Phe Thr Gly Ile Ala Cys Gly
 1490 1495 1500
 Lys Ser Met Met Ser Ser Val Ser Leu Met Gly Gly Arg Gly Gly Val
 1505 1510 1515 1520
 Pro Leu Tyr Asp Arg Asn His Val Thr Gly Ala Ser Ser Ser Ser
 1525 1530 1535
 Ser Ser Thr Lys Ala Thr Leu Tyr Pro Pro Ile Leu Asn Pro Pro Pro
 1540 1545 1550
 Ser Pro Ala Thr Asp Pro Ser Leu Tyr Asn Met Asp Met Phe Tyr Ser
 1555 1560 1565
 Ser Asn Ile Pro Ala Thr Ala Arg Pro Tyr Arg Pro Tyr Ile Ile Arg
 1570 1575 1580
 Gly Met Ala Pro Pro Thr Thr Pro Cys Ser Thr Asp Val Cys Asp Ser
 1585 1590 1595 1600
 Asp Tyr Ser Ala Ser Arg Trp Lys Ala Ser Lys Tyr Tyr Leu Asp Leu
 1605 1610 1615
 Asn Ser Asp Ser Asp Pro Tyr Pro Pro Pro Thr Pro His Ser Gln
 1620 1625 1630
 Tyr Leu Ser Ala Glu Asp Ser Cys Pro Pro Ser Pro Ala Thr Glu Arg
 1635 1640 1645
 Ser Tyr Phe His Leu Phe Pro Pro Pro Pro Ser Pro Cys Thr Asp Ser
 1650 1655 1660
 Ser
 1665

10

20

<210> 764
 <211> 1613
 <212> PRT
 <213> Homo sapiens

<400> 764
 Met Gly Ala Val Leu Arg Ser Leu Leu Ala Cys Ser Phe Cys Val Leu
 1 5 10 15
 Leu Arg Ala Ala Pro Leu Leu Leu Tyr Ala Asn Arg Arg Asp Leu Arg
 20 25 30
 Leu Val Asp Ala Thr Asn Gly Lys Glu Asn Ala Thr Ile Val Val Gly
 35 40 45
 Gly Leu Glu Asp Ala Ala Val Asp Phe Val Phe Ser His Gly Leu
 50 55 60
 Ile Tyr Trp Ser Asp Val Ser Glu Glu Ala Ile Lys Arg Thr Glu Phe
 65 70 75 80
 Asn Lys Thr Glu Ser Val Gln Asn Val Val Val Ser Gly Leu Leu Ser
 85 90 95
 Pro Asp Gly Leu Ala Cys Asp Trp Leu Gly Glu Lys Leu Tyr Trp Thr
 100 105 110

30

Asp Ser Glu Thr Asn Arg Ile Glu Val Ser Asn Leu Asp Gly Ser Leu
 115 120 125
 Arg Lys Val Leu Phe Trp Gln Glu Leu Asp Gln Pro Arg Ala Ile Ala
 130 135 140
 Leu Asp Pro Ser Ser Gly Phe Met Tyr Trp Thr Asp Trp Gly Glu Val
 145 150 155 160
 Pro Lys Ile Glu Arg Ala Gly Met Asp Gly Ser Ser Arg Phe Ile Ile
 165 170 175
 Ile Asn Ser Glu Ile Tyr Trp Pro Asn Gly Leu Thr Leu Asp Tyr Glu
 180 185 190
 Glu Gln Lys Leu Tyr Trp Ala Asp Ala Lys Leu Asn Phe Ile His Lys
 195 200 205
 Ser Asn Leu Asp Gly Thr Asn Arg Gln Ala Val Val Lys Gly Ser Leu
 210 215 220
 Pro His Pro Phe Ala Leu Thr Leu Phe Glu Asp Ile Leu Tyr Trp Thr
 225 230 235 240
 Asp Trp Ser Thr His Ser Ile Leu Ala Cys Asn Lys Tyr Thr Gly Glu
 245 250 255
 Gly Leu Arg Glu Ile His Ser Asp Ile Phe Ser Pro Met Asp Ile His
 260 265 270
 Ala Phe Ser Gln Gln Arg Gln Pro Asn Ala Thr Asn Pro Cys Gly Ile
 275 280 285
 Asp Asn Gly Gly Cys Ser His Leu Cys Leu Met Ser Pro Val Lys Pro
 290 295 300
 Phe Tyr Gln Cys Ala Cys Pro Thr Gly Val Lys Leu Leu Glu Asn Gly
 305 310 315 320
 Lys Thr Cys Lys Asp Gly Ala Thr Glu Leu Leu Leu Leu Ala Arg Arg
 325 330 335
 Thr Asp Leu Arg Arg Ile Ser Leu Asp Thr Pro Asp Phe Thr Asp Ile
 340 345 350
 Val Leu Gln Leu Glu Asp Ile Arg His Ala Ile Ala Ile Asp Tyr Asp
 355 360 365
 Pro Val Glu Gly Tyr Ile Tyr Trp Thr Asp Asp Glu Val Arg Ala Ile
 370 375 380
 Arg Arg Ser Phe Ile Asp Gly Ser Gly Ser Gln Phe Val Val Thr Ala
 385 390 395 400
 Gln Ile Ala His Pro Asp Gly Ile Ala Val Asp Trp Val Ala Arg Asn
 405 410 415
 Leu Tyr Trp Thr Asp Thr Gly Thr Asp Arg Ile Glu Val Thr Arg Leu
 420 425 430
 Asn Gly Thr Met Arg Lys Ile Leu Ile Ser Glu Asp Leu Glu Glu Pro
 435 440 445
 Arg Ala Ile Val Leu Asp Pro Met Val Gly Tyr Met Tyr Trp Thr Asp
 450 455 460
 Trp Gly Glu Ile Pro Lys Ile Glu Arg Ala Ala Leu Asp Gly Ser Asp
 465 470 475 480
 Arg Val Val Leu Val Asn Thr Ser Leu Gly Trp Pro Asn Gly Leu Ala
 485 490 495
 Leu Asp Tyr Asp Glu Gly Lys Ile Tyr Trp Gly Asp Ala Lys Thr Asp
 500 505 510

10

20

30

Lys Ile Glu Val Met Asn Thr Asp Gly Thr Gly Arg Arg Val Leu Val
 515 520 525
 Glu Asp Lys Ile Pro His Ile Phe Gly Phe Thr Leu Leu Gly Asp Tyr
 530 535 540
 Val Tyr Trp Thr Asp Trp Gln Arg Arg Ser Ile Glu Arg Val His Lys
 545 550 555 560
 Arg Ser Ala Glu Arg Glu Val Ile Ile Asp Gln Leu Pro Asp Leu Met
 565 570 575
 Gly Leu Lys Ala Thr Asn Val His Arg Val Ile Gly Ser Asn Pro Cys
 580 585 590
 Ala Glu Glu Asn Gly Gly Cys Ser His Leu Cys Leu Tyr Arg Pro Gln
 595 600 605
 Gly Leu Arg Cys Ala Cys Pro Ile Gly Phe Glu Leu Ile Ser Asp Met
 610 615 620
 Lys Thr Cys Ile Val Pro Glu Ala Phe Leu Leu Phe Ser Arg Arg Ala
 625 630 635 640
 Asp Ile Arg Arg Ile Ser Leu Glu Thr Asn Asn Asn Asn Val Ala Ile
 645 650 655
 Pro Leu Thr Gly Val Lys Glu Ala Ser Ala Leu Asp Phe Asp Val Thr
 660 665 670
 Asp Asn Arg Ile Tyr Trp Thr Asp Ile Ser Leu Lys Thr Ile Ser Arg
 675 680 685
 Ala Phe Met Asn Gly Ser Ala Leu Glu His Val Val Glu Phe Gly Leu
 690 695 700
 Asp Tyr Pro Glu Gly Met Ala Val Asp Trp Leu Gly Lys Asn Leu Tyr
 705 710 715 720
 Trp Ala Asp Thr Gly Thr Asn Arg Ile Glu Val Ser Lys Leu Asp Gly
 725 730 735
 Gln His Arg Gln Val Leu Val Trp Lys Asp Leu Asp Ser Pro Arg Ala
 740 745 750
 Leu Ala Leu Asp Pro Ala Glu Gly Phe Met Tyr Trp Thr Glu Trp Gly
 755 760 765
 Gly Lys Pro Lys Ile Asp Arg Ala Ala Met Asp Gly Ser Glu Arg Thr
 770 775 780
 Thr Leu Val Pro Asn Val Gly Arg Ala Asn Gly Leu Thr Ile Asp Tyr
 785 790 795 800
 Ala Lys Arg Arg Leu Tyr Trp Thr Asp Leu Asp Thr Asn Leu Ile Glu
 805 810 815
 Ser Ser Asn Met Leu Gly Leu Asn Arg Glu Val Ile Ala Asp Asp Leu
 820 825 830
 Pro His Pro Phe Gly Leu Thr Gln Tyr Gln Asp Tyr Ile Tyr Trp Thr
 835 840 845
 Asp Trp Ser Arg Arg Ser Ile Glu Arg Ala Asn Lys Thr Ser Gly Gln
 850 855 860
 Asn Arg Thr Ile Ile Gln Gly His Leu Asp Tyr Val Met Asp Ile Leu
 865 870 875 880
 Val Phe His Ser Ser Arg Gln Ser Gly Trp Asn Glu Cys Ala Ser Ser
 885 890 895
 Asn Gly His Cys Ser His Leu Cys Leu Ala Val Pro Val Gly Gly Phe
 900 905 910

10

20

30

Val Cys Gly Cys Pro Ala His Tyr Ser Leu Asn Ala Asp Asn Arg Thr
 915 920 925
 Cys Ser Ala Pro Thr Thr Phe Leu Leu Phe Ser Gln Lys Ser Ala Ile
 930 935 940
 Asn Arg Met Val Ile Asp Glu Gln Gln Ser Pro Asp Ile Ile Leu Pro
 945 950 955 960
 Ile His Ser Leu Arg Asn Val Arg Ala Ile Asp Tyr Asp Pro Leu Asp
 965 970 975
 Lys Gln Leu Tyr Trp Ile Asp Ser Arg Gln Asn Met Ile Arg Lys Ala
 980 985 990
 Gln Glu Asp Gly Ser Gln Gly Phe Thr Val Val Val Ser Ser Val Pro
 995 1000 1005
 Ser Gln Asn Leu Glu Ile Gln Pro Tyr Asp Leu Ser Ile Asp Ile Tyr
 1010 1015 1020
 Ser Arg Tyr Ile Tyr Trp Thr Cys Glu Ala Thr Asn Val Ile Asn Val
 1025 1030 1035 1040
 Thr Arg Leu Asp Gly Arg Ser Val Gly Val Val Leu Lys Gly Glu Gln
 1045 1050 1055
 Asp Arg Pro Arg Ala Ile Val Val Asn Pro Glu Lys Gly Tyr Met Tyr
 1060 1065 1070
 Phe Thr Asn Leu Gln Glu Arg Ser Pro Lys Ile Glu Arg Ala Ala Leu
 1075 1080 1085
 Asp Gly Thr Glu Arg Glu Val Leu Phe Phe Ser Gly Leu Ser Lys Pro
 1090 1095 1100
 Ile Ala Leu Ala Leu Asp Ser Arg Leu Gly Lys Leu Phe Trp Ala Asp
 1105 1110 1115 1120
 Ser Asp Leu Arg Arg Ile Glu Ser Ser Asp Leu Ser Gly Ala Asn Arg
 1125 1130 1135
 Ile Val Leu Glu Asp Ser Asn Ile Leu Gln Pro Val Gly Leu Thr Val
 1140 1145 1150
 Phe Glu Asn Trp Leu Tyr Trp Ile Asp Lys Gln Gln Gln Met Ile Glu
 1155 1160 1165
 Lys Ile Asp Met Thr Gly Arg Glu Gly Arg Thr Lys Val Gln Ala Arg
 1170 1175 1180
 Ile Ala Gln Leu Ser Asp Ile His Ala Val Lys Glu Leu Asn Leu Gln
 1185 1190 1195 1200
 Glu Tyr Arg Gln His Pro Cys Ala Gln Asp Asn Gly Gly Cys Ser His
 1205 1210 1215
 Ile Cys Leu Val Lys Gly Asp Gly Thr Thr Arg Cys Ser Cys Pro Met
 1220 1225 1230
 His Leu Val Leu Leu Gln Asp Glu Leu Ser Cys Gly Glu Pro Pro Thr
 1235 1240 1245
 Cys Ser Pro Gln Gln Phe Thr Cys Phe Thr Gly Glu Ile Asp Cys Ile
 1250 1255 1260
 Pro Val Ala Trp Arg Cys Asp Gly Phe Thr Glu Cys Glu Asp His Ser
 1265 1270 1275 1280
 Asp Glu Leu Asn Cys Pro Val Cys Ser Glu Ser Gln Phe Gln Cys Ala
 1285 1290 1295
 Ser Gly Gln Cys Ile Asp Gly Ala Leu Arg Cys Asn Gly Asp Ala Asn
 1300 1305 1310

10

20

30

Cys Gln Asp Lys Ser Asp Glu Lys Asn Cys Glu Val Leu Cys Leu Ile
 1315 1320 1325
 Asp Gln Phe Arg Cys Ala Asn Gly Gln Cys Ile Gly Lys His Lys Lys
 1330 1335 1340
 Cys Asp His Asn Val Asp Cys Ser Asp Lys Ser Asp Glu Leu Asp Cys
 1345 1350 1355 1360
 Tyr Pro Thr Glu Glu Pro Ala Pro Gln Ala Thr Asn Thr Val Gly Ser
 1365 1370 1375
 Val Ile Gly Val Ile Val Thr Ile Phe Val Ser Gly Thr Val Tyr Phe
 1380 1385 1390
 Ile Cys Gln Arg Met Leu Cys Pro Arg Met Lys Gly Asp Gly Glu Thr
 1395 1400 1405
 Met Thr Asn Asp Tyr Val Val His Gly Pro Ala Ser Val Pro Leu Gly
 1410 1415 1420
 Tyr Val Pro His Pro Ser Ser Leu Ser Gly Ser Leu Pro Gly Met Ser
 1425 1430 1435 1440
 Arg Gly Lys Ser Met Ile Ser Ser Leu Ser Ile Met Gly Gly Ser Ser
 1445 1450 1455
 Gly Pro Pro Tyr Asp Arg Ala His Val Thr Gly Ala Ser Ser Ser Ser
 1460 1465 1470
 Ser Ser Ser Thr Lys Gly Thr Tyr Phe Pro Ala Ile Leu Asn Pro Pro
 1475 1480 1485
 Pro Ser Pro Ala Thr Glu Arg Ser His Tyr Thr Met Glu Phe Gly Tyr
 1490 1495 1500
 Ser Ser Asn Ser Pro Ser Thr His Arg Ser Tyr Ser Tyr Arg Pro Tyr
 1505 1510 1515 1520
 Ser Tyr Arg His Phe Ala Pro Pro Thr Thr Pro Cys Ser Thr Asp Val
 1525 1530 1535
 Cys Asp Ser Asp Tyr Ala Pro Ser Arg Arg Met Thr Ser Val Ala Thr
 1540 1545 1550
 Ala Lys Gly Tyr Thr Ser Asp Leu Asn Tyr Asp Ser Glu Pro Val Pro
 1555 1560 1565
 Pro Pro Pro Thr Pro Arg Ser Gln Tyr Leu Ser Ala Glu Glu Asn Tyr
 1570 1575 1580
 Glu Ser Cys Pro Pro Ser Pro Tyr Thr Glu Arg Ser Tyr Ser His His
 1585 1590 1595 1600
 Leu Tyr Pro Pro Pro Pro Ser Pro Cys Thr Asp Ser Ser
 1605 1610

<210> 765

<211> 207

<212> PRT

<213> Artificial Sequence

<220>

<223> cytoplasmic domain and portion of transmembrane domain

<400> 765

Arg Val Val Cys Gln Arg Tyr Ala Gly Ala Asn Gly Pro Phe Pro His

10

20

30

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

10

20

30

【図 1 A】

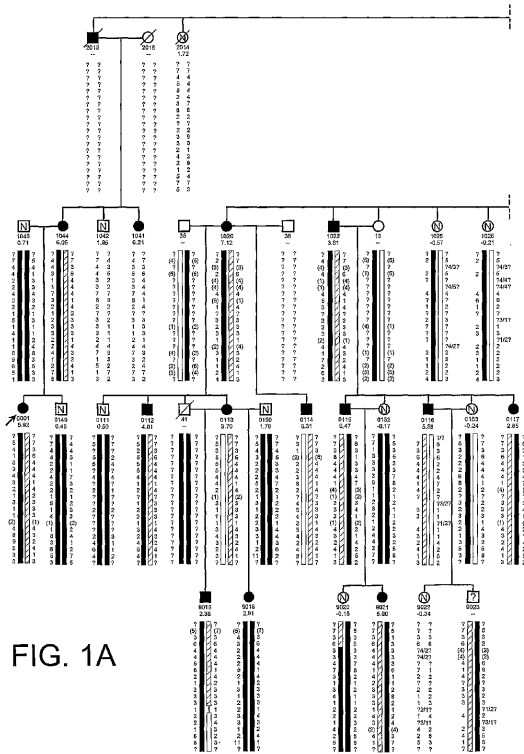


FIG. 1A

【図 1 B】

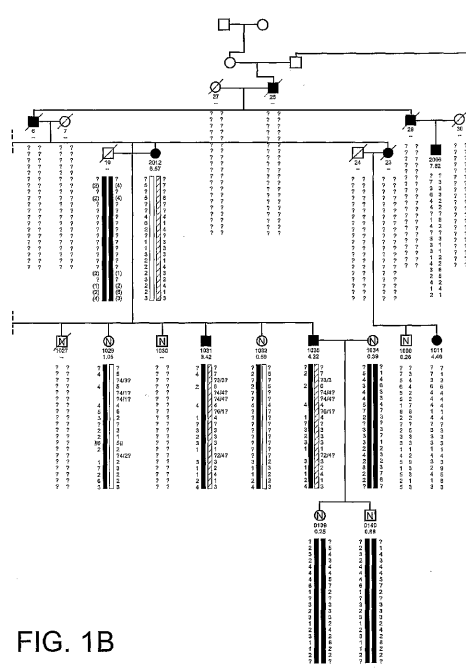


FIG. 1B

【図 2 A】

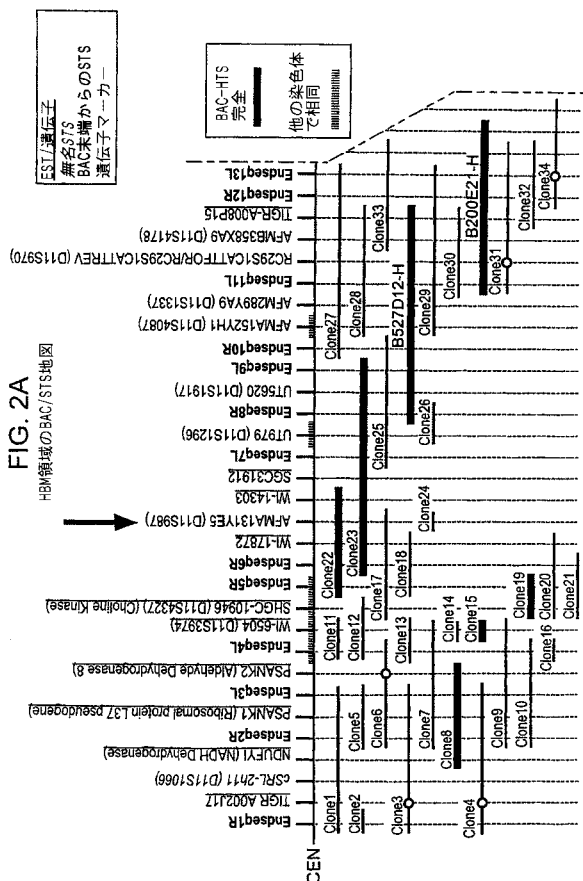


FIG. 2A

【図 2 B】

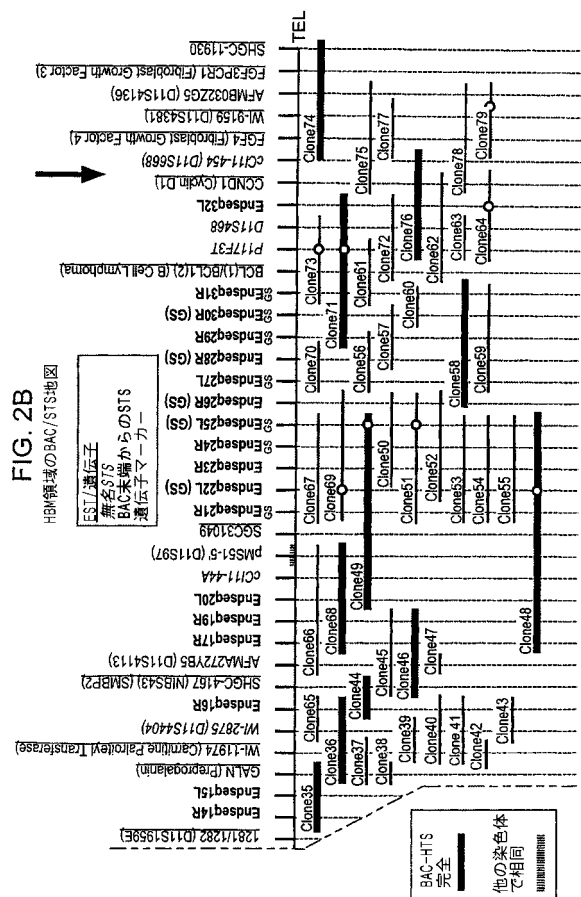


FIG. 2B

【図 3 A】

Exon 1
 ACTAAAGCGCCGCCGCGCCATGGAGCCCGAGTGAGCGCGGCGCGG
 GCCCGTCCGGCCGCCGGAACAATGGAGGCAAGCGCCGCCCGGCGCGC
 GTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCGGCTGTGCGGCTGC
 CCGGCCCCGCGCGCGC

Exon 2 Coordinates: 527d12_Contig308G 30944-30549
 gccccacagCTTCGCCCTCTGCTATTGCGCAACGCCCGGACGTACGGCT
 GGTGGACGCCCGCGGACTCAAGCTGGAGTCCACCATCGTGGTACGCGG
 CCTGGAGGATGCGGCGCGAGTGGACTTCCAGTTTCCAAAGGAGCGGTG
 TACTGGACAGACGTAGCGGAGGAGGCCATCAAGCAGACCTACCTGAACC
 AGACGGGGGCCCGCTGCAGAACCTGGTTCATCTCCGGCTTGGTCTCTCC
 CGACGGCTTCGCTGCGACTGGGTGGGCAAGAAGCTGACTGGACGGA
 CTCAGAGACCAACCGCATCGAGGTGGCCAACCTCAATGGCACATCCCGG
 AAGGTGCTCTTCTGGCAGGACCTTGACACGCGAGGGCCATCGCCTTGG
 ACCCGCTCACGGtagaacctgtg

... 9408 nt ...

Exon 3 Coordinates: 527d12_Contig308G 21141-20945
 cccgtacagGTACATGTACTGGACAGACTGGGCTGAGACGCCCGGATTGA
 GCGGGCAGGGATGGATGGACAGCACCCGGAAGATCATTTGGACTCGGA
 CATTATCTGGCCCAATGGACTGACCATCGCACTGGAGGAGCAGAAGCTC
 TACTGGGTGACGCCAAGCTCAGCTTATCCACCGTGGCCAACCTGGACG
 GCTCGTTCGgttagtccac

... 6094 nt ...

Exon 4 Coordinates: 527d12_Contig308G 15047-14850
 tccctgactgengCAGAAGGTGGTGGAGGGCAGCCTGACGCACCCCTTCGCCC
 TGACGCTCTCCGGGGACACTCTGTACTGGACAGACTGGCAGACCCGCTC
 CATCCATGCTTGCACAACAGCGCACTGGGGGGAAGAGGAAGAGATCTCTG
 AGTGCCTTATACTCACCCATGGACATCCAGGTGCTGAGCCAGGAGCGCG
 AGCCTTCTgtgengtccgg

... 1827 nt ...

Exon 5 Coordinates: 527d12_Contig308G 13220-13088
 ttctcagTCCACATCGCTGTGAGGAGGACAATGGCGGCTGCTCCCACTGT
 GCTGTCTGTCCCAAGCGAGCCTTCTACACATGCGCTGCCCAAGCGG
 TGTGCAGCTGCAGGACAACGGCAGGACGTGTAAAGCAGgtgagggttggaag

FIG. 3A

【図 3 C】

Exon 10 Coordinates: 527d12_Contig309G 31358-31582
 cctgtgcagACCATCAGCCGCGCCTTATGAACCGGAGCTCGGTGGAGCAC
 GTGGTGGAGTTTGGCCTTGACTACCCCGAGGGCATGGCCGTTGACTGGA
 TGGGCAAGAACCTCTACTGGGCCGACACTGGGACCAACAGAATCGAAGT
 GGCGCGGTGGACGGGCAAGTTCGGCAAGTCTCTGTGGAGGGACIT
 GGACAACCCGAGGTGCTGGCCCTGGATCCCAACCAAGGgtgaagtgtttgcctgc

.....1297 nt.....

Exon 11 Coordinates: 527d12_Contig309G 32879-33064
 gtgccttcagCTACATCTACTGGACCGAGTGGGGCGGCAAGCCGAGGATCGT
 CGGGGCTTCATGGAGCGGACCAACTGCATGACCTGCTGGGACAGAGTG
 GGCGGGCCAAAGCACTCAGCAATGACTACGCTGACCGGCTCTACT
 GGACCGACCTGGACCAACATGATGCTAGTGTCTCAACATGCTGGgtgagg
 ccggct

.....2069 nt.....

Exon 12 Coordinates: 527d12_Contig309G 35133-35454
 gttgttcagGTCAGGAGCGGGTCTGATTGCCGACGATCTCCCGCACCCGT
 TCGGTCTGACGCAGTACAGCGATTATATCTACTGGACAGACTGGAATCT
 GCACAGATTGAGCGGGCCGACAAGACTAGCGGGCGGAACCGCACCCCTC
 ATCCAGGGCCACCTGGACTTCTGTGATGGACATCTCTGGTGTTCACCTCT
 CCCGCCAGGATGGCCTCAATGACTGTATGCACAACACGGGCAGTGTGG
 GCAGCTGTGCCTTGCCATCCCGCGCGGCCACCGCTGCGGCTGCCCTCA
 CACTACACCTTGGACCCGACAGCCGCAACTGCAGCCgtgaagtgcctatgtt

.....2006 nt.....

Exon 13 Coordinates: 527d12_Contig309G 37460-37659
 gctctctatCGCCCAACCACTTCTTGCTGTACGCAAGAAATCTGCCATCAGT
 CGGATGATCCGGGACGACGACGAGCCCGGATCTCATCTGCCCTTGC
 ATGCACTAGGAACGCTCAAAGCCATGCACTATGACCCAATGGACAAGTT
 CATCTACTGGGTGGATGGCGGCCAGAACAATCAAGCGAGCCAAGGACGAC
 GGGACCCAGgcagggtgcctgtg

.....6965 nt.....

FIG. 3C

【図 3 B】

... 20923 nt ...

Exon 6 Coordinates: 527d12_Contig309G 7705-8100
 ctccnagGAGCCGAGGAGGTGCTGCTGCTGGCCCGGCGGACGACCTACG
 GAGGATCTCTCTGGACACGCGGACTTCACCGACATCGTCTGCAGGTG
 GACGACATCCGCGACGCCATTGCCATCGACTACGACCCCTAGAGGGCT
 ATCTTACTTGGACAGATGACAGGTGCGGGCCATCCGAGGGCGTACTT
 GGACGGGTCTGGGGCGCAGACGTGGTCAACACCGAGATCAACGACCC
 CGATGGCATCGCGGTGCACTGGGTGGCCCCGAAACCTTACTTGGACCGAC
 ACGGGCACGGACCCGATCGAGGTGACGCGCTCAACGGCACCTCCCGCA
 AGATCTGTGTGCGGAGGACCTGGACGAGCCCCGAGCCATCGCACTGCA
 CCCCCTGATGGgttaagacgggc

..... 3211 nt

Exon 7 Coordinates: 527d12_Contig309G 11311-11482
 ttctctccagCCTCATGTACTGGACAGACTGGGGAGAGAACCCTAAATCGAG
 TGTGCCAATTGGATGGGCAGGAGCGGCGTGTGCTGGTCAATGCTCCC
 TCGGTTGGCCCAACGGGCTGGCCCTGGACCTGCAGGAGGGGAAGCTCT
 ACTGGGGAGACGCCAAGACAGACAAGATCGAGgtgaagtcctgttg

..... 13445 nt

Exon 8 Coordinates: 527d12_Contig309G 24927-25143
 ccgtctgcagGTGATCAATGTTGATGGGACGAAGAGGCGGACCCCTCCTGGAG
 GACAAGCTCCCGCACATTTTCGGGTTCACGCTGTGTGGGCACTTCATCT
 ACTGGACTGACTTGGACGCCCGGACATCGAGCGGGTGCACAAGGTCAA
 GGCCACCGCGGACGTATCATTTGACCAAGTTCGCCGACCTGATGGGGCTC
 AAAGCTGTGAATGTGGCCAAGGTCTGgtgagtcgggggggc

....2826 nt

Exon 9 Coordinates: 527d12_Contig309G 27969-28256
 gttegttcagGAACCAACCGTGTGCGGACAGGAACGGGGGTGCAGCCACC
 TGTGCTTCTTTCACACCCACGCAACCCGTTGTGGTGCCTTCCGCTTCCGCT
 GGAGCTGCTGAGTGACATGAAGACCTGCATCGTGCCTGAGGCGCTTCTTG
 GTCITTCACAGCAGAGCGGCCATCCACAGGATCTCCCTCGAGACCAATA
 ACAACGACGTGGCCCTCCGCTCACGGGCGTCAAGGAGGCGCTCAGCCCT
 GGACTTGTATGTGTCCAACAACACATCTACTGGACAGACGTACGCTGT
 AAGgtgagctgggc

.....3102.....

FIG. 3B

【図 3 D】

Exon 14 Coordinates: 527d12_Contig309G 44624-44832
 ctttgtctacagCCCTTGTGTTTGACCTCTCTGAGCCAAGGGCAAAACCCAGACA
 GGCAGCCCCACGACCTCAGCATCGACATCTACAGCCGACACTGTTCTGT
 GACGTGCGAGGGCCACCAATACCATCAAGCTCCACAGGCTGAGCGGGGA
 GCCATGGGGGTGTTGCTGCTGGGGACCCGACCAAGCCCAAGGGCCATC
 GTCTCAACGCGGAGCGAGGgttagaggccaac

.....1404 nt.....

Exon 15 Coordinates: 527d12_Contig309G 46236-46427
 ccacccctccagGTACCTGTACTTACCAACATGCAGGACCGGGCAGCCAAGA
 TCGAAACGCGAGCCCTGGAGCGGACCGAGCCGAGGTCTCTTCAACAC
 CGGCTCATCCGCTCTGTCCTGTTGGTGTGGTGTGGAACAACATGCGCAAG
 CTGTCTTGGTGGACGCGGACCTGAAGCGCATTGAGAGCTGTGACCTGT
 CAGgtacgcgcocgg

.....686 nt.....

Exon 16 Coordinates: 527d12_Contig309G 47113-47322
 ggcgtgtcagGGGCCAACCGCTGACCTGGAGGACGCCAACATCGTGCAGC
 CTCTGGGCTGACCATCCTTGGGAAGCATCTCTACTGGATCGAACGCCCA
 GCAGCAGATGATCGAGCGTGTGGAGAAACACCCGGGGACAAGCGGAC
 TCGCATCCAGGGCGGTGTGCCCACTTCACTGGCATTCATGCACTGGAG
 GAAGTCAGCCTGGAGGAGTCTgtacgtgggggc

.....3884 nt.....

Exon 17 Coordinates: 527d12_Contig309G 51206-51331
 tttgttgtcagCAGCCACCACTGTGCCCCGTGACAATGGTGGCTGTCCACAT
 CTGTATTGCCAAGGCTGATGGGACACACGCTGCTATGCCCACTGCCAC
 CTCGTGCTCTCTCAGAACCTGTGACCTGTGGAGgttaggtgacataggto

.....3905 nt.....

Exon 18 Coordinates: 527d12_Contig309G 55236-55472
 gttctctctgtccctccocccagAGCCGCCCCACCTGCTCCCGGACCAAGTTGCATGTG
 CCACAGGGGAGATCGACTGTATCCCGGGGCGTGGCGTGTGACGGCTT
 TCCCGAGTGCATGACCAAGCGGACGAGGAGGGCTGCCCGTGTGCTCC
 GCGGCCAGTTCCCTTCCGCGCGGGGTCAAGTGTGTGACCTGCGGCTGC
 GCTGCGACGGCGAGGCAGACTGTCAAGGACCGCTCAGACGAGGTGGACT
 GTGACGgttaggcctcc

.....3052 nt.....

FIG. 3D

【図 3 E】

Exon 19 Coordinates: 527d12_Contig309G 58524-58634
 tctcttcgagCCATCTGCCTGCCAACCAAGTTCCGGTGTGCGAGCGGCCAGTG
 TGTCTCTACTCAAAACAGCATGCGACTCTCTCCCCAGCTGTATCGACGGCT
 CCGACGAGCTCATGTGTGgtgagccagctt

.....1448 nt.....

[illegible]

...1095 nt....

Exon 21 Coordinates: 527d12_Contig309G 61414-61552
cttcctgcagGCATCGCATGCGGAAAGTCCATGATGAGCTCCGTGAGCCTGA
TGGGGGGCGGGGGGGGGTGCCTCTACGACCGGAACCAACGTCACG
GGGCCTCGTCCAGCAGCTCGTCCAGCACGAAGGCCACGCTGATACCGCC
Ggtgagggggggg

.....6513 nt.....

Exon 22 Coordinates: 527d12_Contig309G 68065-68162
 ttgctctctcagATCTTGAAACCGCGCCCTCCCCGGCCACGGACCCCTCCCT
 GTACAACATGGACATGTTCTACTCTTCAAACATTCGGGCCACTGCGAGAC
 CGTACAGgtaggacatccccctcag

.....2273 nt.....

FIG. 3E

【図 3 F】

[illegible]

FIG. 3F

【図 4】

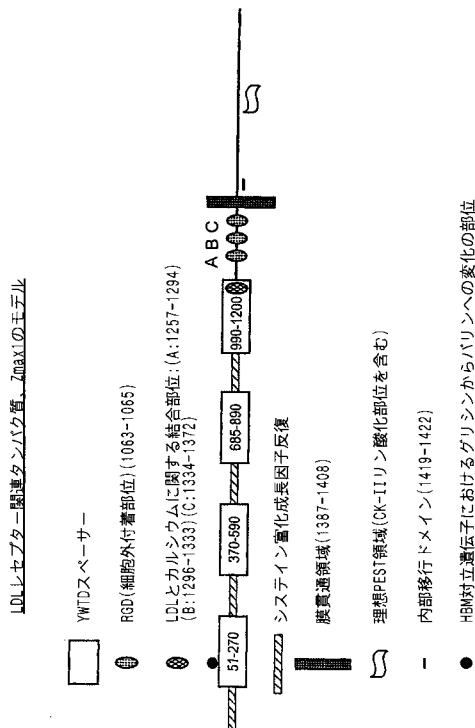


FIG. 4

【図 5】

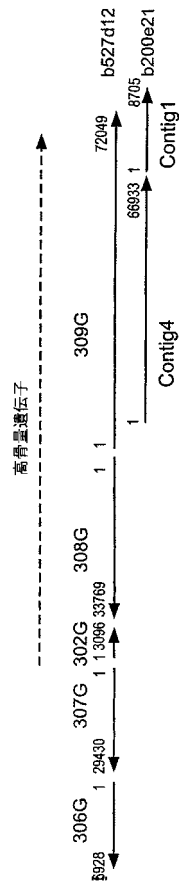


FIG. 5

【図 6 E】

FIG. 6E

2221 TGACTACCCGAGGCGATGCGCTTGACTTGATGGCAGAACTCTACTTGGCGCGACAC 2280
 718 D Y P E G M A V D W M G K N L Y W A D T 737
 2281 TGGGACCAACAGATGAAAGTGGCGGCTGGACGGGCAAGTTCGGGCAAGTCTCTGTTG 2340
 738 G T N R I E V A R L D G Q F R Q V L V W 757
 2341 GAGGCACTTGGCAACCCGAGGTGCTGGCCCTGGATCCCAACCAAGGCTACATCTACTG 2400
 758 R D L D N P R S L A L D P T K G Y I Y W 777
 2401 GACCGACTGGCGCGCAACCGAGGATCTCTGGGCTTCATGGAAGGACCAACTGCAT 2460
 778 T E E W G G K P R I V R A F M D G T N C M 797
 2461 GACCTGGTGACAGGTGGCGCGGCGACGACCTCACCAATGACTACCTGACGACG 2520
 798 T L V D K V G R A N D L T I D Y A D Q R 817
 2521 CCTCTACTGGACCAACTGGACACCAACATGATCGAGTCGTCCCAACATGCTGGGTCAGGA 2580
 818 L Y W T D L D T N M I E S S N M L G Q E 837
 2581 GCGGTCGTGATTCGCGCATCTCCCGACCCGTTTCGCTTCGACCAAGTACAGCGATTGA 2640
 838 R V V I A C D D L P H F F G L T Q Y S D Y 857
 2641 TATCTACTGACACACTGGATCTGCACAGCATTTAGCGGCGCGACAGACTAGCGCGG 2700
 858 I Y W T D W N L H S I E R A D K T S G R 877
 2701 GAACGCAACCTTCATCCAGGGCCACTTGGACTTCGTGATGGAATCTCTGGTGTTCACATC 2760
 878 N R T L I I Q G H L D F V M D I L V F H S 897

【図 6 F】

FIG. 6F

2761 CTCGCGCAGATGGCTCAATGATGTATGCAACAAACGGGCAAGTGGGCACTGTG 2820
 898 S R Q D G L N D C M H N N G Q C G Q L C 917
 2821 CTTTGCCATCCCGGGGCGCACGCGCTGGCGCTCGGCTCACACTACACCTGGACCCGAC 2880
 918 L A I P G H R C G C A S H Y I L D P S 937
 2881 CAGCGCAACTGAGCGCGCCCAACCACTTCTTGCTGTTCAAGCAAGAAATCTGCCATCAG 2940
 938 S R N C S P P T F L L F S Q K S A I S 957
 2941 TGGGATATCCGGACCCAGACAGCCCGGATCTCATCTGCCCCCTGCATGGACTGAG 3000
 958 R M I P D D Q H S P D L I L P L H G L R 977
 3001 GAACTTCAAAGCCATGACTATGACCCACTGACAGTTCATCTACTGCGTGGATGGCG 3060
 978 N V K A I D Y D P L D K F I Y W D G R 997
 3061 CAGAAATCAAGGAGCCCAAGACGAGGAGCCAGGCGCTTGTGTTGACCTCTGTGAG 3120
 998 Q N I K R A K D D G T Q P F V L T S L S 1017
 3121 CCAAGGCCAAACCCAGACAGGCGAGCCCGACCTTCAGCATCGACATCTACAGCCGAC 3180
 1018 Q G Q N P D R Q P H D L S I D I Y S R T 1037
 3181 ACTGTTCTGAGGTGAGGCGACCAATACATCAACTCCACAGGCTGAGCGGGAAC 3240
 1038 L F W T C E A T N T I N V H R L S G E A 1057
 3241 CATGGGGTGTGCTGCTGGGACCGCGGCAAGCCAGGGCCATCGCTGTCAACGGGA 3300
 1058 M G V V L R G D R D K P R A I V V N A E 1077

【図 6 G】

FIG. 6G

3301 CGAGGTACTCTGTACTTCCACACATGCMGGGACCGGGACCCAGATCGAAGCGCGACG 3360
 1078 R G Y L Y F T N M Q D R A A K I E R A A 1097
 3361 CTTGAGCGGACCGAGCGCGAGTCTCTTACACACCGGCGCTCATCCGCGCTGTGGCCCT 3420
 1098 L D G T E R E V L F T T G L I R P V A L 1117
 3421 GTGTGTGACAAACACACTGGGCAAGCTGTCTTGGGTGGACCGGACCTTGAAGCGCATGA 3480
 1118 V V D N T L G K L F W V D A D L K R I E 1137
 3481 GAGCTGTGACTCTGAGGGCCACACCCCTGACCTTGGAGGAGCCCAACATCTGAGCC 3540
 1138 S C D L S G A N R L E D A N I V Q P 1157
 3541 TCTGGCCCTGACCATCTTGGCAAGCATCTCTACTGATCGACCGCCAGCAGAGATGAT 3600
 1158 L G L T I L G K H L Y W I D R Q Q M I 1177
 3601 CGAGCTGTGGAGAAACACCGGGACAGCGGATCTCGCATCCAGGGCGCTTCGCCCA 3660
 1178 E R V E K T T G D K R T R I Q G R V A H 1197
 3661 CTTGACTGTGATCCATGATGAGGAGTCAAGCTGGAGGAGTTCACAGCCACCCATG 3720
 1198 L T G I H A V E V S L E F S A H P C 1217
 3721 TGCCCTGACAAATGTTGGTCTCCCACTCTGTAATTCGCAAGGTTGATGGGACACACG 3780
 1218 A R D N G G C S H I C I A K G D G T P R 1237
 3781 GTGCTCATGCCAGTCAACTCGTCTCTCGAGAACCTTGCTGACCTGTGGAGCGCGCC 3840
 1238 C S C P V H L V L L Q N L L T C G E P P 1257

【図 6 H】

FIG. 6H

3841 CACCTCTCCCGACCAAGTTTCATGTGCCACAGGGGAGTCGACTGTATCCCGGGGC 3900
 1258 T C S P D Q F A C A T G E I D C I P G A 1277
 3901 CTGGCCCTGTGACGGCTTCCCGAGTGCATGACAGAGCGACGAGGAGGCTGCCCGCT 3960
 1278 W R C D G F P E C D D Q S D E E G C P V 1297
 3961 GTGCTCCGCGCCAGTTTCCCTTGGCGCGGGGTCAGTGTGTGAGACTCGCGCTGCGCTG 4020
 1298 C S A A Q F P C A R G Q C V D L R L R C 1317
 4021 CGAGCGGAGCGAGACTGTACGAGCCGCTCAGACAGGTGAGCTGTGACGCCATCTGCT 4080
 1318 D G E A D C Q D R S D E V D C D A I C L 1337
 4081 GCCCAACAGATTCGGTGTGGAGCGCCAGTGTCTCTCATCAAAACAGCAGTCCGACTC 4140
 1338 P N Q F R C A S G Q C V L I K Q Q C D S 1357
 4141 CTTCCCGGACTGTATCGACGCTCCGACGAGCTCATGTGTGAATCAACAAAGCGCGCTC 4200
 1358 F P D C I D G S D E L M C E I T K P S 1377
 4201 AGACGACAGCCCGCCACAGCAAGTGCATCGGCGCGCTCATTTGGATGATCTCTCTCT 4260
 1378 D D S P A H S A I G P V I G I I S L 1397
 4261 CTTGCTCATGAGTGTGTATTTGTGTGCCAGCGCTGTGTGTCGACGCTATCGGG 4320
 1398 F V M G G V Y F V C Q R V V C Q R Y A G 1417
 4321 GGCACAGGGCCCTTCCGCAACAGTATGTACAGCGGACCCCGCAGCTGCCCTCAATTT 4380
 1418 A N G P F P H E Y V S G T P H V P L N F 1437

【図 6 I】

FIG. 6I

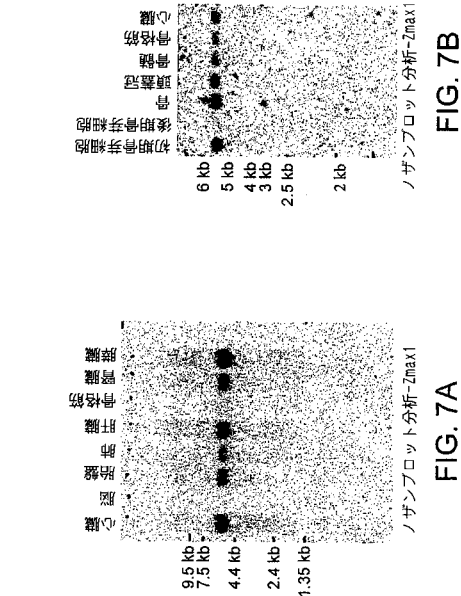
4381	CATAGCCCGGGCGGTTCCACGATGGCCCTTCACAGGATCGCATCGGGAAGTCCAT	4440
1438	I A P G S Q H G P F T G I A C G K S M	1457
4441	GATGAGCTCCGTAAGCTGATGGGGGCGGGCGGGTGCCCTCTACGACCGAACA	4500
1458	M S S V S L M G R G G V P L Y D R N H	1477
4501	CGTACACGGGGCTCTGTCACGAGCTCGTCCAGCAGGAGCCACGCTGTACCGCGAT	4560
1478	V T G A S S S S S T K A T L Y P P I	1497
4561	CTTGACCCCGGCGCTCCCGGCGACGACCCCTGCTGTACAAATGACATGTTCTA	4620
1498	L N P P S P A T D P S L Y N M D M F Y	1517
4621	CTCTTCAAAATTCGGGCGACTGCGAGACCGTACAGGCCCTACATCATTCGAGAGTGC	4680
1518	S S N I P A T A R P Y R P Y I I R G M A	1537
4681	GCCCCGAGCGCGCTGCGAGCACCGACGCTGTGACAGCGACTACAGCGCCAGCGCTG	4740
1538	P P T T P C S T D V C D S D Y S A S R W	1557
4741	GAAGGCCACGAAGTACTGGAATTTGAATCGGACTCAGACCCCTATCCACCCACC	4800
1558	K A S K Y Y L D L N S D S D P Y P P P	1577
4801	CACGCCCCACAGCCAGTACCTGTGCGCGGAGACAGCTGCCGCCCTCGCCCGCACCGA	4860
1578	T P H S Q Y L S A E D S C P P P A T E	1597
4861	GAGGAGTACTTCCATCTCTTCCGCGCCCTCCGCTCCCGCTGCAAGGACTCATCTGAC	4920
1598	R S Y F H L F P P P P S P C T D S S	1615

【図 6 J】

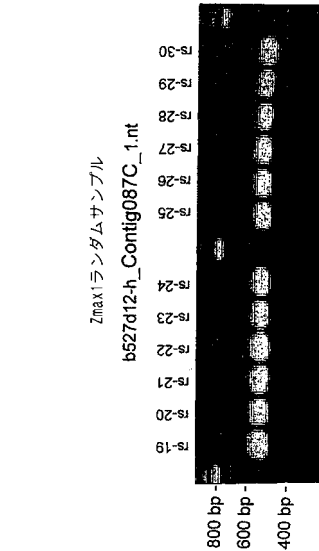
FIG. 6J

4921	TGGCGCGGCCACTCTGGCTCTCTGTGCCCTGTAAATAGTTTAAATGACAAAGA	4980
4981	AAAAATATATTTTATGATTTAAAAATTAATATTAATTTGGGATTTAAACATGAGAAA	5040
5041	TGTAACCTGTGATGGGTGGGAGGCTGGGAGACTTTGTACAGTGGAGAAATTTAT	5100
5101	AAACTTAATTTGTAAAAA	5120

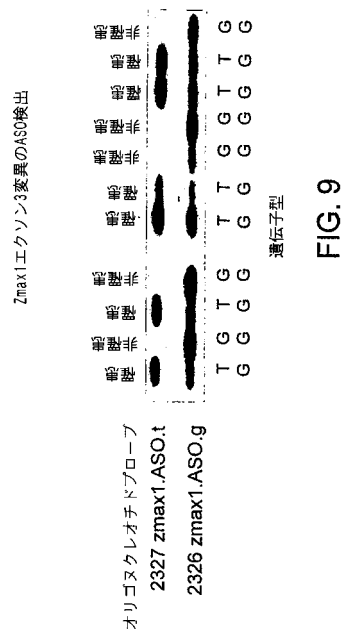
【図 7】



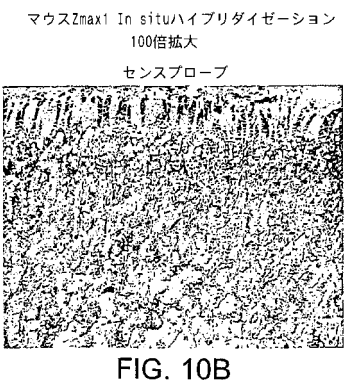
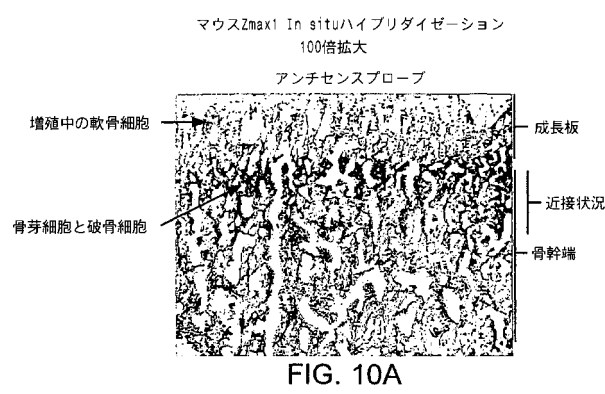
【図 8】



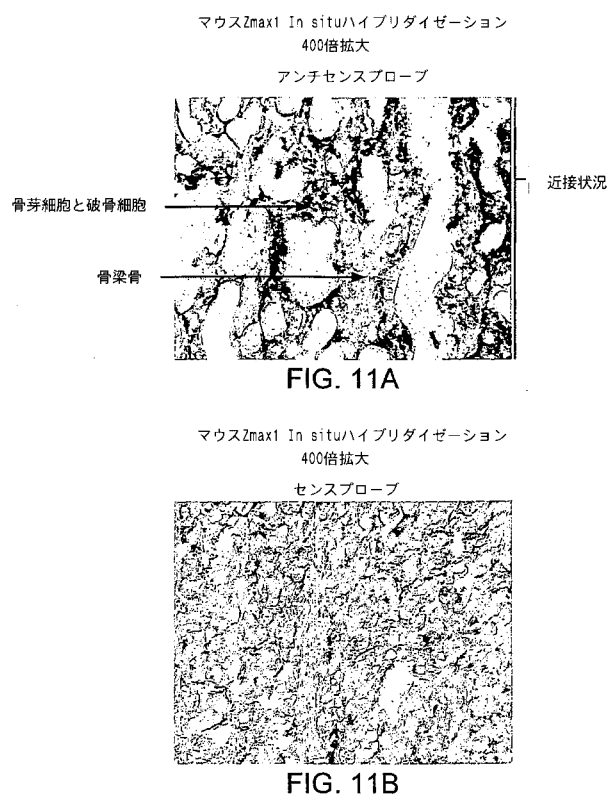
【図 9】



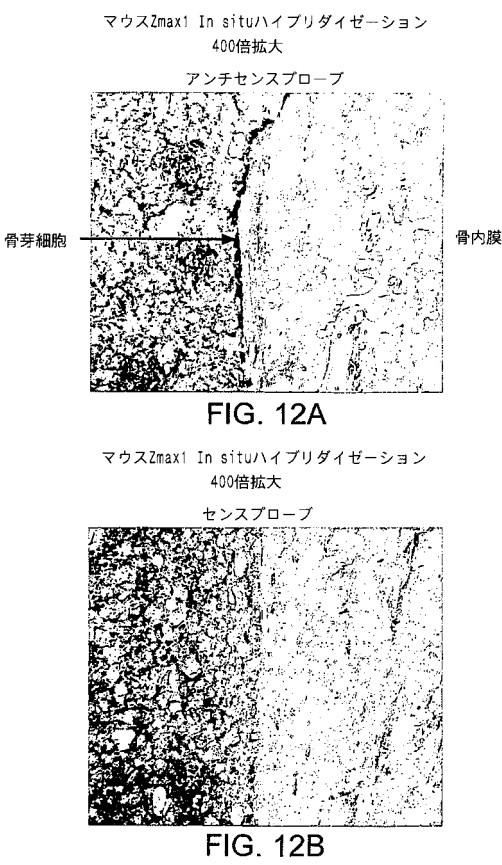
【図 10】



【図 11】



【図 12】



【図 17】

HBM: トランスジェニックマウスのための構造物
発現の確認
HOB-02-02細胞*への一過性トランスフェクション

CMV β アクチン		I型コラーゲン	
HBM X 1,000	Zmax1 X 1,000	HBM X 10	Zmax1 X 10
*空のベクターをトランスフェクトされた細胞におけるZmax1と比較した増加倍数			

FIG. 17

【図 18】

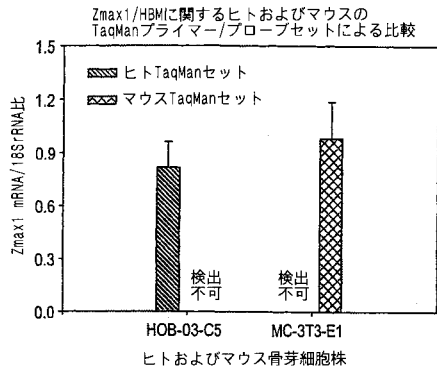


FIG. 18

【図 19】

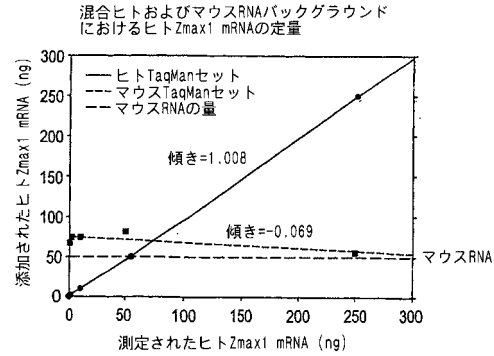


FIG. 19

【図 20】

HBM: トランスジェニックマウス
TaqMan分析によるmRNA発現*

株	組織						
	肋骨	大腿骨	心臓	生殖腺	脳	腎臓	肝臓
HBMMCB A							
2	7-10		20-90	2-30	6-11	5-9	<1
13	1-2		6-7	3-4	5-6	<1	<1
18	10-11						
HBMMTIC							
19	7-8	19-20	1				1
35	1	1	0				0

*HOB-03-C5におけるZmax1に対して

FIG. 20

【図 21 A】

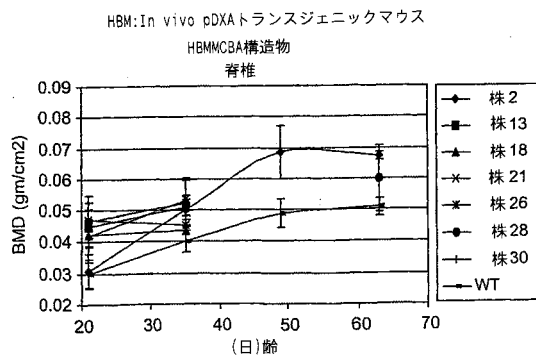


FIG. 21A

【図 21 C】

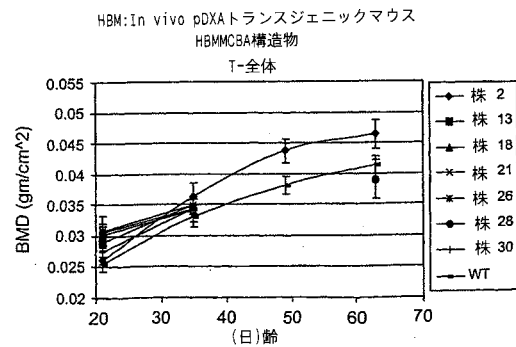


FIG. 21C

【図 21 B】

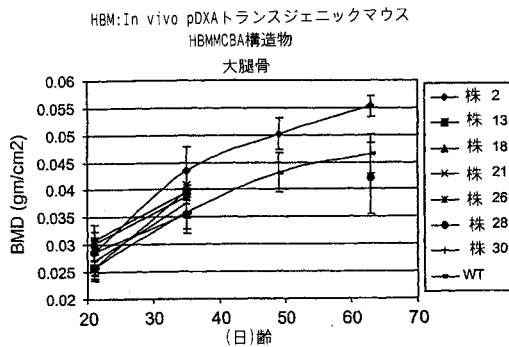


FIG. 21B

【図 21 D】

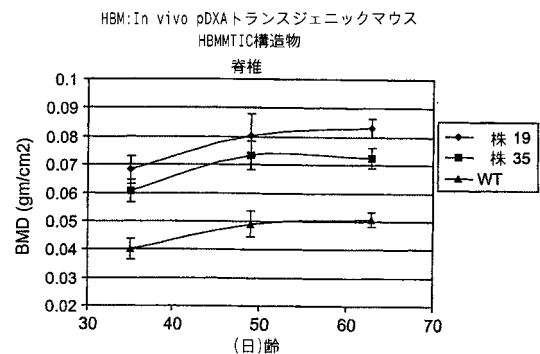
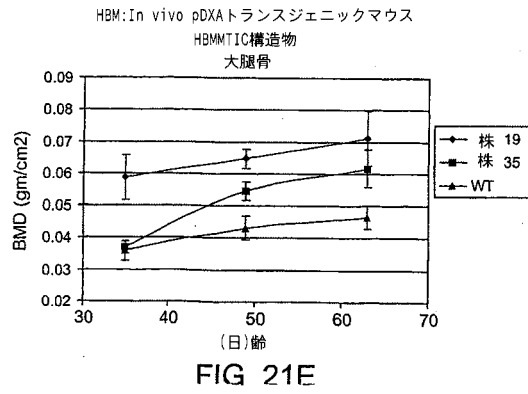
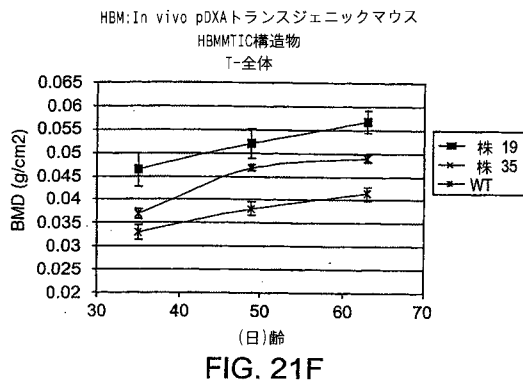


FIG. 21D

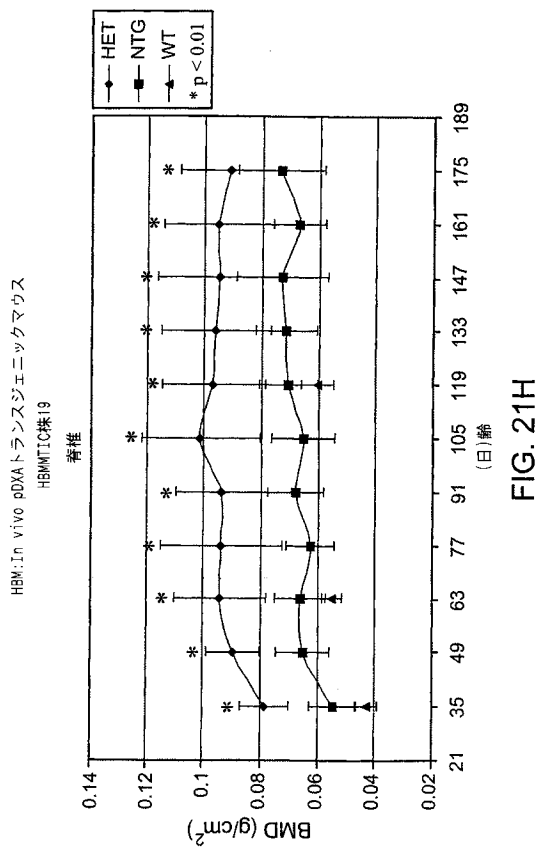
【図 2 1 E】



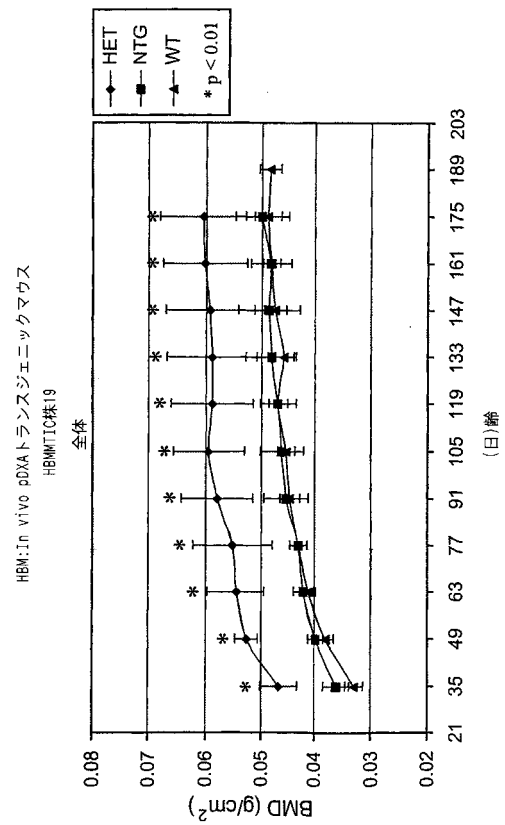
【図 2 1 F】



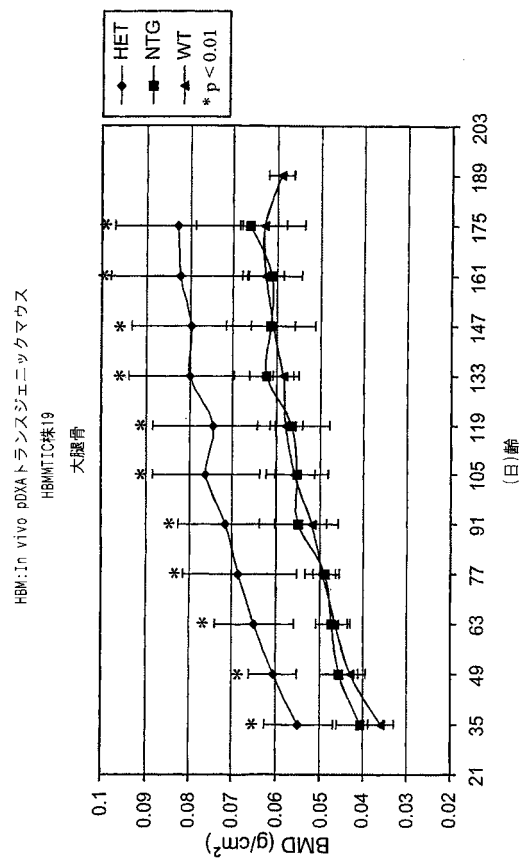
【図 2 1 H】



【図 2 1 G】



【図 2 1 I】



【図 2 1 J】

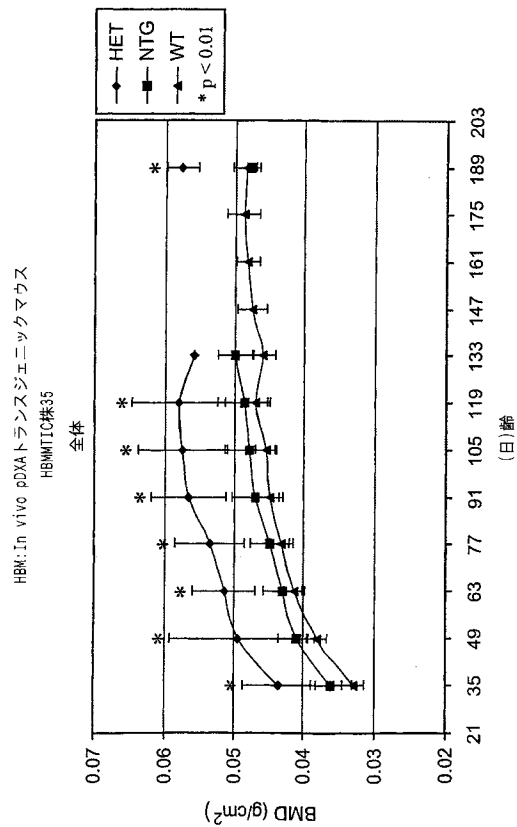


FIG. 21J

【図 2 1 K】

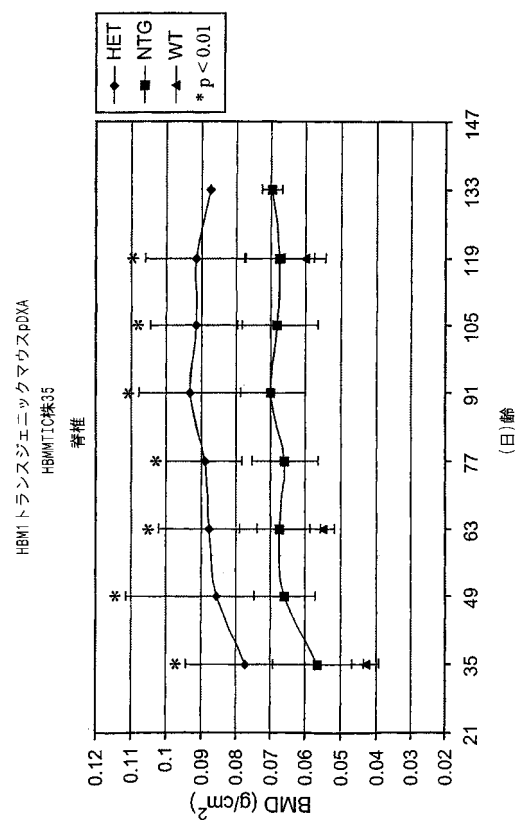


FIG. 21K

【図 2 1 L】

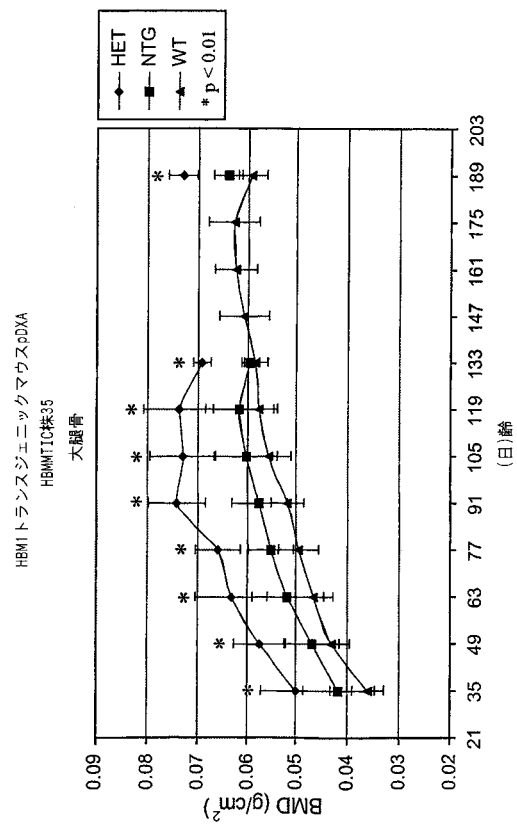


FIG. 21L

【図 2 2】

HBM transgenic mice: in vivo pDXA*
5-week-old mice vs WT for BMD% change

株	n	大腿骨	脊椎	全体	腰椎mRNA
HBM:MCBA					
2	11	21	24	10	x7-10
13	4	11	27	5	x1-2
18	1	8	9	3	x10-11
21	1	10	12	6	0
28	15	0	30	4	x1
30	7	4	32	4	0
HBM:MTIC					
19	5	63	70	41	x7-8
35	1	4	47	6	x1

*CV: 大腿骨(2.7%); 脊椎(6.4%); 全体(1.7%)

FIG. 22

【図 2 3】

HBM transgenic mice: in vivo pDXA*
9-week-old mice vs WT for BMD% change

株	n	大腿骨	脊椎	全体
HBM:MCBA				
2	3	19	32	12
HBM:MTIC				
19	2	52	64	37
35	3	32	43	18

*CV: 大腿骨(2.7%); 脊椎(6.4%); 全体(1.7%)

FIG. 23

【図 2 4 A】

FIG. 24A

HBMGI_2AS

```
TCTAGACTCG AGGCGGCGCG CCATTGTGCA CTAAGCGCC GCCGCGCGC CATGGAGCCC
GAGTGAGCGC GCGCGGCGCG CGTCCGCGCG CCGACAACA TGGAGGCAGC GCCGCCCGGG
CCGCCGTGGC CGCTGCTGCT GCTGCTGCTG CTGCTGCTGG CGCTGTGCGG CTGCCCGGCC
CCCGCCGCGG CCTCGCGCT CCTGCTATT GCCAACCGCC GGGACGTACG GCTGTGGAC
GCCGCGGGAG TCAAGCTGGA GTCCACCATC GTGTTCAGCG GCCTGGAGGA TGCGCCGCA
GTGGACTTCC AGTTTTCCAA GGGAGCCGTG TACTGGACAG ACGTGAGCGA GGAAGCCATC
AAGCAGACCT ACCTGAACCA GACGGGGGCC GCCGTGCAGA ACGTGGTCAT CTCCGGCCTG
GTCTCTCCCG ACGGCTCGC CTGCGACTGG GTGGGCAAGA AGCTGTACTG GACGGACTCA
GAGACCAACC GCATCGAGGT GGGCAACCTC AATGGACAT CCGGAAGGT GCTCTTCTGG
CAGGACCTTG ACCAGCCGAG GGCCATCGCC TTGGACCCCG CTCACGGTA CATGTAAGTGG
ACAGACTGGG TTGAGAGCC CCGGATTGAG CGGGCAGGGA TGGATGGCAG CACCGGAAG
ATCAATTGTG ACTCGGACAT TTAAGTCCG AATGGACTGA CCATCGACCT GGAAGGACAG
AAGCTCTACT GGGCTGACGC CAAGCTCAGC TTACTCCACC GTGCCAAGCT GGAAGGCTCG
TTCCGGCAGA AGGTGGTGG GGGCAGCTCG ACGACCCCT TCGCCCTGAC GCTCTCCGGG
GACACTCTGT ACTGCAGACA CTGGCAGACC CGCTCCATCC ATGCCGTCAA CAAGCGCACT
GGGGGGAAGA GGAAGAGAT CCTGAGTCC CTCTACTCAC CCATGGACAT CCAAGTGCTG
AGCCAGGAGG GGCAGCCCTT CTTCACACT CGCTGTGAGG AGGACAATGG CGGCTGTCC
CACCTGTGCC TGTGTGCCCC AAGCGAGCT TTCTACACAT GCGCTGTGCC CACGGGTGTG
CAGCTGCAGG ACAAGCGCAG GACGTGTAA GAGGAGCGCG AGGAGGTGCT GCTGTGGCG
CGGCGGACGG ACCTACGGAG GATCTGCTG GACACGCCGG ACTTCACCGA CATGCTGTG
CAGGTGGACG ACATCCGGGA CGCCATTGCC ATCGACTAGC ACCCGCTAGA GGGGTATGTC
TACTGGACAG ATGACGAGGT GCGGGCAGC CGCAGGGCGT ACCTGGACGG GTCTGGGCGC
CAGACGCTGG TCAACACCGA GATCAACGAC CCCGATGGCA TCGCGTGA GCTGGTGGCG
CGAAACCTCT ACTGACCGA CACGGGCAGC GACCGCATCG AGGTGACCG CCTCAACGCG
```

【図 2 4 B】

FIG. 24B

HBMGI_2AS

```
ACCTCCCGCA AGATCCTGGT GTCGAGGAC CTGGACGAGC CCCGAGCCAT CGCACTGCAC
CCCGTGATGG GCCTCATGTA CTGGACAGAC TGGGAGAGA ACCCTAAAT CGAGTGTCG
AACTTGGATG GGCAGGAGCG GCGTGTGCTG GTCAATGCCT CCCTCGGGTG GCCCAACGGC
CTGGCCCTGG ACCTGCAGGA GGGGAAGCTC TACTGGGAG ACGCCAAGAC AGACAAGATC
GAGGTGATCA ATGTTGATGG GACGAAGAGG CGGACCTCTC TGGAGACAA GCTCCCGCAC
ATTTTCGGGT TCACGCTGCT GGGGGAATTC ATCTACTGGA CTGACTGGCA GCGCGCAGC
ATCGAGCGGG TGCACAAGGT CAAGGCCAGC CGGGACGTCA TCATTGACCA GCTGCCGAC
CTGATGGGGC TCAAGCTGT GAATGTGGCC AAGTCGTGG GAACCAACCC GTGTGCGGAC
AGGAACGGGG GGTGCAGCCA CCTGTGCTTC TTACACCCCG ACGCAACCCG GTGTGGCTGC
CCCATCGGCC TGGAGTGCT GAGTGACATG AAGACCTGCA TCGTGCCTGA GGCCTCTTG
GTCTTCACCA GCAGAGCCGC CATCCACAGG ATCTCCCTCG AGACCAATAA CAACGACGTG
GCCATCCGCG TCACGGGCGT CAAGGAGGCC TCAGCCCTGG ACTTTGATGT GTCCAACAAC
CACATCTACT GGACAGACGT CAGCCTGAAG ACCATCAGCC GCGCCTTCAT GAACGGGAGC
TCGGTGGAGC ACGTGGTGA GTTTGGCTTT GACTACCCCG AGGGCATGGC GGTGACTGG
ATGGGCAAGA ACCTCTACTG GCGCAGACT GGGACCAACA GAATCGAAGT GCGCGGGCTG
GACGGGCGAG TCCGGCAAGT CCTCGTGTGG AGGGAATGG ACAACCCGAG GTGCTGTGGC
CTGGATCCCA CCAAGGGCTA CATCTACTGG ACCGAGTGG GCGGCAAGCC GAGGATCGTG
CGGGCTTCA TGGACGGGCA CAACTGCATG ACGCTGTGG ACAAGTGG GCGGCCAATG
GACCTACCA TTGACTACG TGACCAAGCG CTCTACTGSA CCGACCTGGA CACCAACATG
ATCGAGTCGT CCAACATGCT GGGTCAGGAG CGGGTCGTA TTGCCGACGA TCTCCGCGAC
CCGTTCGGTC TGACGAGTA CAGGATTAT ATCTACTGSA CAGACTGGA TCTGCACAGC
ATTGAGCGGG CCGACAAGAG TAGCGCGCG AACCACCC TCATCCAGGG CCACCTGGAG
TTGCTGATGG ACATCCTGGT GTTCCACTCC TCCCGCAGG ATGGCTCAA TGACTGTATG
```

【図 2 4 C】

FIG. 24C

HBMGI_2AS

```
CACAACAACG GGCAGTGTGG GCAGTGTGCT CTTGCCATCC CCGCGGCCA CCGCTGCGCG
TGGCGCTCAC ACTACACCTT GGACCCAGC AGCCGCAACT GCAGCCCGCC CACCACTTTC
TTGCTTTCA GCCAGAAATC TGCCATCAGT CGGATGATCC CGGACGACCA GCACAGCCCG
GATCTCATCT TGGCCCTGCA TGGACTGAGG AACGTCAAAG CCATCGACTA TGACCACCTG
GACAAGTTCA TCTACTGGGT GGATGGGCGC CAGAACATCA AGCGAGCCAA GGAAGCGGG
ACCCAGCCCT TTGTTTTGAC CTCTGTGAGC CAAGGCCAAA ACCCAGACAG GCAGCCCAAC
GACCTCAGCA TGCACATCTA CAGCCGAGCA CTGTTCTGGA CGTGCGAGGC CACCAATACC
ATCAACGTCC ACAGGCTGAG CCGGGAGGCC ATGGGGGTGG TGCTGCGTGG GGAACCGCAC
AAGCCAGGG CCAATCTGCT CAACCGGAGG CGAGGGTACC TGATTTTAC CAACATGACG
GACCGGGCAG CCAAGATCGA ACGCGAGCC CTGGACGGCA CCGAGCGCGA GGTCTCTTC
ACCACCGGCC TCATCGGCCG TGTGGCCCTG GTGGTAGACA ACACACTGGG CAAGCTGTT
TGGTGGAGC CGGACCTGAA GCGATTGAG AGCTGTGACC TGTGAGGGC CAACCGCCTG
ACCTGGAGG AGCCCAACAT CGTGCAGCT CTGGGCTGA CCATCTTGG CAAGCATCTC
TACTGGATCG ACCCGACGCA GCAGATGATC GAGCGTGTGG AGAAGACCAC CGGGACAAAG
CGGACTCGCA TCCAGGGCGG TGTGCGCCAC CTCACTGGCA TCCATGCAGT GGAAGGAATC
AGCCTGGAGG AGTTCTCAGC CACCCCATGT GCGCGTGACA ATGGTGGCTG CTCCACATC
TGTATTGCCA AGGGTGATGG GACACCAGG TGCTCATGCC CAGTCCACCT CGTGTCTCTG
CAGAACCTGC TGACCTGTGG AGAGCCGCC ACCTGCTCC CGGACAGTT TGATGTGCC
ACAGGGGAGA TCGACTGTAT CCCCAGGGCC TGCGCTGTG ACGCTTTCC CGAGTGGCAT
GACCAGAGCG ACGAGGAGGG CTGCGCGTGG TGCTCGCGG CCAAGTTTCC CTGCGCGGG
GGTCAGTGTG TGGACTGCG CCTGCTGCT GACGGCGAGG CAGACTGTCA GGACCGCTCA
GACGAGGCGG ACTGTGACGC CATCTGCTG CCCAACCAAT TCCGGTGTG GAGCGGCCAG
TGTGTCTCA TCAACACGA GTGCGACTCC TTCCCGACT GTATCGAGCG CTCCGACGAG
CTCATGTGTG AAATACCAA GCGGCCCTCA GACGACAGCC GCGCCACAG CAGTGCCATC
```

【図 2 4 D】

FIG. 24D

HBMGI_2AS

```
GGGCGCGTCA TTGGCATCAT CCTCTCTCTC TTGTCATGG GTGGTGTCTA TTTTGTGTGC
CAGCGCGTGG TGTGCCAGCG CTATGCGGGG GCCAACGGGC CTTTCCCGCA CGAGTATGTC
AGCGGGACCC CGCAGCTGCC CCTCAATTTT ATAGCCCCGG GCGGTTCCCA GCATGGCCCC
TTCACAGGCA TCGCATCGGG AAGTCCATG ATGAGCTCCG TGAGCTGAT GGGGGGCCGG
GGCGGGGTGC CCCTCTACGA CCGGAACCAAC GTCACAGGGG CCTCGTCCAG CAGCTGCTCC
AGCACGAAGG CCACGCTGTA CCGCGCGATC CTGAACCCCG CGCCCTCCCC GGCACGGAC
CCCTCCCTGT ACAACATGGA CATGTTCTAC TCTTCAACA TTCCGGCCAC TGCGAGACCG
TACAGGCCCT ACATCATTCG AGGAATGGCG CCCCCGACGA CGCCCTGCA GACCGACGTG
TGTGACAGCG ACTACAGCGC CAGCGCTGG AAGGCCAGCA AGTACTACTT GGAATTGAAC
TCGGACTCAG ACCCCTATCC ACCCCACCCC ACGCCCCACA GCCAGTACTT GTCGGCGGAG
GACAGCTGCC CGCCCTCGCC CGCCACCGAG AGGAGCTACT TCCATCTCTT CCGCGCCCT
CGTCCCCCT GCACGGACTC ATCTGACCT CCGCGGGGCC ACTGTGCTT CTCTGTGCC
CTGTAAATAG TTTTAAATAT GAACAAGAA AAAAATATAT TTTATGATTT AAAAAATAA
TATAATTGGG ATTTTAAAA CATGAGAAAT GTGAAGTGTG ATGGGGTGG GAGGGCTGGG
AGAACTTTGT ACAGTGGAGA AATATTTATA AACTTAATTT TGTAAACAG AAAAAATAA
AAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
AAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA GCGGCCCG
```

【図 2 5 A】

FIG. 25A

ZMAXGI_3AS

TCTAGACTCG AGGCGGCGC CATTTGTGCA CTAAGCGCC GCCGCCGCG CATGGAGCCC
GAGTGAGCGC GGGCGGGGCC CGTCCGGCCG CCGGACAACA TGGAGGCAGC GCCGCCCGGG
CCGCCGTGGC CGCTGCTGCT GCTGCTGCTG CTGCTGCTGG CGCTGTGCGG CTGCCCGGCC
CCGCCCGCGG CCTCGCCGCT CCTGCTAITT GCCAACCGCC GGGACGTACG GCTGTGGTAC
GCCGGCGGAG TCAAGCTGGA GTCCACCATC GTGCTACGCG GCCTGGAGGA TGGGCCGCA
GTGGACTTCC AGTTTTCCAA GGGAGCCGTG TACTGGACAG ACGTGAAGCA GGAAGCCATC
AAGCAGACCT ACCTGAACCA GACGGGGGCC GCCGTGCAAG ACGTGGTAT CTCCGGCCTG
GTCTCTCCCG ACGGCCCTCG CTGCGACTGG GTGGGCAAGA AGCTGTACTG GACGGACTCA
GAGACCAACC GCATCGAGGT GGCCAACCTC AATGGCACAT CCCGGAAGGT GCTCTTCTGG
CAGGACCTTG ACCAGCCGAG GGCATCGCG TTGGACCCCG CTCACGGGTA CATGTACTGG
ACAGACTGGG GTGAGACGCC CGGATTAGG CGGGACGGGA TGGATGGCAG CACCCGGAAG
ATCATTGTGG ACTCGACAT TTAAGTGCCC AATGGACTGA CCATCGACCT GGAAGAGCAG
AAGCTCTACT GGGCTGACGC CAAGCTCAGC TTCACTCAC GTGCCAACCT GGACGGCTCG
TTCCGGCAGA AGGTGGTGGG GGGCAGCCTG ACGACCCCTC TCGCCCTGAC GCTCTCCGGG
GACACTCTGT ACTGGACAGA CTGGCAGACC CGCTCCATCC ATGCCGTCAA CAAGCGCAT
GGGGGGAAGA GGAAGAGGAT CTTGAGTGCC CTCTACTCAC CCATGGACAT CCAGTGTGCTG
AGCCAGGAGC GGCAGCTTTT CTTCACACT CGCTGTGAGG AGGACAATGG CGGCTGCTCC
CACCTGTGCG TGCTGTCCCC AAGCGAGCCT TTCTACACAT GCGCCTGCC CACGGGTGTG
CAGCTGCAGG ACAACGGCAG GACGTGTAAG GCAGGAGCCG AGGAGGTGCT GCTGCTGGCC
CGCGGGACGG ACCTACGGAG GATCTCGCTG GACACGCCGG ACTTCACCGA CATCGTGTG
CAGGTGGACG ACATCCGGCA GCCCAATTGC ATCGACTAGC ACCCGCTAGA GGGCTATGTC
TACTGGACAG ATGACGAGGT GCGGCCCATC CGCAGGGCGT ACCTGGACGG GTCTGGGGCG
CAGACGCTGG TCAACACCGA GATCAACGAC CCGCATGGCA TCGCGGTGCA CTGGCTGGCC
CGAAACCTCT ACTGGACCGA CACGGGACG GACCGCATCG AGGTGACGCG CCTCAACGCC

【図 2 5 B】

FIG. 25B

ZMAXGI_3AS

ACCTCCCGCA AGATCCTGGT GTCGGAGGAC CTGGACGAGC CCCAGCCAT CGCACTGCAC
CCCGTGATGG GCCTCATGTA CTGGACAGAC TGGGAGAGA ACCCTAAAT CGAGTGTGCC
AACTTGGATG GGCAGGAGCG GCGTGTGCTG GTCAATGCC CTCTCGGGTG GCCCAACGGC
CTGGCCCTGG ACCTGCAGGA GGGGAAGCTC TACTGGGGAG ACGCAAGAC AGACAAGATC
GAGGTGATCA ATGTTGATGG GACGAAGAGG CGGACCTCTC TGGAGGACAA GCTCCCGCAC
ATTTTCGGGT TCACGCTGCT GGGGGACTTC ATCTACTGGA CTGACTGGCA GCGCCGACG
ATCGAGCGGG TGCACAAGGT CAAGGCCAGC CGGGACGTCA TCATTGACCA GCTGCCCGAC
CTGATGGGGC TCAAGCTGT GAATGTGGCC AAGGTCTCGC GAACCAACCC GTGTGCGGAC
AGGAACGGGG GGTGCAGCCA CCTGTGCTTC TTCACACCCC ACGCAACCCG GTGTGGCTGC
CCCATCGGCC TGGAGCTGCT GAGTGACATG AAGACTGCA TCGTGCCTGA GGCCTTCTTG
GTCTTCACCA GCAGAGCCCG CATCCACAGG ATCTCCCTCG AGACCAATAA CAACGACGTG
GCCATCCCG TCACGGGCGT CAAGGAGGCC TCAGCCCTGG ACTTTGATGT GTCCAACAAC
CACATCTACT GGCACAGAGT CAGCCTGAAG ACCATCAGCC GCGCCTTCAAT GAACGGGAGC
TCGGTGGAGC ACGTGGTGA GTTTGGCCTT GACTACCCCG AGGCACTGGC CGTTGACTGG
ATGGGCAAGA ACCTCTACTG GGCCGACACT GGGACAACA GAATCGAAGT GGCAGCGCTG
GACGGGCAGT TCCGGCAAGT CCTCGTGG AGGGACTTGG ACAACCCGAG GTCGCTGGCC
CTGGATCCCA CCAAGGGCTA CATCTACTGG ACCGAGTGGG GCGGCAAGCC GAGGATCGTG
CGGGCCTTCA TGGACGGGAC CAACTGCATG ACGCTGGTGG ACAAGGTGGG CCGGGCCAAC
GACCTCACA TTGACTACGC TGACCAGCGC CTCTACTGGA CCGACCTGGA CACCAACATG
ATCGAGTCTG CCAACATGCT GGGTCAGGAG CGGGTCTGGA TTGCGGACGA TCTCCCGCAC
CCGTCGCTG TGACGAGTA CAGCGATTAT ATCTACTGGA CAGACTGGAA TCTGCACAGC
ATTGAGCGGG CCGACAAGAC TAGCGGCCGG AACCCGACCC TCATCCAGGG CCACCTGGAC
TTCTGTGATG ACATCCTGGT GTTCCACTCC TCCCGCCAGG ATGGCTCTAA TGACTGTATG

【図 2 5 C】

FIG. 25C

ZMAXGI_3AS

CACAACAACG GGCAGTGTGG GCAGCTGTGC CTTCCTATCC CCGCGGCCA CCGCTGGGG
TGCGCCTCAC ACTACACCTT GGACCCAGC AGCCGCAACT GCAGCCGCGC CACCACCTTC
TTGCTGTTCA GCCAGAAATC TGCATCAGT CGGATGATCC CGGACGACCA GCACAGCCCG
GATCTCATCC TGCCCTTGCA TGGACTGAGG AAGTCAAA GCATGACTA TGACCCACTG
GACAAGTTCA TCTACTGGGT GGAATGGGCG CAGAATCA AGCGAGCCAA GGACGACGGG
ACCCAGCCCT TTGTTTTGAC CTCTCTGAGC CAAGGCCAAA ACCCAGACAG GCAGCCCCAC
GACCTCAGCA TCGACATCTA CAGCCGGACA CTGTCTGGA CGTGCAGGGC CACCAATACC
ATCAAGCTCC ACAGGCTGAG CCGGGAAGCC ATGGGGGTGG TGCTGCGTGG GGACCGGAC
AAGCCAGGG CCATGCTGCT CAACGCGGAG CGAGGGTACC TGTACTTCA CAACATGCAAG
GACCGGGCAG CCAAGATCGA ACGCGAGCCG CTGGACGGCA CCGAGCGCGA GGTCTCTTTC
ACCACGGGCC TCATCCGCC CTGTGCCCTG GTGGTAGACA ACACACTGGG CAAGCTGTTC
TGGGTGGACG CGGACCTGAA GCGCATGAG AGCTGTGACC TGTCAAGGGC CAACCGCTG
ACCTGGAGG ACGCCAACAT CGTGACGCT CTGGGCTGGA CCATCCTTGG CAAGCATCTC
TACTGGATCG ACGCCAGCA GCAGATGATC GAGCGTGTGG AGAAGCACAC CCGGGAACAAG
CGGACTCGCA TCCAGGGCCG TGTCGCCCAT CTCACTGGCA TCCATGCAGT GGAGGAAGTC
AGCCTGGAGG AGTTCTCAGC CCACCCATGT GCCCGTGACA ATGGTGGCTG CTCCACATC
TGTATTGCCA AGGGTGTGAG GACACCAGG TGCTCATGCC CAGTCCACCT CGTGCTCTCG
CAGAACCTGC TGACCTGTGG AGAGCCGCC ACCTGCTCCC CGGACCAAGT TGCATGTGCC
ACAGGGGAGA TCGACTGTAT CCGCGGGGCC TGCGCTGTG ACGGCTTTCC CGAGTGGCAT
GACCAGAGCG ACGAGGAGGG CTGCCCGCTG TGCTCCGCCG CCAAGTCCCC TGCGGCGGG
GGTCAGTGTGT GGACCTGCGC CTGCGCTGG ACGCGAGGG AGACTGTGAG GACCGCTCA
GACGAGGGGA CTGTGACGCC ATCTGCTGC CCAACAGTT CCGGTGTGCG AGCGGCCAG
TGTGCTCTCAT CAAACAGCAG TGCGACTCT TCCCGACTG TATCGACGGC TCCGACGAG
CTCATGTGTGA AATCAACAA CCGCCCTCAG ACGACAGCCC GCGCCACAGC AGTGCCATC

【図 2 5 D】

FIG. 25D

ZMAXGI_3AS

GGGCCCGTCA TTGGCATCAT CCTCTCTCTC TTGCTCATGG GTGGTGCTTA TTTTGTGTGC
CAGCGCGTGG TGTGCCAGCG CTATGCGGGG GCCAACGGGC CCTTCCCGCA CGAGTATGTC
AGCGGGACCC CGCAGTGCC CCTCAATTC ATAGCCCCGG GCGGTGCCA GCATGGCCCC
TTCACAGGCA TCGCATGCGG AAGTCCATG ATGAGCTCCG TGAGCCTGAT GGGGGGCCGG
GGCGGGGTGC CCTCTACGA CCGGAACCA GTCACAGGG CCTCTCCAG CAGCTGTCTC
AGCAGCAAGG CCACGCTGTA CCGCCGATC CTGAACCCGC CGCCCTCCCC GGCACGGAC
CCCTCCCTGT ACAACATGGA CATGTTCTAC TCTTCAACA TTCCGGCCAC TGCAGAGCCG
TACAGGCCCT ACATCATTCG AGGAATGGCG CCCCCGACGA CGCCCTGCAG CACCGACGTG
TGTGACAGCG ACTACAGCGC CAGCCGCTGG AAGGCCAGCA AGTACTACCT GGATTTGAAC
TCGGACTCAG ACCCTATCC ACCCCACCC ACGCCCCACA GCCATACCT GTCGGCGGAG
GACAGTGC CCCTCTGCC CGCCACCGAG AGGAGCTACT TCCATCTCTT CCGCCCCCT
CCGTCCTCTC GCACGGACTC ATCTGACTC CCGCGGGGCC ACTCTGGCTT CTCTGTGCC
CTGTAATAGT TTTTAAATAT GAACAAGAA AAAAATATAT TTTATGATTT AAAAAATAA
TATAATTGGG ATTTTAAAAA CATGAGAAAT GTGAACCTGT ATGGGGTGG CAGGGCTGG
AGAACTTTGT ACAGTGAGGA AATAATTATA AACTTAATTT TGTAACACAG AAAAAAAAA
AAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
AAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA

【図 2 6 A】

FIG. 26A

ヒトおよびマウスZmax1(LRP5)アミノ酸配列のアラインメント
上=ヒト、下=マウス

```

1 MEA...APPGPPWLLLLLLLLLALCGCPAPAAASPLLLFANRRDRLV 46
1 METAPTRAPPPPPPLLLLL...LYCSLVFAAASPLLLFANRRDRLV 45
47 DAGGVKLESTIVVSGLEDAAAVDFQFSKGAVYMTDVSEEAIKQTYLNQTG 96
46 DAGGVKLESTIVASGLEDAAAVDFQFSKGAVYMTDVSEEAIKQTYLNQTG 95
97 AAVQNVVISGLVSPDGLACDWGKKLYWTDSETNRIEVANLNGTSRKVL 146
96 AAAQNVISGLVSPDGLACDWGKKLYWTDSETNRIEVANLNGTSRKVL 145
147 WQDLQDPAIALDPAGHYMYTDMGETPRIERAGMDGSTRKIIVDSDIYW 196
146 WQDLQDPAIALDPAGHYMYTDMGEAPRIERAGMDGSTRKIIVDSDIYW 195
197 PNGLTIDLEEQLYADAKLSFIHRANLDGSRQKVVVEGSLTHPFALTLS 246
196 PNGLTIDLEEQLYADAKLSFIHRANLDGSRQKVVVEGSLTHPFALTLS 245
247 GDTLYWTDWQTRSIHACNKRGTGGKKEILSALYSPMDIQVLSQERQPFH 296
246 GDTLYWTDWQTRSIHACNKRGTGGKKEILSALYSPMDIQVLSQERQPFH 295
297 TRCEDNNGGCSHLCLLSPSEPFYTACPTGVQLQDNGRTCKAGAEVLL 346
296 TPCEEDNNGGCSHLCLLSPREPFYTACPTGVQLQDNGRTCKAGAEVLL 345
347 ARRTDLRRISLDTPDFTDIVLQVDDIRHAIADYDPLEGYVYMTDDEVRA 396
346 ARRTDLRRISLDTPDFTDIVLQVDDIRHAIADYDPLEGYVYMTDDEVRA 395
397 IRRAYLDGSGAQTIVNTEINDPDGIAVDWVARNLYMTDTGTDRIEVRNL 446
396 IRRAYLDGSGAQTIVNTEINDPDGIAVDWVARNLYMTDTGTDRIEVRNL 445
447 GTSRKILVSEDLDEPRAIALHPVGLMYMTDNGENPKIECANLDQERRV 496
446 GTSRKILVSEDLDEPRAIALHPVGLMYMTDNGENPKIECANLDQERRV 495
497 LVNASLGWPNGLALDLQEGKLYWGDAKTDKIEVINVDGTRKRTLEDKLP 546
496 LVNTSLGWPNGLALDLQEGKLYWGDAKTDKIEVINVDGTRKRTLEDKLP 545
547 HIFGFTLLGDFTYMTDQRRSIERVHKVKSADVIIDQLPDLMLKAVNV 595
546 HIFGFTLLGDFTYMTDQRRSIERVHKVKSADVIIDQLPDLMLKAVNV 595
597 AKVVGTPNCPADRNGGCSHLCFETPHATRCGCPGLELLSDMKTCIVPEAF 646
596 AKVVGTPNCPADRNGGCSHLCFETPHATRCGCPGLELLSDMKTCIVPEAF 645
647 LVFTSRAAIHRISLETNNNOVAIPLTGKVEASALDFDVSNNHIYMTDVS 696
646 LVFTSRATHIRISLETNNNOVAIPLTGKVEASALDFDVSNNHIYMTDVS 695

```

【図 2 6 C】

FIG. 26C

ヒトおよびマウスZmax1(LRP5)アミノ酸配列のアラインメント

```

1447 PFTGIACGKSMSSSVSLMGGRGVPLYDRNHVTGASSSSSSSKATLYPP 1496
1446 PFPGIPCSKSMSSSVSLMGGRGVPLYDRNHVTGASSSSSSSKATLYPP 1495
1497 ILNPPSPATDPSLYNMDMFYSSNIPATARPYPYIRGMAPPTPCSTO 1546
1496 ILNPPSPATDPSLYNVDFYSSGIPATARPYPYIRGMAPPTPCSTO 1545
1547 VCDSDYSASRWKASKYYLDLNSDDPYPPPTPHSQYLSAEDSCPPSPAT 1596
1546 VCDSDYSISRWKSKYYLDLNSDDPYPPPTPHSQYLSAEDSCPPSPAT 1595
1597 ERSYFHLFPFPPSPCTDSS 1615
1596 ERSYCHLFPFPPSPCTDSS 1614

```

【図 2 6 B】

FIG. 26B

ヒトおよびマウスZmax1(LRP5)アミノ酸配列のアラインメント

```

697 KTISRFAFMNGSSVEHVFGLDYPEGMAVDWMGNLYWADTGTNRIEVAR 746
696 KTISRFAFMNGSSVEHVFGLDYPEGMAVDWMGNLYWADTGTNRIEVAR 745
747 LDGQFRQVLVWRDLNPRSLALDPTKGYIYMTWGGKPRIVRAFMGTNC 796
746 LDGQFRQVLVWRDLNPRSLALDPTKGYIYMTWGGKPRIVRAFMGTNC 795
797 MTLVDKVGRANDLTIDYADQRLYMTDLOTNMISSNMLGQERVVIADDL 846
796 MTLVDKVGRANDLTIDYADQRLYMTDLOTNMISSNMLGQERVVIADDL 845
847 HPFGLTQYSOIYMTDWNLSIERADKTSGRNRTLQGHDFVMDILVFH 896
846 YPFGLTQYSOIYMTDWNLSIERADKTSGRNRTLQGHDFVMDILVFH 895
897 SSRQDGLNDCHMNGQCGQLCLAIPEGGRGCSASHYTLDPSSRNCSPPT 946
896 SSRQDGLNDCHMNGQCGQLCLAIPEGGRGCSASHYTLDPSSRNCSPPT 945
947 FLLFSQKSAISRMIPDDQHSDDLPLHLGLRNKVAIDYDPLDKFIYWDG 996
946 FLLFSQKFAISRMIPDDQLSPDLPLHLGLRNKVAIDYDPLDKFIYWDG 995
997 RQNIKRAKDDGTQPFVLTSLSQGQNPDRQPHDLSDIYSRTLFWTCEATN 1046
996 RQNIKRAKDDGTQPSHLTSPSQLSFDRQPHDLSDIYSRTLFWTCEATN 1045
1047 TINVHRLSGEAMGVVLRGDRDKPRAIVVNAERGVLFTNMQDRAAKIERA 1096
1046 TINVHRLDGDAMGVVLRGDRDKPRAIVVNAERGVLFTNMQDRAAKIERA 1095
1097 ALDGTREVLFTTGLIRPVAVVDNTLGLKFWVDADLKRIESCDLSGANR 1146
1096 SLDGTREVLFTTGLIRPVAVVDNALGLKFWVDADLKRIESCDLSGANR 1145
1147 LTLEDANIVQPLGLTILGKHLIYWDROQQHIERVEKTTGDKRTRVQGRV 1196
1146 LTLEDANIVQVGLTVLGRHLYWIDROQQHIERVEKTTGDKRTRVQGRV 1195
1197 HLTGIHAVEEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSCPVLVL 1246
1196 HLTGIHAVEEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSCPVLVL 1245
1247 LQNLTLGCEPPTCSPDQFACATGEIDCIPGAWRCDFPECAQDSDEEGCP 1296
1246 LQNLTLGCEPPTCSPDQFACTGEIDCIPGAWRCDFPECAQDSDEEGCP 1295
1297 VCSAAQFPFCARGQVDRLRCDGEADCDQRSDEADCAICLPNQFRCA5G 1346
1296 VCSA5QFPFCARGQVDRLRCDGEADCDQRSDEADCAICLPNQFRCTSG 1345
1347 QCVLIKQCCDSFPDCIDGSDLMCEITKPPSDPSHA5SAGPVIGIILS 1396
1346 QCVLIKQCCDSFPDCADGSDLMCEINKPPSDDIPHA5SAGPVIGIILS 1395
1397 LFVMGVYFVQCRVVCQRYAGANGPFPHYVSGTPHVPNLFIAPGGSQH 1446
1396 LFVMGVYFVQCRVVCQRYTGASGPFPHYVGGAPHVPLNFIAPGGSQH 1445

```

【図 2 7 A】

FIG. 27A

ヒトLRP5およびLRP6アミノ酸配列のアラインメント

上=LRP5、下=LRP6
下線=抗体産生のために標識された領域
※類似性:83.542 %同一性:70.901

```

1 MEAAPGPPWLLLLLLLLLALCGCPAPAAASPLLLFANRRDRLVLDAGG 50
1 .....MGAVLRSLLA.CSFVLLRAAPLLLYANRRDLRLVDATN 38
51 VKLESTIVVSGLEDAAAVDFQFSKGAVYMTDVSEEAIKQTYLNQTGA 106
39 GKENATIVVSGLEDAAAVDFVPSHGLIYWSOVSEEAIRTEFNKTE.SVQ 87
101 NVVISGLVSPDGLACDWGKKLYWTDSETNRIEVANLNGTSRKVLWQDL 156
88 NVVISGLVSPDGLACDWGKKLYWTDSETNRIEVSNLDGSLRKVLWQEL 137
151 DQPAIALDPAGHYMYTDMGETPRIERAGMDGSTRKIIVDSDIYWPNGL 200
138 DQPAIALDPSSGFMWTDWGEVPKIERAGMDGSSRFIINSEIYWPNGL 187
201 TIDLEEQLYADAKLSFIHRANLDGSRQKVVVEGSLTHPFALTLSGDTL 250
188 TIDLEEQLYADAKLNFHNSNLDGTRQAVVKGSLPHPFALTLEFIDL 237
251 YWTDWQTRSIHACNKRGTGGKKEILSALYSPMDIQVLSQERQPFHTRCE 300
238 YWTDWQTRSIHACNKRGTGGKKEILSALYSPMDIQVLSQERQPFHTRCE 287
301 EDNNGGCSHLCLLSPSEPFYTACPTGVQLQDNGRTCKAGAEVLLARRT 350
288 EDNNGGCSHLCLLSPSEPFYTACPTGVQLQDNGRTCKAGAEVLLARRT 337
351 DLRRISLDTPDFTDIVLQVDDIRHAIADYDPLEGYVYMTDDEVRAIRRA 400
338 DLRRISLDTPDFTDIVLQVDDIRHAIADYDVEGYVYMTDDEVRAIRRS 387
401 YLDGSGAQTIVNTEINDPDGIAVDWVARNLYMTDTGTDRIEVRNLNGTS 450
388 YLDGSGAQTIVNTEINDPDGIAVDWVARNLYMTDTGTDRIEVRNLNGTS 437
451 KILVSEDLDEPRAIALHPVGLMYMTDNGENPKIECANLDQERRVLNVA 500
438 KILVSEDLDEPRAIALHPVGLMYMTDNGENPKIECANLDQERRVLNVA 487
501 SLGWPNGLALDLQEGKLYWGDAKTDKIEVINVDGTRKRTLEDKLPHFIF 550
488 SLGWPNGLALDLQEGKLYWGDAKTDKIEVINVDGTRKRTLEDKLPHFIF 537
551 FTLLGDFTYMTDQRRSIERVHKVKSADVIIDQLPDLMLKAVNAKV 600
538 FTLLGDFTYMTDQRRSIERVHKVKSADVIIDQLPDLMLKATNVHVRVI 587
601 GTNPACADRNGGCSHLCFETPHATRCGCPGLELLSDMKTCIVPEAFVFT 650
588 GTNPACADRNGGCSHLCLYRPGQLRCACPIGLELLSDMKTCIVPEAFLLFS 637

```

【図 27B】

FIG. 27B

ヒトLRP5およびLRP6アミノ酸配列のアラインメント

```

651 SRAATHRISLETNNNDVAIPLTGVKEASALDFDVSNHHIYWTQVSLKTI 700
638 RRADIRRIISLETNNNDVAIPLTGVKEASALDFDVTONRIYWTQVSLKTI 687
701 RAFMNGSSVHVVEFGLDYPEGHVDWNGKLYWADTGTNRIEVARLDGQ 750
688 RAFMNGSALEHVVEFGLDYPEGHVDWNGKLYWADTGTNRIEVSKLDGQ 737
751 FRQVLVWRDLNPRSLALDPTKGYIYWTWGGKPRIVRAFMGDTNCHTLV 800
738 HRQVLVWKDLSPRALALDPAEGFHYWTWGGKPKIDRAAMGSERTTLV 787
801 DKVGRANDLTIYADQRLYWTDLDTNMISSNMLGQERVVIADDLPHPF 850
788 PNVGRANGLTIYAKRRLYWTDLDTNMISSNMLGLNREVIADDLPHPF 837
851 LTQYSYIYWTDMNLSHIERADKTSGRNRTLIQGHLDVYHDLVHSSRQ 900
838 LTQYQDIYWTDSRRSIERANKTSGRNRTLIQGHLDVYHDLVHSSRQ 887
901 DGLNDCMHNNGQCQLCLAIPEGGRGCGASHYTIQPSRNCSPPTTFL 949
888 SGWNECASSNGCHSLCLAVPVGGFVCGCPAHVSLNADNRTCSAPTFL 937
950 FSQKSAISRMIPODQSPDLPLHLGNVKAIDYDPLQKFIYWDGRQN 999
938 FSQKSAINRMVIDEQSPDILPLHSLRNVAIDYDPLQKLYWIDSRQN 987
1000 .IKRAKDDGTOPE.VLTSLSOGQNPDRQPHOLSIDIYSRTLFWTCEATN 1047
988 MIRKAQEDGSGFTVVVSVSPQNLEIQPYDLSIDIYSRYIYWTCEATN 1037
1048 INVHRLSGEAMGVVLRGRDQKPRALVVAERGLYFTNHQDRAAKIERAA 1097
1038 INVTRLGGRSVGVVLKGEQDRPRAIVVNPEKGYMYFTNLQERSPKIERAA 1087
1098 LDGTEREVLFTTGLIRPVALVVDNTLGKLFVVDADLKRIESCDSLGNRL 1147
1088 LDGTEREVLFFSGLSKPIALALDSRLGKLFVWDSLRRIESSDSLGNRI 1137
1148 TLEDANIVQPLGLTILGKHLIWDROQQMIERVEKTTGDKRTRIQGRVAH 1197
1138 VLEDNLIQPVGLTVFENWLIWDKQQMIKIDHTGREGRKQVQARIAQ 1187
1198 LTGIHAVEEVSLEEFSAHPCARDNGGCSHICIAKGDTGPRCSQVHLVL 1247
1188 LSDIHAVKELNLQERYQHPCAQDNGGCSHICLVKGDGTTTRCSCPMHLVL 1237
1248 QNLLTCGEPTTCSPOFACATGETICTPQAMRCOGFPECDDQSDPEEGCPV 1297
1238 QDELSCGEPTTCSPOFTCTGETICTPQAMRCOGFTECEDHSDENLCPV 1287
1298 CSAAOFPFCARGCVDLRLRCDGEADCQDRSDEADCAICLPNQFRCASGQ 1347
1288 CSSESQFCASGQCIDGALRCNGDANQDKSDEKNEVLCLIDQFRCANQQ 1337
1348 CVLTKQOCDSPFPCIDGSDDELMEITKPPSDSPAHSAGIPVIGIILSL 1397
1338 CIGKHKKCDHNVDCSKSDLEDCYTEEP...APQATNTVGSVIGIVITI 1384

```

【図 27C】

FIG. 27C

ヒトLRP5およびLRP6アミノ酸配列のアラインメント

```

1398 FVMGGVYFVQVRVVCQRYAGANGPFPHEVVS.GTPHVLNFIAPGGSQHG 1446
1385 FVSGTVYFICQRMCLPRMKGDGETMTNDVYVHGASVPLGYVPHPSLSG 1434
1447 PFTGIACGKSMSSVSLMGRRGVPLYDRNHVTGASSSSSSSKATLYPP 1496
1435 SLPGHSRGKSHISSLSIMGSSGPP.YDRAHVTGASSSSSSSKGTYPFA 1483
1497 ILNPPPSPATDPSLYNMDMEYSSNIPATAR..PYRPHYIRGHAPPTTCS 1544
1484 ILNPPPSPATERSHYTMEFGYSSNSPSTHRSYSYRPHFAPPTTCS 1533
1545 TDVCDSDYASRWKAS...KYLDLNSDSDPYPPPTPHSQYLSAE. 1587
1534 TDVCDSDYAPSRMTSVATAKGYTSDLYDSEVPVPPPTPRSQYLSAEEN 1583
1588 .DSCPPSPATERSY.FHLFPPPPSPCTDSS 1615
1584 YESCPPSPYTERSYSHLYPPPPSPCTDSS 1613

```

【図 28】

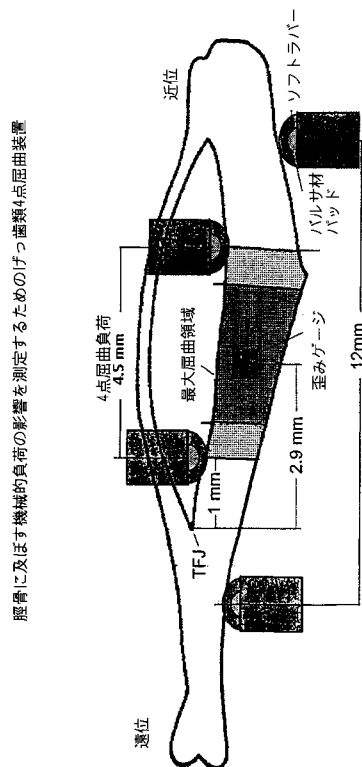


FIG. 28

【図 29】

HBM1における骨形成・骨吸収に対する応答

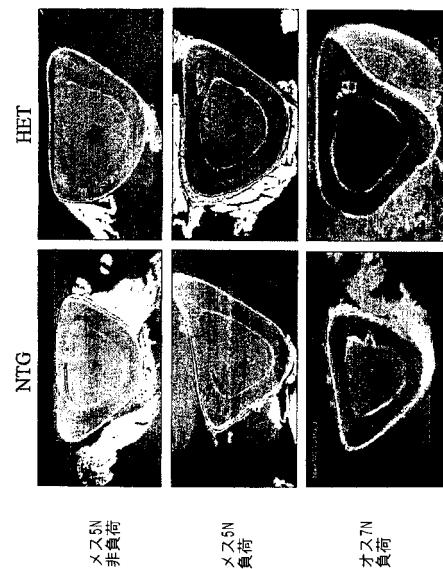


FIG. 29

【手続補正書】

【提出日】平成16年1月14日(2004.1.14)

【手続補正1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

配列番号2の位置582におけるチミンを少なくとも含み、配列番号2の少なくとも15個の連続するヌクレオチドを含む配列に機能的に連結され、動物細胞および／またはヒト細胞におけるタンパク質の発現を導くことができるプロモーター領域を含む核酸を含む体細胞および／または生殖細胞を有するトランスジェニック動物。

【請求項2】

配列番号4をコードする配列を含み、配列番号4の位置171のバリンに相当するバリンのコードンを少なくとも含む核酸であって、動物細胞および／またはヒト細胞におけるタンパク質の発現を導くことができる、機能的に連結されたプロモーター領域をさらに含む核酸を含む体細胞および／または生殖細胞を有するトランスジェニック動物。

【請求項3】

核酸が配列番号2を含む請求項1または2に記載のトランスジェニック動物。

【請求項4】

配列番号1の少なくとも15個の連続するヌクレオチドを含む配列に機能的に連結された、動物細胞および／またはヒト細胞におけるタンパク質の発現を導くプロモーター領域を含む核酸を含む体細胞および／または生殖細胞を有する、骨密度変調の研究のためのトランスジェニック動物であって、骨量が、骨密度、骨力、骨梁数、骨の大きさおよび骨組織結合性の中から選択される1つ以上のパラメーターについて、同一種の非トランスジェニック動物と比較して、変調されているトランスジェニック動物。

【請求項5】

プロモーター領域が、CMV、RSV、SV40およびEF-1 α 、CMV β アクチン、ヒストン、I型コラーゲン、TGF β 1、SX2、c fos/c jun、Cbfa1、Fra/Jun、Dlx5、オステオカルシン、オステオポンチン、骨シアロタンパク質、またはコラゲナーゼプロモーター領域である請求項1から4のいずれかに記載のトランスジェニック動物。

【請求項6】

プロモーター領域が、骨特異的プロモーター領域である請求項1から4のいずれかに記載のトランスジェニック動物。

【請求項7】

プロモーター領域が、CMV β アクチンプロモーター領域である請求項1から4のいずれかに記載のトランスジェニック動物。

【請求項8】

プロモーター領域が、I型コラーゲンプロモーター領域である請求項1から4のいずれかに記載のトランスジェニック動物。

【請求項9】

請求項1から8のいずれかに記載のトランスジェニック動物の胚。

【請求項10】

トランスジェニック動物が、ヒトHBMタンパク質を発現する請求項1から3のいずれかに記載のトランスジェニック動物。

【請求項11】

ヒトHBMタンパク質が、骨組織において最も多く発現される請求項10に記載のトランスジェニック動物。

【請求項 1 2】

H B M 表現型を表す請求項 1 から 3 のいずれかに記載のトランスジェニック動物。

【請求項 1 3】

骨量が、骨密度、骨力、骨梁数、骨の大きさおよび骨組織結合性の中から選択される 1 つ以上のパラメーターについて、同一種の非トランスジェニック動物と比較して、変調されている請求項 1 から 3 のいずれかに記載のトランスジェニック動物。

【請求項 1 4】

ヒト H B M タンパク質が発現され、繁殖力があり、ヒト H B M 遺伝子をその子孫に伝える請求項 1 から 3 のいずれかに記載のトランスジェニック動物。

【請求項 1 5】

ヒト L R P 5 タンパク質が発現され、繁殖力があり、ヒト H B M 遺伝子をその子孫に伝える請求項 4 に記載のトランスジェニック動物。

【請求項 1 6】

請求項 1 から 1 5 のいずれかに記載のトランスジェニック動物またはその子孫から産生されたトランスジェニック動物。

【請求項 1 7】

L R P 5 をコードする遺伝子に変化を加えたゲノムを有するトランスジェニックマウスであって、該変化が、相同組換えによるジーンターゲティングのための核酸であって、マウス L R P 5 遺伝子に相同である第 1 セクションと、マウス L R P 5 遺伝子の別のセクションに相同である第 2 セクションと、第 1 および第 2 セクション間における L R P 5 遺伝子の一部の操作された欠失、核酸配列の変化または核酸の挿入を含む中央セクションとを含む核酸の、胚性幹細胞または多能性細胞への導入により引き起こされ、該核酸が内因性遺伝子と相同組換えできるトランスジェニックマウス。

【請求項 1 8】

ジーンターゲティングのための核酸の中央セクションが、A T G 開始コドンの操作された欠失、操作されたフレームシフト変異、操作された停止コドン、ネオマイシン抵抗性配列、l o x P 組換え部位、または合成転写中止配列を含む請求項 1 7 に記載のトランスジェニックマウス。

【請求項 1 9】

ジーンターゲティングのための核酸が、マウス L R P 5 遺伝子のイントロン配列およびエクソン配列の双方をさらに含む請求項 1 7 に記載のトランスジェニックマウス。

【請求項 2 0】

ジーンターゲティングのための核酸が、マウス L R P 5 遺伝子のアミノ酸配列の位置 1 7 0 においてグリシンからバリンへの変化をコードするコドンを含み、変化した遺伝子が H B M タンパク質をコードする請求項 1 7 に記載のトランスジェニックマウス。

【請求項 2 1】

前記変化が、L R P 5 遺伝子が発現しないようにする L R P 5 遺伝子の破壊である請求項 1 7 に記載のトランスジェニックマウス。

【請求項 2 2】

L R P 6 をコードする遺伝子に変化を加えたゲノムを有するトランスジェニックマウスであって、該変化が、相同組換えによるジーンターゲティングのための核酸であって、マウス L R P 6 遺伝子に相同である第 1 セクションと、マウス L R P 6 遺伝子の別のセクションに相同である第 2 セクションと、第 1 および第 2 セクション間における L R P 6 遺伝子の一部の操作された欠失、核酸配列の変化または核酸の挿入を含む中央セクションとを含む核酸の、胚性幹細胞または多能性細胞への導入により引き起こされ、該核酸が内因性遺伝子と相同組換えでき、該トランスジェニック動物が変調された W n t 活性、D k k 活性、脂質レベルまたは骨量を有するトランスジェニックマウス。

【請求項 2 3】

前記変化が、L R P 6 遺伝子が発現しないようにする L R P 6 遺伝子の破壊である請求項 2 2 に記載のトランスジェニックマウス。

【請求項 2 4】

繁殖力があり、前記変化した遺伝子をその子孫に伝える請求項 2 2 に記載のトランスジェニックマウスまたはその子孫。

【請求項 2 5】

繁殖力があり、前記変化した遺伝子をその子孫に伝える請求項 1 7 から 2 4 のいずれかに記載のトランスジェニックマウス。

【請求項 2 6】

繁殖力があり、前記変化した遺伝子をその子孫に伝えるトランスジェニックマウスであって、子孫が、野生型と比較して、骨密度、骨力、骨梁数、骨の大きさおよび骨組織結合性の中から選択される少なくとも 3 つのパラメーターにより示される変調された骨量の表現型を表す請求項 1 7 から 2 4 のいずれかに記載のトランスジェニックマウス。

【請求項 2 7】

マウス胚性幹細胞をマウス胚盤胞へ導入することで産生される請求項 1 7 から 2 6 のいずれかに記載のトランスジェニックマウス。

【請求項 2 8】

請求項 1 7 から 2 7 のいずれかに記載のトランスジェニックマウスから産生されたトランスジェニックマウス。

【請求項 2 9】

請求項 1 から 2 8 のいずれかに記載のトランスジェニック動物からなる動物の第 1 群および対照動物の第 2 群を含む骨密度変調の研究のための動物モデル。

【請求項 3 0】

対照動物群が、ヒト L R P 5 をコードする核酸を含む細胞を有するトランスジェニック動物を含む請求項 2 9 に記載の動物モデル。

【請求項 3 1】

骨量決定要素の研究方法であって、

(a) 請求項 1 から 2 8 のいずれかに記載のトランスジェニック動物の第 1 群を供するステップと、

(b) 該トランスジェニック動物における少なくとも 1 つの骨発達パラメーターを測定するステップと、
を含む方法。

【請求項 3 2】

骨量のモデレーターの研究方法であって、

(a) 請求項 1 から 2 8 のいずれかに記載のトランスジェニック動物の第 1 群を供するステップと、

(b) 被験化合物を投与するステップと、

(c) 被験化合物を投与されたトランスジェニック動物における少なくとも 1 つの発達パラメーターを測定するステップと、
を含む方法。

【請求項 3 3】

骨量の研究方法であって、

(a) 請求項 1 から 2 8 のいずれかに記載のトランスジェニック動物の第 1 群を供するステップと、

(b) 実験手順を施すステップと、

(c) 実験手順を施された動物における少なくとも 1 つの発達パラメーターを測定するステップと、
を含む方法。

【請求項 3 4】

被験化合物が、注射投与、経口投与、坐薬投与、移植片投与、または局所的投与される請求項 3 2 に記載の方法。

【請求項 3 5】

実験手順が、卵巣切除、骨負荷の制限、および骨負荷の増強の中から選択される請求項 33 に記載の方法。

【請求項 36】

トランスジェニック動物群が HBM を発現し、HBM を発現し、被験化合物を投与されたトランスジェニック動物において骨量密度が変調されている請求項 32 に記載の方法。

【請求項 37】

トランスジェニックマウス群が HBM を発現し、HBM を発現し、被験化合物を投与されたトランスジェニックマウスにおいて骨量密度が上昇している請求項 36 に記載の方法。

【請求項 38】

被験化合物が、ホルモン、成長因子、ペプチド、RNA、DNA、ミネラル、ビタミン、天然生成物、または合成有機化合物を含む請求項 32 に記載の方法。

【請求項 39】

実験手順が、外科的手順、遺伝子治療手順、薬物治療手順、食事療法、または運動を含む請求項 33 に記載の方法。

【請求項 40】

HBM が骨障害に及ぼす影響を研究する方法であって、

(a) 骨障害の表現型を有する動物の胚を供するステップと、

(b) 請求項 1 から 4 のいずれかに記載の核酸を該胚の第 1 の群に導入するステップと、

(c) 胚を偽妊娠マウスに移すステップと、

(d) 得られたマウスの少なくとも 1 つの発達パラメーターを測定するステップと、を含む方法。

【請求項 41】

骨形成／再吸収の代理マーカーを同定する方法であって、

(a) 請求項 29 に記載の骨発達の動物モデルを供するステップと、

(b) 該動物における代理マーカー候補を定量的に測定するステップと、

(c) トランスジェニック動物群の測定値を対照動物群の測定値と比較するステップと、

を含む方法。

【請求項 42】

HBM が心臓障害に及ぼす影響を研究する方法であって、

(a) 請求項 1 から 28 のいずれかに記載のトランスジェニック動物の第 1 群を供するステップと、

(b) 被験化合物を投与されたトランスジェニック動物の少なくとも 1 つの心臓健康パラメーターを測定するステップと、を含む方法。

【請求項 43】

HBM が心臓障害に及ぼす影響を研究する方法であって、

(a) 請求項 1 から 28 のいずれかに記載のトランスジェニック動物の第 1 群を供するステップと、

(b) 被験化合物を投与するステップと、

(c) 被験化合物を投与されたトランスジェニック動物の少なくとも 1 つの心臓健康パラメーターを測定するステップと、

を含む方法であって、被験化合物が、ホルモン、成長因子、ペプチド、RNA、DNA、ミネラル、ビタミン、天然生成物、または合成有機化合物を含む方法。

【請求項 44】

心臓健康パラメーターが、血清脂質濃度である請求項 43 に記載の方法。

【請求項 45】

心臓保護処置から骨量変調効果をスクリーニングする方法であって、

請求項 3 1 に記載の方法を含み、心臓保護処置を前記動物の第 1 群のサブグループに投与するステップをさらに含む方法。

【請求項 4 6】

請求項 1 から 2 8 のいずれかに記載のトランスジェニック動物またはその子孫に由来する単離された細胞。

【請求項 4 7】

骨芽細胞または破骨細胞である請求項 4 6 に記載の単離された細胞。

【請求項 4 8】

内因性 L R P 5 の発現が、相同組換えまたは非相同組換えにより導入された変化した遺伝子の調節配列により変調されるトランスジェニック動物。

【請求項 4 9】

心臓保護処置から骨量変調効果をスクリーニングする方法であって、

請求項 1 から 2 8 のいずれかに記載のトランスジェニック動物の第 1 群を供するステップと、

心臓保護処置を該動物の第 1 群のサブグループに投与するステップと、

少なくとも処置されたマウスにおいて少なくとも 1 つの骨変調パラメーターを測定するステップと、

を含む方法。

【請求項 5 0】

前記核酸が、配列番号 4 の位置 1 3 3 0 で少なくともアラニンからバリンへの置換をさらにコードする請求項 1 から 1 6 のいずれかに記載のトランスジェニック動物。

【請求項 5 1】

相同組換えによるジーンターゲティングのための核酸であって、マウス L R P 5 遺伝子に相同である第 1 セクションと、マウス L R P 5 遺伝子の別のセクションに相同である第 2 セクションと、第 1 および第 2 セクション間における L R P 5 遺伝子の一部の操作された欠失、核酸配列の変化または核酸の挿入を含む中央セクションとを含み、内因性遺伝子と相同組換えできる核酸。

【請求項 5 2】

中央セクションが、A T G 開始コドンの操作された欠失、操作されたフレームシフト変異、操作された停止コドン、ネオマイシン抵抗性配列、l o x P 組換え部位、または合成転写中止配列を含む請求項 5 1 に記載の核酸。

【請求項 5 3】

マウス L R P 5 遺伝子のイントロン配列およびエクソン配列の双方をさらに含む請求項 5 1 に記載の核酸。

【請求項 5 4】

マウス L R P 5 遺伝子のアミノ酸配列の位置 1 7 0 でグリシンからバリンへの変化をコードするコドンを含み請求項 5 1 に記載の単離された核酸。

【請求項 5 5】

L R P 5 をコードする遺伝子に変化を加えたゲノムを有するトランスジェニックマウスを産生する方法であって、

(a) 請求項 5 1 に記載の核酸を供するステップと、

(b) 該核酸をマウス胚性幹細胞に導入するステップと、

(c) 該核酸を含む胚性幹細胞を選択するステップと、

(d) ステップ (c) の胚性幹細胞をマウス胚盤胞に導入するステップと、

(e) ステップ (d) の胚盤胞を偽妊娠マウスに移すステップと、

(f) 移した胚盤胞を、該核酸に関してキメラであるマウスに発達させるステップと、を含む方法。

【請求項 5 6】

胚性幹細胞の導入が、マイクロインジェクションによるものである請求項 5 5 に記載の方法。

【請求項 57】

(a) キメラマウスと野生型マウスと交配させ、前記変化に関してヘテロ接合性であるマウスを得るステップと、

(b) 該ヘテロ接合性マウスを交配させ、前記変化に関してホモ接合性であるマウスを産生するステップと、

をさらに含む請求項 56 に記載の方法。

【請求項 58】

骨量に関連する遺伝子を同定する方法であって、

(a) 請求項 29 に記載の骨発達の動物モデルを供するステップと、

(b) 該動物における遺伝子発現のプロファイルを測定するステップと、

(c) 該動物の第 1 の群の測定値を動物の対照群の測定値と比較するステップと、

を含む方法。

【手続補正書】

【提出日】平成16年6月18日(2004.6.18)

【手続補正 1】

【補正対象書類名】明細書

【補正対象項目名】0037

【補正方法】変更

【補正の内容】

【0037】

本発明の別の態様によると、単離された核酸配列（配列番号 736～744）：

(1)

5'-GTCAGCCTGGAGGAGTTCTCA-3'；

5'-TCACCCTTGGCAATACAGATGT-3'；および

6-FAM-5'-CCCACCCATGTGCCCCGTGACA-3'；または

(2)

5'-CGTGATTGCCGACGATCTC-3'；

5'-TTCCGGCCGCTAGTCTTGT-3'；および

6-FAM-5'-CGCACCCGTTCTGGTCTGACGCAGTAC-3'

を含む、ヒトLRP5 mRNAまたはHBM mRNAを定量するための試薬セットが提供される。

【手続補正 2】

【補正対象書類名】明細書

【補正対象項目名】0141

【補正方法】変更

【補正の内容】

【0141】

蛍光標識オリゴヌクレオチドプライマーおよび未標識オリゴヌクレオチドプライマーを用いてDNAを遺伝子型解析した。Integrated DNA Technologies, Inc.(Coralville、アイオワ州)から標識および未標識オリゴヌクレオチドを入手した。マイクロサテライト遺伝子型解析のためのその他の全ての試薬はPerkin Elmer-Applied Biosystems, Inc.(「PE-ABI」)(Norwalk、コネチカット州)から購入した。個々のPCR反応を、PE-ABIに記載されるように、AmpliTaq(商標)DNAポリメラーゼを用いて各々のマーカに関して実施した。反応物を脱イオン化ホルムアミド、ブルー・デキストランおよびTAMRA 350サイズ・スタンダード(PE-ABI)を含有する負荷バッファー3.5 μlに加えた。95℃で5分間加熱してDNAを変性させた後、モデル377 DNAシーケンサー(PE-ABI、Foster City、カリフォルニア州)の操作マニュアルに記載されているように、サンプルを負荷し、電気泳動した。ゲル電気泳動の後、PE-ABI GENESCAN(登録商標)およびGENOTYPER(商標)ソフトウェアを用いてデータを分析した。最初に、GENESCAN(

商標) ソフトウェア内で、分析の第1ステップの前にレーントラッキングを手作業で最適化した。ゲルレーンデータを抽出した後、各レーンの標準的な曲線プロファイルを試験し、直線性およびサイズコーリングを確認した。これらのパラメーターのいずれかで問題を有したレーンを再トラッキングし、確認した。一度全レーンをトラッキングし、サイズスタンダードを正確に同定すると、対立遺伝子同定のためにデータをGENOTYPER (商標) にインポートした。対立遺伝子の呼び出し (ピニング) を迅速にするために、Dr.Guy Van Campのインターネットウェブサイト(alt.www.uia.ac.be/u/dnalab/ld.html)からのプログラム・Linkage Designerを使用した。このプログラムはGENOTYPER (商標) により作製したデータの家系ドロ잉プログラムCyrillic(バージョン2.0、Cherwell Scientific Publishing Limited, Oxford、英国)へのインポート、およびそれに続くLINKAGE (Lathropら、Am.J.Hum.Genet.37:482-498(1985))プログラムを用いる連鎖解析を大いに促進する。

【手続補正3】

【補正対象書類名】明細書

【補正対象項目名】0144

【補正方法】変更

【補正の内容】

【0144】

本発明では、発明者はHBM間隔をマーカーD11S987とGTC HBM マーカー5の間の領域に狭めることを報告している。これらの2つのマーカーは元来の分析(D11S905およびD11S937)からデリミッティング・マーカーの間にあり、そして互いにおよそ3cmである。マーカーD11S4191、D11S1883、D11S1785、D11S4113、D11S4136、D11S4139(Dibら、Nature 380:152-154(1996)、FGF3 (Polymeropolousら、Nucl.Acids Res.18:7468(1990)) (遺伝子マーカーに関する情報はゲノムデータベースのインターネットサイト、dwww.gdv.org/で見出すことができる)、ならびにマーカー、GTC HBM マーカー1、GTC HBM マーカー2、GTC HBM マーカー3、GTC HBM マーカー4、GTC HBM マーカー5、GTC HBM マーカー6、およびGTC HBM マーカー7からの遺伝子型データを用いて間隔を狭めることができる。

【手続補正4】

【補正対象書類名】明細書

【補正対象項目名】0178

【補正方法】変更

【補正の内容】

【0178】

2. クロスマッチプログラムを用いて配列内のBACベクター配列を「マスキング」した(Phil Green、chimera.biotech.washington.edu/UWGC)。前記で詳記したショットガンライブラリー構築によりショットガンライブラリーにいくつかのBACベクターが残るので、このプログラムを用いてBACコンティグの配列をBACベクターと比較し、次の工程の前にいずれかのベクター配列をマスキングした。マスキングされた配列を配列ファイルで「X」と印を付け、続く分析の間不活性のままであった。

【手続補正5】

【補正対象書類名】明細書

【補正対象項目名】0208

【補正方法】変更

【補正の内容】

【0208】

大腿骨および頭蓋冠骨組織から志向的にクローン化されたcDNAライブラリーを当業者に周知の方法(例えばSoares、in Automated DNA Sequencing and Analysis、Adams、FieldsおよびVenter編、Academic Press、ニューヨーク州、110-114頁、(1994))により構築した。最初に骨をハンマーで断片に破碎し、小片を液体窒素中で凍結し、組織パルペライ

ザー (Spectrum Laboratory Products) で粉末まで小さくした。ポリトロンホモジナイザー (Brinkman Instruments) を用いて標準的な酸グアニジニウム・チオシアネート・フェノール・クロロホルム抽出バッファー (例えば Chomczynski および Sacchi, Anal. Biochem. 162:156-159 (1987)) で粉末化された骨をホモジナイズすることにより、粉末化された骨から R N A を抽出した。加えて、ヒト脳および肺全 R N A を Clontech から購入した。ダイナビーズ d T を用いて、製造者の推奨に従って (Dynal, Inc.) ポリ A R N A を全 R N A から単離した。配列: 5'-AACTGGAAGAATTTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTTTTT-3' (SEQ ID NO:27) のオリゴヌクレオチドプライマーを用いて第 1 鎖 c D N A 合成を開始した。このプライマーを c D N A の 3' 末端で NotI 制限部位 (下線部) を導入した。「ワンチューブ」c D N A 合成キットを用いて、製造者 (Life Technologies, Bethesda, メリーランド州) に記載されるように、第 1 および第 2 鎖合成を実施した。二本鎖 D N A を T 4 ポリヌクレオチドキナーゼで処理して、分子の末端が平滑であり (Soares, in Automated DNA Sequencing and Analysis, Adams, Fields および Venter 編, Academic Press, ニューヨーク州, 110-114 頁 (1994)), 次に平滑末端 c D N A が B i o g e l カラム (Huynh ら, DNA Cloning, 1 巻, Glover 編, IRL Press, Oxford, 49-78 頁, (1985)) により、または s i z e - s e p 4 0 0 セファロースカラム (Pharmacia, カタログ番号 27-5105-01) を用いて大きさに選択されることを確実にした。4 0 0 塩基対またはそれ以上の長さの c D N A のみを次の行程で使用した。次いで当分野で周知の方法により (Soares, 1994), E c o R I アダプター (配列: 5' OH-A A T T C G G C A C G A G-OH 3' (配列番号 28) および 5' p-C T C G T G C C G-OH 3' (配列番号 29)) を二本鎖 c D N A にライゲートした。次いで NotI で消化することにより c D N A の 3' 末端から E c o R I アダプターを除去した (Soares, 1994)。次いで c D N A をプラスミドベクター pBluescript (登録商標) I I K S + にライゲートし (Stratagene, La Jolla, カリフォルニア州)、そして当業者に周知のエレクトロポレーション法によりライゲートされた材料を大腸菌 D H 1 0 B または D H 1 2 S に形質転換した (Soares, 1994)。37℃で一晩成長させた後、Mega-prep キット (Qiagen, Chatsworth, カリフォルニア) に関して指示されるように処理することによりプレートを擦り取った後、大腸菌コロニーから D N A を回収した。1 次形質転換体の全数の一部を計数し、そして平均のインサートの大きさおよび c D N A インサートを含まないプラスミドのパーセンテージを決定することにより c D N A ライブラリーの品質を推定した。さらなる c D N A ライブラリー (ヒト全脳、心臓、腎臓、白血球、および胎児脳) を Life Technologies, Bethesda, メリーランド州から購入した。

【手続補正 6】

【補正対象書類名】明細書

【補正対象項目名】0231

【補正方法】変更

【補正の内容】

【0231】

ほぼ中央で多型性を有する対立遺伝子特異的オリゴヌクレオチド (A S O) を設計した。オリゴヌクレオチドは 5' 末端でリン酸不含であり、G i b c o B R L から購入した。オリゴヌクレオチドの配列 (配列番号 631~632) は、

2326 ZMAX.ASO.g: AGACTGGGGTGGAGACGC

2327 ZMAX.ASO.t: CAGACTGGGTTGAGACGCC

である。

【手続補正 7】

【補正対象書類名】明細書

【補正対象項目名】0259

【補正方法】変更

【補正の内容】

【0259】

本明細書で論じる酵母 2 ハイブリッド実験で用いるアミノ酸配列は細胞質ドメイン全体

および膜貫通ドメインの一部から成り、これを以下に示す（アミノからカルボキシへの配向）（配列番号 8 1 2）：

【化 1】

RVVQRYAGA NGPFPHEYVS GTPHVPLNFI APGGSQHGPF TGIACGKSMM SSVSLMGGRG
GVPLYDRNHV TGASSSSSSS TKATLYPPIL NPPSPATDP SLYNMDMFYS SNIPATVRPY
RPYIIRGMAP PTPCSTDVC DSDYSASRWK ASKYLDLNS DSDPYPPPT PHSQYLSAED
SCPPSPATER SYFHLFPPPP SPCTDSS

【手続補正 8】

【補正対象書類名】明細書

【補正対象項目名】0 3 2 6

【補正方法】変更

【補正の内容】

【0 3 2 6】

CHO 発現ベクター pHTOP-Fc を Sal I および Not I で消化した。アクリルアミドゲル切片からの電気溶出により、介在配列を精製して残りのベクターから除去した。Sal I は pHTOP-Fc において 5' を内在性のミツバチ・メリチンシグナル配列まで切断し、Not I はちょうど 5' でコード化配列 IgG1-Fc まで切断する。得られた Sal I-Not I pHTOP-Fc ベクターはシグナル配列が除去されており、Not I クローニング部位は IgG-Fc への 5' 融合の作製を施しやすい。pCMVSPORT6 における LRP5 全長および pCMVSPORT6 における HBM 全長を別個に Xma I（シグナル配列をコードする ORF の領域内で切断する）および BamH I（ORF 内で切断する）で消化して LRP5 および HBM の 2286 bp 5' フラグメントを作製した。HBM から LRP5 を区別する変異体はこのフラグメントに在る。別に、LRP5 DNA を BamH I および Sac I で消化して LRP5 と HBM 遺伝子の双方に共通する 1800 bp 3' フラグメントを単離した。一緒にするとこれらの 2 個のフラグメントは HBM および LRP5 細胞外ドメインのコード化配列、少しのシグナル配列の最初の 6 個のアミノ酸のコード化配列ならびに細胞外ドメインの末端の前の 18 個の終末アミノ酸を構成し、本発明者は Kyte-Doolittle プロットから配列（配列番号 7 4 5）「SPAHSS」の末端まで推定した。

【手続補正 9】

【補正対象書類名】明細書

【補正対象項目名】0 3 2 7

【補正方法】変更

【補正の内容】

【0 3 2 7】

合成二本鎖を設計して、イニシエーターメチオニンおよび Kozak 配列を含む元来の Xma I 部位の 5' の LRP5/HBM シグナル配列コード化配列を再生した。Sal I（5'）および Xma I（3'）付着末端を含有するようにこの二本鎖を設計して末端を適応させて前記した遺伝子フラグメントを pHTOP-Fc ベクターに適応させた。以下に提示するような 2 つの部分的相補性オリゴヌクレオチド（配列番号 7 4 6 ~ 7 4 7）からこの合成二本鎖を構築した。

5'-TCGACCACCATGGAGGCAGCGCCGC-3'

3'-GGTGGTACCTCCGTCGCGGCGGGCC-5'

【手続補正 10】

【補正対象書類名】明細書

【補正対象項目名】0 3 2 8

【補正方法】変更

【補正の内容】

【0 3 2 8】

第2の合成二本鎖を設計して元来のS a c I部位から配列「. S P A H S S」のセリンに続く細胞外ドメインの推定される末端までを再生し、またI g G－F cの下流へのフレーム内融合を可能にするクロニング部位をコードした。S a c I（5'）およびN o t I（3'）付着末端を含有するようにこの二本鎖を設計して前記した遺伝子フラグメントをp H T O P－F cベクターに適應させた。以下に提示するような2つの部分的相補性オリゴヌクレオチド（配列番号748～749）からこの合成二本鎖を構築した。

5'-CATGTGTGAAATCACCAAGCCGCCCTCAGACGACAGCCCGGCCACAGCAGTGGC-3'

3'-TCGAGTACACACTTTAGTGGTTCGGCGGGAGTCTGCTGTCTCGGGCCGGGTGTCGTCACCGCCGG-5'

【手続補正11】

【補正対象書類名】明細書

【補正対象項目名】0340

【補正方法】変更

【補正の内容】

【0340】

【表7】

アミノ酸	アミノ酸	H/M*差異	コメント	配列番号：
171-187	VETPRIERAGMDGSTRK	5	HBM含有多型性	750
264-278	NKRTGGKRKEILSAL	3	細胞外	751
290-301	ERQPFFHTRCEE	2	EGF-Iに隣接細胞外	752
532-546	VDGTKRRTLLEDKLP	5	細胞外	753
901-915	DGLNDCMHNNNGQCGQ	2	EGF-III内細胞外	754
1010-1021	PFVLTSLSQGQN	6	細胞外、ヒト特異的	755
1415-1429	YAGANGPFPHEYVSG	3	細胞質内	756
1452-1464	ACGKSMMSSVSLM	5	細胞質内、ヒト特異的	757
1556-1573	RWKASKYYLDLNSDSDPY	1	細胞質内	758
888-902	SGWNECASSNGHCSH		LRP6特異的	759
1308-1321	NGDANCQDKSDEKN		LRP6特異的	760

*ヒトとマウス配列間のH/M差異

【手続補正12】

【補正対象書類名】明細書

【補正対象項目名】0344

【補正方法】変更

【補正の内容】

【0344】

一度s c F v分子を単離すると、これを試薬として免疫化学において用いて種々のヒトの正常および疾患組織におけるL R P 5タンパク質発現を検出した。L R P 5の細胞外ドメインのペプチド1000－1021（I K R A K D D G T Q P F V L T S L S Q G Q N）（配列番号761）に対する3つのファージクローンの詳細なs c F V抗体免疫組織化学分析により心筋、腎臓、肺および肝臓での陽性染色が示された。前立腺癌における発現もまた検出した。これらの結果はm R N A組織分布プロファイルおよび公開されたL R P 5 m R N A局在性(Kimら、J.Biochem.124:1072-1076(1998))の報告と合致する。得られたファージクローンはプールから生じ、これをシーケンシングして分子のF v領域にお

ける変異体の可能性を同定した。一度同定すると、次いで無傷および機能的免疫グロブリン γ (IgG) 分子の最終的な組み立てのためにCOS細胞に同時トランスフェクトするため、適当なscFVを可変性重鎖および可変性軽鎖DNA構築物にサブクローン化することができる。次いで当分野で公知なように、さらに特異性および反応性に関して細胞により発現されるIgGを特徴付けすることができる。

【手続補正13】

【補正対象書類名】明細書

【補正対象項目名】0349

【補正方法】変更

【補正の内容】

【0349】

例えばLRP5をpcDNA3.1 (Invitrogen、Carlsbad、カリフォルニア州)にクローン化した。これを用いて³⁵S標識インビトロ翻訳(Promega、Madison、ウィスコンシン州)LRP5を作製した。ペプチド免疫原：RWKASKYYLDLNSDSDPY (配列番号758)を志向する抗体(10 μ g/ml)3109および3110を、10 μ g/ml特異的ペプチド(すなわちRWKASKYYLDLNSDSDPY) (配列番号758)もしくは非特異的ペプチド(すなわちSGWNECASSNGHCSH) (配列番号759)のいずれかの存在下かまたはペプチド不含でインビトロ翻訳産物20 μ lと組み合わせ、そして4℃で1.5時間インキュベートした。次いでプロテインAセファロースをサンプル(予め網状赤血球ライゼートで約1.5時間遮断した)に加え、そしてサンプルを4℃で1時間振盪した。プロテインAセファロースをPBS 0.5mlで3回洗浄した。続いて結合タンパク質を4~12%グラジエントNuPAGEゲル(Invitrogen)上で製造者の指示書に従って分離した。ゲルを80℃で30分間乾燥させ、そして次にKodak X-OMAT-ARフィルムに24から48時間暴露した。特異的プローブが³⁵S標識LRP5免疫沈澱タンパク質に関していずれかの抗体と有意に競合するのが観察された。競合は非特異的タンパク質では観察されなかった。

【手続補正14】

【補正対象書類名】明細書

【補正対象項目名】0373

【補正方法】変更

【補正の内容】

【0373】

4つの独立した粒子クローンからLRP5構築物を作製した。これらのクローンをLRP5特異的プライミングしたcDNAライブラリーから単離した。部分的LRP5 cDNAクローンは内部探索シーケンシングクローンセットに、オリゴdTプライミングHelix細胞cDNAライブラリーから単離されたL236B P0049E08として存在した。このクローンを5'プライミング末端でトランケートした。全長cDNAを作製するのに必要なより多くの5'プライム含有フラグメントを単離するために、Clontechヒト肝臓ポリA mRNA (カタログ番号6510~1、ロット番号9060032)およびcDNA合成およびプラスミドクローニングキット(カタログ番号18248~013)のLife Technologies SuperScript(登録商標)プラスミドシステムからLRP5遺伝子特異的cDNAライブラリーを作製した。このライブラリーをL401と称した。製造者のプロトコルを以下の修飾を伴って実施した。1)第1および第2鎖合成反応の双方におけるDEPC処理水を α -³²P-dCTPと置き換えた。2)およそ1kb間隔でLRP5遺伝子に特異的であるように選択されたオリゴヌクレオチドを用いて逆転写をプライミングした。公のデータベースに対してプログラムBLASTを用いてこれらの配列を調査してLRP5遺伝子特異性を確認した。3)2つの別個の逆転写反応を実施した。最初の反応(A)は以下の、より多くのLRP5遺伝子の3'領域、(配列番号762~765) 47114: 5'-CGTACGTAAAGCGGCCGCTTGGCAATACAGATGTGGGA-3', 47116: 5'-CGTACGTAAAGCGGCCGCAGTAGCTCCTCTCGGTGGC-3', 47118: 5'-CGTACGTAAAGCGGCCGCTCATCATGGACTTTCCG-3', およ

び47120: 5'-CGTACGTAAAGCGGCCGCGCACTGCTGTTTGATGAGG-3'にアニーリングするオリゴヌクレオチドでプライミングした。第2の反応(B)は前記した4つのオリゴヌクレオチド、および(配列番号766~768)47108: 5'-CGTACGTAAAGCGGCCGCGAGTGTGGAAGAAAGGCTGC-3', 47110: 5'-CGTACGTAAAGCGGCCGCGCAGTAGAGCTTCCCCTCCTGC-3', および47112: 5'-CGTACGTAAAGCGGCCGCGTCCATCACGAAGTCCAGGT-3'を使用した。全てのオリゴヌクレオチドはNot I リンカー配列を含有し、0.02 µg/µlの濃度で使用した。4) 双方の逆転写反応物に由来するSal I 適合cDNAを1%アガロース、0.1 µg/ml 臭化エチジウム、1x TAEゲル上の電気泳動によりサイズ分画化した。0.6から8.0 kbの臭化エチジウム染色cDNAをゲルから切除した。製造者のプロトコルを用いてエレクトロエレーション(ISC0 Little Blue Tank Electroelutor(登録商標))によりアガロースゲルからcDNAを精製した。次いで反応AおよびBに由来する精製cDNAと一緒にプールした。5) サイズ分画化Sal I 適合cDNAをNot I-Sal I 消化pBluescript(登録商標)(Stratagene, La Jolla, カリフォルニア州)にライゲートした。

【手続補正15】

【補正対象書類名】明細書

【補正対象項目名】0375

【補正方法】変更

【補正の内容】

【0375】

ライブラリスクリーニングに使用するための2つのプローブを、鋳型としてライブラリーL401 100 ngを用いるポリメラーゼ連鎖反応(PCR)により作製した。標準的なPCR技術を用いた。反応混合物は各オリゴプライマー 10ピコモル; 0.2 mM 各dATP、dTTP、dCTPおよびdGTP(PE Applied Biosystemsカタログ番号N808-0260); Expand (商標) High Fidelity Taq DNAポリメラーゼ 1.5単位、および1x PCR反応バッファー(Roche Molecular Biochemicals、カタログ番号731-641; 10 mM Tris-HCl、1.5 mM MgCl₂、50 mM KCl、pH 8.3)を含有した。混合物を99℃で1分間インキュベートし、続いて96℃で15秒間、50℃で30秒間、72℃で1分間の30サイクル、72℃で7分間の最終インキュベーションを行った(MJResearch DNA Engine(登録商標)Tetrat PCT-225)。LRP5遺伝子のNot I-Sal Iフラグメントを増幅するオリゴ(配列番号769~770) 107335: 5'-CAGCGGCCTGGAGGATGC-3' および107338: 5'-CGGTCCAGTAGAGGTTTCG-3'を用いて第1鎖を作製した。LRP5遺伝子のSac I-Kpn Iフラグメントを増幅するオリゴ(配列番号771~772) 107341: 5'-CATCAGCCGCGCCTTCATG-3' および107342: 5'-CCTGCATGTTGGTGAAGTAC-3'を用いて第2鎖を作製した。Qiaquick (登録商標) キットおよび製造者のプロトコル(Qiagenカタログ28106)を用いて双方のPCR産物を精製し、次いで製造者のプロトコルに従ってベクターpCRII-TOPO (登録商標) (Invitrogenカタログ番号K4600)にサブクローン化した。精製されたプラスミドDNA(標準的な分子生物学技術を使用)の制限消化および続くDNA配列分析(ABI Prism BigDye Terminator Cycle(登録商標)シーケンシング、カタログ番号4303154、ABI377機器)により陽性サブクローンを同定した。各々の配列確認されたpCRII-TOPO (登録商標) クローンのEcoRI制限消化(New England Biolabs、カタログ番号R0101L)によりプローブDNAを単離した。1%アガロース、0.1 µg/ml 臭化エチジウム、1x TAEゲル上のゲル電気泳動により制限フラグメントをサイズ分画化した。インサートDNAをゲルから切除し、製造者のプロトコルに従ってClontech NucleoSpin(登録商標)核酸精製キット(カタログ番号K3051-2)を用いて精製した。Prime-It II(登録商標)ランダムプライマー標識キットおよびプロトコル(Stratagene、カタログ番号300385)を用いて、精製されたDNAフラグメント(25~50 ng)をRedivue (登録商標) (α³²P)-dCTP(Amersham Pharmacia、カタログ番号AA0005)で標識した。組み込まれなかったdCTPをAmershamのNICK (登録商標) カラムおよびプロトコル(カタログ番号17-0855-02)で除去した。

【手続補正16】

【補正対象書類名】明細書

【補正対象項目名】0380

【補正方法】変更

【補正の内容】

【0380】

1) 塩基1-1366:LRP5 cDNA構築物、GTC.Zmax1 13からXbaI-SalIフラグメントが得られた。GTC.Zmax1 13は、LRP5の全ORFを含有する5057BPインサートを含有する。クローンを平滑末端化し、そしてpSTBlue-1のEcoRV部位にクローン化した。pBluescript (商標) II誘導体の骨ランダムプライミングcDNAライブラリーのスクリーニングに由来する5'クローンを、pBluescript (商標) IIの骨dTプライミングcDNAライブラリーからのPCR産物に由来する3'クローンと融合させることによりこのクローンを作製した。DNAzymeポリメラーゼを用いてLRP5特異的順行性プライマー(配列番号773) 5'-GCCCGAACAACCTCTACTGGACCGAC-3'および逆行性プライマー(配列番号774) 5'-GCCCAACCCCAATCACAGTTCACATT-3'を用いてPCRを実施した。得られた3.7kb PCR産物をPCR-XL-TOPOにクローン化した。全長クローン作製するために、5'および3'プラスミドをGibco/BRLのDM1 (dam-) に形質転換した。5'プラスミドをXbaIで消化し、そして3'プラスミドをHindIIIで消化した。消化されたプラスミドをT4ポリメラーゼで埋めて平滑末端を作製し、そしてBclIで切断した。1.7kb 5'フラグメントおよび3.5kb 3'フラグメントをゲル精製し、一緒にライゲートし、そしてpSTBlue1のEcoRV部位にクローン化した。これにより塩基100で始まるコード化配列を伴う短い5'UTRが提供される。さらに、これは5'多重クローニング部位でいくつかのさらなる制限部位を担持する。このフラグメントは位置558でGenbankアクセッション番号AF077820配列に相対する、結果的にA(AF077820)からGに変化するDNA多型性をも含有し;この変異はアミノ酸の差異には至らない(Pro)。

【手続補正17】

【補正対象書類名】明細書

【補正対象項目名】0386

【補正方法】変更

【補正の内容】

【0386】

HBM変異G171Vの作製

PCRを用いて、アミノ酸171でグリシンからバリンへの予測されるアミノ酸変化に至るHBM変異を全長ヒトLRP5 cDNA(プラスミドZMAXGI 3AS)に導入して、位置611のGをTに変化させた。LRP5遺伝子の内因性NotI部位をランキングするオリゴ(配列番号775) 107335:(5'-CAGCGGGCCTGGAGGATGC-3')および(配列番号776) 49513:(5'-CGGGTACATGTACTGGACAGCTGATTAGC-3')を用いてHBM変異の導入を行った。この方法によりPCR産物の3'末端で新規PvuII部位が作製される。該産物の5'末端にScaI部位を導入するオリゴを用いて第2のPCR反応を完了し、3'プライミング末端でLRP5の内因性SalI部位を含有する。Qiaquick(登録商標)手順(Qiagen Inc.)を用いてPCR産物を精製し;前記したベクターpCRII-TOPO(Invitrogen)にサブクローン化した。全て前記のとおり、プラスミドDNAを単一の細菌性コロニーから精製し、制限消化および続く配列分析により分析した。配列確認されたpCRII-TOPOクローンを各々NotI-PvuIIおよびScaI-SalIで制限消化した。得られたDNAフラグメントを前記のとおりサイズ分画化し、精製した。次いでこれらの2つのフラグメントをNotI-SalI

消化により調製したベクター、pB l u e s c r i p t（登録商標）にサブクローン化した。P v u I IおよびS c a I産物の双方は二本鎖核酸を消化するのに用いた場合、平滑末端を生じる。従って、得られたライゲートされたフラグメントはP v u I IまたはS c a I部位のいずれかを再生することができず、新たに導入されたH B M変異を除いてコンセンサスL R P 5配列のみを含有する。全長L R P 5遺伝子に変異を導入するために、この得られたプラスミドをM s c IおよびS a l Iで消化したが、一方L R P 5プラスミドG T C . Z m a x 1 1 3のX b a I - M s c I消化によりL R P 5の5'領域が得られた。これらの2つのフラグメントと一緒にX b a I - S a l I消化したpB l u e s c r i p t（登録商標）にライゲートし、結局類似の1. 3 6 6 k b X b a I - S a l Iフラグメントを作製した。違いはこの構築物が前記したH B M変異を含有するというだけである。次いでL R P 5遺伝子に関して正確に前記のとおり、この新たに作製されたH B M変異を含むX b a I - S a l Iフラグメントの置換を伴う全長H B M c D N Aをp C M V S P O R T 6 . 0に集めた。D N A配列分析により全c D N Aインサートを確認し、H B M変異の導入を確認した。

【手続補正18】

【補正対象書類名】明細書

【補正対象項目名】0397

【補正方法】変更

【補正の内容】

【0397】

以下のプライマーおよびプローブセットを用いて、製造者(PE Applied Biosystems、Foster City、カリフォルニア州)に記載されるように、ヒト細胞から単離されたRNAのT a q M a n（登録商標）定量的逆転写酵素ポリメラーゼ連鎖反応(R T - P C R)分析を実施した。

ヒトZ m a x 1 - 1 / H B M : (配列番号 777~779)

順行性プライマー: 5' - G T C A G C C T G G A G G A G T T C T C A - 3'

逆行性プライマー: 5' - T C A C C C T T G G C A A T A C A G A T G T - 3'

プローブ: 6 - F A M - 5' - C C C A C C C A T G T G C C C G T G A C A - 3'

【手続補正19】

【補正対象書類名】明細書

【補正対象項目名】0398

【補正方法】変更

【補正の内容】

【0398】

PE Applied BiosystemsのヒトG A P D Hコントロールキットでの多重プロトコルを用いて、実験的プライマー/プローブセットの結果をヒトG A P D Hレベルに正規化した。以下のプライマーおよびプローブセットを用いて、製造者(PE Applied Biosystems、Foster City、カリフォルニア州)に記載されるように、ネズミ細胞および組織から単離されたRNAの種特異的T a q M a n（登録商標）定量的R T - P C R分析を実施した。

ヒトZ m a x 1 - 1 / H B M - 1 : (配列番号 780~782)

順行性プライマー: 5' - C G T G A T T G C C G A C G A T C T C - 3'

逆行性プライマー: 5' - T T C C G G C C G C T A G T C T T G T - 3'

プローブ: 6 - F A M - 5' - C G C A C C C G T T C G G T C T G A C G C A G T A C - 3'

マウスZ m a x 1 - 1 / H B M - 1 : (配列番号 783~785)

順行性プライマー: 5' - C T T T C C C C A C G A G T A T G T T G G T - 3'

逆行性プライマー: 5' - A A G G G A C C G T G C T G T G A G C - 3'

プローブ: 6 - F A M - 5' - A G C C C C T C A T G T G C C T C T C A A C T T C A T A G - 3'

【手続補正20】

【補正対象書類名】明細書

【補正対象項目名】0403

【補正方法】変更

【補正の内容】

【0403】

初めに、マウスLRP5遺伝子のエクソン5つを含むマウスゲノムDNA BACクローン473P5 (Genbankアクセッション番号AZ095413) のゲノムDNAを用いて骨ジーンターゲットベクターを構築した。Research Genetics(Huntsville、アラバマ州)によりエクソン3に関するポリメラーゼ連鎖反応(PCR)スクリーニングを用いてマウス129SvJゲノムBACライブラリーからこのクローンを単離した。順行性プライマーの配列(5'-GAGCGGGCAGGGATGGATGG-3') (配列番号786) および逆行性プライマーの配列(5'-AGGTGGCACGGTGGATGAAGC-3') (配列番号787) を用いてPCRによりエクソン3を増幅し; 以下の熱循環条件を用いた。95℃で0.5分間、55℃で1分間および72℃で1分間の30サイクル。このクローンのマウスLRP5遺伝子との同一性を、鋳型としてBACクローンDNAを用いてエクソン3をシーケンシングすることにより確認した。pGEM-T-easy T/Aクローニングキットを用いてPCR産物をクローン化した。

【手続補正21】

【補正対象書類名】明細書

【補正対象項目名】0405

【補正方法】変更

【補正の内容】

【0405】

双方のZmax1 KI/KO標的化ベクター(AおよびB)を、エクソン3でG170Vアミノ酸置換をコードするGからTへのヌクレオチド置換するように修飾した。鋳型として相同性の短いアームの野生型配列を用いて、重複PCR変異誘発によりこれらの修飾を導入した。加えて、Zmax1 KI/KO(B)標的化ベクターの1kbの短いアームを、5'末端PmeI制限認識部位を含むように修飾した。配列(5'-AAGCTTGTTTTAAACTGGGCAATGGTGGCACCATGGTTGTAAAT-3') (配列番号788) の順行性プライマーおよび配列:

【化2】

(5'-GGGCTTCC~~A~~CCCAGTCAGTCCAGTACATGTACCT-3')

(配列番号789) の逆行性(変異原性)プライマーを用いて5'重複フラグメントを作製した。30サイクルで利用した熱循環条件は95℃で0.5分間、55℃で1分間および72℃で1分間である。配列:

【化3】

(5'-CTGACTGGG~~T~~TGGAAGCACCCCGGATCGAGC-3')

(配列番号790) の順行性プライマーおよび配列(5'-GAATTCAATCGGTACCTGTGCGGCGCTTCAATTG-3') (配列番号791) の逆行性(変異原性)プライマーを用いて3'フラグメントを作製した。30サイクルで利用した熱循環条件は95℃で0.5分間、55℃で1分間および72℃で1分間である。最終重複PCRは鋳型としてPCR反応の5'フラグメントおよび3'フラグメント各々1mlを利用し、各々5'および3'フラグメントの順行性および逆行性プライマーならびに同一の熱循環パラメーターを用いて増幅を実施した。pGEM-T-easy T/Aクローニングキットを用いて最終PCR産物をクローン化した。変異誘発されたエクソン3を600

b p B s m B I - X b a I フラグメントとして Z m a x 1 K I / K O (B) から切除し、Z m a x 1 K I / K O (A) に移した。

【手続補正 2 2】

【補正対象書類名】明細書

【補正対象項目名】0 4 0 7

【補正方法】変更

【補正の内容】

【0 4 0 7】

相同性の標的化ベクター領域をフランキングする、および直ぐ隣接するゲノムフラグメントの P C R クローニングにより Z m a x 1 K I / K O (B) の外側プローブを調製する。Z m a x 1 K I / K O (B) に用いられる 5 ' 外側プローブは配列 (5 ' - T G A G A T G T C C T G T C T G T G G C - 3 ') (配列番号 7 9 2) の順行性プライマーおよび配列 (5 ' - T C C T T T C C T T C C C T A C A G T T G - 3 ') (配列番号 7 9 3) の逆行性プライマーを用いて作製された 4 9 8 b p フラグメントである。これらのプライマーを用いて 3 0 サイクルで利用した熱循環条件は、9 5 °C で 0 . 5 分間、5 5 °C で 1 分間および 7 2 °C で 1 分間である。3 ' 外側プローブは配列 (5 ' - C C T A A G G A T C T C C T T G T G T C T G T G G - 3 ') (配列番号 7 9 4) の順行性プライマーおよび配列 (5 ' - C T G C A G C A G G T C A G T A G C C T G C - 3 ') (配列番号 7 9 5) の逆行性プライマーを用いて作製された 6 0 0 b p フラグメントである。これらのプローブを用いて 3 0 サイクルで利用した熱循環条件は、9 5 °C で 0 . 5 分間、5 5 °C で 1 分間および 7 2 °C で 1 分間である。双方のプローブはゲノムサザン分析において L R P 5 遺伝子に特異的である。p G E M - T - e a s y T / A クローニングキットを用いて P C R 産物をクローン化した。

【手続補正 2 3】

【補正対象書類名】明細書

【補正対象項目名】0 4 0 8

【補正方法】変更

【補正の内容】

【0 4 0 8】

L R P 5 m R N A 構造および転写レベルのリボヌクレアーゼ保護分析のためのプローブをエクソン 3 からエクソン 4 を含有する c D N A フラグメントを P C R クローニングすることにより調製した。P C R 反応は鋳型として完全 c D N A、配列 (5 ' - T G A G A T G T C C T G T C T G T G G C - 3 ') (配列番号 7 9 6) の順行性プライマー、配列 (5 ' - T C C T T T C C T T C C C T A C A G T T G - 3 ') (配列番号 7 9 7) の逆行性プライマーおよび以下の 3 0 サイクルのための熱循環条件を使用した。9 5 °C で 0 . 5 分間、5 5 °C で 1 分間および 7 2 °C で 1 分間。p G E M - T - e a s y T / A クローニングキットを用いて P C R 産物をクローン化した。

【手続補正 2 4】

【補正対象書類名】明細書

【補正対象項目名】0 4 1 3

【補正方法】変更

【補正の内容】

【0 4 1 3】

以下のプライマーセットを遺伝子型解析に使用する。

H B M M C B A :

5 ' プライマー : 2 9 6 塩基対フラグメント

順行性 : 5 ' - G C T T C T T G G C G T G T G A C C G G C G - 3 ' (配列番号 7 9 8)

逆行性 : 5 ' - G C C G C A C A G C G C C A G C A G C A G C - 3 ' (配列番号 7 9 9)

3'プライマー：400塩基対フラグメント

順行性：5'－C A C C C A C A G C C A G T A－3' (配列番号 800)

逆行性：5'－A T T T G C C C T C C C A T A T G T C C T T C C－3' (配列番号 801)

H B M M T I C：

5'プライマー：382塩基対フラグメント

順行性：5'－T T C C T C C C A G C C C T C C T C C A T C A G－3' (配列番号 802)

逆行性：5'－G C C G C A C A G C G C C A G C A G C A G C－3' (配列番号 803)

3'プライマー：524塩基対フラグメント

順行性：5'－G A A T G G C G C C C C C G A C G A C－3' (配列番号 804)

逆行性：5'－G C T C C C A T T C A T C A G T T C C A T A G G－3' (配列番号 805)

【手続補正25】

【補正対象書類名】明細書

【補正対象項目名】0538

【補正方法】変更

【補正の内容】

【0538】

【図1】図1は遺伝子連鎖研究に用いられた個体の家系を示す。各々の個体の下に、ID番号、脊髄BMDに関するzスコアがあり、対立遺伝子が、染色体11上の重要なマーカーを示す。黒ぬりの記号は「患っている」個体を表す。「N」を含む記号は「患っていない」個体である。37人の個体からのDNAについて遺伝子型を同定した。クエションマークは遺伝子型が同定されなかった未知の遺伝子型または個体を示す。

【図2】図2は11q13.3におけるHBM領域のBAC／STS含量物理地図を表す。STSマーカーは遺伝子、EST、マイクロサテライト、ランダム配列、およびBACエンドシーケンスを長い水平方向の線の上に記す。GDBに存在するマーカーに関しては、同一の命名法を用いている。遺伝子座の名前(D11S####)は利用可能な場合、第1次名の後に括弧の中に列挙する。BACエンドシーケンスから誘導したSTSをBAC名と共に列挙し、クロンの左および右末端に関して各々LまたはRと続ける。2つの大きな矢印はHBMの重要な領域を定義する遺伝子マーカーを示す。STSの下で水平方向の線は、9倍の適用範囲のBACライブラリーのPCR基盤のスクリーニングにより同定されたBACクロンを示す。白丸はライブラリースクリーニング中、マーカーが対応するBACライブラリアドレスを増幅しなかったことを示している。クロン名は以下の慣用法を用いる。BACに関してはB、プレート、列および囲みアドレス、続いてHがHBMプロジェクトを示す(すなわちB36F16-H)。

【図3】図3A～3Fはイントロン配列の側面に位置しているLRP5のゲノム構造を示す。翻訳は下線を付したエクソン1の「ATG」により開始される。HBM遺伝子の多型性の部位はエクソン3であり、下線を付した「G」により表され、それによりこのヌクレオチドはHBM遺伝子の「T」である。mRNAの3'非翻訳領域にはエクソン23内で下線を付している(エクソン1、配列番号40；エクソン2、配列番号41；エクソン3、配列番号42；エクソン4、配列番号43；エクソン5、配列番号44；エクソン6、配列番号45；エクソン7、配列番号46；エクソン8、配列番号47；エクソン9、配列番号48；エクソン10、配列番号49；エクソン11、配列番号50；エクソン12、配列番号51；エクソン13、配列番号52；エクソン14、配列番号53；エクソン15、配列番号54；エクソン16、配列番号55；エクソン17、配列番号56；エクソン18、配列番号57；エクソン19、配列番号58；エクソン20、配列番号59；

エクソン 2 1、配列番号 6 0；エクソン 2 2、配列番号 6 1およびエクソン 2 3、配列番号 6 2）。

【図 4】図 4 は Y W T D スペーサー、細胞外付着部位、L D L およびカルシウムの結合部位、システイン富化成長因子リピート、膜貫通領域、C K - I I リン酸化部位を伴う理想的な P E S T 領域および内部移行ドメインを含む L R P 5 のドメイン構成を示す。図 4 はまた H B M タンパク質で生じるグリシンのバリンへの変化の部位を示す。シグナルペプチドはアミノ酸 1 ～ 3 1 に位置し、細胞外ドメインはアミノ酸 3 2 ～ 1 3 8 5 に位置し、膜貫通セグメントはアミノ酸 1 3 8 6 ～ 1 4 1 3 に位置し、細胞質ドメインはアミノ酸 1 4 1 4 ～ 1 6 1 5 に位置する。

【図 5】図 5 は H B M 遺伝子に関連する B A C コンティグ B 5 2 7 D 1 2 および B 2 0 0 E 2 1 の略図である。

【図 6】図 6 A ～ 6 J は野生型遺伝子 L R P 5 のヌクレオチド（配列番号 1）およびアミノ酸（配列番号 3）配列を示す。ヌクレオチド 5 8 2 でのグアニンのチミンへの塩基対置換の位置（配列番号 2、4）には下線を付している。この対立遺伝子変異体は H B M 遺伝子である。H B M 遺伝子は 1 7 1 位置でグリシンのバリンへのアミノ酸置換を有するタンパク質をコードする。5' 非翻訳領域（U T R）は塩基 1 から 7 0 の境界であり、3' U T R は塩基 4 9 1 6 - 5 1 2 0 の境界である。

【図 7】図 7 A および 7 B は種々組織における L R P 5 の発現を示すノーザンブロット分析である。

【図 8】図 8 は P C R 産物分析である。

【図 9】図 9 は L R P 5 エクソン 3 変異の対立遺伝子特異的オリゴヌクレオチド検出である。

【図 1 0】図 1 0 は 1 0 0 倍拡大した、センスおよびアンチセンスプローブを用いるインサイチュ・ハイブリダイゼーションによるマウス L R P 5 の細胞局在化である。

【図 1 1】図 1 1 は 4 0 0 倍拡大した、センスおよびアンチセンスプローブを用いるインサイチュ・ハイブリダイゼーションによるマウス L R P 5 の細胞局在化である。

【図 1 2】図 1 2 は 4 0 0 倍拡大した、センスおよびアンチセンスプローブを用いる、骨内膜での骨芽細胞のインサイチュ・ハイブリダイゼーションによるマウス L R P 5 の細胞局在化である。

【図 1 3】図 1 3 は M C - 3 T 3 細胞での L R P 5 発現のアンチセンス阻止を示す。

【図 1 4】図 1 4 は、野生型 L R P 5 対立遺伝子（左パネル、G 特異的オリゴ；5 5℃、洗浄）と比較した H B M 対立遺伝子（右パネル、T 特異的オリゴ；5 8℃、洗浄）の希少性を説明する L R P 5 エクソン 3 対立遺伝子特異的オリゴヌクレオチド（A S O）アッセイを示す。右パネルで現れる陽性スポットは陽性対照である。

【図 1 5】図 1 5 は接着斑情報伝達における L R P 5（Z m a x 1）の潜在的役割を表すモデルを示す。

【図 1 6】図 1 6 は内因性マウス L R P 5 のノックアウトまたは H B M 多型性の条件的ノックインのための 2 つの L R P 5 ジーンターゲティングベクターを模式的に表す。B、X、および R は各々 D N A B A C 4 7 3 5 P 5 における B a m H I、X b a I および E c o R I 部位を示す。エクソン 3、4 および 5 を黒色の長方形で示す。A G → T 塩基変化は H B M 多型性を作製するためにエクソン 3 の塩基 2 4 で操作される。ネオマイシン抵抗性遺伝子および合成休止配列を含有する L o x P フランキングされるカセットならびに E S 細胞クロンのスクリーニングおよび特徴付けに用いられるプローブの位置をも示す。

【図 1 7】図 1 7 はトランスジェニック（すなわち H B M M C B A および H B M M T I C）および野生型（すなわち Z m a x W T C B A および Z m a x W T T I C）プラスミド構築物により発現を確認する。これらの構築物を H O B - 0 2 - 0 2 細胞に一過性にトランスフェクトし、m R N A レベルを T a q M a n（登録商標）定量 P C R を用いて決定した。H B M M C B A および Z m a x W T C B A を表の左のコラムで示し（すなわち C M V β アクチン）ならびに H B M M T I C および Z m a x W T T I C を表の右のコラムで示す（すなわち I 型コラーゲン）。

【図18】図18はヒトおよびマウスTaqMan（登録商標）プライマー／プローブセット間の比較を示す。HOB（HOB-03-C5）およびマウス（MC-3T3-E1）骨芽細胞mRNAをプローブおよびプライマーを用いて分析した。

【図19】図19はTaqMan（登録商標）プライマー／プローブセットを用いて、ヒトおよびマウスRNA混合物バックグラウンドで発現されたヒトZmax-1 mRNAの定量を示す。結果は添加したヒトLRP5 mRNA（ng）対測定したヒトLRP5 mRNA（ng）を表す。

【図20】図20はTaqMan（登録商標）により分析されたmRNA発現に基づくトランスジェニックマウスにおけるHBMの発現を表す。

【図21】図21A～Cは脊椎（図21A）、大腿骨（図21B）および全身（図21C）においてHBMMCBA構築物を発現する種々トランスジェニックマウス株の分析を表す。図21D～Fは脊椎（図21D）、大腿骨（図21E）および全身（図21F）においてHBMMTIC構築物を発現する種々トランスジェニックマウス株の分析を表す。

図21G～Lは脊椎、大腿骨および全身において17週を通じてHBMMTIC構築物を発現するトランスジェニックマウス株の分析を表す。

【図22】図22はインビボpDXA分析^{*}を用いて、5週でのHBMトランスジェニックマウスにおけるBMD（すなわちHBMMCBAおよびHBMMTIC構築物）の変化を表す。BMD変化を、これもまた5週齢のみである野生型動物と比較して提示する。

【図23】図23はインビボpDXA分析^{*}を用いて、9週でのHBMトランスジェニックマウスにおけるBMD（すなわちHBMMCBAおよびHBMMTIC構築物）の変化を表す。BMD変化を、これもまた9週齢のみである野生型動物と比較して提示する。

【図24】図24（A～D）はHBMGI 2ASベクターインサートの配列（配列番号806）を提示する。

【図25】図25（A～D）はZMAXGI 3ASベクターインサートの配列（配列番号807）を提示する。

【図26】図26（A～C）はヒト（配列番号808）およびマウス（配列番号809）LRP5アミノ酸配列のアラインメントを提示する。

【図27】図27（A～C）はヒトLRP5（配列番号810）およびLRP6（配列番号811）アミノ酸配列のアラインメントを提示する。

【図28】図28はマウスの骨成長に及ぼす負荷の影響を試験するための装置を説明する。

【図29】図29は骨負荷のHBMトランスジェニックおよび非トランスジェニックマウスの骨成長に及ぼす影響の組織形態学的説明を提示する。

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US02/14876										
A. CLASSIFICATION OF SUBJECT MATTER IPC(7) : A01K 67/00, 67/027; C12N 15/00; G01N 33/00 US CL : 536/23.1; 800/3, 13, 18, 21 According to International Patent Classification (IPC) or to both national classification and IPC												
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) U.S. : 536/23.1; 800/3, 13, 18, 21 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) Please See Continuation Sheet												
C. DOCUMENTS CONSIDERED TO BE RELEVANT												
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.										
A	US 5,691,153 (RECKER et al.) 25 November 1997 (25.11.1997), entire document.	17-26, 48, 49, 52-58										
A, P	BAFICO et al. Novel mechanism of Wnt signalling inhibition mediated by Dickopf-1 interaction with LRP6/Arrow. Nature Cell Biology. July 2001, Vol. 3, pages 683-686, entire document.	17-26, 48, 49, 52-58										
A, P	GONG et al. LDL receptor-related protein 5 (LRP5) affects bone accrual and eye development. Cell. 16 November 2001, Vol. 107, pages 513-523, entire document.	17-26, 48, 49, 52-58										
A	HEY et al. Cloning of a novel member of the low-density lipoprotein receptor family. Gene. 1998, Vol. 216, pages 103-111, entire document.	17-26, 48, 49, 52-58										
A, P	MITCHELL et al. Functional analysis of secreted and transmembrane proteins critical to mouse development. Nature Genetics. July 2001, Vol. 28, pages 241-249, entire document.	17-26, 48, 49, 52-58										
A	PINSON et al. An LDL-receptor-related protein mediates Wnt signalling in mice. Nature. 28 September 2000, Vol. 407, pages 535-538, entire document.	17-26, 48, 49, 52-58										
A	TAMAL et al. LDL-receptor-related proteins in Wnt signal transduction. Nature. 28 September 2000, Vol. 407, pages 530-535, entire document.	17-26, 48, 49, 52-58										
☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.												
* Special categories of cited documents: <table border="0"> <tr> <td>"A" document defining the general state of the art which is not considered to be of particular relevance</td> <td>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</td> </tr> <tr> <td>"E" earlier application or patent published on or after the international filing date</td> <td>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</td> </tr> <tr> <td>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</td> <td>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</td> </tr> <tr> <td>"O" document referring to an oral disclosure, use, exhibition or other means</td> <td>"&" document member of the same patent family</td> </tr> <tr> <td>"P" document published prior to the international filing date but later than the priority date claimed</td> <td></td> </tr> </table>			"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention	"E" earlier application or patent published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone	"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art	"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family	"P" document published prior to the international filing date but later than the priority date claimed	
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention											
"E" earlier application or patent published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone											
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art											
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family											
"P" document published prior to the international filing date but later than the priority date claimed												
Date of the actual completion of the international search 30 June 2002 (30.06.2002)		Date of mailing of the international search report 06 AUG 2002										
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231		Authorized officer Anne Marie Baker, Ph.D. Telephone No. (703)308-0196										
Facsimile No.												

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/14876

C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WEHRLE et al. arrow encodes an LDL-receptor-related protein essential for Wingless signalling. Nature. 28 September 2000, Vol. 407, pages, 527-530, entire document.	17-26, 48, 49, 52-58

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/14876

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claim Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☒ Claim Nos.: 1-4 and 15
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
Please See Continuation Sheet
3. ☒ Claim Nos.: 5-14, 16, 27-47, 50, 51, 59
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest ☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/14876

Continuation of Box I Reason 2:

A sequence listing in computer readable form was provided but was technically flawed. Thus, the claims could not be searched on the basis of the SEQ ID Nos. The remaining claims (Claims 17-26, 48, 49, and 52-58) were searched to the extent possible without the Sequence Listing using keywords based on the description.

Continuation of B. FIELDS SEARCHED Item 3:

WEST

Dialog (file: medicine)

search terms: LRP5, LRP6, bone mass, bone density, Zmax1, LR3, HBM, gene, transgen?, mouse, murine

フロントページの続き

(51)Int.Cl. ⁷	F I	テーマコード (参考)
C 1 2 Q 1/02	C 1 2 Q 1/68	Z
C 1 2 Q 1/68	G 0 1 N 33/15	Z
G 0 1 N 33/15	G 0 1 N 33/50	X
G 0 1 N 33/50	G 0 1 N 33/50	Z
	C 1 2 N 5/00	B

(31)優先権主張番号 60/361,293

(32)優先日 平成14年3月4日(2002.3.4)

(33)優先権主張国 米国(US)

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT, BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN, TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EC,EE,ES,FI,GB,GD,GE, GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,OM,PH,P L,PT,RO,RU,SD,SE,SG,SI,SK,SL,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VN,YU,ZA,ZM,ZW

(特許庁注：以下のものは登録商標)

U N I X

(72)発明者 ベックス, フレデリック・ジェイムズ, ザ・サード
アメリカ合衆国ペンシルヴァニア州19073, ニュータウン・スクエア, レイクショア・ドライ
ヴ 4

(72)発明者 ボーディン, ピーター・ヴァン・ネスト
アメリカ合衆国ペンシルヴァニア州19083, ヘイヴァータウン, アカデミー・レイン 174
2

(72)発明者 アスキュー, ロジャー・ジー
アメリカ合衆国マサチューセッツ州01921, ボックスフォード, ハイ・リッジ・ロード 11

Fターム(参考) 2G045 AA25 AA40 BB03 BB20 CB01 CB13 CB17 DA12 DA13 DA14
DA77 GC06
4B024 AA01 AA11 AA20 BA80 CA04 CA20 DA02 DA03 EA04 FA02
GA11 HA11 HA20
4B063 QA01 QQ02 QQ79 QQ91 QR72 QR80 QS31 QX10
4B065 AA91X AA91Y AC14 AC20 BA02 CA24 CA43 CA44 CA46
4C084 AA02 AA17 NA14 ZA962 ZC112