

(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2004-510408

(P2004-510408A)

(43) 公表日 平成16年4月8日(2004.4.8)

(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 6 1 K 39/00	A 6 1 K 39/00	H
A 6 1 K 45/00	A 6 1 K 45/00	
A 6 1 K 48/00	A 6 1 K 48/00	
A 6 1 K 51/00	A 6 1 P 35/00	
	審査請求 未請求 予備審査請求 有 (全 177 頁) 最終頁に続く	

(21) 出願番号	特願2002-500720 (P2002-500720)	(71) 出願人	502160545 ディアデクサス インコーポレーテッド アメリカ合衆国 カリフォルニア州 94 080、サウス サン フランシスコ、オ イスター ポイント ブルバード 343
(86) (22) 出願日	平成13年5月29日 (2001.5.29)	(74) 代理人	100102842 弁理士 葛和 清司
(85) 翻訳文提出日	平成14年11月26日 (2002.11.26)	(72) 発明者	マシーナ, ロベルト エー. アメリカ合衆国 カリフォルニア州 95 136、サン ホセ、クレッセンド アベ ニュー 4118
(86) 国際出願番号	PCT/US2001/017583	(72) 発明者	チェン, セイユ アメリカ合衆国 カリフォルニア州 94 404、フォスター シティ、ミラ ス トリート 160
(87) 国際公開番号	W02001/092528		
(87) 国際公開日	平成13年12月6日 (2001.12.6)		
(31) 優先権主張番号	60/207,383		
(32) 優先日	平成12年5月26日 (2000.5.26)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 大腸癌の診断、モニタリング、ステージング、イメージングおよび処置の方法

(57) 【要約】

本発明は、CSGポリペプチド、このポリペプチドをコードするポリヌクレオチド、特にこのポリヌクレオチドを発現することによる、上記ポリペプチドの製造方法、ならびに上記ポリペプチドのアゴニストおよびアンタゴニストに関する。本発明はさらに、研究、診断および臨床技術に一部関連する用途のための、かかるポリヌクレオチド、ポリペプチド、アゴニストおよびアンタゴニストを用いるための方法に関する。

【特許請求の範囲】**【請求項 1】**

C S G であって、

(a) 配列番号 1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21 または 22 のポリヌクレオ

チド、またはそれらの変異体、

(b) 配列番号 1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21 または 22 のポリヌクレオチドにより発現されるタンパク質またはそれらの変異体、あるいは

(c) 配列番号 1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21 または 22 のアンチセンス配列に、ストリンジェントな条件下にてハイブリダイズすることが可能なポリヌクレオチド、
を含む、前記 C S G。 10

【請求項 2】

患者における大腸癌の存在を診断する方法であって、

(a) 患者における細胞、組織または体液における請求項 1 に記載の C S G レベルを測定すること、および

(b) 正常ヒト対照由来の細胞、組織または体液における C S G レベルと、測定された C S G レベルとを比較すること、ここで、正常ヒト対照に対する患者における測定 C S G レベルの変化は、大腸癌の存在に関連する、

を含む、前記方法。 20

【請求項 3】

患者における大腸癌の転移を診断する方法であって、

(a) 転移したことがわかっていない大腸癌を有する患者を同定すること、

(b) 該患者由来の細胞、組織または体液の試料中の、請求項 1 に記載の C S G レベルを測定すること、および

(c) 正常ヒト対照の細胞、組織または体液における C S G レベルと、測定 C S G レベルとを比較すること、ここで、正常ヒト対照に対する患者における測定 C S G レベルの増加は、転移した癌に関連する、

を含む、前記方法。 30

【請求項 4】

大腸癌を有する患者における大腸癌をステージングする方法であって、

(a) 大腸癌を有する患者を同定すること、

(b) 患者由来の細胞、組織または体液の試料中の請求項 1 に記載の C S G レベルを測定すること、および

(c) 正常ヒト対照の細胞、組織または体液における C S G レベルと、測定 C S G レベルとを比較すること、ここで、正常ヒト対照に対する患者における測定 C S G レベルの増加は、進行している癌に関連し、測定 C S G レベルの減少は、消退または寛解している癌に関連する、

を含む、前記方法。 40

【請求項 5】

転移の発生に関して、患者における大腸癌をモニタリングする方法であって、

(a) 転移したことがわかっていない大腸癌を有する患者を同定すること、

(b) 患者由来の細胞、組織または体液の試料中の請求項 1 に記載の C S G レベルを定期的に測定すること、および

(c) 正常ヒト対照の細胞、組織または体液における C S G レベルと、定期的に測定した C S G レベルとを比較すること、ここで、正常ヒト対照に対する患者において定期的に測定した C S G レベルのいずれか 1 つの増加が、転移した癌に関連する、

を含む、前記方法。

【請求項 6】

50

患者における大腸癌のステージの変化をモニタリングする方法であって、

(a) 大腸癌を有する患者を同定すること、

(b) 患者由来の細胞、組織または体液における請求項 1 に記載の CSG レベルを定期的に測定すること、および

(c) 正常ヒト対照の細胞、組織または体液における CSG レベルと、定期的に測定した CSG レベルとを比較すること、ここで、正常ヒト対照に対する患者において定期的に測定した CSG レベルのいずれか 1 つの増加は、ステージが進行している癌に関連し、減少が、ステージが後退しているかまたは寛解している癌に関連する、を含む、前記方法。

【請求項 7】

大腸癌のイメージングおよび処置における使用のための潜在的治療薬を発見する方法であって、化合物を、請求項 1 に記載の CSG に結合する能力、または該化合物の非存在下における CSG に対しその発現を減少させる能力に関してスクリーニングすることを含み、ここで、CSG に結合する化合物の能力または CSG の発現を減少させる能力は、該化合物が大腸癌のイメージングおよび処置において有用であることを示す、前記方法。

【請求項 8】

請求項 1 に記載の CSG によりコードされるポリペプチドを特異的に結合する抗体。

【請求項 9】

請求項 8 に記載の抗体を患者に投与することを含む、患者における大腸癌のイメージングの方法。

【請求項 10】

抗体が、常磁性イオンまたは放射性同位体で標識されている、請求項 9 に記載の方法。

【請求項 11】

請求項 1 に記載の CSG の発現または活性を下方調節する化合物を患者に投与することを含む、患者における大腸癌の処置の方法。

【請求項 12】

請求項 1 に記載の CSG を発現する標的細胞に対する免疫応答を誘発する方法であって、免疫応答が標的細胞に対して惹起されるように、免疫原性刺激量の CSG ポリペプチドをヒト患者に送達することを含む、前記方法。

【請求項 13】

CSG ポリペプチドが、配列番号 1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21 または 22 のポリヌクレオチドによりコードされている、請求項 12 に記載の方法。

【請求項 14】

請求項 1 に記載の CSG を含む大腸癌処置用ワクチン。

【発明の詳細な説明】

【0001】

発明の技術分野

本発明は、部分として、新たに同定されたポリヌクレオチドおよびポリペプチド、該ポリヌクレオチドおよびポリペプチドの変異体ならびに誘導体、該ポリヌクレオチドおよびポリペプチドならびにそれらの変異体および誘導体を作製する方法、該ポリペプチドのアゴニストおよびアンタゴニスト、ならびに癌、特に大腸癌の検出、診断、モニタリング、ステージング、予後判定、イメージングおよび処置のための、該ポリヌクレオチド、ポリペプチド、変異体、誘導体、アゴニストおよびアンタゴニストの使用に関する。特に、これらおよびその他の観点において、本発明は、以後大腸特異遺伝子または「CSG」と称する大腸特異ポリヌクレオチドおよびポリペプチドに関する。

【0002】

発明の背景

大腸の癌は、腸管に限局していれば、高度に処置可能で、多くの場合治癒する疾患である。それは、米国において最も頻繁に診断される悪性腫瘍の 1 つであるとともに、2 番目に

10

20

30

40

50

多い癌死の原因である。初期治療は手術であり、それにより約50%の患者が治癒する。しかしながら、術後の再発が大きな問題であり、多くの場合それが最終的な死因となる。大腸癌の予後判定は、腸管壁を介した腫瘍の侵入度およびリンパ節転移(nodal involvement)の有無に明白に関係している。これら2つの特徴は、この疾患のために開発された全てのステージングシステムの基礎を形成している。処置の決定は、通常、ステージングのための旧デューク(older Duke's)分類または改変Astler-Coller(MAC)分類を参照してなされる。

【0003】

腸管閉塞および腸管穿孔が、大腸癌患者における不良な予後の指標である。処置前の癌胎児性抗原(CEA)および糖鎖抗原19-9(CA19-9)の上昇した血清レベルもまた、予後に関してネガティブな意味を有する。

10

上診時における70歳を超える年齢は、標準的な治療にとって禁忌ではない。許容し得る罹患率および死亡率、ならびに長期の生存が、この患者集団で達成されている。

【0004】

本疾患が高頻度であり(年間約160,000件の新たな結腸および直腸癌の症例)、ハイリスク集団が分かっており、一次病変の増殖が緩徐であること証明されており、早期の病変での生存が良好であり、そしてスクリーニング検査が比較的容易で正確なため、大腸癌のスクリーニングは、50歳以上の全ての成人、特に結腸直腸癌の一等親血縁者の日常的なケアの一部となるべきである。

【0005】

大腸癌の検出、診断、モニタリング、ステージング、および予後判定に用いられる手順は、患者の予後にとって決定的に重要である。例えば、早期大腸癌と診断された患者は、一般的に、遠隔転移性大腸癌と診断された患者の5年生存率と比較して、ずっと高い5年生存率を有する。大腸癌を早期に検出するための、より感受性が高く特異的な新たな診断方法が明らかに必要である。

20

【0006】

初期治療後およびアジュバント療法中に、大腸癌患者を綿密にモニタリングして、治療に対する応答を測定し、持続性もしくは再発性の転移の疾患を検出する。大腸癌、その再発および進行を検出する際に、より感受性が高く特異的な大腸癌マーカーが明らかに必要である。

30

【0007】

大腸癌を管理する際の別の重要なステップは、患者の疾患のステージを決定することである。ステージの決定は、潜在的な予後判定価値を有し、最適治療を設計するための基準を提供する。一般に、大腸癌の病理学的ステージングは、より正確な予後判定を与えるため、臨床的ステージングよりも好ましい。しかしながら、臨床的ステージングは、病理学的評価用組織を得るための侵襲的方法によらないので、少なくとも病理学的ステージングと同程度に正確である場合には好ましいであろう。大腸癌のステージングは、異なる浸潤のステージを識別することができる、細胞、組織または体液における新規マーカーを検出することにより改良されるであろう。

【0008】

したがって、かかる癌が転移しているか否かを決定するための、より高感度で正確な、ヒトにおける大腸癌をステージングする方法、および転移の発生に関して、ヒトにおける転移していない大腸癌の進行をモニタリングする方法が強く求められている。

40

【0009】

本発明では、本明細書中でCSGと呼ばれる大腸特異遺伝子による大腸癌の検出、診断、モニタリング、ステージング、予後判定、イメージングおよび処置の方法が提供される。本発明に関しては、CSGはとりわけ配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22のポリヌクレオチド配列を含む遺伝子により発現される未変性(native)タンパク質を表す。本明細書中では、「CSG」はまた、遺伝コードの縮重に起因して、配列番号1

50

、 2、 3、 4、 5、 6、 7、 8、 9、 10、 11、 12、 13、 14、 15、 16、 17、 18、 19、 20、 21または22に比べヌクレオチド配列中に変化を含むが、依然として同一タンパク質をコードするポリヌクレオチドをも意味する。

【0010】

あるいは、本明細書中で用いられるCSGが意味するものは、配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22のポリヌクレオチド配列を含む遺伝子にコードされる未変性mRNA、配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22のポリヌクレオチド配列を含む遺伝子のレベル、または配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22のアンチセンス配列に、ストリンジェントな条件下でハイブリダイズ可能なポリヌクレオチドのレベルを意味する。

10

【0011】

本発明の他の目的、特徴、利点および側面は、以下の説明より当業者に明らかになるだろう。しかしながら、以下の説明および具体的な例は、発明の好ましい態様を示すものではあるが、例示のみを目的として描写されたものであることが理解されるべきである。以下の記述を読むことにより、そして本開示のその他の部分を読むことにより、開示の発明の精神および範囲内の各種変更および改良は当業者に容易に明らかになるだろう。

【0012】

発明の概要

これらの目的、およびその他の目的に対し、本発明の目的は、配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22のポリヌクレオチド、配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22のポリヌクレオチドにより発現されるタンパク質、あるいはそのタンパク質を発現するそれらの変異体、またはストリンジェントな条件下にて配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22のアンチセンス配列にハイブリダイズすることができるポリヌクレオチド、を含むCSGを提供することである。

20

30

【0013】

本発明の別の目的は、細胞、組織または体液中のCSGレベルの変化を、正常ヒト対照の、好ましくは同種類の細胞、組織または体液中のCSGのレベルと比較して分析することにより、大腸癌の存在を診断する方法であって、正常ヒト対照に対する患者におけるCSGレベルの変化が大腸癌に関連する方法を提供することである。

【0014】

さらに、転移した大腸癌を有することが疑われるヒト患者を同定し、かかる患者由来の細胞、組織または体液試料をCSGについて分析し、かかる細胞、組織または体液中のCSGレベルを正常ヒト対照の、好ましくは同種類の細胞、組織または体液中のCSGレベルと比較することにより、転移をしたことがわかっていない大腸癌を有する患者における転移性大腸癌を診断する方法であって、正常ヒト対照に対する患者におけるCSGレベルの増加が、転移をした大腸癌に関連する方法が提供される。

40

【0015】

発明はさらに、大腸癌を有するヒト患者を同定し、かかる患者由来の細胞、組織または体液試料を分析し、かかる細胞、組織または体液中のCSGレベルを、正常ヒト対照試料の、好ましくは同種類の細胞、組織または体液中のCSGレベルと比較することにより、かかる癌を有するヒトにおいて大腸癌をステージングする方法であって、正常ヒト対照に対する患者におけるCSGレベルの増加が、消退または寛解している癌に関連する方法を提供する。

【0016】

50

さらに、転移の発生について、大腸癌を有するヒトにおいて大腸癌をモニタリングする方法が提供される。この方法は、転移したことがわかっていないかかる癌を有する患者を同定することと、かかる患者由来の細胞、組織または体液試料をCSGについて定期的に分析することと、かかる細胞、組織または体液中のCSGレベルを、正常ヒト対照試料の、好ましくは同種類の細胞、組織または体液中のCSGレベルと比較することとを含み、ここで、正常ヒト対照に対する患者におけるCSGレベルの増加が、転移した癌に関連している。

【0017】

さらに、大腸癌を有するヒトにおけるCSGレベルを観察することにより、大腸癌を有するヒトの大腸癌のステージの変化をモニタリングする方法が提供される。この方法は、かかる癌を有するヒト患者を同定することと、かかる患者由来の細胞、組織または体液試料をCSGについて定期的に分析することと、かかる細胞、組織または体液中のCSGレベルを、正常ヒト対照試料の、好ましくは同種類の細胞、組織または体液中のCSGレベルと比較することとを含み、ここで、正常ヒト対照に対する患者におけるCSGレベルの増加が、進行している癌に関連し、CSGレベルの減少が、消退または寛解している癌に関連する。

10

【0018】

さらに、大腸癌のイメージングおよび処置に使用するための、CSGを標的とする新たな治療薬を設計する方法を提供する。例えば、一態様では、CSGを標的とする抗体、またはかかる抗体の断片のような治療薬を用い、疾患もしくは症状を検出または診断することを目的とし、患者におけるCSGの局在を処置、検出またはイメージングすることができる。この態様では、正常組織に比して検出される標識抗体量の増加が腫瘍転移または成長の指標となるであろう。かかる抗体は、ポリクローナル、モノクローナルまたはオムニクローナルであることができ、あるいは分子生物学的技術により調製できる。

20

【0019】

本発明にて、および本明細書全体にわたって用いられる場合、「抗体」という用語はまた、SELEXと呼ばれ、当業者に既知である*in vitro*進化プロトコルに由来するものなどのアダプターおよび一本鎖オリゴヌクレオチドを含むとされる。抗体は放射性同位体および常磁性金属を含むが、これらに限定されない各種の検出可能でかつ治療的な標識で標識することができる。CSGの濃度および/または活性を下げる小分子および抗体などの治療薬もまたCSGの過剰発現を特徴とする疾患の処置に使用できる。かかる薬剤は本明細書中の教示に従い容易に同定できる。

30

【0020】

本発明の他の目的、特徴、利点および側面は、以下の説明から当業者に明らかになるだろう。しかしながら、以下の説明および具体的な例は、発明の好ましい態様を示しているが、例示のみを目的としていると理解されるべきである。以下の説明を読み、および本開示のその他部分を読むことから、開示の発明の精神および範囲内の各種変更および改良は当業者に容易に明らかになるだろう。

【0021】

用語集

以下の例示の説明は本明細書中、特に例中に頻繁に使用されるある種の用語の理解を助けるために提供される。説明は、便宜的に提供されるものであり、発明を限定するものではない。

40

【0022】

単離されたとは、その自然状態から「人工的」に変更されること、すなわち自然界で生じる場合、それはその本来の環境から変更されたか、または取り出されたか、あるいはその両方を意味する。

その用語が本明細書中で使用されるように、例えば、生きている動物に、その自然な状態で自然に存在しているポリヌクレオチドまたはポリペプチドは「単離された」ものではないが、その自然状態の共存物質から分離された同一ポリヌクレオチドまたはポリペプチド

50

は「単離された」ものである。例えば、ポリヌクレオチドに関しては、単離されたという用語は、それが天然に存在する染色体および細胞から分離されていることを意味する。

【0023】

単離の一部として、または単離に続いて、かかるポリヌクレオチドを、例えば、突然変異誘発、融合タンパク質を形成させること、および宿主内での増殖または発現のために、DNAなどの他のポリヌクレオチドに結合させることができる。単離ポリヌクレオチドは、単独で、またはベクターなどの他のポリヌクレオチドと結合させて、培養液または生体全体において、宿主細胞に導入することができる。培養液または生体全体において、宿主細胞へ導入する場合、その用語が本明細書中で用いられるように、かかるDNAは、自然の形態または環境にないことから、依然単離されている。同様に、ポリヌクレオチドおよびポリペプチドは、例えば、培地配合物、ポリヌクレオチドまたはポリペプチドの、例えば細胞への導入用の溶液、化学または酵素反応のための組成物もしくは溶液などの組成物中に現れてもよいが、それらは、天然に存在する組成物ではなく、そこで、それらは、本明細書中で用いられるような用語の意味において、単離ポリヌクレオチドまたはポリペプチドのままである。

10

【0024】

オリゴヌクレオチドは、比較的短いポリヌクレオチドを指す。多くの場合、この用語は一本鎖デオキシリボヌクレオチドを指すが、とりわけ、同様に一本鎖または二本鎖リボヌクレオチド、RNA:DNAハイブリッド、および二本鎖DNAも指すことができる。一本鎖DNAプローブオリゴヌクレオチドのようなオリゴヌクレオチドは、自動オリゴヌクレオチド合成装置にて行なわれるものなどの化学的方法により合成される。しかしながら、オリゴヌクレオチドは、例えば *in vitro* 組換えDNA媒介技術にて、ならびに細胞および生体でのDNA発現による、各種のその他の方法により作製することができる。

20

【0025】

まず化学的に合成したDNAは、典型的には、5'リン酸塩なしで得られる。かかるオリゴヌクレオチドの5'末端は、組換えDNA分子形成に典型的に用いられるDNAリガーゼを使用する連結反応によるホスホジエステル結合形成の基質ではない。かかるオリゴヌクレオチドの連結が望ましい場合には、キナーゼおよびATPを使用するものなどの標準的技術によりリン酸塩を付加することができる。

30

【0026】

化学的に合成したオリゴヌクレオチドの3'末端は、遊離の水酸基を一般的に有しており、T4DNAリガーゼのようなりガーゼの存在下にて、容易に別のオリゴヌクレオチドなどの別のポリヌクレオチドの5'リン酸塩とホスホジエステル結合を形成するだろう。周知のとおり、この反応は、望ましい場合、連結前に他のポリヌクレオチドの5'リン酸塩を除去することで選択的に阻止することができる。

【0027】

ポリヌクレオチドは一般に、任意のポリリボヌクレオチドあるいはポリデオキシリボヌクレオチドを指し、未修飾RNAもしくはDNA、ならびに修飾RNAもしくはDNAを包含する。したがって、例えば、本明細書中で用いるポリヌクレオチドは、とりわけ一本鎖および二本鎖DNA、一本鎖および二本鎖領域の混合物であるDNA、一本鎖および二本鎖RNA、ならびに一本鎖および二本鎖領域の混合物であるRNA、一本鎖、あるいはより一般的には二本鎖または一本鎖および二本鎖領域の混合物であり得るDNAおよびRNAを含むハイブリッド分子を表す。

40

【0028】

さらに、本明細書中で用いるポリヌクレオチドは、RNAまたはDNA、あるいはRNAとDNAの両方を含む三重鎖領域を指す。かかる領域内の鎖は同一分子または異なる分子由来であってもよい。この領域には、1または2以上の分子全てを含んでもよいが、より一般的にはいくつかの分子の領域のみが含まれる。三重らせん領域の分子の1つは、オリゴヌクレオチドであることが多い。

50

【0029】

本明細書中で使用される場合、ポリヌクレオチドという用語にはまた、1または2以上の修飾塩基を含有する上記のようなDNAまたはRNAが含まれる。すなわち、安定性またはその他の理由から修飾された主鎖を持つDNAまたはRNAは、本明細書中でその用語が意図されるような「ポリヌクレオチド」である。さらに、イノシンなどの稀な塩基あるいはトリチル化塩基などの修飾塩基を含むDNAまたはRNAも、2例のみ挙げたが、本明細書中でその用語が使用されているポリヌクレオチドである。

【0030】

当業者に既知の有益な用途を提供する様々な修飾が、DNAおよびRNAに対してなされることが理解されるだろう。本明細書中で使用されるポリヌクレオチドという用語には、化学的、酵素的または代謝的に修飾された形のポリヌクレオチド、ならびにとりわけ単純および複雑な細胞を含む、ウイルスおよび細胞に特徴的なDNAおよびRNAの化学形態が包含される。

10

【0031】

ポリペプチドは、本明細書中で使用される場合、以下に記載の全てのポリペプチドが含まれる。ポリペプチドの基本構造は既知であり、当該分野の多くの教本およびその他の刊行物に記載されている。この状況では、本明細書中ではその用語は、ペプチド結合により直鎖状に相互に結合された2または3以上のアミノ酸を含む任意のペプチドまたはタンパク質を指すのに用いられる。本明細書中で用いる場合、その用語は一般には例えばペプチド、オリゴペプチドおよびオリゴマーとも称される短鎖型、および当該分野において一般にタンパク質と呼ばれる、多くの型がある長鎖型の両方を指す。

20

【0032】

ポリペプチドが20種類の天然アミノ酸と一般に呼ばれる20種類のアミノ酸以外のアミノ酸を含むことが多々あること、および末端アミノ酸を含む多くのアミノ酸が既定のポリペプチドにおいて、プロセッシングやその他の翻訳後修飾のような自然のプロセスにより、または当該技術分野で既知の化学的修飾技術によって修飾されてもよいことが理解されるだろう。ポリペプチド中に自然に生ずる一般的修飾でさえ、あまりに多すぎて、本明細書中にその全体を列挙することはできないが、それらは基礎的テキスト、および詳細なモノグラフ、ならびに多くの研究文献に十分に記載されており、それらは当業者にとって既知である。

30

【0033】

本発明のポリペプチド中に存在する既知の修飾は、代表的な少数のもの挙げると、アセチル化、アシル化、ADP-リボシル化、アミド化、フラビンの共有結合、ヘム部分の共有結合、ヌクレオチドまたはヌクレオチド誘導体の共有結合、脂質または脂質誘導体の共有結合、ホスホチジルイノシトールの共有結合、架橋、環化、ジスルフィド結合形成、脱メチル化、共有架橋形成、システイン形成、ピログルタメート形成、ホルミル化、ガンマ-カルボキシル化、グリコシル化、GPIアンカー形成、ヒドロキシル化、ヨウ化、メチル化、ミリストイル化、酸化、タンパク質分解性プロセッシング、リン酸化、プレニル化、ラセミ化、セレノイル化、硫酸化、アルギニル化などの転移RNAを介したタンパク質へのアミノ酸付加、およびユビキチン化を含む。

40

【0034】

かかる修飾は当業者に既知であり、科学文献に詳細記載されている。グリコシル化、脂質付着、硫酸化、グルタミン残基のガンマ-カルボキシル化、ヒドロキシル化およびADP-リボシル化を含むがこれらに限定されない幾つかの特に一般的な修飾は、最も基本的なテキスト、例えばPROTEINS STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993)に記載されている。

【0035】

例えばPOSTTRANSLATIONAL COVALENT MODIFICATION

50

N OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York (1983)の中の、Wold, F., Posttranslational Protein Modifications: Perspectives and Prospects, pgs, 1-12、Seifter et al., Analysis for protein modifications and nonprotein cofactors, Meth. Enzymol. 182: 626-646 (1990)、およびRattan et al., Protein Synthesis: Posttranslational Modification and Aging, Ann. N.Y. Acad. Sci. 663: 48-62 (1992)などの、多くの詳しいレビューが本件に利用できる。

10

【0036】

本発明のポリペプチドは必ずしも全体が線状ではないということは理解されるであろう。むしろ、ポリペプチドはユビキチン化の結果により分枝されていてもよく、そして、それらは、一般に自然のプロセッシング現象を含む翻訳後の現象、および自然には起こらない人工的な操作によりもたらされる現象の結果として、分枝を伴った、または伴わない、環状であってもよい。環状、分枝状、および分枝環状ポリペプチドは、非翻訳的自然プロセスにより、ならびに完全な合成的方法によっても合成され得る。

【0037】

修飾は、ペプチド主鎖、アミノ酸側鎖およびアミノまたはカルボキシル末端を含むポリペプチド中のいかなる場所でも起こり得る。実際に、共有結合的修飾によるポリペプチド中のアミノまたはカルボキシル基、あるいは両方の遮蔽は天然および合成ポリペプチドにおいて一般的であり、かかる修飾は、本発明のポリペプチド中にも同様に存在し得る。例えば、タンパク質分解プロセッシング前に大腸菌で作られたポリペプチドのアミノ末端残基は、大部分が不変的にN-ホルミルメチオニンであろう。

20

【0038】

ポリペプチド中に見られる修飾は、多くの場合それが作製される条件を反映するであろう。例えば、宿主内でクローニングした遺伝子を発現することにより作製されたポリペプチドでは、大部分の修飾の性質および程度が宿主細胞の翻訳後修飾能力およびポリペプチドのアミノ酸配列中に存在する修飾シグナルにより決定されるであろう。例えば、周知のように、グリコシル化は大腸菌などの細菌性宿主ではたいてい起こらない。

30

【0039】

したがって、グリコシル化が望まれる場合には、ポリペプチドはグリコシル化用宿主、一般には真核生物細胞で発現させるべきである。しばしば昆虫細胞は、哺乳動物細胞と同一の翻訳後グリコシル化を行う。したがって、とりわけ自然なパターンのグリコシル化を有する哺乳動物タンパク質を効率的に発現させるために、昆虫細胞発現系が開発されている。他の修飾にも同様の考察が適用される。

【0040】

同タイプの修飾が、所定のポリペプチド中の複数の部位にて同一程度または様々な程度で存在し得ることが理解されるだろう。また所定のポリペプチドは多くの型の修飾を含んでもよい。

40

一般に、本明細書中で使用される場合には、ポリペプチドという用語は、かかる全ての修飾、特に宿主細胞中にてポリヌクレオチドを発現させることにより合成されたポリペプチド中に存在する修飾を包含する。

【0041】

本明細書中で用いられる場合、ポリヌクレオチドまたはポリペプチドの変異体は、それぞれ参照(reference)ポリヌクレオチドまたはポリペプチドとは異なるポリヌクレオチドまたはポリペプチドである。

【0042】

変異体ポリヌクレオチドに関して、一般にその差は、参照ヌクレオチドと変異体ヌクレオチドの配列が全体としては極めて類似しており、多くの領域で同一であるように制限され

50

る。したがって、変異体のヌクレオチド配列の変化はサイレント (s i l e n t) であり得る。すなわち、それらはポリヌクレオチドによりコードされるアミノ酸を変化させないだろう。変化がこの型のサイレントな変化に限定される場合、変異体は参照と同一のアミノ酸配列を持つポリペプチドをコードするだろう。あるいは、変異体のヌクレオチド配列中の変化は、参照ポリヌクレオチドによりコードされるポリペプチドのアミノ酸配列を改変してもよい。かかるヌクレオチドの変化は、参照配列によりコードされるポリペプチド中にアミノ酸の置換、付加、欠失、融合および切断を生じ得る。

【 0 0 4 3 】

変異体ポリペプチドに関して、一般にその差は、参照体と変異体の配列が全体としては極めて類似しており、多くの領域で同一であるように制限される。例えば、変異体ポリペプチドと参照ポリペプチドは、1または2以上の、置換、付加、欠失、融合および切断によりアミノ酸配列が異なってもよく、それらは、いかなる組み合わせで存在してもよい。

10

【 0 0 4 4 】

本明細書中で用いる受容体分子は、本発明の C S G ポリペプチドと特異的に結合または相互作用する分子を指し、好ましい古典的な受容体だけでなく、発明のポリペプチドと特異的に結合または相互作用するその他分子も包含する（それらはまたそれぞれ「結合分子」および「相互作用分子」、ならびに「C S G 結合または相互作用分子」とも呼ばれる）。

【 0 0 4 5 】

本発明のポリペプチドと、受容体分子または結合分子または相互作用分子を含むかかる分子との結合は、本発明のポリペプチドに限定され得る（それは非常に高度に好ましい）か、または本発明のポリペプチドに高度に特異的であり得る（それは高度に好ましい）か、または本発明のポリペプチドを含むタンパク質のグループに高度に特異的であり得る（それは好ましい）か、あるいはその内の少なくとも1つが本発明のポリペプチドを含むタンパク質の複数のグループに特異的であり得る。

20

【 0 0 4 6 】

受容体はまた、本発明のポリペプチドに結合する、抗体および抗体由来試薬のように、非天然でもあってもよい。

【 0 0 4 7 】

発明の詳細な説明

本発明は、とりわけ以下に詳述するように、本明細書中で C S G と称する、新規大腸特異ポリペプチドおよびポリヌクレオチドに関する。

30

【 0 0 4 8 】

ポリヌクレオチド

本発明の一側面によれば、C S G ポリペプチドをコードする単離 C S G ポリヌクレオチドが提供される。

配列番号 1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21および22に記載のポリヌクレオチド配列などの本明細書中に提供された情報を用いて、C S G をコードする本発明のポリヌクレオチドは、出発原料としてヒト腫瘍細胞由来の m R N A を用いた c D N A のクローニング法などの標準的クローニングおよびスクリーニング手法を用いて得てもよい。

40

【 0 0 4 9 】

本発明のポリヌクレオチドは、m R N A などの R N A の形態、あるいは、例えば、クローニングにより得た、もしくは化学合成技術により、またはその組み合わせにより生産された c D N A やゲノム D N A を含む D N A の形態であってもよい。D N A は二本鎖でも一本鎖でもよい。一本鎖 D N A は、センス鎖としても知られるコード鎖であるか、またはそれはアンチセンス鎖とも呼ばれる非コード鎖でもよい。

【 0 0 5 0 】

ポリペプチドをコードするコード配列は、配列番号 1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または2

50

2のポリヌクレオチドのコード配列と同一であってもよい。またそれは、遺伝コードの冗長性(縮重)の結果として、配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22によりコードされると同じポリペプチドをコードする、異なる配列を持つポリヌクレオチドであってもよい。

【0051】

これらのポリペプチドをコードする、配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22などの本発明のポリヌクレオチドは、それ自身成熟ポリペプチド用コード配列を含んでもよい。本発明のポリヌクレオチドはまた、成熟ポリペプチド用コード配列およびプレタンパク質、またはプロタンパク質またはプレプロタンパク質配列などの、リーダー配列あるいは分泌配列をコードする配列などのさらなるコード配列含んでもよい。

10

【0052】

本発明のポリヌクレオチドはまた、成熟ポリペプチドをコードする配列を、上記のさらなるコード配列を伴って、または伴わずに、追加の、非コード配列と一緒に含んでもよい。本発明のポリヌクレオチドに組込んでもよい追加の非コード配列の例は、例えば、転写、スプライシングおよびポリアデニル化シグナルを含むmRNAプロセッシング、mRNAのリボソーム結合および安定性の役割を果たす転写された非翻訳配列などのイントロン、ならびに非コード5'および3'配列、ならびに追加の機能を提供するようなアミノ酸をコードする、追加のコード配列を含むが、これらに限定されない。

20

【0053】

したがって、例えば、ポリペプチドは、融合ポリペプチドの精製を容易にするペプチドなどの、マーカー配列と融合してもよい。発明のこの側面のある好ましい実施例では、マーカー配列はとりわけpQEベクター(Qiagen, Inc.)に提供されているタグのようなヘキサ-ヒスチジンペプチドであり、その多くが市販されている。Gentz et al. (Proc. Natl. Acad. Sci., USA 86: 821-824 (1989))に記載されるように、例えばヘキサ-ヒスチジンは融合タンパク質の便利な精製を提供する。HAタグは、インフルエンザ血球凝集素タンパク質由来のエピトープである(Wilson et al., Cell 37: 767 (1984))。

30

【0054】

上記によれば、本明細書中で使用される「ポリペプチドをコードするポリヌクレオチド」という用語は、本発明のポリペプチドをコードする配列、特に配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22を含むポリヌクレオチドを包含する。この用語は、ポリペプチドをコードする単一連続領域または不連続領域(例えばイントロンにより中断された)を、さらにコードおよび/または非コード配列を含有してもよいさらなる領域とともに含むポリヌクレオチドを包含する。

【0055】

本発明は、さらにCSGポリペプチドの断片、類縁体および誘導体をコードする、本明細書中で上述したポリヌクレオチドの変異体に関する。ポリヌクレオチドの変異体は、天然の対立遺伝子変異体のような天然型変異体であってもよく、またはそれは天然に生ずることが知られていない変異体でもよい。かかる非天然型のポリヌクレオチド変異体は、ポリヌクレオチド、細胞または生体に適用されるものを含む突然変異誘発技術により作製され得る。

40

【0056】

この観点の変異体としては、ヌクレオチド置換、欠失または付加により上記ポリヌクレオチドとは異なる変異体である。置換、欠失または付加は、1または2以上のヌクレオチドを包含する。変異体はコードまたは非コード領域、あるいはその両方が変更され得る。コード域内の変更は、保存的または非保存的なアミノ酸置換、欠失または付加を生じてもよい。

50

【0057】

この観点における本発明の特に好適な態様としては、配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22を含むCSGポリヌクレオチドによってコードされたのと同じアミノ酸配列を有するポリペプチドをコードするポリヌクレオチド、それらの変異体、類縁体、誘導体および断片、ならびに変異体、類縁体および誘導体の断片がある。さらにこの観点で特に好適なものは、ポリペプチドの変異体、類縁体、誘導体および断片、ならびに該断片の変異体、類縁体および誘導体をコードするCSGポリヌクレオチドであって、その中で幾つか、少数の、5~10、1~5、1~3、2、1個または0個のアミノ酸残基が任意の組合せで置換、欠失または付加されたものである。

10

【0058】

これらの中で特に好ましいものは、CSGの特性および活性を変化させない、サイレントな置換、付加および欠失である。またこの点で特に好ましいものは、保存的な置換である。最も高度に好ましいのは、置換のない、配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22によってコードされるポリペプチドと同じアミノ酸配列を有するポリペプチドをコードするポリヌクレオチドである。

【0059】

本発明のさらに好ましい態様は、配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22のポリヌクレオチドと少なくとも70%同一であるCSGポリヌクレオチド、およびかかるポリヌクレオチドに相補的であるポリヌクレオチドである。より好ましいのは、配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22のポリヌクレオチドと少なくとも80%同一である領域を含むCSGポリヌクレオチドである。

20

【0060】

この点に関して、それ(the same)と少なくとも90%同一であるCSGポリヌクレオチドが特に好ましく、これらの特に好ましいCSGポリヌクレオチドの中でも、少なくとも95%であるものが特に好ましい。さらに、少なくとも95%であるものの中でも、少なくとも97%であるものが非常に好ましく、これらの中でも、少なくとも98%および99%であるものが特に非常に好ましく、少なくとも99%であるものが最も好ましい。

30

【0061】

さらに、この点に関して特に好適な態様は、配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22のポリヌクレオチドによってコードされた成熟ポリペプチドと実質的に同じ生物学的機能または活性を保持するポリペプチドをコードするポリヌクレオチドである。

【0062】

本発明はさらに、上記CSG配列とハイブリダイズするポリヌクレオチドに関する。この観点では、本発明は特に、ストリンジェントな条件下にて本明細書中で上述したポリヌクレオチドにハイブリダイズするポリヌクレオチドに関する。本明細書中で用いる場合、「ストリンジェントな条件」という用語は、少なくとも95%、好ましくは少なくとも97%の同一性が配列間に存在する場合のみ、ハイブリダイゼーションが起こるのであることを意味する。

40

【0063】

発明のポリヌクレオチドアッセイに関して本明細書中でさらに記載するように、例えば本明細書記載の本発明のポリヌクレオチドは、CSGをコードする完全長cDNAおよびゲノムクローンを単離するため、および、これらCSGと高い配列類似性を有する他遺伝子のcDNAおよびゲノムクローンを単離するために、cDNAおよびゲノムDNA用ハイブリダイゼーションプローブとして使用してもよい。かかるプローブは一般に、少なくと

50

も15塩基を含むだろう。好ましくは、かかるプローブは、少なくとも30塩基を含み、少なくとも50塩基を含んでもよい。

【0064】

例えば、本発明のCSGのコード領域は、既知のDNA配列から合成されたオリゴヌクレオチドプローブを用いてスクリーニングすることにより単離してもよい。本発明の遺伝子の配列に相補的な配列を有する標識オリゴヌクレオチドは、どのライブラリーの成員にプローブがハイブリダイズするかを決定するために、ヒトcDNA、ゲノムDNAまたはmRNAのライブラリーをスクリーニングするのに用いられる。

【0065】

本発明のポリヌクレオチドおよびポリペプチドは、特にポリヌクレオチドアッセイに関して、本明細書中にさら記載するように、ヒトの疾患に対する治療薬や診断薬発見のための研究用試薬および材料として用いられ得る。 10

【0066】

ポリヌクレオチドは、成熟タンパク質およびさらなるアミノまたはカルボキシル末端アミノ酸、あるいは成熟ポリペプチド内部のアミノ酸であるポリペプチド（例えば成熟形態が、1つよりも多いポリペプチド鎖を有する場合）をコードしてもよい。かかる配列は特に、前駆体からのタンパク質の成熟形態へのプロセシングの役割を果たしてもよく、タンパク質輸送を促進してもよく、タンパク質の半減期を延長または短縮してもよく、あるいはアッセイまたは生産に関するタンパク質の操作を容易にしてもよい。一般に*in situ*での場合に、さらなるアミノ酸は、細胞酵素により成熟タンパク質から加工される。 20

【0067】

1または2以上のプロ配列に融合したポリペプチドの成熟形態を持つ前駆体タンパク質は、ポリペプチドの不活性体であり得る。プロ配列が取り除かれると、かかる不活性前駆体は一般に活性化される。活性化の前にプロ配列の幾らかまたは全てが取り除かれてもよい。一般に、かかる前駆体は、プロタンパク質と呼ばれる。

【0068】

まとめると、本発明のポリヌクレオチドは、成熟タンパク質、成熟タンパク質にリーダー配列を加えたもの（プレタンパク質と称してもよい）、プレタンパク質のリーダー配列ではない1または2以上のプロ配列を有する成熟タンパク質の前駆体、またはプロタンパク質の前駆体であり、ポリペプチドの活性かつ成熟体を生ずるプロセシング工程中に一般には取り除かれる、1または2以上のプロ配列を有するプレプロタンパク質をコードし得る。 30

【0069】

ポリペプチド

本発明はさらに、CSGポリペプチド、好ましくは配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22のポリヌクレオチドによりコードされるポリペプチドに関する。本発明はまた、これらポリペプチドの断片、類縁体および誘導体にも関する。本発明のポリペプチドを参照する場合、「断片」、「誘導体」および「類縁体」という用語は、かかるポリペプチドと本質的に同一の生物学的機能または活性を保持するポリペプチドを意味する。したがって、類縁体は、プロタンパク質部分の切断により活性化され、活性成熟ポリペプチドを産生することができるプロタンパク質を含む。 40

【0070】

本発明のポリペプチドは、組換えポリペプチド、天然ポリペプチドまたは合成ポリペプチドであってもよい。ある好ましい態様では、それは組換えポリペプチドである。

【0071】

本発明のポリペプチドの断片、誘導体、または類縁体は、(i)1または2以上のアミノ酸残基が保存的あるいは非保存的アミノ酸残基（好ましくは保存的アミノ酸残基）で置換され、かかる置換アミノ酸残基は遺伝的コードによりコードされるものであってもよく、またはそうでなくてもよいもの、(ii)1または2以上のアミノ酸残基が置換基を含む 50

もの、(i i i) 成熟ポリペプチドが別の化合物、例えばポリペプチドの半減期を延長する化合物（例えばポリエチレングリコール）と融合しているもの、または(i v) 成熟ポリペプチドに、リーダーまたは分泌配列あるいは成熟ポリペプチドもしくはプロタンパク質の精製に利用される配列などのさらなるアミノ酸が融合されているものであってもよい。かかる断片、誘導体および類縁体は、本明細書中の教示により、当業者の範囲内であるとみなされる。

【 0 0 7 2 】

好ましい変異体としては、保存的アミノ酸置換によって参照体から変化したものである。かかる置換は、ポリペプチド中の所定のアミノ酸を同様の特性を持つ別のアミノ酸により置換するものである。典型的に保存的置換とし捉えられるのは、脂肪族アミノ酸 A l a、V a l、L e u および I l e 間のあるものから別のものへの置換、ヒドロキシル残基 S e r と T h r の相互交換、酸性残基 A s p と G l u の交換、アミド残基 A s n と G l n 間の置換、塩基性残基 L y s と A r g の交換、ならびに芳香族残基 P h e、T y r 間の置換である。

10

【 0 0 7 3 】

本発明のポリペプチドおよびポリヌクレオチドは、好ましくは単離形態で提供され、好ましくは均一になるまで精製される。

【 0 0 7 4 】

本発明のポリペプチドは、配列番号 1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21 または 22 のポリヌクレオチドによりコードされるポリペプチド（特に、成熟ポリペプチド）ならびに 1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21 または 22 のポリヌクレオチドによりコードされるポリペプチドと少なくとも 75% の類似性（好ましくは少なくとも 75% の同一性）、より好ましくは少なくとも 90% の類似性（好ましくは少なくとも 90% の同一性）、さらに好ましくは少なくとも 95% の類似性（好ましくは少なくとも 95% の同一性）を有するポリペプチドを包含する。また一般に少なくとも 30 個のアミノ酸、そしてより好ましくは少なくとも 50 個のアミノ酸を含むかかるポリペプチドの部分が包含される。

20

【 0 0 7 5 】

当該技術分野にて既知のとおり、2つのポリペプチド間の「類似性」は、アミノ酸配列を比較し、一方のポリペプチド配列のその保存的アミノ酸置換と第2のポリペプチドの置換とを比較することで決定される。

30

本発明のポリペプチドの断片または部分は、ペプチド合成により対応する完全長ポリペプチドを産生するために用いてもよい。したがって、断片は完全長ポリペプチド産生のための中間体として使用してもよい。本発明のポリヌクレオチドの断片または部分は、本発明の完全長ポリヌクレオチドの合成するに使用してもよい。

【 0 0 7 6 】

断片

また本発明のこの側面の好ましい別態様としては、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21 または 22 のポリヌクレオチドによりコードされるポリペプチドの断片含むポリペプチドがある。この観点では、断片は上記 C S G ポリペプチド、およびそれらの変異体または誘導体のアミノ酸配列と部分的には完全に同一であるが、全てが同一ではないアミノ酸配列を有するポリペプチドである。

40

【 0 0 7 7 】

かかる断片は、「遊離体 (f r e e - s t a n d i n g) 」であってもよく、すなわち他のアミノ酸またはポリペプチドの一部でもなく、融合されていなくてもよく、またそれらが大型ポリペプチドの中に含まれてもよく、それらは、その大型ポリペプチドの一部または領域を形成する。大型ポリペプチド内に含まれる場合には、本明細書中に記載の断片は最も好ましくは単一の連続領域を形成する。しかしながら、複数の断片が単一の大型ポリ

50

ペプチド内に含まれてもよい。

【0078】

例えば、ある好ましい態様は、宿主内での発現用に設計された前駆体ポリペプチド内に含まれ、CSG断片のアミノ末端に融合された異種のプレおよびプロポリペプチド領域、および該断片のカルボキシル末端に融合されたさらなる領域を有する本発明のCSGポリペプチドの断片に関する。したがって、本明細書中に意図する意味の一側面における断片は、CSGポリペプチドに由来する融合ポリペプチドまたは融合タンパク質の1または2以上の部分に関する。

【0079】

本発明のポリペプチド断片の代表例として、約15～約139個のアミノ酸を有する上記断片が挙げられ得る。この状況においては、「約」とは、特に引用される範囲、および一方の極値または両極値においていくつか、数個、5、4、3、2または1個のアミノ酸だけ広い、または狭い範囲を含む。この観点において非常に好ましいものは、引用範囲から一極値または両極値が5アミノ酸程度増減する範囲である。特に非常に好ましいものは、引用範囲から一極値または両極値が3アミノ酸程度増減する範囲である。特に好ましいものは、引用範囲から一極値または両極値が1アミノ酸増減する範囲、または、付加または欠失のない引用範囲である。この観点においてすべてのうち最も非常に好ましいのは、約15～約45アミノ酸からなる断片である。

【0080】

本発明の特に好ましい断片としては、CSGポリペプチドの切断突然変異体がある。切断突然変異体としては、アミノ末端を含む連続残基（すなわち連続領域、部分、または一部）またはカルボキシル末端を含む連続残基の欠失、あるいは二重切断突然変異体の場合には、一方がアミノ末端を含み、もう一方がカルボキシル末端を含む2つの連続残基の欠失を除く、配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22のポリヌクレオチドによりコードされるアミノ酸配列を有するCSGポリペプチド、またはそれらの変異体もしくは誘導体が挙げられる。本明細書記載のサイズ範囲を有する断片もまた、切断断片の好ましい態様であり、一般には断片の中で特に好ましい。

【0081】

また本発明のこの側面において好ましいのは、本発明のCSGポリペプチドの構造的または機能的特質を特徴とする断片である。この観点における本発明の好ましい態様としては、本発明のCSGポリペプチドのアルファヘリックスおよびアルファヘリックス形成領域（「アルファ領域」）、ベータシートおよびベータシート形成領域（「ベータ領域」）、回転（turn）および回転形成（turn-forming）領域（「回転領域」）、コイルおよびコイル形成領域（「コイル領域」）、親水性領域、疎水性領域、アルファ両親媒性領域、ベータ両親媒性領域、フレキシブル領域、表面形成領域および高抗原性指数（high antigenic index）領域を含む断片が挙げられる。

【0082】

上記のタイプの領域は、配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22のポリヌクレオチドによりコードされるアミノ酸配列の分析によって定法により同定される。好ましい領域は、ガルニエ-ロブソン（Garnier-Robson）アルファ領域、ベータ領域、回転領域およびコイル領域、チョウ-ファスマン（Chou-Fasman）アルファ領域、ベータ領域および回転領域、カイト-ドリトル（Kyte-Doolittle）親水性領域および疎水性領域、アイゼンベルク（Eisenberg）アルファおよびベータ両親媒性領域、カルプルス-シュルツ（Karpplus-Schulz）フレキシブル領域、エミニ（Emini）表面形成領域およびジャムセン-ウォルフ（Jamson-Wolf）高抗原性指数領域を含む。

【0083】

この観点において非常に好ましい断片としては、上記特徴の幾つかのような複数の構造的

特徴を組み合わせるCSGの領域を含む断片である。この点に関し、CSGポリペプチドの選択された残基により規定される領域は、その全てが回転領域、親水性領域、フレキシブル領域、表面形成領域および高抗原性指数領域に高度に特徴的なアミノ酸構成により特徴付けられており、特に非常に好ましい領域である。

【0084】

かかる領域は、大型ポリペプチド内に含まれてもよく、または上記のような本発明の好ましい断片そのものでもよい。この項で用いられる「約」という用語は、断片に関して一般的に記載した上記の意味を有することが理解されるであろう。

【0085】

さらに好ましい領域は、CSGポリペプチドの活性を媒介する領域である。この観点における最も非常に好ましいものは、同様の活性または改良された活性を有するか、あるいは望ましくない活性が低減された断片を含む、CSGポリペプチドの化学的、生物学的、またはその他の活性を有する断片である。この観点において非常に好ましいのは、関連ポリペプチドの活性領域と配列または位置について、あるいは配列および位置の両方について相同的であり、大腸特異結合タンパク質を含む断片である。これら観点において特に好ましい断片としては、上記の切断突然変異体である。

10

【0086】

本発明はまた、上記断片をコードするポリヌクレオチド、断片をコードするポリヌクレオチドにハイブリダイズするポリヌクレオチド、特にストリンジェントな条件下にてハイブリダイズするもの、および、該断片をコードするポリヌクレオチドを増幅するためのPCRプライマーなどのポリヌクレオチドに関することが理解されるであろう。これらの観点において、好ましいポリヌクレオチドは上記の好ましい断片に対応するものである。

20

【0087】

融合タンパク質

本発明の1つの態様において、本発明のCSGポリペプチドは、好ましくは他のタンパク質に融合される。これらの融合タンパク質は様々な用途に用いることができる。例えば、本発明のポリペプチドをHisタグ、HAタグ、プロテインA、IgGドメインおよびマルトース結合タンパク質に融合すると精製が容易になる(EP A 394, 827, Traunecker et al., Nature 331: 84-86 (1988)も参照)。同様にIgG-1、IgG-3およびアルブミンへの融合は、*in vivo*での半減期を延長する。

30

【0088】

本発明のポリペプチドに融合された核局在シグナルは、特定の細胞内局在性についてタンパク質を標的にすることができ、一方共有結合ヘテロダイマーまたはホモダイマーは融合タンパク質の活性を増加または低下させることができる。融合タンパク質はまた1つより多い機能を持つキメラ分子を作り出すこともできる。最後に、融合タンパク質は、非融合タンパク質に比べ、融合タンパク質の溶解度および/または安定性を高めることができる。これら上記の全てのタイプの融合タンパク質は、よく知られた手順にしたがって作製することができる。

【0089】

例えば、CSGポリペプチドを、以下の手順を介してIgG分子に融合することができる。簡単に述べると、IgG分子のヒトFc部分を、配列の5'および3'末端に及ぶプライマーを用い、PCR増幅する。これらプライマーはまた、発現ベクター、好ましくは哺乳動物発現ベクターへのクローニングを容易にするために好都合な制限酵素部位を有する。

40

【0090】

例えば、pC4(アクセッション番号209646)を用いる場合、ヒトFc部分はBamHIクローニング部位に連結できる。この手順では、3' BamHI部位は破壊されなければならない。次に、ヒトFc部分を含むベクターはBamHIにより再度制限処理され、これによりベクターは直線化され、本発明のCSGポリヌクレオチドがこのBamHI

50

I 部位に連結される。このポリヌクレオチドは終止コドンなしにクローニングされることが好ましい。さもなければ融合タンパク質は産生されない。

【0091】

天然に存在するシグナル配列を用いて分泌タンパク質を産生する場合には、pC4は第2シグナルペプチドを必要としない。あるいは、天然に存在するシグナル配列を使用しない場合には、ベクターは異種シグナル配列を含むように改変することができる（例えば、WO 96/34891参照）。

【0092】

診断アッセイ

本発明はまた、ヒト患者のCSGレベルと、正常ヒト対照でのCSGのレベルとを比較することにより、癌を検出、診断、モニタリング、ステージング、および予後判定する定量的および定性的両方の診断アッセイならびに方法に関する。本発明に関して、CSGレベルが意味するものは、とりわけ、配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22のポリヌクレオチド配列を含む遺伝子により発現される未変性タンパク質である。

10

【0093】

本明細書中で「CSG」とはまた、遺伝コードの縮重に起因して、配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22と比較した場合にヌクレオチド配列に変化を含むが、依然として同一のタンパク質をコードするポリヌクレオチドを意味する。検出される未変性タンパク質は、完全体、分解産物、分子複合体または化学的修飾体であってもよい。

20

【0094】

あるいは、本明細書中で用いられる場合、CSGが意味するものは、配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22のポリヌクレオチド配列によりコードされる未変性mRNA、配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22のポリヌクレオチド配列を含む遺伝子のレベル、またはストリンジェントな条件下にて配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22のアンチセンス配列にハイブリダイズが可能なポリヌクレオチドのレベルである。

30

【0095】

かかるレベルは、好ましくは、正常および異常レベルの決定を含む、細胞、組織および/または体液の少なくとも1つについて測定される。したがって、例えば正常対照体液、細胞、または組織試料に比較したCSGタンパク質の過剰発現を診断するための本発明による診断アッセイは、大腸癌の存在の診断に使用され得る。

【0096】

本発明のすべての方法は、CSGと同様に、任意に他の癌マーカーのレベルを決定することを含んでもよい。CSGの以外の本発明に有用な他の癌マーカーは、試験される癌に依存しており、当業者に既知である。

40

【0097】

本発明は、好ましくは同種類の正常ヒト対照由来の細胞、組織または体液中のCSGレベルと比較した、細胞、組織または体液中のCSGレベルの変化を分析することにより、大腸癌の存在を診断する方法であって、正常ヒト対照に対する患者におけるCSGレベルの増加が、大腸癌の存在に関連する方法を提供する。

【0098】

本発明を限定するものではないが、典型的に、定量的診断アッセイにおいて、試験される患者が癌を有することを示す陽性の結果は、CSGなどの癌マーカーの細胞、組織または体液レベルが、正常ヒト対照の好ましくは同一細胞、組織または体液中により、少なくとも2倍高く、最も好ましくは少なくとも5倍高いものである。

50

【0099】

本発明はまた、転移の発生に関して、いまだ転移していない大腸癌を有する患者において、転移性大腸癌を診断する方法も提供する。本発明の方法では、転移した大腸癌を持つと疑われる（しかしそれまでに転移したことはわかっていない）ヒト癌患者が同定される。このことは当業者に既知の各種手段により達成される。

【0100】

本発明では、細胞、組織または体液中のCSGレベルの存在を決定することは、特に転移をしていない大腸癌と転移をしている大腸癌とを区別するのに有用である。従来技術は、転移をしている大腸癌と転移をしていない大腸癌とを区別することが難しく、適切な処置の選択はかかる知識に依存することが多い。

10

【0101】

本発明では、かかる細胞、組織または体液にて測定される癌マーカーレベルはCSGであり、正常ヒト対照の好ましくは同種類の細胞、組織または体液中のCSGレベルと比較される。すなわち、観察される癌マーカーがまさに血清中のCSGである場合、このレベルは好ましくは、正常ヒト対照の血清中のCSGレベルと比較される。正常ヒト対照に対する患者におけるCSGの増加は、転移を有する大腸癌に関連する。

【0102】

本発明を限定するものではないが、典型的に、定量的診断アッセイにおいて試験またはモニターされる患者における癌が転移していることを示す陽性結果は、CSGなどの癌マーカーの細胞、組織または体液レベルが、正常患者の好ましくは同一細胞、組織または体液中よりも、少なくとも2倍高く、最も好ましくは少なくとも5倍高いものである。

20

【0103】

本明細書中で用いる正常ヒト対照は、癌を有しないヒト患者および/または患者由来の非癌性試料を含み、転移を診断またはモニタリングする方法では、正常ヒト対照はまた、好ましくは、信頼性の高い方法により転移していない大腸癌を有することが決定されているヒト患者由来の試料を含んでもよい。

【0104】

ステージング

本発明はまた、ヒト患者における大腸癌をステージングする方法も提供する。この方法は、かかる癌を持つヒト患者を同定すること、およびかかるヒト患者由来の細胞、組織または体液をCSGについて分析することを含む。次に、患者において決定されたCSGレベルを、正常ヒト対照の好ましくは同種類の細胞、組織または体液中のCSGレベルと比較し、ここで、正常ヒト対照に対するヒト患者のCSGレベルの増加が進行している癌に関連し、CSGレベルの低下（それでも真の正常レベルよりは高い）が、消退または寛解している癌に関連する。

30

【0105】

モニタリング

さらにかかる癌を持つヒト患者において、転移の発生に関して大腸癌をモニタリングする方法を提供する。この方法は、転移がしていることがわかっていないかかる癌を持つヒト患者を同定すること、かかる患者由来の細胞、組織または体液を定期的にCSGレベルについて分析すること、およびヒト患者において決定されたCSGレベルを、正常ヒト対照の好ましくは同種類の細胞、組織または体液中のCSGレベルと比較することを含み、ここで正常ヒト対照に対するヒト患者のCSGレベルの増加が、転移している癌に関連する。この方法では、正常ヒト対照試料にはまた、以前の患者試料も含まれ得る。

40

【0106】

さらに本発明により提供されるのは、大腸癌のステージの変化を、かかる癌を有するヒト患者においてモニタリングする方法である。該方法は、かかる癌を有するヒト患者を同定すること、かかるヒト患者由来の細胞、組織または体液をCSGについて定期的に分析すること、およびヒト患者において決定されたCSGレベルを、正常ヒト対照の好ましくは同種類の細胞、組織または体液中のCSGレベルと比較することを含み、ここで正常ヒト

50

対照に対するヒト患者におけるCSGレベルの増加が、ステージが進行している癌に関連し、CSGレベルの減少が、ステージが後退しているかまたは寛解している癌に関連する。この方法において、正常ヒト対照試料には、以前の患者試料も含まれる得る。

【0107】

転移の発生に関しての患者のモニタリングは定期的であり、好ましくは4半期ごとに実施される。しかしながら、これは、癌、特定患者および癌のステージに依存して、より高頻度で、またはより低頻度で実施されてもよい。

【0108】

予後判定試験および臨床試験モニタリング

本明細書中に記載の方法はさらに、CSGレベル増加に関連した疾患または障害を発症する対象、またはその危険性を有する対象を同定するための予後判定アッセイとしても利用できる。本発明は、試験試料をヒト患者より得て、CSGを検出する方法を提供する。正常ヒト対照に比べ高レベルのCSGの存在は、癌、特に大腸癌を発症する危険性があるヒト患者のための診断である。

10

【0109】

本発明のCSGの発現または活性を低下させる治療薬の有効性は、臨床試験中における、または、ヒト細胞におけるものなどの*in vitro*スクリーニングアッセイにおける、ヒト患者のCSGの発現レベルを分析することによりモニタリングすることもできる。このようにして、遺伝子発現パターンは、ヒト患者、場合によっては細胞の、試験される薬剤に対する生理学的応答を示すマーカーとして利用できる。

20

【0110】

遺伝子病変および突然変異の検出

本発明の方法はまた、CSGにおける遺伝子病変、または突然変異の検出にも使用することができ、これにより遺伝子病変を持つヒトが大腸癌に対する危険性を有するか、または大腸癌を有するかどうかを決定することができる。遺伝子病変は例えば、本発明のCSGからの1または2以上のヌクレオチドの欠失および/または付加および/または置換の存在、CSGの染色体再配列、CSGの異常修飾(例えばゲノムDNAのメチル化パターン)、CSGのmRNA転写物の非野生型スプライシングパターンの存在、CSGの対立遺伝子の欠損、および/またはCSGタンパク質の不適切な翻訳後修飾の存在を確認することにより検出することができる。本発明のCSGにおけるかかる病変を検出する方法は、当業者に既知である。

30

【0111】

例えば、1つの態様においては、本発明のCSGポリヌクレオチドに対応する遺伝子の変化は、所定の表現型(疾患など)を示す全集団または個々の患者からのRNAの単離によって決定される。次に、当該技術分野にて既知のプロトコルを用いて、これらのRNA試料からcDNAが生成される。これらの技法を詳述した多くの実験室マニュアルの代表例である、例えば、Sambrook et al. (MOLECULAR CLONING: A LABORATORY MANUAL, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989))を参照。

40

【0112】

次に、cDNAを、配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22において所定の領域を取り囲むプライマーを用いるPCR用の鋳型として使用する。PCR条件は、典型的には、Sidransky, D., et al., Science 252: 706 (1991)に記載される緩衝溶液を用いて、95 °Cでの30秒間、52~58 °Cでの60~120秒、および70 °Cでの60~120秒の35サイクルから構成される。Sequithe Thermポリメラーゼ(Epicentre Technologies)を用いて、T4ポリヌクレオチドキナーゼで5'末端にて標識したプライマーを使用して、PCR産物を配列決定する。

50

【0113】

選択エキソンのイントロン-エキソン境界もまた決定され、ゲノムPCR産物を分析して、結果を確認する。次に、突然変異が疑われるPCR産物をクローニングして、配列決定して、直接シーケンスの結果を確認する。Holton T. A. and Graham, M. W., *Nucleic Acids Research*, 19: 1156 (1991)に記載されるように、PCR産物をT尾状(T-tailed)ベクターにクローニングし、T7ポリメラーゼ(United States Biochemical)を用いて配列決定する。病気に冒されていない個体に存在しない突然変異により、病気に冒された個体を同定する。

【0114】

ゲノム再配列もまた、ポリヌクレオチドに対応する遺伝子における変化を決定する方法として観察できる。この方法では、ゲノムクローンを、ジゴキシゲニンデオキシ-ウリジン5'-三リン酸塩(Boehringer Mannheim)でニックトランスレーションし、Johnson, C, et al., *Methods Cell Biol.* 35: 73-99 (1991)に記載されるように、FISHを実施する。標識プローブを用いたハイブリダイゼーションは、対応するゲノム座への特異的ハイブリダイゼーションのための、大過剰のヒトDNAを用いて行われる。

【0115】

4, 6-ジアミジノ-2-フェニルインドールおよびヨウ化プロピジウムで染色体を対比染色し、CおよびRバンドの組合せを生じさせる。正確なマッピングのための整列画像(aligned image)は、冷却電荷結合素子カメラ(Photometrics, Tucson, AZ)および可変励起波長フィルターを併合したトリプルバンドフィルターセット(Chroma Technology, Brattleboro, VT)を用いて得られる(Johnson, C, et al., *Genet. Anal. Tech. Appl.*, 8: 75 (1991))。

【0116】

画像収集、分析および染色体分画長測定は、I See Graphicalプログラムシステム(Inovision Corporation, Durham, NC)を用いて行なう。プローブによりハイブリダイズされるゲノム領域の染色体変化は、挿入、欠失および転座として同定される。これらの変化は、関連疾患用診断マーカーとして用いられる。

【0117】

アッセイ技術

患者由来の試料における、本発明のCSGなどの遺伝子の発現レベル(タンパク質レベルを含む)を決定するのに使用可能なアッセイ技術は、当業者に既知である。かかるアッセイ方法としては、ラジオイムノアッセイ、逆転写酵素PCR(RT-PCR)アッセイ、免疫組織化学アッセイ、in situハイブリダイゼーションアッセイ、競合結合アッセイ、ウエスタンブロット分析、ELISAアッセイおよびプロテオミクスアプローチ、二次元ゲル電気泳動法(2D電気泳動)およびマススペクトロメトリーまたはタンパク質相互作用プロファイリングのような非ゲルベースのアプローチが挙げられるが、これらに限定されない。これらの中でも、ELISAは生物学的液体中の遺伝子の発現タンパク質の診断に好まれることが多い。

【0118】

ELISAアッセイはまず、販売元から容易に入手できない場合には、CSGに特異的な抗体、好ましくはモノクローナル抗体を調製することを含む。さらにCSGに特異的に結合するレポーター抗体が一般には調製される。レポーター抗体は、検出可能な試薬、例えば放射性、蛍光または酵素試薬、例えば、西洋ワサビペルオキシダーゼ酵素あるいはアルカリホスファターゼに結合される。

【0119】

ELISAを行なうには、CSGに特異的な抗体を、抗体を結合する固体支持体、例えば

10

20

30

40

50

ポリスチレンディッシュ上でインキュベートする。次に、ディッシュの上のフリーなタンパク質結合部位をウシ血清アルブミンなどの非特異的タンパク質とインキュベートすることにより覆う。次に、分析する試料をこのディッシュにてインキュベートし、その間にCSGはポリスチレンディッシュに結合した特異抗体に結合する。未結合試料は緩衝液により洗い流される。CSGに対し特異的であり、かつ西洋ワサビペルオキシダーゼなどの検出可能試薬に結合しているレポーター抗体をディッシュに載せ、その結果レポーター抗体はCSGに結合している任意のモノクローナル抗体に結合する。

【0120】

次に、未結合のレポーター抗体を洗い流す。次に、比色基質を含むペルオキシダーゼ活性用試薬をディッシュに添加する。CSG抗体に結合した固定化ペルオキシダーゼが着色反応産物を生ずる。所定時間中に発色した色の量は、試料中に存在するCSGタンパク質の量に比例する。定量的な結果は典型的に、標準曲線を参照することで得られる。

10

【0121】

競合アッセイもまた使用されるが、この場合CSGに特異的な抗体は、固体支持体に結合され、標識CSGおよび宿主由来の試料を固体支持体の上を通過させる。固体支持体に結合した検出される標識量を試料中のCSG量に相関させることができる。

【0122】

また、本発明のCSGの核酸配列の全て、または一部をハイブリダイゼーションプローブとして使用して、核酸法を、大腸癌に関するマーカーとしてのCSGのmRNAの検出に使用することもできる。ポリメラーゼ連鎖反応(PCR)およびその他の核酸法、例えばリガーゼ連鎖反応(LCR)および核酸配列ベースの増幅法(NASBA)は、各種悪性腫瘍の診断およびモニタリングのための悪性細胞の検出に使用できる。

20

【0123】

例えば、逆転写酵素PCR(RT-PCR)は、数千の他のmRNA種の複雑な混合物中において特異的なmRNA集団の存在を検出するのに使用可能な強力な技術である。RT-PCRでは、mRNA種はまず、酵素、逆転写酵素を用いて相補的DNA(cDNA)に逆転写され、次にcDNAは標準的PCR反応で増幅される。したがって、RT-PCRは増幅により、単一のmRNA種の存在を示すことができる。したがって、mRNAがそれを産生する細胞に非常に特異的である場合、RT-PCRを用いて、特定のタイプの細胞の存在を同定することができる。

30

【0124】

固体支持体上に配列された(すなわちグリidding)クローンまたはオリゴヌクレオチドへのハイブリダイゼーションは、その遺伝子の発現の検出、およびその発現レベルの定量的両方に使用できる。このアプローチでは、CSG遺伝子をコードするcDNAは、基板上に固定される。この基板はガラス、ニトロセルロース、ナイロンまたはプラスチックを含むが、これらに限定されない任意の適切なタイプであり得る。CSG遺伝子をコードするDNAの少なくとも一部がこの基板に結合され、次にRNAまたはRNAの相補的DNA(cDNA)コピーであり得る、当該組織から単離された分析体とインキュベートされる。

【0125】

基板に結合したDNAと分析体との間のハイブリダイゼーションは、分析体の放射性標識または蛍光標識、またはハイブリッドを検出するために設計した二次分子を含むが、これらに限定されない複数の手段により検出および定量することができる。遺伝子発現レベルの定量は、分析体からのシグナル強度を既知の標準物質から決定されたシグナル強度と比較することで行うことができる。標準物質は標的遺伝子の*in vitro*転写により得ることができ、その収量を定量し、次にその物質を使って標準曲線を作成する。

40

【0126】

プロテオミクスアプローチのうち、2D電気泳動法が当業者に既知の技術である。血清などの試料からの個別のタンパク質の単離は、通常ポリアクリルアミドゲル上において、タンパク質を異なる特性により順次分離することで達成される。まずタンパク質は、電流を

50

用い大きさにより分離される。電流は全てのタンパク質に均一に作用し、その結果小型のタンパク質ほど大型のタンパク質に比べゲル上を遠くへ移動する。

【0127】

第二の次元は、第一の次元に対し垂直方向に電流を作用させ、タンパク質の大きさに基づいてではなく、各タンパク質の持つ特異的電荷により分離する。異なる配列を持つ2つのタンパク質は、大きさと電荷の両方に基づいて同一であることはないので、2D分離の結果は、各タンパク質が特有のスポットを占有する方形のゲルである。このスポットの化学的または抗体プローブによる分析、あるいは続くタンパク質のマイクロシーケンシングにより、試料中の所定のタンパク質の相対量を明らかにすること、および該タンパク質を同定することができる。

10

【0128】

上記試験は、患者より得たホモジネートまたは可溶化組織などの、各種細胞、体液および/または組織抽出物由来の試料について行うことができる。組織抽出物は、組織生検および剖検材料から常法により得られる。本発明に有用な体液には、血液、尿、唾液またはそれらの任意の他の体分泌物または派生物が含まれる。血液とは、全血、血漿、血清または血液の任意の派生物を含むことを意味している。

【0129】

CSGの *in vitro* ターゲティング / 大腸癌治療

このCSGの同定はまた、癌、特に大腸癌のイメージングおよび処置のための新規治療物質の合理的設計において有用である。例えば、一側面では、CSGに特異的に結合する抗体を産生し、大腸癌を患うことが疑われる患者に *in vivo* で使用することができる。CSGを特異的に結合する抗体は、診断および/または治療目的で大腸癌が疑われる患者に注射することができる。したがって、本発明の別の側面は、ヒト患者に対し有効量の抗体を投与することにより、かかる治療を必要としている患者の大腸癌の発症を防止し、治療するための方法に関する。

20

【0130】

「有効量」とは、腫瘍上に発現した標的抗原に結合し、外科的除去のための腫瘍の退縮または腫瘍の消失をもたらすのに必要な抗体量または濃度を意味する。過剰発現されたCSGへの抗体の結合は、かかるCSGを発現する癌細胞に死をもたらすと考えられる。*in vivo* 診断および処置のための抗体の調製および使用は、当該技術分野で既知である。例えば、インジウム-111で標識された抗体-キレーターは、癌胎児抗原発現腫瘍のラジオイムノシンチグラフィイメージングでの使用について記載されている (Sumerdon et al., Nucl. Med. Biol. 1990 17: 247-254)。

30

【0131】

特に、これらの抗体-キレーターは、再発性結腸直腸癌が疑われる患者における腫瘍検出に使用されている (Griffin et al., J. Clin. Onc. 1991 9: 631-640)。磁気共鳴イメージングにおける使用のための、標識として常磁性イオンを有する抗体も記載されている (Lauffer, R. B. Magnetic Resonance in Medicine 1991 22: 339-342)。CSGに対する抗体は、同様に使用することができる。CSGを特異的に結合する標識抗体は、患者の疾患状態の診断またはステージングを目的として、大腸癌が疑われる患者に注射することができる。

40

【0132】

使用される標識は、使用するイメージング様式に応じ選択されるであろう。例えば、インジウム-111、テクネチウム-99mまたはヨウ素-131などの放射性標識体が平面スキャンまたは単光子放射型コンピュータ断層撮影法 (SPECT) に使用できる。フッ素-19などの陽電子放射標識は、陽電子放射断層撮影法に使用できる。ガドリニウム (Gd) またはマンガン (Mn) などの常磁性イオンは、磁気共鳴イメージング (MRI) に使用できる。正常組織のイメージングと比べた標識の存在から、癌の播種を決定する

50

ことができる。器官または組織内の標識量からはまた、その器官または組織中の癌の有無を決定できる。

【0133】

*in vivo*の方法で使用可能な抗体には、ポリクローナル抗体、モノクローナル抗体、およびオムニクローナル抗体、ならびに分子生物学的技術により調製される抗体が含まれる。抗体断片およびアプタマー、ならびにSELEXと呼ばれる、当業者に既知のインビトロ進化プロトコルより得た一本鎖オリゴヌクレオチドも利用できる。

【0134】

スクリーニングアッセイ

本発明はまた、CSGタンパク質に結合するか、またはCSGタンパク質の発現あるいは活性に対する調節効果を持つモジュレータを同定するための方法を提供する。CSGタンパク質の発現または活性を低下させるモジュレータは、大腸癌の処置に有用であると考えられる。かかるスクリーニングアッセイは、当業者に既知であり、細胞ベースアッセイおよび無細胞アッセイが含まれるが、これらに限定されない。

【0135】

コンピュータイメージングによりCSGの領域に特異的に結合することが推測される小分子もまた、大腸癌のイメージングおよび処置での使用のために、設計、合成、および試験することができる。さらに、分子ライブラリーを、本発明で同定されたCSGに対する分子の結合能力について評価することにより、潜在的な抗癌剤をスクリーニングすることができる。ライブラリーにてCSGに結合できるものとして同定された分子は、大腸癌の処置における使用を目的としたさらなる評価のための重要な候補体である。好ましい態様では、これら分子は、細胞中のCSG発現および/または活性を下方調節するだろう。

【0136】

養子免疫治療およびワクチン

癌の養子免疫治療は、細胞が、定着した腫瘍の消退を直接または間接的に媒介することを目的として、抗腫瘍応答性を持つ免疫細胞を、担腫瘍宿主に投与する治療的アプローチである。リンパ球、特にTリンパ球の輸血は、この治療分類に入るものであり、国立癌研究所(NCI)の研究者等は、末梢血リンパ球または腫瘍浸潤リンパ球(TIL)、皮下リンパ節の生検材料由来のT細胞培養物の自己輸血を複数のヒト癌患者の処置に使用している(Rosenberg, S. A., 米国特許第4,690,914号、1987年9月1日発行; Rosenberg, S. A., et al., 1988, N. Engl. J. Med. 319: 1676-1680)。

【0137】

本発明は、熱ショックタンパク質(hsp)の非共役複合体ありまたはなしで、抗原性CSG分子に対し感作されたマクロファージを用いた、原発性および転移性大腸癌の防止および/または処置のための養子免疫治療の組成物および方法に関する。CSGの抗原性または免疫原性は、CSGタンパク質またはその断片の抗体を産生する能力、またはナイーブなエフェクター細胞を教育する能力(エフェクター細胞は、次に抗原(またはエピトープ)発現標的細胞を溶解する)から容易に確認できる。

【0138】

癌細胞は、定義によれば、異常であり、正常組織には存在しないので、免疫系により外来物質として認識されるべきタンパク質を含む。しかしながら、免疫系はしばしばこの異常に気づかず腫瘍の攻撃をしないようである。癌細胞により産生される外来性CSGは、それらの存在を示すのに使用できる。CSGは腫瘍抗原と呼ばれる短い断片に分解され、細胞表面に提示される。

【0139】

これらの腫瘍抗原はクラスIおよびクラスIIの2つの型が存在するMHCと呼ばれる分子により、細胞表面に保持または提示される。MHCクラスI分子と会合する腫瘍抗原は、細胞障害性T細胞によって認識されるが、一方抗原-MHCクラスII複合体はヘルパー細胞と呼ばれる第2のT細胞サブセットにより認識される。これらの細胞はサイトカイ

10

20

30

40

50

ンを分泌し、これが腫瘍の増殖を遅らせるか、または停止させ、別の型の白血球、B細胞が腫瘍細胞に対する抗体を作るのを助ける。

【0140】

養子免疫治療では、T細胞またはその他の抗原提示細胞（APC）は、腫瘍特異的CSG抗原を用いて、体外（*ex vivo*）で刺激される。次に、刺激された細胞は、患者に再注入され、そこでそれらは癌細胞を攻撃する。研究から、細胞障害性およびヘルパーT細胞の両方を使用すると、いずれか一方のサブセットのみを使用した場合に比べはるかに効果的であることが示されている。さらに、CSG抗原は、米国特許第5,985,270号に記載されているように、熱ショックタンパク質と複合体を形成し、APCを刺激し得る。

10

【0141】

APCは、マクロファージ、樹状細胞、Bリンパ球およびそれらの組合せを含むが、これらに限定されない当該技術分野で既知の抗原提示細胞から選択でき、好ましくはマクロファージである。細胞がその個体に対して自己のものである好ましい使用では、養子移植のための免疫細胞のドナー選択の問題を回避するために、リンパ球、マクロファージ、またはその他のAPC等の自己の免疫細胞が用いられる。自己の免疫細胞の使用により回避できる他の問題は、処置が効を奏さない場合は致命的となる移植片対宿主病である。

【0142】

遺伝子治療を用いた養子免疫治療では、CSGのDNAを従来の遺伝子治療と同様にしてエフェクター細胞内に導入することができる。これによりエフェクター細胞が抗原性タンパク質を産生するように操作されると、腫瘍細胞に対するこれら細胞の細胞障害性が高められ、結果として養子免疫治療が改善される。

20

【0143】

本発明のCSG抗原はまた、大腸癌ワクチン構成成分としても有用である。ワクチンは、免疫刺激量のCSG抗原を含む。免疫刺激量とは、大腸癌の改善または処置を目的としてレシピエント内に所望の免疫応答を惹起することができる抗原量を表す。有効量は、当業者に既知の標準的方法により経験的に決定され得る。

【0144】

CSG抗原は、所望の型の免疫応答、例えば抗体および/または細胞媒介性の応答を誘発するために設計された多くのワクチン配合物のいずれか一つとして提供され得る。かかる配合物は、当該技術分野で既知であり、例えば米国特許第5,585,103号に記載されている配合物を含むが、これに限定されるものではない。免疫応答を刺激するのに使用される本発明のワクチン配合物はまた、薬学的に許容され得るアジュバントも含むことができる。

30

【0145】

ベクター、宿主細胞、発現

本発明は、本発明のポリヌクレオチドを含むベクター、本発明のベクターで遺伝的に操作された宿主細胞、および組換え技術による本発明のポリペプチドの産生にも関する。

【0146】

宿主細胞は、CSGポリヌクレオチドを組み込み、本発明のCSGポリペプチドを発現するように遺伝的に操作することができる。例えば、CSGポリヌクレオチドは、感染、形質導入、トランスフェクション、トランスベクション（*transvection*）および形質転換の周知技術を用いて、宿主細胞内に導入され得る。CSGポリヌクレオチドは、単独または他のポリヌクレオチドとともに導入されてもよい。かかる他のポリヌクレオチドは、単独で導入されてもよく、共導入されてもよく、または本発明のCSGポリヌクレオチドと結合して導入されてもよい。

40

【0147】

例えば、本発明のCSGポリヌクレオチドは、例えば哺乳動物細胞における、コトランスフェクションおよび選択に関する標準的技術を用いて、選択可能マーカートをコードする他の分離したポリヌクレオチドとともに、宿主細胞内にトランスフェクトされてもよい。こ

50

の場合、ポリヌクレオチドは一般に、宿主細胞ゲノム内に安定的に組み込まれるであろう。

【0148】

あるいは、CSGポリヌクレオチドは、宿主内での増殖用の選択可能マーカールを含むベクターに結合されてもよい。ベクター構築物は、上記技術により宿主細胞内に導入されてもよい。一般に、プラスミドベクターは、リン酸カルシウム沈殿物などの沈殿物中に、または荷電脂質との複合体中に、DNAとして導入される。エレクトロポレーションもまた宿主内へのCSGポリヌクレオチド導入に利用され得る。

【0149】

ベクターがウイルスの場合には、それは*in vitro*でパッケージングされてもよく、またはパッケージング細胞内に導入され、パッケージングされたウイルスが細胞内に形質導入されてもよい。当業者により日常的に行なわれている、よく知られた広範な技術が、本発明のこの側面にしたがってCSGポリヌクレオチドを作製し、およびCSGポリヌクレオチドを細胞内に導入するのに好適である。かかる技術は、本明細書中に既述したSambrook et al.などの参考文献に十分に概説されている。

【0150】

本発明で用いてもよいベクターは、例えばプラスミドベクター、一本鎖または二本鎖ファージベクター、一本鎖または二本鎖RNAあるいはDNAウイルスベクターを含む。かかるベクターは、細胞内へのDNAおよびRNA導入に関する既知の技術により、細胞内にポリヌクレオチドとして、好ましくはDNAとして導入され得る。ベクターは、ファージおよびウイルスベクターの場合、好ましくは、感染および形質導入に関する既知の技術によりパッケージされたウイルス、またはキャプシド形成されたウイルスとして細胞内に導入され得る。ウイルスベクターは、複製性または複製欠損性であり得る。後者の場合、ウイルスの増殖は一般に、補完宿主細胞(*complementing host cell*)内でのみ生ずる。

【0151】

本発明のポリヌクレオチド及びポリペプチドの発現のための好ましいベクターは、発現されるべきポリヌクレオチドに作動可能に連結された、宿主細胞での発現に効果的なシス作用性の制御領域を含むベクターを包含するが、これに限定されない。適切なトランス作用性因子は、宿主により供給されるか、補完ベクター(*complementing vector*)により供給されるか、または宿主導入時にベクター自体により供給される。

【0152】

この観点でのある好ましい態様において、ベクターは特異的発現を提供する。かかる特異的発現は、誘発性発現またはある型の細胞でのみの発現、もしくは誘発性および細胞特異的発現の両方であり得る。誘発性ベクターのうち特に好ましいのは、温度および栄養添加物といった操作が容易である環境因子により発現が誘発できるベクターである。原核生物および真核生物宿主での使用のための構成的および誘発的発現ベクターを含めた本発明のこの態様に適切な各種ベクターが既知であり、当業者により常法として使用されている。

【0153】

操作された宿主細胞は、通常の栄養培地中で培養でき、培地はとりわけプロモーターを活性化すること、形質転換体を選択すること、または遺伝子を増幅することに関して適切に改良してもよい。発現のために選択された宿主細胞に従来使用されている温度、pH等の培養条件は、一般的に本発明のCSGポリペプチドの発現に適切であろう。

【0154】

本発明のCSGポリペプチドの発現には非常に多様な発現ベクターが使用可能である。かかるベクターには、染色体、エピソームおよびウイルス由来ベクターを含む。ベクターは、バクテリアプラスミド、バクテリオファージ、酵母エピソーム、酵母染色体要素、パキウイルス、SV40などのパポバウイルス、ワクシニアウイルス、アデノウイルス、鶏痘ウイルス、仮性狂犬病ウイルスおよびレトロウイルスなどのウイルス、ならびに、それらの組合せ、例えば、コスミドおよびファージミドなどの、プラスミドおよびバクテ

リオファージ遺伝的要素に由来するものなどに由来してもよい。これらは全て本発明のこの側面による発現に使用され得る。一般に、宿主内でポリペプチドを発現させるためのポリヌクレオチドの維持、増殖、または発現に適切な任意のベクターをこの観点での発現に使用してもよい。

【0155】

適切なDNA配列は、各種既知の、かつ日常的な技術のいずれかによりベクター内に挿入され得る。一般に発現のためのDNA配列は、DNA配列および発現ベクターを1または2以上の制限エンドヌクレアーゼで切断し、次にT4DNAリガーゼを使い制限断片を一つに結合させることで、発現ベクターに結合される。この目的に使用できる制限および結合の手法は当業者にとって既知かつ常法である。この観点に適切な手法、および別の方法を用い発現ベクターを構築する手法は、当業者にとって既知かつ常法であり、本明細書の他所に引用したSambrook et al. に詳細記載されている。

10

【0156】

発現ベクター内のDNA配列は、例えばmRNA転写に対するプロモーターを含む適切な発現制御配列に作動可能に連結されている。代表的なプロモーターは、よく知られた少数のプロモーターを挙げると、ファージラムダPLプロモーター、大腸菌lac、trpおよびtacプロモーター、SV40初期および後期プロモーター、ならびにレトロウイルスLTRのプロモーターを含む。本発明のこの側面での使用に適切である多くの記載外のプロモーターが既知であり、本発明での議論および実施例に例示される様式によって当業者により容易に使用され得ることが理解されるだろう。

20

【0157】

一般に発現構築体物は転写の開始および終止に関する部位、および転写領域内には翻訳のためのリボソーム結合部位を含む。構築物により発現される成熟転写体のコード部分は、開始部位に翻訳を開始するAUGを、および翻訳されるポリペプチドの末端に適切に位置する終止コドンを含む。

さらに、構築物は、発現を制御しかつ発現をもたらす制御領域を含んでもよい。一般には、多くの一般的に実施されている手法によれば、かかる領域は、とりわけリプレッサー結合部位やエンハンサーのように、転写を制御することで作動するであろう。

【0158】

増殖および発現用のベクターは、一般には選択マーカを含むだろう。かかるマーカはまた増幅にも適切であり、あるいはベクターはこの目的のためのさらなるマーカを含んでもよい。この観点では、好ましくは発現ベクターは1または2以上の選択マーカ遺伝子を含み、形質転換した宿主細胞の選択に関する表現型の特徴を提供する。好ましいマーカは、真核生物細胞培養についてはジヒドロ葉酸還元酵素あるいはネオマイシン耐性を、大腸菌やその他の細菌の培養についてテトラサイクリンあるいはアンピシリン耐性遺伝子を含む。

30

【0159】

本明細書の他所に記載のような適切なDNA配列、ならびに適切なプロモーター、およびその他の適切な制御配列を含むベクターは、その中での所望ポリペプチドの発現に適切である各種既知の技術を用い、適切な宿主内に導入され得る。

40

【0160】

適切な宿主の代表例は、大腸菌、放線菌およびネズミチフス菌細胞などの細菌細胞、酵母細胞などの真菌細胞、ショウジョウバエのS2およびスポドプテラ(Spodoptera)のsf9細胞などの昆虫細胞、CHO、COSおよびボーズ黒色腫細胞などの動物細胞、ならびに植物細胞を含む。非常に多様な発現構築物のための宿主がよく知られており、当業者は本開示により、本発明のこの側面に従って、CSGポリペプチドの発現に適した宿主を容易に選択することができるだろう。

【0161】

より具体的には本発明はまた、発現構築物などの、1または2以上の上記配列を含む組換え構築物も包含する。構築物には、本発明のかかるCSG配列が挿入されている、プラス

50

ミドまたはウイルスベクターなどのベクターが含まれる。配列は前向き、または逆向きに挿入され得る。この観点におけるある好ましい態様では、構築物は、例えばプロモーターなどの、配列に作動可能に連結された制御配列をさらに含む。多数の適切なベクターおよびプロモーターが当業者に知られており、本発明での使用に適切なベクターが多く市販されている。

【0162】

以下に、市販のベクターを例示する。細菌での使用にとりわけ好ましいベクターは、Q i a g e n より入手可能な p Q E 7 0、p Q E 6 0 および p Q E - 9、S t r a t a g e n e より入手可能な p B S ベクター、ファージスクリプト (P h a g e s c r i p t) ベクター、ブルースクリプトベクター、p N H 8 A、p N H 1 6 A、p N H 1 8 A、p N H 4 6 A、および P h a r m a c i a より入手可能な p t r c 9 9 a、p K K 2 2 3 - 3、p K K 2 3 3 - 3、p D R 5 4 0、p R I T 5 である。とりわけ好ましい真核細胞ベクターは、S t r a t a g e n e より入手可能な P W L N E O、p S V 2 C A T、p O G 4 4、p X T 1 および p S G、および P h a r m a c i a より入手可能な p S K V 3、p B P V、p M S G および p S V L である。

10

【0163】

これらベクターは、本発明のこの側面での使用に関し、当業者が利用可能な多くの市販されている既知のベクターから例示することのみを目的とし掲載されている。本開示を読むことにより、宿主内での、本発明の C S G ポリヌクレオチドまたはポリペプチドの導入、維持、増殖および/または発現に適切な任意のその他のプラスミドまたはベクターが、本発明のこの側面に使用できることが当業者に理解されるであろう。

20

【0164】

プロモーター領域は、候補プロモーター断片、すなわちプロモーターを含み得る断片を導入するための、1または2以上の部位の下流の、クロラムフェニコールアセチルトランスフェラーゼ (「 c a t 」) 転写単位などの、プロモーター領域を欠くレポーター転写単位を含むベクターを用いて、所望遺伝子より選択することができる。

【0165】

周知のように、c a t 遺伝子上流にある制限部位にて、プロモーターを含む断片をベクターに導入すると、標準的な C A T アッセイにより検出することができる C A T 活性が発生する。この目的に適切なベクターは既知であり、容易に入手できる。かかる2種類のベクターは p K K 2 3 2 - 8 および p C M 7 である。すなわち、本発明の C S G ポリヌクレオチドの発現のためのプロモーターには、既知、かつ容易に入手可能なプロモーターだけでなく、レポーター遺伝子を使用した上記技術により容易に得ることができるプロモーターも含まれる。

30

【0166】

とりわけ知られている、本発明によるポリヌクレオチドおよびポリペプチドの発現に適切な細菌プロモーターは、大腸菌 l a c i と l a c Z およびプロモーター、T 3 および T 7 プロモーター、g p t プロモーター、ラムダ P R、P L プロモーター、および t r p プロモーターである。本観点に適切な、とりわけ知られている真核生物プロモーターは、C M V 最初期プロモーター、H S V チミジンキナーゼプロモーター、初期および後期 S V 4 0 プロモーター、ラウス肉腫ウイルス (「 R S V 」) のものなどの、レトロウイルス L T R プロモーター、およびマウスメタロチオネイン - I プロモーターなどのメタロチオネインプロモーターである。

40

【0167】

宿主細胞での発現に適切であるベクターおよびプロモーターを選択することは既知の手法であり、発現ベクターの構築、宿主内へのベクターの導入および宿主での発現に必要な技術は、当該技術分野で日常の技術である。

本発明はまた、上記構築物を含む宿主細胞にも関する。宿主細胞は哺乳動物細胞などの高等真核生物細胞、または酵母細胞などの下等真核生物細胞である。あるいは、宿主細胞は細菌細胞などの原核細胞である。

50

【0168】

宿主細胞内への構築物の導入は、リン酸カルシウムトランスフェクション、DEAE-デキストラン媒介性トランスフェクション、陽イオン性脂質媒介性トランスフェクション、エレクトロポレーション、形質導入、感染またはその他の方法により行うことができる。かかる方法は、Davis et al., BASIC METHODS IN MOLECULAR BIOLOGY, (1986)などの多くの標準的研究室マニュアルに記載されている。

宿主細胞内の構築物は従来の様式で用いられ、組換え配列にコードされる遺伝子産物を産生できる。あるいは、本発明のCSGポリペプチドは従来のペプチド合成装置により合成的に作製することができる。

10

【0169】

成熟タンパク質は、適切なプロモーターの制御下において、哺乳動物細胞、酵母、細菌またはその他の細胞中で発現させることができる。無細胞翻訳系は、本発明のDNA構築物に由来するRNAを利用することで、かかるタンパク質の産生に利用することができる。原核生物および真核生物宿主とともに使用することに適したクローニングおよび発現ベクターは、本明細書の他所で引用されたSambrook et al.に記載されている。

【0170】

一般に、組換え発現ベクターは複製起点、下流の構造配列の転写を指示するための高発現遺伝子由来プロモーター、およびベクター暴露後にベクター含有細胞を単離できるようにする選択マーカを含むだろう。特に適切なプロモーターは、3-ホスホグリセレートキナーゼ(「PGK」)などの糖分解酵素、 α -因子、酸性ホスファターゼ、および熱ショックタンパク質をコードする遺伝子に由来するプロモーターである。選択マーカは、大腸菌のアンピシリン耐性遺伝子およびS. cerevisiaeのtrp1遺伝子を含む。

20

【0171】

高等真核生物による本発明のCSGポリペプチドをコードするDNAの転写は、ベクター内にエンハンサー配列を挿入することで高められる。エンハンサーはDNAのシス作用性要素であり、通常約10~300塩基対(bp)であり、所定の宿主細胞タイプにおいてプロモーターの転写活性を高める。エンハンサーの例は、bp100~270にある複製起点の後方に位置するSV40エンハンサー、サイトメガロウイルス初期プロモーターエンハンサー、複製起点の後方にあるポリオーマエンハンサーおよびアデノウイルスエンハンサーを含む。

30

【0172】

本発明のCSGポリペプチドの異種構造配列をコードする本発明のポリヌクレオチドは、標準的技術を用いることで、発現のためのプロモーターと作動可能に連結されるようにベクター内に挿入される。ポリヌクレオチドは、転写開始部位がリボソーム結合部位に対し適切に5'側に位置するように配置される。

【0173】

リボソーム結合部位は発現されるべきポリペプチドの翻訳を開始するAUGに対して5'側になる。一般に、通常AUGである1つの開始コドンから開始し、リボソーム結合部位と開始AUGの間に位置するような他のオープンリーディングフレームは存在しない。また一般に、ポリペプチドの末端には翻訳停止コドンがあり、さらに転写領域の3'末端に適切に配置されたポリアデニル化シグナルおよび転写終止シグナルが存在するだろう。

40

【0174】

小胞体内腔内、細胞周辺腔、または細胞外環境に翻訳タンパク質を分泌するために、発現ポリペプチド中に適切な分泌シグナルを組み込んでもよい。シグナルはポリペプチドに対し同種でも、あるいは異種シグナルでもよい。

【0175】

ポリペプチドは融合タンパク質などの修飾形態として発現されてもよく、分泌シグナルだけでなくさらなる異種機能領域を含んでもよい。すなわち、例えば宿主細胞内、精製中、

50

あるいはその後の取り扱いや保存中における安定性および持久性を改善するために、さらなるアミノ酸領域、特に荷電アミノ酸をポリペプチドのN末端に付加してもよい。

【0176】

精製を容易にするためにポリペプチドに領域を付加してもよい。かかる領域は、ポリペプチドの最終調製前に取り除かれるだろう。分泌または排出を起こさせるために、あるいは安定性を向上させ、とりわけ精製を容易にするために、ポリペプチドにペプチド部分を加えることは、当該分野において一般的であり、日常的な技術である。

【0177】

本発明によるCSGポリヌクレオチドおよびポリペプチドの増殖、維持または発現に適切な原核動物宿主は、大腸菌 (*Escherichia coli*)、枯草菌 (*Bacillus subtilis*) やネズミサルモネラ菌 (*Salmonella typhimurium*) を含む。シュードモナス属 (*Pseudomonas*)、放線菌 (*Streptomyces*) およびブドウ球菌 (*Staphylococcus*) の様々な種はこの観点における適切な宿主である。さらに、当業者に既知のその他の多くの宿主も、この観点において使用できるだろう。

10

【0178】

代表的であるが、非限定的な例としては、細菌での応用に有用な発現ベクターは、選択マーカーおよび既知のクロニングベクターであるpBR322の遺伝要素を含む市販プラスミド由来の細菌性の複製起点を含むことができる。かかる市販ベクターは、例えばpKK223-3 (*Pharmacia Fine Chemicals, Uppsala, Sweden*) およびGEM1 (*Promega Biotec, Madison, Wisconsin, USA*) を含む。

20

【0179】

これらのpBR322「主鎖 (backbone)」部分は適切なプロモーターおよび発現されるべき構造配列と合わせられる。適切な宿主株を形質転換し、宿主株を選択したプロモーターが誘発可能な適切な細胞密度まで増殖させた後に、適切な手段でそれを誘発し (例えば温度変化、または化学的誘発物質への暴露)、細胞はさらなる期間培養される。

【0180】

次に、細胞は一般には、遠心分離により集められ、物理的または化学的手段により破壊され、そして得られた粗製抽出物はさらなる精製のために保持される。タンパク質の発現に用いた微生物細胞は凍結融解サイクル、超音波処理、機械的破壊、または細胞溶解剤の使用を含む任意の従来の方法により破壊することができ、かかる方法は当業者に既知である。

30

【0181】

同様に、発現には様々な哺乳動物細胞培養系が使用できる。典型的な哺乳動物発現系は、Gluzman et al., *Cell* 23: 175 (1981) に記載されている、サル腎臓線維芽細胞のCOS-7細胞系である。適合するベクターを発現させることができるその他の哺乳動物細胞系は、例えばC127、3T3、CHO、HeLa、ヒト腎臓293およびBHK細胞系を含む。

【0182】

哺乳動物発現ベクターは複製起点、適切なプロモーターおよびエンハンサー、さらに発現に必要な任意のリボソーム結合部位、ポリアデニル化部位、スプライスドナーおよびアクセプター部位、転写終止配列、および5'フランキング非転写配列を含むだろう。この観点におけるある好ましい態様では、SV40スプライス部位に由来するDNA配列およびSV40ポリアデニル化部位がこれらのタイプの必須非転写遺伝要素として利用される。

40

【0183】

CSGポリペプチドは、硫酸アンモニウムまたはエタノール沈殿、酸抽出、陰イオンまたは陽イオン交換クロマトグラフィー、ホスホセルロースクロマトグラフィー、疎水相互作用クロマトグラフィー、アフィニティークロマトグラフィー、ヒドロキシアパタイトクロマトグラフィーおよびレクチンクロマトグラフィーを含む、既知の方法により組換え細

50

胞培養物より回収および精製され得る。最も好ましくは、精製には高速液体クロマトグラフィ（「HPLC」）が用いられる。単離または精製中にポリペプチドが変性させられた場合には、活性立体構造を再生するためにタンパク質のリフォールディングに関する既知の技術が利用され得る。

【0184】

本発明のCSGポリペプチドは、天然精製産物、化学合成産物、および例えば細菌、酵母、高等植物、昆虫および哺乳動物を含む原核生物または真核生物宿主から組換え技術により生産された産物を含む。組換え体の産生手法に用いた宿主に応じて、本発明のCSGポリペプチドはグリコシル化されてもよく、あるいはグリコシル化されなくてもよい。さらに、本発明のCSGポリペプチドはさらに幾つかの例では宿主媒介性プロセスの結果としての初期修飾メチオニン残基を含んでもよい。

10

【0185】

CSGポリヌクレオチドおよびポリペプチドは本発明によれば、各種用途、特にCSGの化学的および生物学的な特性を利用する用途に使用され得る。さらなる用途は、細胞、組織および生物体の障害の診断および処置に関する。本発明のこれらの側面は、以下の議論によってさらに説明される。

【0186】

ポリヌクレオチドアッセイ

上記においていくらか詳細に述べたとおり、本発明はまた、例えば診断薬としての相補的ポリヌクレオチドを検出することを目的としたCSGポリヌクレオチドの使用に関する。機能不全に関連したCSGの突然変異体の検出は、CSGの過少発現、過剰発現または発現の変化の結果生ずる疾患または疾患に対する感受性、例えば遺伝性大腸癌に対する感受性の診断を追加または決定することができる診断ツールを提供するだろう。

20

【0187】

ヒトCSG遺伝子に突然変異を有する個体は、各種技術によりDNAレベルで検出されるだろう。診断用の核酸は、患者細胞から、例えば、血液、尿、唾液、組織生検および剖検材料から得られる。ゲノムDNAは検出に直接用いてもよく、または分析前にPCRを用い酵素的に増幅してもよい(Saikiet al., Nature, 324: 163-166 (1986))。RNAまたはcDNAもまた同様に使用される。

【0188】

例えば、配列番号1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21または22のCSGポリヌクレオチドに相補的なPCRプライマーは、CSG発現および突然変異の同定および分析に使用できる。例えば欠失および挿入は、正常遺伝子型と比較したときの増幅産物のサイズ変化により検出できる。点突然変異は、増幅DNAを放射線標識したCSG RNAに、または代替的に、放射線標識したCSGアンチセンスDNA配列にハイブリダイズすることで同定できる。完全一致配列は、RNase A消化または融解温度の差により、不一致二本鎖と識別することができる。

30

【0189】

参照遺伝子と突然変異を有する遺伝子との配列の差はまた、直接DNA配列分析からも示される。さらにクローニングしたDNA断片をプローブとして使用して、特定のDNA断片を検出してもよい。かかる方法の感度は、PCRまたはその他の増幅法を適切に利用することで大きく高めることができる。例えば、配列分析用プライマーは、変法PCRにより作成された二本鎖PCR産物または一本鎖鋳型分子とともに用いられる。配列決定は、放射線標識ヌクレオチドを用いた従来の方法、または蛍光タグを用いた自動配列決定方法によって実施される。

40

【0190】

DNA配列の差に基づく遺伝子検査は、変性剤が存在する、または存在しないゲル中でのDNA断片の電気泳動移動度の変化を検出することで達成され得る。小配列の欠失および挿入は、高解像度ゲル電気泳動法により視覚化することができる。異なる配列のDNA断

50

片は変性ホルムアミド勾配ゲル上で識別されるが、その場合、異なるDNA断片の移動度はそれらの特異的融解温度または部分融解温度に従ってゲル中の異なる点で遅くなる（例えば、Myers et al., Science, 230: 1242 (1985) 参照）。

【0191】

特異的位置における配列の変化はまた、RNaseおよびS1保護または化学的切断法（例えばCotton et al., Proc. Natl. Acad. Sci., USA, 85: 4397-4401 (1985)）などのヌクレアーゼ保護アッセイによっても明らかとなり得る。

したがって、特異的DNA配列の検出は、ハイブリダイゼーション、RNase保護、化学的切断、直接DNA配列分析または制限酵素の使用（例えば制限酵素長多型（「RFLP」））およびゲノムDNAのサザンブロットティングといった方法により達成できる。従来型のゲル電気泳動法およびDNA配列分析法に加えて、突然変異は、in situ分析によっても検出できる。

【0192】

染色体アッセイ

本発明のCSG配列はさらに、染色体同定にも有用である。染色体上の特定部位を同定することが求められており、実際の配列データに基づいた（反復多型）少数の染色体マーキング試薬が、現在染色体位置をマーキングするために利用できる。本発明の各CSG配列は、個々のヒト染色体上の特定位置に特異的に標的化され、それとハイブリダイズすることができる。したがって、CSGは、配列を疾患関連遺伝子と関連させる上での重要な第1段階である、DNAの染色体へのマッピングに用いることができる。

【0193】

この観点に関するある好ましい態様では、本明細書中に開示されたcDNAは本発明のCSGのゲノムDNAをクローニングするのに用いられる。これは、各種既知の技術と、一般に市販されているライブラリーを用いて達成できる。このゲノムDNAは、その目的のための既知の技術を使用したin situ染色体マッピングに用いられる。

【0194】

幾つかの例では、配列は、PCRプライマー（好ましくは15~25bp）をcDNAから調製することで、染色体にマッピングできる。遺伝子の3'非翻訳領域のコンピューター分析を使用することで、ゲノムDNA中の1つより多いエクソンに及ばないプライマーが迅速に選択されて、すなわち増幅工程を複雑にする。次にこれらのプライマーは、個々のヒト染色体を含む体細胞ハイブリッドのPCRスクリーニングに用いられる。プライマーに対応したヒト遺伝子を含むハイブリッドだけが増幅断片を生ずる。

【0195】

体細胞ハイブリッドのPCRマッピングは、特定DNAを特定の染色体に割りふる迅速な手法である。同一オリゴヌクレオチドプライマーとともに本発明を使用すれば、特定の染色体またはより大型ゲノムクローンのプールに由来する断片のパネルを使用し、同様にサブローカリゼーションを行うことができる。その染色体へのマッピングに同様に使用できるその他のマッピング戦略は、in situハイブリダイゼーション、標識フロー選別染色体によるプレスクリーニング、および染色体特異的cDNAライブラリーの構築のためのハイブリダイゼーションによる前選別を含む。

【0196】

展開した中期染色体へのcDNAクローンの蛍光in situハイブリダイゼーション（「FISH」）は、1工程で正確な染色体位置を提供するために使用され得る。この技術は50または60bpといった短いcDNAで用いることができる。この技術は、Verma et al. (HUMAN CHROMOSOMES: A MANUAL OF BASIC TECHNIQUES, Pergamon Press, New York (1988)) によって記載されている。

【0197】

いったん配列を染色体位置に正確にマッピングできれば、染色体における配列の物理的位置と遺伝地図データとを関連付けることができる。かかるデータは、例えばジョンズホプキンス (Johns Hopkins) 大学、ウエルチ (Welch) 医学図書館よりオンラインで入手可能な V. McKusick, MENDELIAN INHERITANCE IN MANに見いだされる。次に、連鎖分析 (物理的に近接する遺伝子の随伴遺伝) により、同一の染色体領域にマッピングされた遺伝子と疾患との間の関連性が同定される。

【0198】

次に、病気に罹った個体と罹っていない個体との間の cDNA またはゲノム配列の差を決定する必要がある。もし変異が病気にかかった個体の一部または全部に観察されるが、いずれの正常個体にも観察されない場合、その変異はおそらく病気の原因物質である可能性がある。

10

【0199】

現時点の物理的マッピング技術および遺伝子マッピング技術の解像度では、疾患関連染色体領域に正確に位置決定される cDNA は、50 ないし 500 の潜在的原因遺伝子のうちの 1 個であろう。(これは 1 メガ塩基マッピングの解像度、および 20 kb 当たり 1 つの遺伝子と仮定)。

【0200】

ポリペプチドアッセイ

上記でいくらか詳細に記載したとおり、本発明はまた、正常および異常レベルの決定を含む、細胞および組織、および例えば血液や尿などの生物学的液体中の CSG ポリペプチドのレベルを検出する定量および診断アッセイなどの診断アッセイに関する。したがって、例えば正常組織試料と比較した CSG ポリペプチドの過剰発現または過少発現を検出する本発明による診断アッセイは、例えば新生物の存在の検出に用いられる。

20

【0201】

宿主由来試料中の、本発明の CSG ポリペプチドなどのタンパク質のレベルの決定に使用できるアッセイ技術は、当業者に既知である。かかるアッセイ法は、ラジオイムノアッセイ、競合結合アッセイ、ウエスタンブロット分析および ELISA アッセイを含む。とりわけ ELISA が好ましい場合が多い。

【0202】

例えば、抗体 - サンドイッチ ELISA を用いて、試料中、好ましくは生物試料中のポリペプチドを検出する。マイクロタイタープレートのウェルを、最終濃度 0.2 ~ 10 $\mu\text{g}/\text{ml}$ で、特異抗体により被覆する。抗体は、モノクローナルまたはポリクローナルであり、本明細書記載の方法により産生される。ウェルへのポリペプチドの非特異的結合が低減するように、ウェルをブロックする。次に、被覆したウェルを、CSG ポリペプチドを含有する試料とともに、室温にて 2 時間より長い時間インキュベートする。

30

【0203】

好ましくは、試料の連続希釈物を用いて、結果を確認すべきである。次に、プレートを脱イオン水または蒸留水で 3 回洗浄して、未結合のポリペプチドを除去する。次に、25 ~ 400 ng の濃度の特異抗体 - アルカリホスファターゼ複合体 50 μl を添加して、室温にて 2 時間インキュベートする。再びプレートを脱イオン水または蒸留水で 3 回洗浄し、未結合の複合体を除去する。

40

【0204】

次に、各ウェルに、4 - メチルウンベリフェリルホスフェート (MUP) または p - ニトロフェニルホスフェート (NPP) 基質溶液 (75 μl) を添加し、プレートを室温にて 1 時間インキュベートする。マイクロタイタープレートリーダーにより、反応を測定する。対照試料の連続希釈物を用いて、標準曲線を作成し、X 軸上にポリペプチド濃度を (対数目盛)、Y 軸上に蛍光または吸光度 (均等目盛) をプロットする。標準曲線を用いて、試料中の CSG ポリペプチド濃度を内挿する。

【0205】

50

抗体

上記でいくらか詳細に記載したとおり、CSGポリペプチド、それらの断片またはその他の誘導体、もしくはその類縁体、あるいはそれらを発現する細胞は、それらに対する抗体を生ずるための免疫原として使用できる。これらの抗体は、ポリクローナルまたはモノクローナル抗体であり得る。本発明はまたキメラ、単鎖およびヒト化抗体ならびにFab断片、またはFab発現ライブラリーの産物も含む。かかる抗体および断片の産生のための、当該技術分野で既知の様々な手法が使用できる。

【0206】

抗体産生のための多様な方法が、Current Protocols, Chapter 2に記載されている。

10

例えば、本発明のCSGポリペプチドを発現する細胞を動物に投与して、ポリクローナル抗体を含有する血清の生産を誘発することができる。好ましい方法では、分泌タンパク質調製物を調製し、それを天然の夾雑物を実質的に含まないように精製する。

【0207】

次に、より高い特異的活性を有するポリクローナル抗血清を産生するために、この調製物を動物に導入する。得られた抗体は、ポリペプチド自体と結合するだろう。このようにすれば、CSGポリペプチドの断片のみをコードする配列でも、これを用いて全未変性ポリペプチドに結合する抗体を生成することができる。次にかかる抗体を用いてそのCSGポリペプチドを発現している組織から、CSGポリペプチドを単離できる。

【0208】

あるいは、モノクローナル抗体を調製することができる。モノクローナル抗体の製造技術の例は、ハイブリドーマ技術(Kohler, G. and Milstein, C., Nature 256: 495-497 (1975))、トリオーマ技術、ヒトB細胞ハイブリドーマ技術(Kozbor et al., Immunology Today 4: 72 (1983))および(Cole et al., MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc. pg. 77-96 (1985))を含むが、これらに限定されない。EBV-ハイブリドーマ技術は、ヒトモノクローナル抗体の製造に有用である。

20

【0209】

ハイブリドーマ技術は、Kohler et al., (Eur. J. Immunol. 6: 511 (1976))、Kohler et al. (Eur. J. Immunol. 6: 292 (1976))およびHammerling et al. (in: Monoclonal Antibodies and T-Cell Hybridomas, Elsevier, N. Y., pp. 563-681 (1981))にも記載されている。一般に、かかる手法は、CSGポリペプチド、またはより好ましくは分泌CSGペプチド発現細胞で、動物(好ましくは、マウス)を免疫することを包含する。

30

【0210】

かかる細胞は、任意の適切な組織培養培地にて培養されてもよいが、10%のウシ胎児血清(約56にて不活性化)添加、ならびに非必須アミノ酸約10g/l、ペニシリン約1,000U/ml、およびストレプトマイシン約100μg/ml添加、アール変法イーグル培地にて細胞を培養することが好ましい。かかるマウスの脾細胞を摘出し、適切な骨髓腫細胞系と融合させる。

40

【0211】

本発明によれば、いかなる適切な骨髓腫細胞系を用いてもよいが、ATCCから入手可能な親骨髓腫細胞系(parent myeloma cell line)(SP20)を用いることが好ましい。融合後、得られたハイブリドーマ細胞は、HAT培地にて選択的に保持され、続いてWands et al. (Gastroenterology 80: 225-232 (1981))に記載されたように、限界希釈によりクローニングされる。次に、かかる選択により得られたハイブリドーマ細胞をアッセイして、ポリペプチドを結合することが可能な抗体を分泌するクローンを同定する。

50

【0212】

あるいは、ポリペプチドに結合することが可能なさらなる抗体を、抗イディオタイプ抗体を用いた二段階手法にて産生することができる。かかる方法は、抗体自身が抗原であるという事実を利用するものであり、したがって、二次抗体に結合する抗体を得ることが可能である。この方法に従って、タンパク質特異抗体を用いて、動物、好ましくはマウスを免疫する。

【0213】

次に、かかる動物の脾細胞を用いて、ハイブリドーマ細胞を生産し、ハイブリドーマ細胞をスクリーニングして、タンパク質特異抗体に結合する能力が、ポリペプチドにより阻止され得る抗体を産生するクローンを同定する。かかる抗体は、タンパク質特異抗体に対する抗イディオタイプ抗体を含み、それを用いて、動物を免疫し、さらなるタンパク質特異抗体の形成を誘発することができる。

10

【0214】

単鎖抗体の生成に関し記載された技術（米国特許第4,946,778号）は本発明の免疫原性ポリペプチド産物に対する単鎖抗体の生成に適用することができる。さらにトランスジェニックマウス、ならびに他の非ヒトトランスジェニック動物もまた、本発明の免疫原性ポリペプチド産物に対するヒト化抗体の発現に使用され得る。

【0215】

本発明の抗体のFabおよびF(ab')₂ならびに他の断片を、本明細書に開示する方法に従って用いてもよいことが理解されるであろう。かかる断片は、典型的に、(Fab断片を産生するための)パインまたは(F(ab')₂断片を産生するための)ペプシンなどの酵素を用いて、タンパク質分解性切断により産生される。あるいは、分泌タンパク質結合断片は、組換えDNA技術の適用により、または合成化学により産生され得る。

20

【0216】

ヒトにおける抗体の*in vivo*使用に関して、「ヒト化」キメラモノクローナル抗体を使用することが好ましい場合がある。かかる抗体は、上述のモノクローナル抗体を産生するハイブリドーマ細胞に由来する遺伝構築物を用いて産生することができる。キメラ抗体を産生する方法は、当該技術分野で既知である（レビューとして、Morrisson, *Science* 229: 1202 (1985)、Oiet al., *BioTechniques* 4: 214 (1986)、Cabilly et al., 米国特許第4,816,567号、Taniguchi et al., EP 171496、Morrisson et al., EP 173494、Neuberger et al., WO 8601533、Robinson et al., WO 8702671、Boulianne et al., *Nature* 312: 643 (1984)、Neuberger et al., *Nature* 314: 268 (1985)を参照)。

30

【0217】

上記抗体は、アフィニティークロマトグラフィーによる単離および/または精製のために、固体支持体に抗体を結合させることで、本発明のCSGポリペプチドを発現しているクローンを単離または同定すること、またはCSGポリペプチドを精製することに使用され得る。上記でさらに詳細に述べられているとおり、CSGに特異的な抗体はまた、癌に苦しむ患者において、腫瘍、特に大腸の癌のイメージングに用いてもよい。かかる抗体はまた、CSGを発現する標的腫瘍に対し治療的に用いてもよい。

40

【0218】

CSG結合分子およびアッセイ

本発明はまた、受容体分子などのCSGを結合する分子の同定に関する方法も提供する。受容体タンパク質などの、CSGを結合するタンパク質をコードする遺伝子は、当業者に既知の様々な方法によって同定できる。例は、リガンドパニング法(Ligand panning)およびFACSソーティングを含むが、これらに限定されない。かかる方法は、例えばColigan et al., *Current Protocols in Immunology* 1(2): Chapter 5 (1991)などの多くの研究用

50

マニュアルに記載されている。

【0219】

発現クローニングは、この目的にも用いることができる。そのためには、本発明のCSGに反応する細胞からポリアデニル化RNAを調製する。このRNAからcDNAライブラリーを作成し、このライブラリーをプールに分ける。次に、このプールを本発明のCSGに反応しない細胞に個別にトランスフェクトする。次に、トランスフェクトした細胞を標識CSGに暴露する。CSGポリペプチドは、放射性ヨウ素化法、または部位特異的プロテインキナーゼに対する認識部位を含めること(inclusion of a recognition site)などの標準的方法を含むが、これらに限定されない多様な既知の技術により標識できる。

10

【0220】

暴露後、細胞は固定され、標識CSGの結合が決定される。これら手法はガラススライド上で好都合に実行される。標識CSGを含むプールが、CSG結合細胞が産生したcDNAを含むとして同定される。次に、これらの陽性体からサブプールが調製され、上記と同様に宿主細胞内にトランスフェクトされ、スクリーニングされる。サブプールプロセスとスクリーニングプロセスを繰り返すことで、受容体分子などの結合分子候補をコードする1または2以上の単一クローンを単離することができる。

【0221】

あるいは標識リガンドは、受容体分子などの、それを結合する分子を発現する細胞から調製された、膜あるいは膜抽出物などの細胞抽出物にフォトアフィニティー結合することができる。架橋した物質はポリアクリルアミドゲル電気泳動(「PAGE」)により分析され、X線フィルムに感光させられる。リガンド-受容体を含む標識複合体は切り出され、ペプチド断片に分解され、タンパク質マイクロシーケンシングにかけることができる。マイクロシーケンシングより得たアミノ酸配列を使用することで、候補受容体分子をコードする遺伝子を同定するためのcDNAライブラリーをスクリーニングするためのユニークまたは縮重オリゴヌクレオチドプローブを設計することができる。

20

【0222】

本発明のポリペプチドはまた、受容体分子などのCSG結合分子の、細胞内または無細胞調製物中でのCSG結合能の評価に使用できる。

【0223】

アゴニストおよびアンタゴニスト-アッセイおよび分子

本発明はまた、細胞においてCSGの作用を高める、または阻止するような化合物をスクリーニングする方法も提供する。本明細書中で用いられる「化合物」は、小有機分子、ペプチド、ポリペプチドおよび抗体、ならびに細胞におけるCSGの作用を高めるまたはアゴナイズする(agonize)または阻止するまたは拮抗する潜在力を有する、任意の他の候補分子を包含するとされる。

30

【0224】

本明細書中で用いる場合、アゴニストは、CSGの天然の生物学的機能を増加させるか、またはCSGに類似の様式で機能する化合物であり、一方アンタゴニストは、本明細書で用いる場合、かかる機能を低下させるまたは排除する化合物である。アゴニストおよび/またはアンタゴニストについてスクリーニングするための多様な既知の方法を、CSGアゴニストまたはアンタゴニストの同定に用いるために適合させることができる。

40

【0225】

例えば、膜などの細胞コンパートメント、または膜調製物などのそれらの調製物は、CSGにより調節されるシグナリングまたは制御経路の分子などの、CSGを結合する分子を発現する細胞から調製され得る。この調製物は、CSGアゴニストまたはアンタゴニストであり得る候補分子の非存在下または存在下にて、標識CSGとインキュベートされる。化合物の結合分子を結合する能力は、標識リガンドの減少した結合となって現れる。自由に結合する化合物、すなわちCSG結合分子の結合において、CSGの効果を誘発しない分子が最も良好なアンタゴニストであろう。

50

【0226】

十分に結合し、CSGと同一または密接に関係した効果を惹起する化合物はアゴニストである。潜在のアゴニストおよびアンタゴニストのCSG類似の効果は、例えば候補分子を細胞または適切な細胞調製物と相互作用させた後に、第2メッセンジャー系の活性を決定すること、およびその効果をCSGの効果、あるいはCSGと同一の効果を惹起する分子の効果と比較することで測定され得る。この観点に関して有用であり得るセカンドメッセンジャー系は、AMPグアニレートシクラーゼ、イオンチャネル、またはホスホイノシチド加水分解セカンドメッセンジャー系を含むが、これらに限定されない。

【0227】

CSGアンタゴニストに関するアッセイの別の例は、競合阻害アッセイに適切な条件の下にてCSGならびに潜在的アンタゴニストを、膜結合CSG受容体分子または組換えCSG受容体分子とを組み合わせる競合アッセイである。CSGは、受容体分子に結合したCSG分子の数を正確に決定し、潜在的アンタゴニストの有効性を評価できるように、例えば放射能で標識することができる。

10

【0228】

潜在的アンタゴニストは、本発明のCSGポリペプチドに結合し、それによりその活性を阻害または無力化する小有機分子、ペプチド、ポリペプチドおよび抗体を含む。潜在的アンタゴニストはまた、小有機分子、ペプチド、CSG誘発活性を誘発することなく、受容体分子などの結合分子上の同一部位に結合し、それによりCSGを結合から排除することによりCSGの作用を阻害する、密接に関連するタンパク質または抗体などのポリペプチドであってもよい。

20

【0229】

潜在的アンタゴニストは、CSGポリペプチドの結合部位に結合し、これを占有することで受容体分子などの細胞結合分子への結合を阻害し、その結果正常な生物学的活性が阻害される小分子を含む。小分子の例としては、小有機分子、ペプチドまたはペプチド様分子が含まれるが、これらに限定されない。

【0230】

その他の潜在的アンタゴニストは、アンチセンス分子を含む。アンチセンス技術は、アンチセンスDNAまたはRNAを通じ、あるいは三重らせん形成を通じて遺伝子発現を制御するのに使用できる。アンチセンス技術は、例えばOkano, J. Neurochem, 56: 560 (1991); OLIGODEOXYNUCLEOTIDES AS ANTISENSE INHIBITORS OF GENE EXPRESSION, CRC Press, Boca Raton, Fla. (1988)の中で論じられている。

30

【0231】

三重らせん形成は、例えばLee et al., Nucleic Acids Research 6: 3073 (1979); Cooney et al., Science 241: 456 (1988)、およびDervan et al., Science 251: 1360 (1991)の中で論じられている。該方法はポリヌクレオチドの相補的DNAまたはRNAへの結合に基づく。例えば本発明の成熟型CSGポリペプチドをコードするポリヌクレオチドの5'コード部分は、約10~40塩基対長のアンチセンスRNAオリゴヌクレオチドの設計に使用され得る。

40

【0232】

DNAオリゴヌクレオチドは、転写物に含まれる遺伝子の領域に相補的であるように設計されており、それによりCSGポリペプチドの転写ならびに生成を妨げる。アンチセンスRNAオリゴヌクレオチドは、in vivoにてmRNAにハイブリダイズし、mRNA分子のCSGポリペプチドへの翻訳を阻止する。また上記オリゴヌクレオチドは、アンチセンスRNAまたはDNAがin vivoで発現され、CSGの産生を阻害するように細胞内に送達されることもできる。

【0233】

50

組成物

本発明はまた、CSGポリヌクレオチドまたはCSGポリペプチド、あるいはそのアゴニストまたはアンタゴニストを含む組成物に関する。

例えば、本発明のCSGポリヌクレオチド、ポリペプチドまたはそのアゴニストまたはアンタゴニストは、細胞、組織または生体への使用のための非無菌的または無菌的な1または2以上のキャリア、例えば対象への投与に適切な薬学的キャリアと併用されてもよい。

【0234】

かかる組成物は、例えば媒体添加物または治療的な有効量の本発明のポリペプチドおよび薬学的に許容し得るキャリアまたは賦形剤を含む。かかるキャリアは、生理食塩水、緩衝生理食塩水、デキストロース、水、グリセロール、エタノール、およびそれらの組み合わせを含んでもよいが、これらに限定されない。配合物は投与様式に適していなければならない。

10

【0235】

本発明の組成物は、個々の患者の臨床状態（特に、ポリペプチドまたは他の化合物単独での処置の副作用）、送達部位、投与方法、投与のスケジューリング、および開業医に既知の他の要素を考慮して、良質の医療のための原則（good medical practice）に沿った様式で、処方されて、投薬される。したがって、本明細書の目的に関する「有効量」は、かかる考慮により決定される。

【0236】

一般的な提案として、1回用量あたり非経口的に投与される分泌ポリペプチドの薬学的総有効量は、約 $1 \mu\text{g}/\text{kg}$ （患者の体重）/日～ $10 \text{mg}/\text{kg}$ （患者の体重）/日の範囲であろうが、上述するように、これは治療上の判断に左右されるであろう。より好ましくは、この用量は、少なくとも $0.01 \text{mg}/\text{kg}/\text{日}$ であり、ヒトに関して最も好ましくは、ホルモンに関して約 $0.01 \sim 1 \text{mg}/\text{kg}/\text{日}$ である。

20

【0237】

連続的に与えられる場合、ポリペプチドまたは他の化合物は、典型的に、1日あたり1～4回の注射により、または連続皮下注入により（例えば、ミニポンプを用いて）、約 $1 \mu\text{g}/\text{kg}/\text{時}$ ～約 $50 \text{mg}/\text{kg}/\text{時}$ の用量速度にて投与される。点滴バッグ（intravenous bag）溶液もまた、用いてもよい。変化を観察するのに必要な処置の期間、および処置後に応答が生じる間隔は、所望の効果に依存して変わるようである。

30

【0238】

本発明の分泌タンパク質を含有する薬学的組成物は、経口的に、経直腸的に、非経口的に、大槽内に、腔内に、腹腔内に、局所的に（パウダー、軟膏、ジェル、滴薬または経皮パッチによるものなど）、口腔に、または口腔もしくは鼻スプレーとして、投与される。「薬学的に許容し得るキャリア」は、無毒性の固形、半固形または液体賦形剤、希釈剤、カプセル用材料または任意のタイプの配合助剤を指す。本明細書で使用する「非経口」という用語は、静脈内、筋内、腹腔内、胸骨内、皮下、および関節内注射ならびに注入を含む投与様式を指す。

【0239】

ポリペプチドまたは他の化合物はまた、持続放出系により適切に投与される。持続放出性組成物の適切な例としては、造形品、例えば、フィルムまたはマイクロカプセルの形態にある半透性ポリマーマトリックスが挙げられる。持続放出性マトリックスとしては、ポリラクチド（米国特許第3,773,919号およびEP 58481）、L-グルタミン酸およびガンマ-エチル-L-グルタメートのコポリマー（Sidman, U. et al., Biopolymers 22: 547-556 (1983)）、ポリ(2-ヒドロキシエチルメタクリレート）（R. Langer et al., J. Biomed. Mater. Res. 15: 167-277 (1981)）、およびR. Langer, Chem. Tech. 12: 98-105 (1982)）、エチレンビニルアセテート（R. Langer et al.）、およびポリ-D-()-3-ヒドロキシ酪酸（EP 133,988）が挙げられる。

40

50

【0240】

持続放出性組成物はまた、リポソームエンタラップポリペプチドを包含する。ポリペプチドまたは他の化合物を含有するリポソームは、既知の方法により調製される (Epstein et al., Proc. Natl. Acad. Sci. USA 82: 3688-3692 (1985)、Hwang et al., Proc. Natl. Acad. Sci. USA 77: 4030-4034 (1980)、EP 52322、EP 36676、EP 88046、EP 143949、EP 142641、日本国特許出願第83-118008号、米国特許第4,485,045号および第4,544,545号、ならびにEP 102324)。通常、リポソームは、小さい(約200~800オングストローム)単層型であり、そこでの脂質含量は、約30(mol%)コレステロールよりも高く、選択比率は、最適な治療に関して調節される。

【0241】

非経口投与に関して、一態様では、ポリペプチドまたは他の化合物は、一般的に所望の純度にて、単位投与量の注射可能な形態(溶液、懸濁液、またはエマルジョン)で、薬学的に許容し得るキャリア、すなわち、使用する投与量および濃度にてレシピエントに無毒性であり、配合物の他の成分と併用可能(compatible)であるものと混合することにより配合される。

例えば、配合物は、好ましくは、酸化剤、およびポリペプチドまたは他の化合物にとって有害であると知られている他の化合物を含まない。

【0242】

一般に、配合物は、液体キャリアまたは微細固体キャリア、あるいはその両方と、均一かつ緊密にポリペプチドまたは他の化合物を接触させることにより調製される。次に、必要な場合には、生成物を所望の配合物に成形する。好ましくは、キャリアは、非経口キャリア、より好ましくは、レシピエントの血液と等張である溶液である。かかるキャリア媒質の例としては、水、生理食塩水、リンゲル溶液、およびブドウ糖溶液が挙げられる。揮発性油およびオレイン酸エチルのような非水性媒質、ならびにリポソームもまた、本発明に有用である。

【0243】

キャリアは適切に、等張性および化学的安定性を高める物質などの、少量の添加剤を含有する。かかる物質は、用いる投与量および濃度にて、レシピエントに対して無毒性であり、リン酸塩、クエン酸塩、コハク酸塩、酢酸、および他の有機酸またはそれらの塩などの緩衝液、アスコルビン酸などの酸化防止剤、低分子量(約10残基未満)ポリペプチド、例えば、ポリアルギニンまたはトリペプチド、血清アルブミン、ゼラチンまたは免疫グロブリンなどのタンパク質、ポリビニルピロリドンなどの親水性ポリマー、グリシン、グルタミン酸、アスパラギン酸またはアルギニンなどのアミノ酸、セルロースまたはその誘導体、グルコース、マンノース、またはデキストリンを含めた単糖類、二糖類、および他の炭化水素、EDTAなどのキレート剤、マンニトールまたはソルビトールなどの糖アルコール、ナトリウムなどの対イオン、および/またはポリソルベート、ポロキサマー、またはPEGなどの非イオン性界面活性剤を含む。

【0244】

ペプチドまたは他の化合物は、典型的に、約0.1mg/ml~100mg/ml、好ましくは1~10mg/mlの濃度で、約3~8のpHにて、かかるビヒクル中で配合される。先述の賦形剤、キャリアまたは安定剤のうちのあるものの使用は、結果としてポリペプチド塩または他の化合物の塩の形成を生じる結果となることは理解されるであろう。

【0245】

治療用投与のために使用されるべきポリペプチドはいずれも無菌であべきである。無菌状態は、滅菌濾過膜(例えば、0.2ミクロンの膜)を通じた濾過により容易に達成される。治療用ポリペプチド組成物は一般に、滅菌アクセスポートを有する容器、例えば、皮下注射針により突き刺すことが可能なストッパーを有する点滴溶液バッグまたはバイアルに入れられる。

10

20

30

40

50

【0246】

ポリペプチドは通常、単用量または複数用量容器（例えば、密封アンプルまたはバイアル）中に、水溶液として、または再構成用凍結乾燥配合物として、保管されるであろう。凍結乾燥配合物の例として、10mlバイアルを滅菌濾過した1%（w/v）ポリペプチド水溶液5mlで充填し、得られた混合物を凍結乾燥する。注入溶液は、注射用静菌水を用いて、凍結乾燥ポリペプチドを再構成することにより調製される。

【0247】

キット

本発明はさらに、本発明の上記組成物の1または2以上の成分が充填された1または2以上の容器を含む、薬学的パックおよびキットに関する。1または2以上のかかる容器には、医薬品または生物製品の製造、使用または販売を管轄する政府機関が定めた形式で書かれた、該機関による、ヒトへの投与に関する製品の製造、使用または販売の承認を示す通知を添付することができる。

10

【0248】

投与

本発明のCSGポリペプチドまたはポリヌクレオチドまたはその他の化合物、好ましくはそのアゴニストまたはアンタゴニストは、単独または治療化合物などのその他の化合物とともに使用してもよい。

薬学的組成物は、例えばとりわけ局所、経口、肛門、膺、静脈内、腹腔内、筋肉内、皮下、鼻内または皮内経路による投与を含む効果的かつ便宜的な方法で投与され得る。

20

【0249】

薬学的組成物は一般に、特定の1または2以上の適応症の処置あるいは予防に有効な量で投与される。一般には、組成物は少なくとも約10μg/体重kgの量で投与される。しかしながら、最適投与量は、適応症、その重症度、投与経路、合併症状等を考慮しながら、各処置の様式および適応症に関する標準的方法により決定されるだろう。

【0250】

個体においてCSGポリペプチドの標準または正常発現レベルの減少により引き起こされる症状は、好ましくは分泌形態にある、本発明のCSGポリペプチド、またはそのアゴニストを投与することにより処置できることが理解されるであろう。したがって、本発明はまた、CSGポリペプチドレベルの増加を必要とする個体を処置する方法であって、かかる個体においてCSGポリペプチドの活性レベルを増加させるためのCSGポリペプチドまたはそのアゴニストの量を含む薬学的組成物を、かかる個体に投与することを含む方法を提供する。例えば、減少したCSGポリペプチドレベルを有する患者は、1日用量0.1~100μg/kgのCSGポリペプチドまたはそのアゴニストを6日間連続で受けてもよい。好ましくは、CSGポリペプチドが投与される場合、それは好ましくは分泌形態においてである。

30

【0251】

本発明の組成物はまた、CSGポリペプチドの増大したレベルを処置するために投与することもできる。例えば、アンチセンス技術を、本発明のCSGポリペプチドの産生を抑制するために用いることができる。この技術は、癌などの様々な病因に起因するポリペプチド、好ましくは分泌形態のもののレベルを減少させる方法の一例である。

40

【0252】

異常に増加したポリペプチドレベルを有すると診断された患者に、0.5、1.0、1.5、2.0および3.0mg/kg/日にて、21日間、アンチセンスポリヌクレオチドを静脈内に投与することができる。この処置が十分に耐えられた場合、処置は7日の休止期間後に繰り返される。また、CSGポリペプチドのアンタゴニストを含む組成物を、患者におけるCSGレベルを減少させるために投与することもできる。

【0253】

遺伝子治療

CSGポリヌクレオチド、ポリペプチド、ポリペプチドであるアゴニストおよびアンタゴ

50

ニストは、しばしば「遺伝子治療」と称される治療様式において、かかるポリペプチドを *in vivo* で発現させることにより、本発明により使用されてもよい。

【0254】

したがって、例えば、患者由来の細胞は、*ex vivo* でポリペプチドをコードする DNA または RNA などのポリヌクレオチドにより操作されてもよく、次に操作された細胞を、ポリペプチドにより処置されるべき患者に提供することができる。例えば、細胞は本発明のポリペプチドをコードする RNA を含むレトロウイルスプラスミドベクターを用いて、*ex vivo* で操作することができる。かかる方法は、当該技術分野で既知であり、本発明でのそれらの使用は本発明中の教示より明らかになるだろう。

【0255】

同様に細胞は、当分野既知の方法により *in vivo* でのポリペプチド発現のため *in vivo* で操作され得る。例えば、本発明のポリペプチドは、上述したとおり、複製欠損レトロウイルスベクターでの発現のために操作され得る。

【0256】

次に、レトロウイルス発現構築物が単離され、本発明のポリペプチドをコードする RNA を含むレトロウイルスプラスミドベクターで形質導入されたパッケージング細胞内に、そのパッケージング細胞が次に所望遺伝子を含む感染性ウイルス粒子を産生するように導入され得る。これらの産生細胞は、*in vivo* で細胞を操作し、*in vivo* にてポリペプチドを発現させることを目的として、患者に投与され得る。本発明のポリペプチドを投与するためのこれらおよびその他の方法は、本発明の教示より当業者に明らかになるはずである。

【0257】

本明細書中で上述したレトロウイルスプラスミドベクターが由来し得るレトロウイルスとしては、モロニー Maus 白血病ウイルス、脾臓壊死ウイルス、ラウス肉腫ウイルス、ハーベイ肉腫ウイルス、トリ白血病ウイルス、テナガザル白血病ウイルス、ヒト免疫不全ウイルスなどのレトロウイルス、アデノウイルス、骨髄増殖性肉腫ウイルス、および哺乳動物腫瘍ウイルスが挙げられるが、これらに限定されない。一態様では、レトロウイルスプラスミドベクターはモロニー Maus 白血病ウイルスより得られる。

【0258】

かかるベクターは、ポリペプチドを発現させるための 1 または 2 以上のプロモーターを含むだろう。好適なプロモーターの選択は、本明細書の教示から当業者に明らかとなるだろう。しかしながら、使用され得る適切なプロモーターは、レトロウイルス LTR、SV40 プロモーター、Miller et al., *Bio techniques* 7: 980-990 (1989) 記載のヒトサイトメガロウイルス (CMV) プロモーター、およびヒストン、RNA ポリメラーゼ III およびベータアクチンプロモーターなどの真核生物細胞性プロモーターを含むが、これらに限定されない。使用され得るその他のウイルスプロモーターは、アデノウイルスプロモーター、チミジンキナーゼ (TK) プロモーター、および B19 パルボウイルスプロモーターを含むが、これらに限定されない。

【0259】

用いられ得るさらなるプロモーターは、呼吸器合胞体ウイルス (RSV) プロモーター、MMT プロモーターなどの誘導性プロモーター、メタロチオネインプロモーター、熱ショックプロモーター、アルブミンプロモーター、ApoA I プロモーター、ヒトグロビンプロモーター、単純ヘルペスチミジンキナーゼプロモーターなどのウイルスチミジンキナーゼプロモーター、レトロウイルス LTR、ベータアクチンプロモーター、およびヒト成長ホルモンプロモーターを含む。プロモーターは、ポリペプチドをコードする遺伝子を制御する未変性プロモーターでもよい。

【0260】

本発明のポリペプチドをコードする核酸配列は、好適なプロモーターの制御下に置かれ得る。

【0261】

10

20

30

40

50

1つの態様において、レトロウイルスプラスミドベクターは、プロデューサー細胞系を形成するパッケージング細胞系の誘導に使用される。トランスフェクトされ得るパッケージング細胞の例としては、PE501、PA317、Y-2、Y-AM、PA12、T19-14X、VT-19-17-H2、YCRE、YCRIP、GP+E-86、GP+envAm12およびMiller, A., Human Gene Therapy 1: 5-14 (1990)記載のDAN細胞系が挙げられるが、これらに限定されない。

【0262】

ベクターは、当該技術分野に既知の任意の手段によりパッケージング細胞内に導入される。かかる手段としては、エレクトポレーション、リポソームの利用、およびCaPO₄沈殿が挙げられるが、これらに限定されない。あるいは、レトロウイルスプラスミドベクターはリポソーム内にカプセル封入されるか、または脂質に結合された後に宿主に投与されてもよい。プロデューサー細胞系は、ポリペプチドをコードする核酸配列を含む感染性レトロウイルスベクター粒子を産生するだろう。

10

【0263】

次に、かかるレトロウイルスベクター粒子を利用して、真核生物細胞に*in vitro*または*in vivo*にて形質導入し得る。形質導入された真核生物細胞は、ポリペプチドをコードする核酸配列を発現するだろう。形質導入され得る真核生物細胞としては、胚性幹細胞、胚性腫瘍細胞、ならびに造血幹細胞、肝細胞、線維芽細胞、筋芽細胞、ケラチノサイト、内皮細胞、および気管支上皮細胞が含まれるが、これらに限定されない。

【0264】

典型的な遺伝子治療の一方法は、CSGポリペプチドまたはそのアゴニストもしくはアンタゴニストを発現することが可能な線維芽細胞を、患者に移植することを伴う。一般に、線維芽細胞は、皮膚生検により対象から得られる。得られた組織を組織培養培地に置き、小片に分離する。組織の小片を組織培養フラスコの湿った表面上へ載せ、各フラスコにおいて約10片を載せる。フラスコを上下ひっくり返し、きつく閉めて、室温で一晩放置する。

20

【0265】

室温にて24時間後、フラスコをひっくり返し、組織片をフラスコの底に固定させたまま、新鮮な培地(例えば、Ham's F12培地、10% FBS、ペニシリンおよびストレプトマイシン加)を添加する。次に、フラスコを37℃にて約1週間インキュベートする。この時点で、新鮮な培地を添加し、続いて数日毎に取り換える。さらに2週間の培養の後、線維芽細胞の単層が現れる。

30

【0266】

単層をトリプシン処理し、大きなフラスコへはがし取る。モロニーマウス肉腫ウイルスの長い末端反復配列に隣接したpMV-7(Kirschmeier, P. T. et al., DNA, 7: 219-25 (1988))をEcoRIおよびHindIIIで消化した後、仔ウシ腸ホスファターゼ(calf intestinal phosphatase)で処理する。アガロースゲル上で線状ベクターを分画し、ガラスビーズを用いて精製する。

【0267】

本発明のCSGポリペプチドまたはそのアゴニストもしくはアンタゴニストをコードするcDNAを、それぞれ5'および3'末端配列に相当するPCRプライマーを用いて増幅することができる。好ましくは、5'プライマーは、EcoRI部位を含有し、3'プライマーは、HindIII部位を含む。等量のモロニーマウス肉腫ウイルス線状主鎖ならびに増幅したEcoRIおよびHindIII断片を、T4DNAリガーゼの存在下、一緒に添加する。得られた混合物を、2つの断片の連結に適した条件下にて保持する。

40

【0268】

次に、連結混合物を用いて、バクテリアHB 101を形質転換した後、ベクターが適切に挿入された所定の遺伝子を有することを確認する目的で、カナマイシンを含有する寒天上に蒔く。両栄養性pA317またはGP+am12パッケージング細胞を、10%の仔

50

ウシ血清 (CS)、ペニシリンおよびストレプトマイシン加ダルベッコ変法イーグル培地 (DMEM) 中で、コンフルエント密度に、組織培養で増殖させる。次に、遺伝子を含む MSV ベクターを培地に添加し、パッケージング細胞にベクターで形質導入する。

【0269】

すると、パッケージング細胞は、遺伝子を含む感染ウイルス粒子を生産する (パッケージング細胞は、今度はプロデューサー細胞と称される)。形質導入したプロデューサー細胞に、新鮮な培地を添加し、続いて、コンフルエントなプロデューサー細胞の 10 cm プレートから、培地を採取する。ミリポアフィルターを通して感染ウイルス粒子を含む使用済培地を濾過し、剥離したプロデューサー細胞を除去し、続いてこの培地を用いて、線維芽細胞を感染させる。線維芽細胞のサブコンフルエントなプレートから、培地を除去し、プロデューサー細胞からの培地とすばやく入れ換える。この培地を除去し、新鮮な培地と入れ換える。

10

【0270】

ウイルスの力価が高い場合、実質的にすべての線維芽細胞は感染し、選択は必要ない。力価が非常に低い場合、neo または his のような選択マーカーを有するレトロウイルスベクターを使用する必要がある。線維芽細胞がいったん効果的に感染したら、線維芽細胞を分析して、タンパク質が産生されるかどうかを決定する。次に、操作した線維芽細胞を、単独で、または cyto dex 3 マイクロキャリアビーズ上にてコンフルエントまで増殖させた後に、宿主に移植する。

【0271】

あるいは、CSG に関連する障害、疾患および症状を処置するために、in vivo 遺伝子治療方法を用いることもできる。遺伝子治療方法は、ポリペプチドの発現を増加または減少させるための、裸の核酸 (DNA、RNA、およびアンチセンス DNA または RNA) 配列の動物への導入に関する。

20

【0272】

例えば、本発明の CSG ポリヌクレオチドまたは、それに対するアゴニストもしくはアンタゴニストをコードする核酸配列は、プロモーターまたは標的組織によって、ポリペプチド発現に必要な任意の他の遺伝要素に作動可能に連結されてもよい。かかる遺伝子治療ならびに送達技術および方法は、当該技術分野で既知であり、例えば、WO 90/11092、WO 98/11779、米国特許第 5,693,622 号、第 5,705,151 号、第 5,580,859 号、Tabata H. et al. (1997) Cardiovasc. Res. 35 (3): 470-479、Chao J et al. (1997) Pharmacol. Res. 35 (6): 517-522、Wolf J. A. (1997) Neuromuscul. Disord. 7 (5): 314-318、Schwartz B. et al. (1996) Gene Ther. 3 (5): 405-411、Tsurumi Y. et al. (1996) Circulation 94 (12): 3281-3290 (参照により本明細書に援用される) を参照されたい。

30

【0273】

ポリヌクレオチド構築物は、組織 (心臓、筋肉、皮膚、肺、肝臓、腸など) の間質腔への注射などの、動物細胞へ注入可能な物質を送達するいかなる方法により送達されてもよい。ポリヌクレオチド構築物は、薬学的に許容し得る液体または水性キャリア中で送達することができる。

40

【0274】

「裸の」ポリヌクレオチド、DNA または RNA という用語は、細胞への侵入を補佐、促進または助長するよう作用する、ウイルス配列、ウイルス粒子、リボソーム配合物、リポフェクチンまたは沈殿剤などを含むいかなる送達ビヒクルも含まない配列を指す。しかしながら、ポリヌクレオチドはまた、当業者に既知の方法により調製され得るリボソーム配合物 (例えば、Felgner P. L. et al. (1995) Ann. NY Acad. Sci. 772: 126-139、および Abdallah B. et al

50

． (1 9 9 5) B i o l . C e l l 8 5 (1) : 1 - 7 に教示されているもの) にて送達されてもよい。

【 0 2 7 5 】

遺伝子治療方法にて使用されるポリヌクレオチドベクター構築物は、好ましくは、宿主ゲノムに組み込まれない構築物であり、またそれらは、複製が可能な配列を含有しないであろう。当業者に既知のいかなる強力なプロモーターも、DNA発現を促進するために使用することができる。他の遺伝子治療技術と異なり、標的細胞へ裸の核酸配列を導入することの1つの主要な利点は、細胞におけるポリヌクレオチド合成の一過性の性質である。研究により、非複製DNA配列が細胞に導入されて、最大6ヶ月の期間、所望のポリペプチドの産生を提供することができることがわかった。

10

【 0 2 7 6 】

ポリヌクレオチド構築物は、筋肉、皮膚、脳、肺、肝臓、脾臓、骨髄、胸腺、心臓、リンパ、血液、骨、軟骨、膵臓、腎臓、胆嚢、胃、腸、精巣、卵巣、子宮、直腸、神経系、眼、腺、および結合組織を含む、動物内組織の間質腔へ送達され得る。組織の間質腔は、細胞間液、器官組織の細網線維間のムコ多糖マトリックス、管または房の壁における弾性線維、線維性組織のコラーゲン線維、または筋細胞を覆う結合組織内もしくは骨小腔中にある同じマトリックスを含む。

【 0 2 7 7 】

循環血漿およびリンパ管のリンパ液により占有される腔も同様である。筋組織の間質腔への送達が好ましい。ポリヌクレオチド構築物は、これらの細胞を含む組織への注射により

20

【 0 2 7 8 】

送達および発現は、未分化またはあまり完全には分化していない細胞、例えば、血液の幹細胞または皮膚線維芽細胞などにて達成され得るが、それらは、好ましくは、分化した永続性の非分裂細胞に送達され、そこで発現される。in vivoの筋細胞は特に、ポリヌクレオチドを取り込み、発現するそれらの能力に優れている。

【 0 2 7 9 】

裸のポリヌクレオチド注射に関して、有効投与量のDNAまたはRNAは、約0.05 μg / 体重kg ~ 約50 mg / 体重kgの範囲であろう。好ましくは、投与量は、0.005 mg / kg ~ 約20 mg / kgであり、より好ましくは約0.05 mg / kg ~ 約5 mg / kgであろう。当然のことながら、当業者が理解するように、この投与量は、注射の組織部位により変化するのである。適切かつ有効な投与量の核酸配列は、当業者により容易に決定することができ、それは、処置する症状および投与経路に依存し得る。

30

【 0 2 8 0 】

好ましい投与経路は、組織の間質腔への注射の非経口経路によるものである。しかしながら、特に肺もしくは気管支組織、咽喉または鼻の粘膜への送達用のエアロゾル配合物の吸入などの、他の非経口経路もまた使用してもよい。さらに、裸のポリヌクレオチド構築物は、血管形成中に、当該手順で使用されるカテーテルにより、動脈へ送達することができる。

【 0 2 8 1 】

筋肉中に注射されたポリヌクレオチドのin vivoでの用量反応効果は以下のように決定される。本発明のポリペプチドをコードするmRNAの産生のための適切な鋳型DNAを、標準的な組換えDNA方法論に従って調製する。鋳型DNAは、環状または線状のいずれであってもよく、それは裸のDNAとして用いられるか、あるいはリポソームと複合体を形成する。次に、マウスの四頭筋に、様々な量の鋳型DNAを注射する。

40

【 0 2 8 2 】

5 ~ 6 週齢の雌および雄Balb / Cマウスを、2.5%アベルチン0.3 mlの腹腔内注射により麻酔する。腹側大腿上に1.5 cmの切開を施し、四頭筋を直接見えるようにする。筋肉の膝への遠位付着部位から約0.5 cmかつ約0.2 cmの深さで、1分にわたり、27ゲージ針を通じて、1 ccの注射器内のキャリア0.1 ml中の鋳型DNAを

50

注射する。将来の位置確認のために注入部位上に縫合糸を設置し、皮膚をステンレス製クリップで閉じる。

【0283】

適切なインキュベーション時間後（例えば、7日）、四頭筋全体を切除することにより、筋肉摘出物を調製する。個々の四頭筋のすべての5番目の15 μm断面を、タンパク質発現のために組織化学的に染色する。タンパク質発現に関するタイムコースは、種々のマウス由来の四頭筋が異なる時間にて採取されることを除いて、同様の様式にて行われ得る。注射後の筋肉中のDNAの永続性は、注射したマウスおよび対照マウスから全細胞DNAおよびHIRT上清を調製した後に、サザンブロット分析により確定されてもよい。

【0284】

マウスでの上記実験の結果を用いて、裸のDNAを用いた、ヒトおよび他の動物における適切な投与量および他の治療パラメータを外挿することができる。

【0285】

非ヒトトランスジェニック動物

本発明のCSGポリペプチドはまた、非ヒトトランスジェニック動物にて発現させることもできる。マウス、ラット、ウサギ、ハムスター、モルモット、ブタ、小型ブタ（micro-pig）、ヤギ、ヒツジ、ウシ、および非ヒト霊長類、例えば、ヒヒ、サル、およびチンパンジーが含まれるが、これらに限定されないあらゆる種の非ヒト動物を用いて、トランスジェニック動物を生み出してもよい。当該技術分野で既知の任意の技法を用いて、動物に導入遺伝子（すなわち、本発明のポリペプチド）を導入して、トランスジェニック動物の創始系統を生産してもよい。

【0286】

かかる技法としては、前核（pronuclear）マイクロインジェクション（Paterson et al., Appl. Microbiol. Biotechnol. 40: 691-698 (1994)、Carver et al., Biotechnology (NY) 11: 1263-1270 (1993)、Wright et al., Biotechnology (NY) 9: 830-834 (1991); およびHoppe et al., 米国特許第4,873,191号(1989)、生殖系列（Van der Putten et al., Proc. Natl. Acad. Sci., USA 82: 6148-6152 (1985)、胚盤胞または胚へのレトロウイルス媒介性遺伝子移入、胚性幹細胞における遺伝子ターゲティング（Thompson et al., Cell 56: 313-321 (1989)、細胞または胚へのエレクトロポレーション（Lo, 1983, Mol Cell. Biol. 3: 1803-1814 (1983)、遺伝子銃を用いた本発明のポリヌクレオチドの導入（例えば、Ulmer et al., Science 259: 1745 (1993)を参照）、胚性多能性幹細胞への核酸構築物の導入および胚盤胞への幹細胞の移し戻し、および精子媒介性遺伝子移入（Lavitrano et al., Cell 57: 717-723 (1989)）が挙げられるが、これらに限定されない。

【0287】

かかる技法のレビューとして、Gordon, "Transgenic Animals", Intl. Rev. Cytol. 115: 171-229 (1989)を参照されたい（それは、参照により全体が本明細書に援用される）。

当該技術分野において既知のいかなる技法、例えば、静止状態へ誘導された培養胚細胞、胎児細胞または成人細胞からの核の、脱核卵母細胞への核移入（Campbell et al., Nature 380: 64-66 (1966)、Wilmut et al., Nature 385: 810813 (1997)）を用いて、本発明のポリヌクレオチドを含有するトランスジェニッククローンを生産してもよい。

【0288】

本発明は、すべての動物細胞において導入遺伝子を持つトランスジェニック動物、ならびに、すべてではないが幾つかの細胞において導入遺伝子を持つ動物、すなわちモザイク動

10

20

30

40

50

物またはキメラ動物を提供する。導入遺伝子は、単一導入遺伝子として、またはコンカテマー（例えば、頭-頭タンデム、もしくは頭-尾タンデム）のような多重コピーとして、組み込まれてもよい。導入遺伝子はまた、例えば、Lasko et al. の教示 (Lasko et al., Proc. Natl. Acad. Sci. USA 89: 6232-6236 (1992)) に従って、特定の細胞タイプに選択的に導入されて、そこで活性化されてもよい。

【0289】

かかる細胞タイプ特異的な活性化に必要な調節遺伝子は、所定の特定細胞タイプに依存し、それは当業者には明らかであろう。ポリヌクレオチド導入遺伝子が内因性遺伝子の染色体部位に組み込まれることが望ましい場合、遺伝子ターゲティングが好ましい。簡潔に述べると、かかる技法を利用する場合、内因性遺伝子に相同的な幾つかのヌクレオチド配列を含有するベクターを、内因性遺伝子のヌクレオチド配列へ、染色体配列との相同組換えにより組み込み、その機能を崩壊させる目的で設計する。

10

【0290】

導入遺伝子はまた、特定の細胞タイプへ選択的に導入されてもよく、したがって、例えば Gu et al. (Science 265: 103-106 (1994)) の教示に従って、その細胞タイプのみで内因性遺伝子を不活性化させる。かかる細胞タイプに特異的な不活性化に必要な調節配列は、所定の特定細胞タイプに依存し、それは当業者には明らかであろう。

【0291】

トランスジェニック動物がいったん創出されれば、組換え遺伝子の発現は、標準的な技法を利用してアッセイされ得る。初期スクリーニングは、導入遺伝子の組み込みが起ったことを確認すべく動物組織を分析するために、サザンプロット分析またはPCR技法により達成され得る。

20

【0292】

トランスジェニック動物の組織における導入遺伝子のmRNA発現レベルはまた、動物から得られた組織試料のノーザンプロット分析、in situハイブリダイゼーション分析、および逆転写酵素PCR (rt-PCR) が含まれるが、これらに限定されない技法を用いて評価してもよい。トランスジェニック遺伝子発現組織の試料はまた、導入遺伝子産物に特異的な抗体を用いて、免疫細胞化学的に、または免疫組織化学的に評価されてもよい。

30

【0293】

創始動物がいったん生産されれば、それらを、繁殖、同系交配、異系交配、または交雑させて、特定の動物コロニーを生産してもよい。かかる繁殖戦略の例としては、別個の系統を樹立するために、1つよりも多い組み込み部位を有する創始動物を異系交配すること、各導入遺伝子の付加的発現の効果が理由で、高いレベルで導入遺伝子を発現する複合遺伝子導入をもたらすために、別個の系統を同系交配させること、発現を増大させ、かつDNA分析による動物のスクリーニングの必要性をなくすために、所定の組み込み部位に関してホモ接合性の動物を生産するために、ヘテロ接合性のトランスジェニック動物を交雑させること、複合ヘテロ接合性またはホモ接合性系統を生産するために、別個のホモ接合性系統を交雑させること、および当該実験モデルに適切な明瞭なバックグラウンドに導入遺伝子を配置させるために繁殖させることが挙げられるが、これらに限定されない。

40

【0294】

本発明のトランスジェニック動物は、本発明のCSGポリペプチドの生物機能を説明し、CSGの異常発現に関連した症状および/または障害を研究し、かかるCSG関連症状および/または疾患を回復するのに効果的な化合物をスクリーニングするのに有用な動物モデル系を含むが、これに限定されない用途を有する。

【0295】

ノックアウト動物

内因性の遺伝子発現はまた、標的 (targeted) 相同組換えを用いて、遺伝子およ

50

びノもしくはそのプロモーターを不活性化または「ロックアウト」することにより低減することができる（例えば、Smithies et al., Nature 317: 230-234 (1985)、Thomas & Capecchi, Cell 51: 503512 (1987)、Thompson et al., Cell 5: 313-321 (1989)を参照。それらはそれぞれ、その全体が、参照により本明細書に援用される）。

【0296】

例えば、内因性CSGポリヌクレオチド配列（遺伝子のコード領域または調節領域）に相同的なDNAに隣接した突然変異体である本発明の非機能性CSGポリヌクレオチド（または完全に非関連のDNA配列）を、選択マーカーおよび/またはネガティブ選択マーカーを用いて、あるいは用いずに使用して、*in vivo*で本発明のポリペプチドを発現する細胞をトランスフェクトすることができる。別の態様では、当該技術分野にて既知の技法を用いて、所定の遺伝子を含むが、発現はしない細胞におけるロックアウトを生成する。

10

【0297】

標的相同組換えを介したDNA構築物の挿入は、結果として標的遺伝子の不活性化を招く。かかるアプローチは、研究および農業の分野に特に適しており、そこで胚性幹細胞の改変を用いて、不活性な標的遺伝子を有する動物の子孫を発生させることができる（例えば、上述のThomas & Capecchi 1987およびThompson 1989を参照）。組換えDNA構築物が、当業者には明らかであろう適切なウイルスベクターを用いて、*in vivo*にて、必要とされる部位に直接的に投与されるか、または標的にされるという条件で、このアプローチはまた、ヒトにおける使用にルーチンで適合される。

20

【0298】

本発明のさらなる態様では、本発明のCSGポリペプチドを発現するように遺伝子操作された細胞、あるいは本発明のCSGポリペプチドを発現しないように遺伝子操作された（例えば、ロックアウト）細胞を、*in vivo*にて患者に投与する。かかる細胞は、患者、またはMHC適合ドナーから得てもよく、線維芽細胞、骨髄細胞、血液細胞（例えば、リンパ球）、脂肪細胞、筋細胞、内皮細胞を含むことができるが、これらに限定されない。

30

【0299】

細胞は、組換えDNA技法を用いて、*in vitro*にて遺伝子操作され、細胞に本発明のポリペプチドのコード配列を導入するか、あるいは例えば形質導入（ウイルスベクター、好ましくは細胞ゲノムに導入遺伝子を組み込むベクターを用いて）またはプラスミド、コスミド、YAC、裸のDNA、エレクトロポレーション、リポソーム等の使用を含むが、これらに限定されないトランスフェクション手法により、本発明のポリペプチドに関するコード配列および/または内因性調節配列を崩壊させる。

【0300】

本発明のCSGポリペプチドのコード配列は、強力な構成もしくは誘導プロモーターまたはプロモーター/エンハンサーの制御下に置くことにより、本発明のCSGポリペプチドの発現、好ましくは分泌を達成することができる。本発明のCSGポリペプチドを発現、好ましくは分泌する操作細胞を全身的に、例えば、循環にて、または腹腔内にて患者に導入することができる。

40

【0301】

あるいは、細胞をマトリックス中に組み込み、体内へ移植することができ、例えば、遺伝子操作した線維芽細胞は、皮膚移植片の一部として移植することができ、または、遺伝子操作した内皮細胞は、リンパ管または血管移植片の一部として移植することができる（例えば、各々その全体が参照により本明細書に援用される米国特許第5,399,349号、および米国特許第5,460,959号を参照）。

【0302】

50

投与されるべき細胞が非自系または非MHC適合細胞である場合、それらは、導入細胞に対する宿主免疫応答の発生を防ぐ既知の技法を用いて、投与され得る。例えば、細胞は、じかに接している細胞外環境との構成成分の交換が可能であるものの、導入細胞が宿主免疫系により認識されることが不可能なカプセル化形態で導入されてもよい。

【0303】

本発明のトランスジェニックおよび「ノックアウト」動物は、本発明のCSGポリペプチドの生物機能を説明し、異常なCSGの発現に関連した症状および/または障害を研究し、かかるCSG関連症状および/または疾患を改善するのに効果的な化合物をスクリーニングするのに有用な動物モデル系を含むが、これらに限定されない用途を有する。

【0304】

例

本発明は以下の例によりさらに詳細に説明される。本例は具体的な態様を参照することで発明を例示するためにのみ提供される。これら例示は、本発明のある特定の側面を記述するものの、開示された発明の範囲を限定あるいは制限するものではない。

【0305】

全ての例は、特に詳細記載された場合を除き、当業者にとって既知、かつ日常的である標準的な技術を用いて実施された。以下の例の日常的な分子生物学的技術は、例えばSambrook et al., MOLECULAR CLONING: A LABORATORY MANUAL, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989)に記載されているとおりに実施された。

【0306】

CSGの同定

CSG(大腸特異遺伝子)の同定は、本明細書中でCLASPと呼ぶデータマイニング、癌情報自動検索パッケージ(Cancer Leads Automatic Search Package)を用いた、Incyte Pharmaceuticals, Palo Alto, CAから入手可能なデータベースであるLIFESeq Gold中のデータの系統的分析により行なわれた。

【0307】

CLASPは以下の工程を行なう。最初に、全ての他の器官に対する標的器官中の対応するESTの存在度レベルに基づき、高発現器官特異遺伝子が選択される。次に、それぞれの高発現器官特異遺伝子の発現レベルが、正常組織、腫瘍組織、および腫瘍または疾患に関連する組織ライブラリーにおいて分析される。候補は、ESTの構成要素の証明、ならびに、腫瘍組織または腫瘍ライブラリーにおける専らのまたはより頻繁な発現に基づいて選択される。

【0308】

こうして、CLASPにより、高発現器官および癌特異遺伝子の同定が可能となる。次に、最終的な手動による詳細な評価を行ない、遺伝子選択の仕上げをする。

CLASP法を用いて、次のIncyte配列が、CSGとして同定された。

【0309】

【表1】

10

20

30

40

配列番号: LSGold 遺伝子ID

1	237623
2	234891
3	262167
4	246508
5	203279
6	983538
7	206344
8	222237
9	118593
10	337950
11	982786
12	398963
13	203640
14	88875
15	230552
16	407124
17	62662
18	230495
19	470880
20	898601
21	29586
22	370788

10

20

【0310】

遺伝子発現の相対的定量

蛍光性タックマン (Taqman) プロブを用いたリアルタイム定量PCRは、Taq DNAポリメラーゼの5' - 3'ヌクレアーゼ活性を利用する定量検出システムである。この方法は、5'をレポーター色素で、下流の3'をクエンチャー色素で標識した内部蛍光オリゴヌクレオチドプロブ (タックマン) を用いる。PCRの間にTaq DNAポリメラーゼの5' - 3'ヌクレアーゼ活性がレポーターを放出し、次にその蛍光がモデル7700シーケンス検出システム (PE Applied Biosystems , Foster City , CA , USA) のレーザー検出器により検出することができる。

30

【0311】

内因性対照の増幅を使用し、反応に添加された試料RNAの量が標準化され、逆転写酵素 (RT) の効率が正規化される。シクロフィリン、グリセルアルデヒド - 3 - リン酸デヒドロゲナーゼ (GAPDH)、または18SリボソームRNA (rRNA) のいずれかがこの内部対照として用いられた。研究される全試料間の相対定量を計算するために、1つの試料についての標的RNAレベルを結果比較のための基礎 (キャリブレーター) として用いた。「キャリブレーター」に対する定量化は、標準曲線法または比較法 (User Bulletin # 2 : ABI PRISM 7700シーケンス検出システム) を用いて得ることができる。

40

【0312】

標的遺伝子の組織分布およびレベルを、正常組織および癌組織の全ての試料について決定した。全RNAは、正常組織、癌組織、および癌とそれに対応する対応隣接組織から抽出した。続いて逆転写酵素を用いて第1鎖cDNAを調製し、各標的遺伝子に特異的なプライマーとタックマンプロブとを用いてポリメラーゼ連鎖反応を行った。結果は、ABI PRISM 7700シーケンス検出装置を用いて分析した。無名数は、キャリブレーター組織と比較した場合の特定組織中にある標的遺伝子の相対発現レベルである。

【0313】

以下のプライマーをリアルタイム定量PCRに用いた。

フォワードプライマー:

50

T G G A A A T A G A T T C A G G G G T C A T (配列番号 2 3)

リバースプライマー :

C G G G T G T A C C T C A C T G A C T T C (配列番号 2 4)

Q - PCRプローブ :

T G T C T T C C G A G A G A A C C A G G C T C C G (配列番号 2 5)

【 0 3 1 4 】

表 1 に記載の無名数は、24 の異なる正常組織における、遺伝子 I D 2 0 3 2 7 9 (本明細書中では、C l n 1 2 9 または配列番号 5 と呼ばれる) の相対発現レベルである。値は全て、正常肝臓 (キャリブレーター) と比較された。これら R N A 試料は、異なる個体由来の特定組織の試料をプールすることにより得られた市販のプールである。

【 0 3 1 5 】

【 表 2 】

表 1 : プール試料における CSG Cln129 発現の相対レベル

組織	正常
副腎	0
膀胱	0
脳	0
子宮頸部	0
大腸	0.7
子宮内膜	0.4
食道	0
心臓	0
腎臓	3.7
肝臓	1
肺	0
乳腺	0.2
筋	0
卵巣	0
膵臓	0
前立腺	0
直腸	23
小腸	1.5
脾臓	0
胃	0.8
精巣	0.1
胸腺	0.4
気管	0
子宮	0

【 0 3 1 6 】

表 1 中の発現の相対レベルは、C l n 1 2 9 m R N A の発現が、正常直腸のプール中で高レベルに (2 3)、そして、腎臓でより低レベルに (3 . 7) 検出されたことを示している。これに対し、C l n 1 2 9 は、分析した他の 2 2 の正常組織プールにおいて、極めて低レベルで発現されていた。さらに、直腸での発現レベルは、腎臓での発現に比べ 6 倍高い。これらの結果は、C l n 1 2 9 m R N A の発現が、直腸組織に高度に特異的であることを証明する。

【 0 3 1 7 】

表 1 の無名数は、異なる個体由来の特定組織試料のプールを分析して得たものである。それらは、表 2 の単一個体の組織試料より得た R N A に由来する無名数とは比較できない。表 2 に記載の無名数は、21 対の対応試料の C l n 1 2 9 の発現の相対レベルである。全ての値が正常肝臓 (キャリブレーター) と比較されている。対応ペアは、特定組織に関する癌試料に由来する m R N A、および単一個体由来の同一組織に関する正常隣接試料に由来する m R N A により形成される。

10

20

30

40

50

【0318】

【表3】

表2：個体試料におけるCSG Cln129発現の相対レベル

試料 ID	組織	癌	正常
ClnAS98	大腸 上行結腸 (C)1	383	24
ClnCM67	大腸 盲腸 (B)2	15	8
ClnCXGA	大腸 直腸 (A)3	85	118
ClnMT38	大腸 S状結腸 (D)4	33	18
ClnRC24	大腸 直腸 (D)5	77	29
ClnRC67	大腸 直腸 (B)6	0.9	15
ClnRS45	大腸 直腸S状結腸 (C)7	161	25
ClnSG27	大腸 S状結腸 (C)8	48	13
ClnSG33	大腸 S状結腸 (C)9	190	100
ClnSG36	大腸 S状結腸 (E)10	186	93
ClnRC89	大腸 直腸 (D)11	0	28
Bld32XK	膀胱 1	0	0
CvxKS52	子宮頸部1	0	0
Endo8XA	子宮内膜 1	0	0.7
Kid106XD	腎臓 1	0	6.7
Liv15XA	肝臓 1	1.7	3.2
Lng47XQ	肺 1	3.4	0
Mam59X	乳腺 1	1.3	0
Pro34B	前立腺 1	0	0
SmInt	小腸 1	5.4	1.7
Utr85XU	子宮 1	0.9	0

0= 陰性

【0319】

11の異なる組織を代表する表2中の42の試料のうち、有意な発現は、大腸、腎臓および小腸組織でのみ見られた。これらの結果は、表1に示した正常試料により得られた組織特異性の結果を確認するものである。表1および表2は、合わせて合計24のヒト組織種中の66の試料を表している。大腸および直腸とは異なる22の組織種を代表する合計42の試料のうち、1つの小腸試料、1つの肺試料、1つの肝臓試料、および1つの腎臓試料のみが、Cln129の発現を示した。

【0320】

同一の個体からの大腸癌試料と正常隣接組織におけるmRNA発現のレベルの比較が表2に示されている。Cln129は、11の癌試料のうち8において(73%) (大腸1、2、4、5、7、8、9、10)正常隣接組織に比べ、より高レベルで発現していた。総じて、高レベルの組織特異性に加え、試験した大腸癌対応試料の73%におけるmRNAの上方調節は、Cln129が大腸癌の診断マーカーであることを示している。

【0321】

本発明は、上述の説明および例に特に記載したものと違うように実施されてもよいことは明らかであろう。上記の教示に鑑みて、本発明の数多くの変更およびバリエーションが可能であり、したがってそれらは、併記の特許請求の範囲内である。

発明の背景、詳細な説明および例において引用した各文献(特許、特許出願、論文、要約、実験室マニュアル、書籍、または他の開示を含む)の開示全体は、参照により本明細書に組み込まれる。さらに、本明細書とともに提出する配列表のハードコピーおよび相当するコンピュータ読取り可能形態の両方が、全体として参照により本明細書中に援用される。

。

【国際公開パンフレット】

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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
6 December 2001 (06.12.2001)

PCT

(10) International Publication Number
WO 01/92528 A2

(51) International Patent Classification: C12N 15/12, C07K 14/07, C12Q 1/68, G01N 33/574, C07K 16/18, A61K 31/00, B93B95, B93B17, A61P 35/00

RECIPON, Hervé [FR/FR], 85 Futuna Avenue, San Francisco, CA 94115 (US)

(21) International Application Number: PCT/US01/17583

(74) Agents: LICATA, Jane, Massey et al., Licata & Tyrrell P.C., 66 E. Main Street, Madison, NJ 08053 (US)

(22) International Filing Date: 29 May 2001 (29.05.2001)

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

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data: 60/207,583 26 May 2000 (26.05.2000) US

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

(71) Applicant for all designated States except US: DI-ADENUS, INC. [US/US], 3505 Octavius Drive, Santa Clara, CA 95054 (US)

(72) Inventors and Applicants (for US only): MACINA, Roberto, A. [AR/US], 4118 Crescendo Avenue, San Jose, CA 95136 (US); CHEN, Sei-Yu [—/US], 160 Mira Street, Foster City, CA 94044 (US); PILLTA, Jason [US/US], Apartment 15, 1240 Dale Avenue, Mountain View, CA 94040 (US); SUN, Yongming [US/US], Apartment 260, 869 S. Winchester Boulevard, San Jose, CA 95128 (US).

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

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



WO 01/92528 A2

(54) Title: METHOD OF DIAGNOSING, MONITORING, STAGING, IMAGING AND TREATING COLON CANCER

(57) Abstract: The invention relates to CSG polypeptides, polynucleotides encoding the polypeptides, methods for producing the polypeptides, in particular by expressing the polynucleotides, and agonists and antagonists of the polypeptides. The invention further relates to methods for utilizing such polynucleotides, polypeptides, agonists and antagonists for applications, which relate, in part, to research, diagnostic and clinical arts.

WO 01/92528

PCT/US01/17583

- 1 -

METHOD OF DIAGNOSING, MONITORING, STAGING, IMAGING AND
TREATING COLON CANCER

FIELD OF THE INVENTION

5 This invention relates, in part, to newly identified polynucleotides and polypeptides; variants and derivatives of the polynucleotides and polypeptides; processes for making the polynucleotides and the polypeptides, and their variants and derivatives; agonists and antagonists of the polypeptides; and
10 uses of the polynucleotides, polypeptides, variants, derivatives, agonists and antagonists for detecting, diagnosing, monitoring, staging, prognosticating, imaging and treating cancers, particularly colon cancer. In particular, in these and in other regards, the invention relates to colon
15 specific polynucleotides and polypeptides hereinafter referred to as colon specific genes or "CSGs".

BACKGROUND OF THE INVENTION

Cancer of the colon is a highly treatable and often
20 curable disease when localized to the bowel. It is one of the most frequently diagnosed malignancies in the United States as well as the second most common cause of cancer death. Surgery is the primary treatment and results in cure in approximately 50% of patients. However, recurrence following
25 surgery is a major problem and often is the ultimate cause of death.

The prognosis of colon cancer is clearly related to the degree of penetration of the tumor through the bowel wall and the presence or absence of nodal involvement. These two
30 characteristics form the basis for all staging systems developed for this disease. Treatment decisions are usually

WO 01/92528

PCT/US01/17583

- 2 -

made in reference to the older Duke's or the Modified Astler-Coller (MAC) classification scheme for staging.

Bowel obstruction and bowel perforation are indicators of poor prognosis in patients with colon cancer. Elevated pretreatment serum levels of carcinoembryonic antigen (CEA) and of carbohydrate antigen 19-9 (CA 19-9) also have a negative prognostic significance.

Age greater than 70 years at presentation is not a contraindication to standard therapies. Acceptable morbidity and mortality, as well as long-term survival, are achieved in this patient population.

Because of the frequency of the disease (approximately 160,000 new cases of colon and rectal cancer per year), the identification of high-risk groups, the demonstrated slow growth of primary lesions, the better survival of early-stage lesions, and the relative simplicity and accuracy of screening tests, screening for colon cancer should be a part of routine care for all adults starting at age 50, especially those with first-degree relatives with colorectal cancer.

Procedures used for detecting, diagnosing, monitoring, staging, and prognosticating colon cancer are of critical importance to the outcome of the patient. For example, patients diagnosed with early colon cancer generally have a much greater five-year survival rate as compared to the survival rate for patients diagnosed with distant metastasized colon cancer. New diagnostic methods which are more sensitive and specific for detecting early colon cancer are clearly needed.

Colon cancer patients are closely monitored following initial therapy and during adjuvant therapy to determine response to therapy and to detect persistent or recurrent disease or metastasis. There is clearly a need for a colon cancer marker which is more sensitive and specific in detecting colon cancer, its recurrence, and progression.

WO 01/92528

PCT/US01/17583

- 3 -

Another important step in managing colon cancer is to determine the stage of the patient's disease. Stage determination has potential prognostic value and provides criteria for designing optimal therapy. Generally, 5 pathological staging of colon cancer is preferable over clinical staging because the former gives a more accurate prognosis. However, clinical staging would be preferred were it at least as accurate as pathological staging because it does not depend on an invasive procedure to obtain tissue for 10 pathological evaluation. Staging of colon cancer would be improved by detecting new markers in cells, tissues, or bodily fluids which could differentiate between different stages of invasion.

Accordingly, there is a great need for more sensitive 15 and accurate methods for the staging of colon cancer in a human to determine whether or not such cancer has metastasized and for monitoring the progress of colon cancer in a human which has not metastasized for the onset of metastasis.

In the present invention, methods are provided for 20 detecting, diagnosing, monitoring, staging, prognosticating, imaging and treating colon cancer via colon specific genes referred to herein as CSGs. For purposes of the present invention, CSG refers, among other things, to native protein expressed by the gene comprising a polynucleotide sequence of 25 SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22. By "CSG" it is also meant herein polynucleotides which, due to degeneracy in genetic coding, comprise variations in nucleotide sequence as compared to SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 30 15, 16, 17, 18, 19, 20, 21 or 22 but which still encode the same protein. In the alternative, what is meant by CSG as used herein, means the native mRNA encoded by the gene comprising the polynucleotide sequence of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 35 21 or 22, levels of the gene comprising the polynucleotide

WO 01/92528

PCT/US01/17583

- 4 -

sequence of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, or 22, or levels of a polynucleotide which is capable of hybridizing under stringent conditions to the antisense sequence of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22.

Other objects, features, advantages and aspects of the present invention will become apparent to those of skill in the art from the following description. It should be understood, however, that the following description and the specific examples, while indicating preferred embodiments of the invention are given by way of illustration only. Various changes and modifications within the spirit and scope of the disclosed invention will become readily apparent to those skilled in the art from reading the following description and from reading the other parts of the present disclosure.

SUMMARY OF THE INVENTION

Toward these ends, and others, it is an object of the present invention to provide CSGs comprising a polynucleotide of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22, a protein expressed by a polynucleotide of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22 or a variant thereof which expresses the protein; or a polynucleotide which is capable of hybridizing under stringent conditions to the antisense sequence of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22.

It is another object of the present invention to provide a method for diagnosing the presence of colon cancer by analyzing for changes in levels of CSG in cells, tissues or bodily fluids compared with levels of CSG in preferably the same cells, tissues, or bodily fluid type of a normal human control, wherein a change in levels of CSG in the patient

WO 01/92528

PCT/US01/17583

- 5 -

versus the normal human control is associated with colon cancer.

Further provided is a method of diagnosing metastatic colon cancer in a patient having colon cancer which is not
5 known to have metastasized by identifying a human patient suspected of having colon cancer that has metastasized; analyzing a sample of cells, tissues, or bodily fluid from such patient for CSG; comparing the CSG levels in such cells, tissues, or bodily fluid with levels of CSG in preferably the
10 same cells, tissues, or bodily fluid type of a normal human control, wherein an increase in CSG levels in the patient versus the normal human control is associated with colon cancer which has metastasized.

Also provided by the invention is a method of staging
15 colon cancer in a human which has such cancer by identifying a human patient having such cancer; analyzing a sample of cells, tissues, or bodily fluid from such patient for CSG; comparing CSG levels in such cells, tissues, or bodily fluid with levels of CSG in preferably the same cells, tissues, or
20 bodily fluid type of a normal human control sample, wherein an increase in CSG levels in the patient versus the normal human control is associated with a cancer which is progressing and a decrease in the levels of CSG is associated with a cancer which is regressing or in remission.

Further provided is a method of monitoring colon cancer
25 in a human having such cancer for the onset of metastasis. The method comprises identifying a human patient having such cancer that is not known to have metastasized; periodically analyzing a sample of cells, tissues, or bodily fluid from
30 such patient for CSG; comparing the CSG levels in such cells, tissue, or bodily fluid with levels of CSG in preferably the same cells, tissues, or bodily fluid type of a normal human control sample, wherein an increase in CSG levels in the patient versus the normal human control is associated with a
35 cancer which has metastasized.

WO 01/92528

PCT/US01/17583

- 6 -

Further provided is a method of monitoring the change in stage of colon cancer in a human having such cancer by looking at levels of CSG in a human having such cancer. The method comprises identifying a human patient having such cancer; periodically analyzing a sample of cells, tissues, or bodily fluid from such patient for CSG; comparing the CSG levels in such cells, tissue, or bodily fluid with levels of CSG in preferably the same cells, tissues, or bodily fluid type of a normal human control sample, wherein an increase in CSG levels in the patient versus the normal human control is associated with a cancer which is progressing and a decrease in the levels of CSG is associated with a cancer which is regressing or in remission.

Further provided are methods of designing new therapeutic agents targeted to a CSG for use in imaging and treating colon cancer. For example, in one embodiment, therapeutic agents such as antibodies targeted against CSG or fragments of such antibodies can be used to treat, detect or image localization of CSG in a patient for the purpose of detecting or diagnosing a disease or condition. In this embodiment, an increase in the amount of labeled antibody detected as compared to normal tissue would be indicative of tumor metastases or growth. Such antibodies can be polyclonal, monoclonal, or omniclonal or prepared by molecular biology techniques. The term "antibody", as used herein and throughout the instant specification is also meant to include aptamers and single-stranded oligonucleotides such as those derived from an *in vitro* evolution protocol referred to as SELEX and well known to those skilled in the art. Antibodies can be labeled with a variety of detectable and therapeutic labels including, but not limited to, radioisotopes and paramagnetic metals. Therapeutic agents such as small molecules and antibodies which decrease the concentration and/or activity of CSG can also be used in the treatment of

WO 01/92528

PCT/US01/17583

- 7 -

diseases characterized by overexpression of CSG. Such agents can be readily identified in accordance with teachings herein.

Other objects, features, advantages and aspects of the present invention will become apparent to those of skill in the art from the following description. It should be understood, however, that the following description and the specific examples, while indicating preferred embodiments of the invention, are given by way of illustration only. Various changes and modifications within the spirit and scope of the disclosed invention will become readily apparent to those skilled in the art from reading the following description and from reading the other parts of the present disclosure.

GLOSSARY

The following illustrative explanations are provided to facilitate understanding of certain terms used frequently herein, particularly in the examples. The explanations are provided as a convenience and are not limitative of the invention.

ISOLATED means altered "by the hand of man" from its natural state; i.e., that, if it occurs in nature, it has been changed or removed from its original environment, or both.

For example, a naturally occurring polynucleotide or a polypeptide naturally present in a living animal in its natural state is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein. For example, with respect to polynucleotides, the term isolated means that it is separated from the chromosome and cell in which it naturally occurs.

As part of or following isolation, such polynucleotides can be joined to other polynucleotides, such as DNAs, for mutagenesis, to form fusion proteins, and for propagation or expression in a host, for instance. The isolated polynucleotides, alone or joined to other polynucleotides such

WO 01/92528

PCT/US01/17583

- 8 -

as vectors, can be introduced into host cells, in culture or in whole organisms. When introduced into host cells in culture or in whole organisms, such DNAs still would be isolated, as the term is used herein, because they would not be in their naturally occurring form or environment. Similarly, the polynucleotides and polypeptides may occur in a composition, such as media formulations, solutions for introduction of polynucleotides or polypeptides, for example, into cells, compositions or solutions for chemical or enzymatic reactions, for instance, which are not naturally occurring compositions, and, therein remain isolated polynucleotides or polypeptides within the meaning of that term as it is employed herein.

OLIGONUCLEOTIDE(S) refers to relatively short polynucleotides. Often the term refers to single-stranded deoxyribonucleotides, but it can refer as well to single- or double-stranded ribonucleotides, RNA:DNA hybrids and double-stranded DNAs, among others.

Oligonucleotides, such as single-stranded DNA probe oligonucleotides, often are synthesized by chemical methods, such as those implemented on automated oligonucleotide synthesizers. However, oligonucleotides can be made by a variety of other methods, including in vitro recombinant DNA-mediated techniques and by expression of DNAs in cells and organisms.

Initially, chemically synthesized DNAs typically are obtained without a 5' phosphate. The 5' ends of such oligonucleotides are not substrates for phosphodiester bond formation by ligation reactions that employ DNA ligases typically used to form recombinant DNA molecules. Where ligation of such oligonucleotides is desired, a phosphate can be added by standard techniques, such as those that employ a kinase and ATP.

The 3' end of a chemically synthesized oligonucleotide generally has a free hydroxyl group and, in the presence of a ligase such as T4 DNA ligase, readily will form a

WO 01/92528

PCT/US01/17583

- 9 -

phosphodiester bond with a 5' phosphate of another polynucleotide, such as another oligonucleotide. As is well known, this reaction can be prevented selectively, where desired, by removing the 5' phosphates of the other polynucleotide(s) prior to ligation.

POLYNUCLEOTIDE(S) generally refers to any polyribonucleotide or polydeoxyribonucleotide and is inclusive of unmodified RNA or DNA as well as modified RNA or DNA. Thus, for instance, polynucleotides as used herein refers to, among other things, single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, polynucleotide, as used herein, refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The strands in such regions may be from the same molecule or from different molecules. The regions may include all of one or more of the molecules, but more typically involve only a region of some of the molecules. One of the molecules of a triple-helical region often is an oligonucleotide.

As used herein, the term polynucleotide is also inclusive of DNAs or RNAs as described above that contain one or more modified bases. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are "polynucleotides" as that term is intended herein. Moreover, DNAs or RNAs comprising unusual bases, such as inosine, or modified bases, such as tritylated bases, to name just two examples, are polynucleotides as the term is used herein.

It will be appreciated that a great variety of modifications have been made to DNA and RNA that serve many useful purposes known to those of skill in the art. The term polynucleotide as it is employed herein embraces such

WO 01/92528

PCT/US01/17583

- 10 -

chemically, enzymatically or metabolically modified forms of polynucleotides, as well as chemical forms of DNA and RNA characteristic of viruses and cells, including simple and complex cells, inter alia.

5 POLYPEPTIDES, as used herein, includes all polypeptides as described below. The basic structure of polypeptides is well known and has been described in innumerable textbooks and other publications in the art. In this context, the term is used herein to refer to any peptide or protein comprising two
10 or more amino acids joined to each other in a linear chain by peptide bonds. As used herein, the term refers to both short chains, which also commonly are referred to in the art as peptides, oligopeptides and oligomers, for example, and to longer chains, which generally are referred to in the art as
15 proteins, of which there are many types. It will be appreciated that polypeptides often contain amino acids other than the 20 amino acids commonly referred to as the 20 naturally occurring amino acids, and that many amino acids, including the terminal amino acids, may be modified in a given
20 polypeptide, either by natural processes such as processing and other post-translational modifications, or by chemical modification techniques which are well known to the art. Even the common modifications that occur naturally in polypeptides are too numerous to list exhaustively here, but they are well
25 described in basic texts and in more detailed monographs, as well as in a voluminous research literature, and they are well known to those of skill in the art.

Modifications which may be present in polypeptides of the present invention include, to name an illustrative few,
30 acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent
attachment of phosphatidylinositol, cross-linking,
35 cyclization, disulfide bond formation, demethylation,

WO 01/92528

PCT/US01/17583

- 11 -

formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination.

Such modifications are well known to those of skill and have been described in great detail in the scientific literature. Several particularly common modifications including, but not limited to, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation are described in most basic texts, such as, for instance PROTEINS STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993). Many detailed reviews are available on this subject, such as, for example, those provided by Wold, F., Posttranslational Protein Modifications: Perspectives and Prospects, pgs. 1-12 in POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York (1983); Seifter et al., Analysis for protein modifications and nonprotein cofactors, Meth. Enzymol. 182: 626-646 (1990) and Rattan et al., Protein Synthesis: Posttranslational Modifications and Aging, Ann. N.Y. Acad. Sci. 663: 48-62 (1992).

It will be appreciated that the polypeptides of the present invention are not always entirely linear. Instead, polypeptides may be branched as a result of ubiquitination, and they may be circular, with or without branching, generally as a result of posttranslation events including natural processing event and events brought about by human manipulation which do not occur naturally. Circular, branched and branched circular polypeptides may be synthesized by non-

WO 01/92528

PCT/US01/17583

- 12 -

translation natural processes and by entirely synthetic methods, as well.

Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. In fact, blockage of the amino and/or carboxyl group in a polypeptide by a covalent modification is common in naturally occurring and synthetic polypeptides and such modifications may be present in polypeptides of the present invention, as well. For instance, the amino terminal residue of polypeptides made in *E. coli*, prior to proteolytic processing, almost invariably will be N-formylmethionine.

The modifications that occur in a polypeptide often will be a function of how it is made. For polypeptides made by expressing a cloned gene in a host, for instance, the nature and extent of the modifications, in large part, will be determined by the host cell posttranslational modification capacity and the modification signals present in the polypeptide amino acid sequence. For instance, as is well known, glycosylation often does not occur in bacterial hosts such as *E. coli*. Accordingly, when glycosylation is desired, a polypeptide can be expressed in a glycosylating host, generally a eukaryotic cell. Insect cells often carry out the same posttranslational glycosylations as mammalian cells. Thus, insect cell expression systems have been developed to express efficiently mammalian proteins having native patterns of glycosylation, *inter alia*. Similar considerations apply to other modifications.

It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications.

In general, as used herein, the term polypeptide encompasses all such modifications, particularly those that

WO 01/92528

PCT/US01/17583

- 13 -

are present in polypeptides synthesized by expressing a polynucleotide in a host cell.

VARIANT(S) of polynucleotides or polypeptides, as the term is used herein, are polynucleotides or polypeptides that
5 differ from a reference polynucleotide or polypeptide, respectively.

With respect to variant polynucleotides, differences are generally limited so that the nucleotide sequences of the reference and the variant are closely similar overall and, in
10 many regions, identical. Thus, changes in the nucleotide sequence of the variant may be silent. That is, they may not alter the amino acids encoded by the polynucleotide. Where alterations are limited to silent changes of this type a variant will encode a polypeptide with the same amino acid
15 sequence as the reference. Alternatively, changes in the nucleotide sequence of the variant may alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Such nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and
20 truncations in the polypeptide encoded by the reference sequence.

With respect to variant polypeptides, differences are generally limited so that the sequences of the reference and the variant are closely similar overall and, in many region,
25 identical. For example, a variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions, fusions and truncations, which may be present in any combination.

RECEPTOR MOLECULE, as used herein, refers to molecules
30 which bind or interact specifically with CSG polypeptides of the present invention and is inclusive not only of classic receptors, which are preferred, but also other molecules that specifically bind to or interact with polypeptides of the invention (which also may be referred to as "binding
35 molecules" and "interaction molecules," respectively and as

WO 01/92528

PCT/US01/17583

- 14 -

"CSG binding or interaction molecules". Binding between polypeptides of the invention and such molecules, including receptor or binding or interaction molecules may be exclusive to polypeptides of the invention, which is very highly preferred, or it may be highly specific for polypeptides of the invention, which is highly preferred, or it may be highly specific to a group of proteins that includes polypeptides of the invention, which is preferred, or it may be specific to several groups of proteins at least one of which includes polypeptides of the invention.

Receptors also may be non-naturally occurring, such as antibodies and antibody-derived reagents that bind to polypeptides of the invention.

15 DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to novel colon specific polypeptides and polynucleotides, referred to herein as CSGs, among other things, as described in greater detail below.

Polynucleotides

20 In accordance with one aspect of the present invention, there are provided isolated CSG polynucleotides which encode CSG polypeptides.

Using the information provided herein, such as the polynucleotide sequences set out in SEQ ID NO:1, 2, 3, 4, 5, 25 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, and 22, a polynucleotide of the present invention encoding a CSG may be obtained using standard cloning and screening procedures, such as those for cloning cDNAs using mRNA from cells of a human tumor as starting material.

30 Polynucleotides of the present invention may be in the form of RNA, such as mRNA, or in the form of DNA, including, for instance, cDNA and genomic DNA obtained by cloning or produced by chemical synthetic techniques or by a combination thereof. The DNA may be double-stranded or single-stranded. 35 Single-stranded DNA may be the coding strand, also known as

WO 01/92528

PCT/US01/17583

- 15 -

the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand.

The coding sequence which encodes the polypeptides may be identical to the coding sequence of the polynucleotides of SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22. It also may be a polynucleotide with a different sequence, which, as a result of the redundancy (degeneracy) of the genetic code, encodes the same polypeptides as encoded by SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22.

Polynucleotides of the present invention, such as SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22, which encode these polypeptides may comprise the coding sequence for the mature polypeptide by itself. Polynucleotides of the present invention may also comprise the coding sequence for the mature polypeptide and additional coding sequences such as those encoding a leader or secretory sequence such as a pre-, or pro- or prepro-protein sequence. Polynucleotides of the present invention may also comprise the coding sequence of the mature polypeptide, with or without the aforementioned additional coding sequences, together with additional, non-coding sequences. Examples of additional non-coding sequences which may be incorporated into the polynucleotide of the present invention include, but are not limited to, introns and non-coding 5' and 3' sequences such as transcribed, non-translated sequences that play a role in transcription, mRNA processing including, for example, splicing and polyadenylation signals, ribosome binding and stability of mRNA, and additional coding sequence which codes for amino acids such as those which provide additional functionalities. Thus, for instance, the polypeptide may be fused to a marker sequence such as a peptide which facilitates purification of the fused polypeptide. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine

WO 01/92528

PCT/US01/17583

- 16 -

peptide, such as the tag provided in the pQE vector (Qiagen, Inc.), among others, many of which are commercially available. As described in Gentz et al. (Proc. Natl. Acad. Sci., USA 86: 821-824 (1989)), for instance, hexa-histidine provides for
5 convenient purification of the fusion protein. The HA tag corresponds to an epitope derived of influenza hemagglutinin protein (Wilson et al., Cell 37: 767 (1984)).

In accordance with the foregoing, the term "polynucleotide encoding a polypeptide" as used herein
10 encompasses polynucleotides which include a sequence encoding a polypeptide of the present invention, particularly SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22. The term encompasses polynucleotides that include a single continuous region or discontinuous
15 regions encoding the polypeptide (for example, interrupted by introns) together with additional regions, that also may contain coding and/or non-coding sequences.

The present invention further relates to variants of the herein above described polynucleotides which encode for
20 fragments, analogs and derivatives of the CSG polypeptides. A variant of the polynucleotide may be a naturally occurring variant such as a naturally occurring allelic variant, or it may be a variant that is not known to occur naturally. Such non-naturally occurring variants of the polynucleotide may be
25 made by mutagenesis techniques, including those applied to polynucleotides, cells or organisms.

Among variants in this regard are variants that differ from the aforementioned polynucleotides by nucleotide
30 substitutions, deletions or additions. The substitutions, deletions or additions may involve one or more nucleotides. The variants may be altered in coding or non-coding regions or both. Alterations in the coding regions may produce conservative or non-conservative amino acid substitutions, deletions or additions.

WO 01/92528

PCT/US01/17583

- 17 -

Among the particularly preferred embodiments of the invention in this regard are polynucleotides encoding polypeptides having the same amino acid sequence encoded by a CSG polynucleotide comprising SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, or 22; variants, analogs, derivatives and fragments thereof, and fragments of the variants, analogs and derivatives. Further particularly preferred in this regard are CSG polynucleotides encoding polypeptide variants, analogs, derivatives and fragments, and variants, analogs and derivatives of the fragments, in which several, a few, 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues are substituted, deleted or added, in any combination. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the CSG. Also especially preferred in this regard are conservative substitutions. Most highly preferred are polynucleotides encoding polypeptides having the amino acid sequences as polypeptides encoded by SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22, without substitutions.

Further preferred embodiments of the invention are CSG polynucleotides that are at least 70% identical to a polynucleotide of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22, and polynucleotides which are complementary to such polynucleotides. More preferred are CSG polynucleotides that comprise a region that is at least 80% identical to a polynucleotide of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, or 22. In this regard, CSG polynucleotides at least 90% identical to the same are particularly preferred, and among these particularly preferred CSG polynucleotides, those with at least 95% are especially preferred. Furthermore, those with at least 97% are highly preferred among those with at least 95%, and among

WO 01/92528

PCT/US01/17583

- 18 -

these those with at least 98% and at least 99% are particularly highly preferred, with at least 99% being the most preferred.

Particularly preferred embodiments in this respect, 5 moreover, are polynucleotides which encode polypeptides which retain substantially the same biological function or activity as the mature polypeptides encoded by a polynucleotide of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22.

10 The present invention further relates to polynucleotides that hybridize to the herein above-described CSG sequences. In this regard, the present invention especially relates to polynucleotides which hybridize under stringent conditions to the herein above-described polynucleotides. As herein used, 15 the term "stringent conditions" means hybridization will occur only if there is at least 95% and preferably at least 97% identity between the sequences.

As discussed additionally herein regarding polynucleotide assays of the invention, for instance, 20 polynucleotides of the invention as described herein, may be used as a hybridization probe for cDNA and genomic DNA to isolate full-length cDNAs and genomic clones encoding CSGs and to isolate cDNA and genomic clones of other genes that have a high sequence similarity to these CSGs. Such probes 25 generally will comprise at least 15 bases. Preferably, such probes will have at least 30 bases and may have at least 50 bases.

For example, the coding region of CSG of the present invention may be isolated by screening using an 30 oligonucleotide probe synthesized from the known DNA sequence. A labeled oligonucleotide having a sequence complementary to that of a gene of the present invention is used to screen a library of human cDNA, genomic DNA or mRNA to determine which members of the library the probe hybridizes with.

WO 01/92528

PCT/US01/17583

- 19 -

The polynucleotides and polypeptides of the present invention may be employed as research reagents and materials for discovery of treatments and diagnostics to human disease, as further discussed herein relating to polynucleotide assays, 5 *inter alia*.

The polynucleotides may encode a polypeptide which is the mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids interior to the mature polypeptide (when the mature form has more than one polypeptide chain, for 10 instance). Such sequences may play a role in processing of a protein from precursor to a mature form, may facilitate/protein trafficking, may prolong or shorten protein half-life or may facilitate manipulation of a protein for assay or production, among other things. As generally is the 15 case in situ, the additional amino acids may be processed away from the mature protein by cellular enzymes.

A precursor protein having the mature form of the polypeptide fused to one or more prosequences may be an inactive form of the polypeptide. When prosequences are 20 removed, such inactive precursors generally are activated. Some or all of the prosequences may be removed before activation. Generally, such precursors are called proproteins.

In sum, a polynucleotide of the present invention may 25 encode a mature protein, a mature protein plus a leader sequence (which may be referred to as a preprotein), a precursor of a mature protein having one or more prosequences which are not the leader sequences of a preprotein, or a preproprotein, which is a precursor to a proprotein, having 30 a leader sequence and one or more prosequences, which generally are removed during processing steps that produce active and mature forms of the polypeptide.

Polypeptides

The present invention further relates to CSG 35 polypeptides, preferably polypeptides encoded by a

WO 01/92528

PCT/US01/17583

- 20 -

polynucleotide of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10,
11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, or 22. The
invention also relates to fragments, analogs and derivatives
of these polypeptides. The terms "fragment," "derivative" and
5 "analog" when referring to the polypeptides of the present
invention means a polypeptide which retains essentially the
same biological function or activity as such polypeptides.
Thus, an analog includes a proprotein which can be activated
by cleavage of the proprotein portion to produce an active
10 mature polypeptide.

The polypeptide of the present invention may be a
recombinant polypeptide, a natural polypeptide or a synthetic
polypeptide. In certain preferred embodiments it is a
recombinant polypeptide.

15 The fragment, derivative or analog of a polypeptide of
or the present invention may be (i) one in which one or more
of the amino acid residues are substituted with a conserved
or non-conserved amino acid residue (preferably a conserved
amino acid residue) and such substituted amino acid residue
20 may or may not be one encoded by the genetic code; (ii) one
in which one or more of the amino acid residues includes a
substituent group; (iii) one in which the mature polypeptide
is fused with another compound, such as a compound to increase
the half-life of the polypeptide (for example, polyethylene
25 glycol); or (iv) one in which the additional amino acids are
fused to the mature polypeptide, such as a leader or secretory
sequence or a sequence which is employed for purification of
the mature polypeptide or a proprotein sequence. Such
fragments, derivatives and analogs are deemed to be within the
30 scope of those skilled in the art from the teachings herein.

Among preferred variants are those that vary from a
reference by conservative amino acid substitutions. Such
substitutions are those that substitute a given amino acid in
a polypeptide by another amino acid of like characteristics.
35 Typically seen as conservative substitutions are the

WO 01/92528

PCT/US01/17583

- 21 -

replacements, one for another, among the aliphatic amino acids Ala, Val, Leu and Ile; interchange of the hydroxyl residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange
5 of the basic residues Lys and Arg and replacements among the aromatic residues Phe, Tyr.

The polypeptides and polynucleotides of the present invention are preferably provided in an isolated form, and preferably are purified to homogeneity.

10 The polypeptides of the present invention include the polypeptide encoded by the polynucleotide of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22 (in particular the mature polypeptide) as well as polypeptides which have at least 75% similarity
15 (preferably at least 75% identity), more preferably at least 90% similarity (more preferably at least 90% identity), still more preferably at least 95% similarity (still more preferably at least 95% identity), to a polypeptide encoded by SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18,
20 19, 20, 21, or 22. Also included are portions of such polypeptides generally containing at least 30 amino acids and more preferably at least 50 amino acids.

As known in the art "similarity" between two polypeptides is determined by comparing the amino acid
25 sequence and its conserved amino acid substitutes of one polypeptide sequence with that of a second polypeptide.

Fragments or portions of the polypeptides of the present invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, the
30 fragments may be employed as intermediates for producing the full-length polypeptides. Fragments or portions of the polynucleotides of the present invention may be used to synthesize full-length polynucleotides of the present invention.

WO 01/92528

PCT/US01/17583

- 22 -

Fragments

Also among preferred embodiments of this aspect of the present invention are polypeptides comprising fragments of a polypeptide encoded by a polynucleotide of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22. In this regard a fragment is a polypeptide having an amino acid sequence that entirely is the same as part but not all of the amino acid sequence of the aforementioned CSG polypeptides and variants or derivatives thereof.

Such fragments may be "free-standing," i.e., not part of or fused to other amino acids or polypeptides, or they may be contained within a larger polypeptide of which they form a part or region. When contained within a larger polypeptide, the presently discussed fragments most preferably form a single continuous region. However, several fragments may be comprised within a single larger polypeptide. For instance, certain preferred embodiments relate to a fragment of a CSG polypeptide of the present comprised within a precursor polypeptide designed for expression in a host and having heterologous pre- and pro-polypeptide regions fused to the amino terminus of the CSG fragment and an additional region fused to the carboxyl terminus of the fragment. Therefore, fragments in one aspect of the meaning intended herein, refers to the portion or portions of a fusion polypeptide or fusion protein derived from a CSG polypeptide.

As representative examples of polypeptide fragments of the invention, there may be mentioned those which have from about 15 to about 139 amino acids. In this context "about" includes the particularly recited range and ranges larger or smaller by several, a few, 5, 4, 3, 2 or 1 amino acid at either extreme or at both extremes. Highly preferred in this regard are the recited ranges plus or minus as many as 5 amino acids at either or at both extremes. Particularly highly preferred are the recited ranges plus or minus as many as 3

WO 01/92528

PCT/US01/17583

- 23 -

amino acids at either or at both the recited extremes. Especially preferred are ranges plus or minus 1 amino acid at either or at both extremes or the recited ranges with no additions or deletions. Most highly preferred of all in this regard are fragments from about 15 to about 45 amino acids.

Among especially preferred fragments of the invention are truncation mutants of the CSG polypeptides. Truncation mutants include CSG polypeptides having an amino acid sequence encoded by a polynucleotide of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22, or variants or derivatives thereof, except for deletion of a continuous series of residues (that is, a continuous region, part or portion) that includes the amino terminus, or a continuous series of residues that includes the carboxyl terminus or, as in double truncation mutants, deletion of two continuous series of residues, one including the amino terminus and one including the carboxyl terminus. Fragments having the size ranges set out herein also are preferred embodiments of truncation fragments, which are especially preferred among fragments generally.

Also preferred in this aspect of the invention are fragments characterized by structural or functional attributes of the CSG polypeptides of the present invention. Preferred embodiments of the invention in this regard include fragments that comprise alpha-helix and alpha-helix forming regions ("alpha-regions"), beta-sheet and beta-sheet-forming regions ("beta-regions"), turn and turn-forming regions ("turn-regions"), coil and coil-forming regions ("coil-regions"), hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions and high antigenic index regions of the CSG polypeptides of the present invention. Regions of the aforementioned types are identified routinely by analysis of the amino acid sequences encoded by the polynucleotides of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16,

WO 01/92528

PCT/US01/17583

- 24 -

17, 18, 19, 20, 21 or 22. Preferred regions include Garnier-Robson alpha-regions, beta-regions, turn-regions and coil-regions, Chou-Fasman alpha-regions, beta-regions and turn-regions, Kyte-Doolittle hydrophilic regions and hydrophilic regions, Eisenberg alpha and beta amphipathic regions, Karplus-Schulz flexible regions, Emini surface-forming regions and Jameson-Wolf high antigenic index regions. Among highly preferred fragments in this regard are those that comprise regions of CSGs that combine several structural features, such as several of the features set out above. In this regard, the regions defined by selected residues of a CSG polypeptide which all are characterized by amino acid compositions highly characteristic of turn-regions, hydrophilic regions, flexible-regions, surface-forming regions, and high antigenic index-regions, are especially highly preferred regions. Such regions may be comprised within a larger polypeptide or may be by themselves a preferred fragment of the present invention, as discussed above. It will be appreciated that the term "about" as used in this paragraph has the meaning set out above regarding fragments in general.

Further preferred regions are those that mediate activities of CSG polypeptides. Most highly preferred in this regard are fragments that have a chemical, biological or other activity of a CSG polypeptide, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Highly preferred in this regard are fragments that contain regions that are homologs in sequence, or in position, or in both sequence and to active regions of related polypeptides, and which include colon specific-binding proteins. Among particularly preferred fragments in these regards are truncation mutants, as discussed above.

It will be appreciated that the invention also relates to polynucleotides encoding the aforementioned fragments, polynucleotides that hybridize to polynucleotides encoding the fragments, particularly those that hybridize under stringent

- 25 -

conditions, and polynucleotides such as PCR primers for amplifying polynucleotides that encode the fragments. In these regards, preferred polynucleotides are those that correspond to the preferred fragments, as discussed above.

5 Fusion Proteins

In one embodiment of the present invention, the CSG polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of the present
10 polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See also EP A 394,827; Trauneker, et al., Nature 331: 84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the
15 the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create
chimeric molecules having more than one function. Finally,
20 fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of these types of fusion proteins described above can be made in accordance with well known protocols.

For example, a CSG polypeptide can be fused to an IgG
25 molecule via the following protocol. Briefly, the human Fc portion of the IgG molecule is PCR amplified using primers that span the 5' and 3' ends of the sequence. These primers also have convenient restriction enzyme sites that facilitate
cloning into an expression vector, preferably a mammalian
30 expression vector. For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. In this protocol, the 3' BamHI site must be destroyed. Next, the vector containing the human Fc portion
is re-restricted with BamHI thereby linearizing the vector,
35 and a CSG polynucleotide of the present invention is ligated

WO 01/92528

PCT/US01/17583

- 26 -

into this BamHI site. It is preferred that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced.

If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e. g., WO 96/34891.)

10 Diagnostic Assays

The present invention also relates to diagnostic assays and methods, both quantitative and qualitative for detecting, diagnosing, monitoring, staging and prognosticating cancers by comparing levels of CSG in a human patient with those of CSG in a normal human control. For purposes of the present invention, what is meant by CSG levels is, among other things, native protein expressed by a gene comprising the polynucleotide sequence of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22. By "CSG" it is also meant herein polynucleotides which, due to degeneracy in genetic coding, comprise variations in nucleotide sequence as compared to SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22 but which still encode the same protein. The native protein being detected may be whole, a breakdown product, a complex of molecules or chemically modified. In the alternative, what is meant by CSG as used herein, means the native mRNA encoded by a polynucleotide sequence of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22, levels of the gene comprising the polynucleotide sequence of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, or 22, or levels of a polynucleotide which is capable of hybridizing under stringent conditions to the antisense sequence of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16,

WO 01/92528

PCT/US01/17583

- 27 -

17, 18, 19, 20, 21, or 22. Such levels are preferably determined in at least one of cells, tissues and/or bodily fluids, including determination of normal and abnormal levels. Thus, for instance, a diagnostic assay in accordance with the invention for diagnosing overexpression of CSG protein compared to normal control bodily fluids, cells, or tissue samples may be used to diagnose the presence of colon cancer.

All the methods of the present invention may optionally include determining the levels of other cancer markers as well as CSG. Other cancer markers, in addition to CSG, useful in the present invention will depend on the cancer being tested and are known to those of skill in the art.

The present invention provides methods for diagnosing the presence of colon cancer by analyzing for changes in levels of CSG in cells, tissues or bodily fluids compared with levels of CSG in cells, tissues or bodily fluids of preferably the same type from a normal human control, wherein an increase in levels of CSG in the patient versus the normal human control is associated with the presence of colon cancer.

Without limiting the instant invention, typically, for a quantitative diagnostic assay a positive result indicating the patient being tested has cancer is one in which cells, tissues or bodily fluid levels of the cancer marker, such as CSG, are at least two times higher, and most preferably are at least five times higher, than in preferably the same cells, tissues or bodily fluid of a normal human control.

The present invention also provides a method of diagnosing metastatic colon cancer in a patient having colon cancer which has not yet metastasized for the onset of metastasis. In the method of the present invention, a human cancer patient suspected of having colon cancer which may have metastasized (but which was not previously known to have metastasized) is identified. This is accomplished by a variety of means known to those of skill in the art.

WO 01/92528

PCT/US01/17583

- 28 -

In the present invention, determining the presence of CSG levels in cells, tissues or bodily fluid, is particularly useful for discriminating between colon cancer which has not metastasized and colon cancer which has metastasized.

5 Existing techniques have difficulty discriminating between colon cancer which has metastasized and colon cancer which has not metastasized and proper treatment selection is often dependent upon such knowledge.

In the present invention, the cancer marker levels measured in such cells, tissues or bodily fluid is CSG, and are compared with levels of CSG in preferably the same cells, tissue or bodily fluid type of a normal human control. That is, if the cancer marker being observed is just CSG in serum, this level is preferably compared with the level of CSG in serum of a normal human control. An increase in the CSG in the patient versus the normal human control is associated with colon cancer which has metastasized.

Without limiting the instant invention, typically, for a quantitative diagnostic assay a positive result indicating the cancer in the patient being tested or monitored has metastasized is one in which cells, tissues or bodily fluid levels of the cancer marker, such as CSG, are at least two times higher, and most preferably are at least five times higher, than in preferably the same cells, tissues or bodily fluid of a normal patient.

Normal human control as used herein includes a human patient without cancer and/or non cancerous samples from the patient; in the methods for diagnosing or monitoring for metastasis, normal human control may preferably also include samples from a human patient that is determined by reliable methods to have colon cancer which has not metastasized.

Staging

The invention also provides a method of staging colon cancer in a human patient. The method comprises identifying a human patient having such cancer and analyzing cells,

WO 01/92528

PCT/US01/17583

- 29 -

tissues or bodily fluid from such human patient for CSG. The CSG levels determined in the patient are then compared with levels of CSG in preferably the same cells, tissues or bodily fluid type of a normal human control, wherein an increase in
5 CSG levels in the human patient versus the normal human control is associated with a cancer which is progressing and a decrease in the levels of CSG (but still increased over true normal levels) is associated with a cancer which is regressing or in remission.

10 *Monitoring*

Further provided is a method of monitoring colon cancer in a human patient having such cancer for the onset of metastasis. The method comprises identifying a human patient having such cancer that is not known to have metastasized;
15 periodically analyzing cells, tissues or bodily fluid from such human patient for CSG; and comparing the CSG levels determined in the human patient with levels of CSG in preferably the same cells, tissues or bodily fluid type of a normal human control, wherein an increase in CSG levels in the
20 human patient versus the normal human control is associated with a cancer which has metastasized. In this method, normal human control samples may also include prior patient samples.

Further provided by this invention is a method of monitoring the change in stage of colon cancer in a human
25 patient having such cancer. The method comprises identifying a human patient having such cancer; periodically analyzing cells, tissues or bodily fluid from such human patient for CSG; and comparing the CSG levels determined in the human patient with levels of CSG in preferably the same cells,
30 tissues or bodily fluid type of a normal human control, wherein an increase in CSG levels in the human patient versus the normal human control is associated with a cancer which is progressing in stage and a decrease in the levels of CSG is associated with a cancer which is regressing in stage or in

WO 01/92528

PCT/US01/17583

- 30 -

remission. In this method, normal human control samples may also include prior patient samples.

Monitoring a patient for onset of metastasis is periodic and preferably done on a quarterly basis. However, this may be done more or less frequently depending on the cancer, the particular patient, and the stage of the cancer.

Prognostic Testing and Clinical Trial Monitoring

The methods described herein can further be utilized as prognostic assays to identify subjects having or at risk of developing a disease or disorder associated with increased levels of CSG. The present invention provides a method in which a test sample is obtained from a human patient and CSG is detected. The presence of higher CSG levels as compared to normal human controls is diagnostic for the human patient being at risk for developing cancer, particularly colon cancer.

The effectiveness of therapeutic agents to decrease expression or activity of the CSGs of the invention can also be monitored by analyzing levels of expression of the CSGs in a human patient in clinical trials or in *in vitro* screening assays such as in human cells. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the human patient, or cells as the case may be, to the agent being tested.

25 Detection of genetic lesions or mutations

The methods of the present invention can also be used to detect genetic lesions or mutations in CSG, thereby determining if a human with the genetic lesion is at risk for colon cancer or has colon cancer. Genetic lesions can be detected, for example, by ascertaining the existence of a deletion and/or addition and/or substitution of one or more nucleotides from the CSGs of this invention, a chromosomal rearrangement of CSG, aberrant modification of CSG (such as of the methylation pattern of the genomic DNA), the presence of a non-wild type splicing pattern of a mRNA transcript of

WO 01/92528

PCT/US01/17583

- 31 -

CSG, allelic loss of CSG, and/or inappropriate post-translational modification of CSG protein. Methods to detect such lesions in the CSG of this invention are known to those of skill in the art.

5 For example, in one embodiment, alterations in a gene corresponding to a CSG polynucleotide of the present invention are determined via isolation of RNA from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated
10 from these RNA samples using protocols known in the art. See, e.g. Sambrook et al. (MOLECULAR CLONING: A LABORATORY MANUAL, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989)), which is illustrative of the many laboratory manuals that detail these techniques. The cDNA is
15 then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22. PCR conditions typically consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C; and 60-120 seconds at
20 70°C, using buffer solutions described in Sidransky, D., et al., Science 252: 706 (1991). PCR products are sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase (Epicentre Technologies). The intron-exon borders of selected exons are
25 also determined and genomic PCR products analyzed to confirm the results. PCR products harboring suspected mutations are then cloned and sequenced to validate the results of the direct sequencing. PCR products are cloned into T-tailed vectors as described in Holton, T. A. and Graham, M. W.,
30 Nucleic Acids Research, 19 : 1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

Genomic rearrangements can also be observed as a method
35 of determining alterations in a gene corresponding to a

WO 01/92528

PCT/US01/17583

- 32 -

polynucleotide. In this method, genomic clones are nick-translated with digoxigenin deoxy-uridine 5'triphosphate (Boehringer Mannheim), and FISH is performed as described in Johnson, C. et al., *Methods Cell Biol.* 35: 73-99 (1991).

5 Hybridization with a labeled probe is carried out using a vast excess of human DNA for specific hybridization to the corresponding genomic locus. Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C-and R-bands. Aligned images for precise

10 mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters (Johnson et al., *Genet. Anal. Tech. Appl.*, 8: 75 (1991)). Image collection, analysis

15 and chromosomal fractional length measurements are performed using the ISee Graphical Program System (Inovision Corporation, Durham, NC). Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations

20 are used as a diagnostic marker for an associated disease.

Assay Techniques

Assay techniques that can be used to determine levels of gene expression (including protein levels), such as CSG of the present invention, in a sample derived from a patient are

25 well known to those of skill in the art. Such assay methods include, without limitation, radioimmunoassays, reverse transcriptase PCR (RT-PCR) assays, immunohistochemistry assays, *in situ* hybridization assays, competitive-binding assays, Western Blot analyses, ELISA assays and proteomic

30 approaches: two-dimensional gel electrophoresis (2D electrophoresis) and non-gel based approaches such as mass spectrometry or protein interaction profiling. Among these, ELISAs are frequently preferred to diagnose a gene's expressed protein in biological fluids.

WO 01/92528

PCT/US01/17583

- 33 -

An ELISA assay initially comprises preparing an antibody, if not readily available from a commercial source, specific to CSG, preferably a monoclonal antibody. In addition a reporter antibody generally is prepared which binds specifically to CSG. The reporter antibody is attached to a detectable reagent such as radioactive, fluorescent or enzymatic reagent, for example horseradish peroxidase enzyme or alkaline phosphatase.

To carry out the ELISA, antibody specific to CSG is incubated on a solid support, e.g. a polystyrene dish, that binds the antibody. Any free protein binding sites on the dish are then covered by incubating with a non-specific protein such as bovine serum albumin. Next, the sample to be analyzed is incubated in the dish, during which time CSG binds to the specific antibody attached to the polystyrene dish. Unbound sample is washed out with buffer. A reporter antibody specifically directed to CSG and linked to a detectable reagent such as horseradish peroxidase is placed in the dish resulting in binding of the reporter antibody to any monoclonal antibody bound to CSG. Unattached reporter antibody is then washed out. Reagents for peroxidase activity, including a colorimetric substrate are then added to the dish. Immobilized peroxidase, linked to CSG antibodies, produces a colored reaction product. The amount of color developed in a given time period is proportional to the amount of CSG protein present in the sample. Quantitative results typically are obtained by reference to a standard curve.

A competition assay can also be employed wherein antibodies specific to CSG are attached to a solid support and labeled CSG and a sample derived from the host are passed over the solid support. The amount of label detected which is attached to the solid support can be correlated to a quantity of CSG in the sample.

WO 01/92528

PCT/US01/17583

- 34 -

Using all or a portion of a nucleic acid sequence of CSG of the present invention as a hybridization probe, nucleic acid methods can also be used to detect CSG mRNA as a marker for colon cancer. Polymerase chain reaction (PCR) and other nucleic acid methods, such as ligase chain reaction (LCR) and nucleic acid sequence based amplification (NASBA), can be used to detect malignant cells for diagnosis and monitoring of various malignancies. For example, reverse-transcriptase PCR (RT-PCR) is a powerful technique which can be used to detect the presence of a specific mRNA population in a complex mixture of thousands of other mRNA species. In RT-PCR, an mRNA species is first reverse transcribed to complementary DNA (cDNA) with use of the enzyme reverse transcriptase; the cDNA is then amplified as in a standard PCR reaction. RT-PCR can thus reveal by amplification the presence of a single species of mRNA. Accordingly, if the mRNA is highly specific for the cell that produces it, RT-PCR can be used to identify the presence of a specific type of cell.

Hybridization to clones or oligonucleotides arrayed on a solid support (i.e. gridding) can be used to both detect the expression of and quantitate the level of expression of that gene. In this approach, a cDNA encoding the CSG gene is fixed to a substrate. The substrate may be of any suitable type including but not limited to glass, nitrocellulose, nylon or plastic. At least a portion of the DNA encoding the CSG gene is attached to the substrate and then incubated with the analyte, which may be RNA or a complementary DNA (cDNA) copy of the RNA, isolated from the tissue of interest. Hybridization between the substrate bound DNA and the analyte can be detected and quantitated by several means including but not limited to radioactive labeling or fluorescence labeling of the analyte or a secondary molecule designed to detect the hybrid. Quantitation of the level of gene expression can be done by comparison of the intensity of the signal from the analyte compared with that determined from known standards.

WO 01/92528

PCT/US01/17583

- 35 -

The standards can be obtained by *in vitro* transcription of the target gene, quantitating the yield, and then using that material to generate a standard curve.

Of the proteomic approaches, 2D electrophoresis is a technique well known to those in the art. Isolation of individual proteins from a sample such as serum is accomplished using sequential separation of proteins by different characteristics usually on polyacrylamide gels. First, proteins are separated by size using an electric current. The current acts uniformly on all proteins, so smaller proteins move farther on the gel than larger proteins. The second dimension applies a current perpendicular to the first and separates proteins not on the basis of size but on the specific electric charge carried by each protein. Since no two proteins with different sequences are identical on the basis of both size and charge, the result of a 2D separation is a square gel in which each protein occupies a unique spot. Analysis of the spots with chemical or antibody probes, or subsequent protein microsequencing can reveal the relative abundance of a given protein and the identity of the proteins in the sample.

The above tests can be carried out on samples derived from a variety of cells, bodily fluids and/or tissue extracts such as homogenates or solubilized tissue obtained from a patient. Tissue extracts are obtained routinely from tissue biopsy and autopsy material. Bodily fluids useful in the present invention include blood, urine, saliva or any other bodily secretion or derivative thereof. By blood it is meant to include whole blood, plasma, serum or any derivative of blood.

In Vivo Targeting of CSG/Colon Cancer Therapy

Identification of this CSG is also useful in the rational design of new therapeutics for imaging and treating cancers, and in particular colon cancer. For example, in one embodiment, antibodies which specifically bind to CSG can be

WO 01/92528

PCT/US01/17583

- 36 -

raised and used *in vivo* in patients suspected of suffering from colon cancer. Antibodies which specifically bind CSG can be injected into a patient suspected of having colon cancer for diagnostic and/or therapeutic purposes. Thus, another aspect of the present invention provides for a method for preventing the onset and treatment of colon cancer in a human patient in need of such treatment by administering to the patient an effective amount of antibody. By "effective amount" it is meant the amount or concentration of antibody needed to bind to the target antigens expressed on the tumor to cause tumor shrinkage for surgical removal, or disappearance of the tumor. The binding of the antibody to the overexpressed CSG is believed to cause the death of the cancer cell expressing such CSG. The preparation and use of antibodies for *in vivo* diagnosis and treatment is well known in the art. For example, antibody-chelators labeled with Indium-111 have been described for use in the radioimmunosciintographic imaging of carcinoembryonic antigen expressing tumors (Sumerdon et al. Nucl. Med. Biol. 1990 17:247-254). In particular, these antibody-chelators have been used in detecting tumors in patients suspected of having recurrent colorectal cancer (Griffin et al. J. Clin. Onc. 1991 9:631-640). Antibodies with paramagnetic ions as labels for use in magnetic resonance imaging have also been described (Lauffer, R.B. Magnetic Resonance in Medicine 1991 22:339-342). Antibodies directed against CSG can be used in a similar manner. Labeled antibodies which specifically bind CSG can be injected into patients suspected of having colon cancer for the purpose of diagnosing or staging of the disease status of the patient. The label used will be selected in accordance with the imaging modality to be used. For example, radioactive labels such as Indium-111, Technetium-99m or Iodine-131 can be used for planar scans or single photon emission computed tomography (SPECT). Positron emitting labels such as Fluorine-19 can be used in positron emission

WO 01/92528

PCT/US01/17583

- 37 -

tomography. Paramagnetic ions such as Gadolinium (III) or Manganese (II) can be used in magnetic resonance imaging (MRI). Presence of the label, as compared to imaging of normal tissue, permits determination of the spread of the cancer. The amount of label within an organ or tissue also allows determination of the presence or absence of cancer in that organ or tissue.

Antibodies which can be used in *in vivo* methods include polyclonal, monoclonal and omniclonal antibodies and antibodies prepared via molecular biology techniques. Antibody fragments and aptamers and single-stranded oligonucleotides such as those derived from an *in vitro* evolution protocol referred to as SELEX and well known to those skilled in the art can also be used.

15 *Screening Assays*

The present invention also provides methods for identifying modulators which bind to CSG protein or have a modulatory effect on the expression or activity of CSG protein. Modulators which decrease the expression or activity of CSG protein are believed to be useful in treating colon cancer. Such screening assays are known to those of skill in the art and include, without limitation, cell-based assays and cell free assays.

Small molecules predicted via computer imaging to specifically bind to regions of CSG can also be designed, synthesized and tested for use in the imaging and treatment of colon cancer. Further, libraries of molecules can be screened for potential anticancer agents by assessing the ability of the molecule to bind to the CSGs identified herein. Molecules identified in the library as being capable of binding to CSG are key candidates for further evaluation for use in the treatment of colon cancer. In a preferred embodiment, these molecules will downregulate expression and/or activity of CSG in cells.

35

WO 01/92528

PCT/US01/17583

- 38 -

Adoptive Immunotherapy and Vaccines

Adoptive immunotherapy of cancer refers to a therapeutic approach in which immune cells with an antitumor reactivity are administered to a tumor-bearing host, with the aim that the cells mediate either directly or indirectly, the regression of an established tumor. Transfusion of lymphocytes, particularly T lymphocytes, falls into this category and investigators at the National Cancer Institute (NCI) have used autologous reinfusion of peripheral blood lymphocytes or tumor-infiltrating lymphocytes (TIL), T cell cultures from biopsies of subcutaneous lymph nodes, to treat several human cancers (Rosenberg, S. A., U.S. Patent No. 4,690,914, issued Sep. 1, 1987; Rosenberg, S. A., et al., 1988, N. England J. Med. 319:1676-1680).

The present invention relates to compositions and methods of adoptive immunotherapy for the prevention and/or treatment of primary and metastatic colon cancer in humans using macrophages sensitized to the antigenic CSG molecules, with or without non-covalent complexes of heat shock protein (hsp). Antigenicity or immunogenicity of the CSG is readily confirmed by the ability of the CSG protein or a fragment thereof to raise antibodies or educate naive effector cells, which in turn lyse target cells expressing the antigen (or epitope).

Cancer cells are, by definition, abnormal and contain proteins which should be recognized by the immune system as foreign since they are not present in normal tissues. However, the immune system often seems to ignore this abnormality and fails to attack tumors. The foreign CSG proteins that are produced by the cancer cells can be used to reveal their presence. The CSG is broken into short fragments, called tumor antigens, which are displayed on the surface of the cell. These tumor antigens are held or presented on the cell surface by molecules called MHC, of which there are two types: class I and II. Tumor antigens in association with MHC class

WO 01/92528

PCT/US01/17583

- 39 -

I molecules are recognized by cytotoxic T cells while antigen-MHC class II complexes are recognized by a second subset of T cells called helper cells. These cells secrete cytokines which slow or stop tumor growth and help another type of white blood cell, B cells, to make antibodies against the tumor cells.

In adoptive immunotherapy, T cells or other antigen presenting cells (APCs) are stimulated outside the body (ex vivo), using the tumor specific CSG antigen. The stimulated cells are then reinfused into the patient where they attack the cancerous cells. Research has shown that using both cytotoxic and helper T cells is far more effective than using either subset alone. Additionally, the CSG antigen may be complexed with heat shock proteins to stimulate the APCs as described in U.S. Patent No. 5,985,270.

The APCs can be selected from among those antigen presenting cells known in the art including, but not limited to, macrophages, dendritic cells, B lymphocytes, and a combination thereof, and are preferably macrophages. In a preferred use, wherein cells are autologous to the individual, autologous immune cells such as lymphocytes, macrophages or other APCs are used to circumvent the issue of whom to select as the donor of the immune cells for adoptive transfer. Another problem circumvented by use of autologous immune cells is graft versus host disease which can be fatal if unsuccessfully treated.

In adoptive immunotherapy with gene therapy, DNA of the CSG can be introduced into effector cells similarly as in conventional gene therapy. This can enhance the cytotoxicity of the effector cells to tumor cells as they have been manipulated to produce the antigenic protein resulting in improvement of the adoptive immunotherapy.

CSG antigens of this invention are also useful as components of colon cancer vaccines. The vaccine comprises an immunogenically stimulatory amount of a CSG antigen.

WO 01/92528

PCT/US01/17583

- 40 -

Immunogenically stimulatory amount refers to that amount of antigen that is able to invoke the desired immune response in the recipient for the amelioration, or treatment of colon cancer. Effective amounts may be determined empirically by standard procedures well known to those skilled in the art.

The CSG antigen may be provided in any one of a number of vaccine formulations which are designed to induce the desired type of immune response, e.g., antibody and/or cell mediated. Such formulations are known in the art and include, but are not limited to, formulations such as those described in U.S. Patent 5,585,103. Vaccine formulations of the present invention used to stimulate immune responses can also include pharmaceutically acceptable adjuvants.

Vectors, host cells, expression

The present invention also relates to vectors which include polynucleotides of the present invention, host cells which are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques.

Host cells can be genetically engineered to incorporate CSG polynucleotides and express CSG polypeptides of the present invention. For instance, CSG polynucleotides may be introduced into host cells using well known techniques of infection, transduction, transfection, transvection and transformation. The CSG polynucleotides may be introduced alone or with other polynucleotides. Such other polynucleotides may be introduced independently, co-introduced or introduced joined to the CSG polynucleotides of the invention.

For example, CSG polynucleotides of the invention may be transfected into host cells with another, separate, polynucleotide encoding a selectable marker, using standard techniques for co-transfection and selection in, for instance, mammalian cells. In this case, the polynucleotides generally will be stably incorporated into the host cell genome.

WO 01/92528

PCT/US01/17583

- 41 -

Alternatively, the CSG polynucleotide may be joined to a vector containing a selectable marker for propagation in a host. The vector construct may be introduced into host cells by the aforementioned techniques. Generally, a plasmid vector is introduced as DNA in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. Electroporation also may be used to introduce CSG polynucleotides into a host. If the vector is a virus, it may be packaged *in vitro* or introduced into a packaging cell and the packaged virus may be transduced into cells. A wide variety of well known techniques conducted routinely by those of skill in the art are suitable for making CSG polynucleotides and for introducing CSG polynucleotides into cells in accordance with this aspect of the invention. Such techniques are reviewed at length in reference texts such as Sambrook et al., previously cited herein.

Vectors which may be used in the present invention include, for example, plasmid vectors, single- or double-stranded phage vectors, and single- or double-stranded RNA or DNA viral vectors. Such vectors may be introduced into cells as polynucleotides, preferably DNA, by well known techniques for introducing DNA and RNA into cells. The vectors, in the case of phage and viral vectors, also may be and preferably are introduced into cells as packaged or encapsidated virus by well known techniques for infection and transduction. Viral vectors may be replication competent or replication defective. In the latter case viral propagation generally will occur only in complementing host cells.

Preferred vectors for expression of polynucleotides and polypeptides of the present invention include, but are not limited to, vectors comprising cis-acting control regions effective for expression in a host operatively linked to the polynucleotide to be expressed. Appropriate trans-acting factors either are supplied by the host, supplied by a

- 42 -

complementing vector or supplied by the vector itself upon introduction into the host.

In certain preferred embodiments in this regard, the vectors provide for specific expression. Such specific
5 expression may be inducible expression or expression only in certain types of cells or both inducible and cell-specific. Particularly preferred among inducible vectors are vectors that can be induced to express by environmental factors that are easy to manipulate, such as temperature and nutrient
10 additives. A variety of vectors suitable to this aspect of the invention, including constitutive and inducible expression vectors for use in prokaryotic and eukaryotic hosts, are well known and employed routinely by those of skill in the art.

The engineered host cells can be cultured in
15 conventional nutrient media which may be modified as appropriate for, *inter alia*, activating promoters, selecting transformants or amplifying genes. Culture conditions such as temperature, pH and the like, previously used with the host cell selected for expression, generally will be suitable for
20 expression of CSG polypeptides of the present invention.

A great variety of expression vectors can be used to express CSG polypeptides of the invention. Such vectors include chromosomal, episomal and virus-derived vectors. Vectors may be derived from bacterial plasmids, from
25 bacteriophage, from yeast episomes, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and from combinations thereof such as those derived from plasmid and
30 bacteriophage genetic elements, such as cosmids and phagemids. All may be used for expression in accordance with this aspect of the present invention. Generally, any vector suitable to maintain, propagate or express polynucleotides to express a polypeptide in a host may be used for expression in this
35 regard.

WO 01/92528

PCT/US01/17583

- 43 -

The appropriate DNA sequence may be inserted into the vector by any of a variety of well-known and routine techniques. In general, a DNA sequence for expression is joined to an expression vector by cleaving the DNA sequence
5 and the expression vector with one or more restriction endonucleases and then joining the restriction fragments together using T4 DNA ligase. Procedures for restriction and ligation that can be used to this end are well known and routine to those of skill. Suitable procedures in this
10 regard, and for constructing expression vectors using alternative techniques, which also are well known and routine to those skill, are set forth in great detail in Sambrook et al. cited elsewhere herein.

The DNA sequence in the expression vector is operatively
15 linked to appropriate expression control sequence(s), including, for instance, a promoter to direct mRNA transcription. Representative promoters include the phage lambda PL promoter, the *E. coli* lac, trp and tac promoters, the SV40 early and late promoters, and promoters of retroviral
20 LTRs, to name just a few of the well-known promoters. It will be understood that numerous promoters not mentioned are also suitable for use in this aspect of the invention and are well known and readily may be employed by those of skill in the manner illustrated by the discussion and the examples herein.

25 In general, expression constructs will contain sites for transcription initiation and termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the mature transcripts expressed by the constructs will include a translation initiating AUG at the
30 beginning and a termination codon appropriately positioned at the end of the polypeptide to be translated.

In addition, the constructs may contain control regions that regulate as well as engender expression. Generally, in accordance with many commonly practiced procedures, such

WO 01/92528

PCT/US01/17583

- 44 -

regions will operate by controlling transcription, such as repressor binding sites and enhancers, among others.

Vectors for propagation and expression generally will include selectable markers. Such markers also may be suitable
5 for amplification or the vectors may contain additional markers for this purpose. In this regard, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed
10 host cells. Preferred markers include dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, and tetracycline or ampicillin resistance genes for culturing in *E. coli* and other bacteria.

The vector containing the appropriate DNA sequence as described elsewhere herein, as well as an appropriate
15 promoter, and other appropriate control sequences, may be introduced into an appropriate host using a variety of well known techniques suitable to expression therein of a desired polypeptide. Representative examples of appropriate hosts include bacterial cells, such as *E. coli*, *Streptomyces* and
20 *Salmonella typhimurium* cells; fungal cells, such as yeast cells; insect cells such as *Drosophila S2* and *Spodoptera Sf9* cells; animal cells such as CHO, COS and Bowes melanoma cells; and plant cells. Hosts for a great variety of expression constructs are well known, and those of skill will be enabled
25 by the present disclosure readily to select a host for expressing a CSG polypeptide in accordance with this aspect of the present invention.

More particularly, the present invention also includes recombinant constructs, such as expression constructs,
30 comprising one or more of the sequences described above. The constructs comprise a vector, such as a plasmid or viral vector, into which such CSG sequence of the invention has been inserted. The sequence may be inserted in a forward or reverse orientation. In certain preferred embodiments in this
35 regard, the construct further comprises regulatory sequences,

WO 01/92528

PCT/US01/17583

- 45 -

including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and there are many commercially available vectors suitable for use in the present invention.

The following vectors, which are commercially available, are provided by way of example. Among vectors preferred for use in bacteria are pQE70, pQE60 and pQE-9, available from Qiagen; pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia. Among preferred eukaryotic vectors are PWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. These vectors are listed solely by way of illustration of the many commercially available and well known vectors that are available to those of skill in the art for use in accordance with this aspect of the present invention. It will be appreciated by those of skill in the art upon reading this disclosure that any other plasmid or vector suitable for introduction, maintenance, propagation and/or expression of a CSG polynucleotide or polypeptide of the invention in a host may be used in this aspect of the invention.

Promoter regions can be selected from any desired gene using vectors that contain a reporter transcription unit lacking a promoter region, such as a chloramphenicol acetyl transferase ("cat") transcription unit, downstream of a restriction site or sites for introducing a candidate promoter fragment; i.e., a fragment that may contain a promoter. As is well known, introduction into the vector of a promoter-containing fragment at the restriction site upstream of the cat gene engenders production of CAT activity detectable by standard CAT assays. Vectors suitable to this end are well known and readily available. Two such vectors are pKK232-8 and pCM7. Thus, promoters for expression of CSG

WO 01/92528

PCT/US01/17583

- 46 -

polynucleotides of the present invention include, not only well known and readily available promoters, but also promoters that readily may be obtained by the foregoing technique, using a reporter gene.

5 Among known bacterial promoters suitable for expression of polynucleotides and polypeptides in accordance with the present invention are the *E. coli* *lacI* and *lacZ* promoters, the T3 and T7 promoters, the *gpt* promoter, the lambda PR, PL promoters and the *trp* promoter. Among known eukaryotic
10 promoters suitable in this regard are the CMV immediate early promoter, the HSV thymidine kinase promoter, the early and late SV40 promoters, the promoters of retroviral LTRs, such as those of the Rous sarcoma virus ("RSV"), and metallothionein promoters, such as the mouse metallothionein-I
15 promoter.

Selection of appropriate vectors and promoters for expression in a host cell is a well known procedure and the requisite techniques for expression vector construction, introduction of the vector into the host and expression in the
20 host are routine skills in the art.

The present invention also relates to host cells containing the above-described constructs. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell. Alternatively,
25 the host cell can be a prokaryotic cell, such as a bacterial cell.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection,
30 electroporation, transduction, infection or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al. BASIC METHODS IN MOLECULAR BIOLOGY, (1986).

Constructs in host cells can be used in a conventional
35 manner to produce the gene product encoded by the recombinant

WO 01/92528

PCT/US01/17583

- 47 -

sequence. Alternatively, CSG polypeptides of the invention can be synthetically produced by conventional peptide synthesizers.

Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook et al. cited elsewhere herein.

Generally, recombinant expression vectors will include origins of replication, a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence, and a selectable marker to permit isolation of vector containing cells after exposure to the vector. Among suitable promoters are those derived from the genes that encode glycolytic enzymes such as 3-phosphoglycerate kinase ("PGK"), α -factor, acid phosphatase, and heat shock proteins, among others. Selectable markers include the ampicillin resistance gene of *E. coli* and the *trp1* gene of *S. cerevisiae*.

Transcription of DNA encoding the CSG polypeptides of the present invention by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 base pairs (bp) that act to increase transcriptional activity of a promoter in a given host cell-type. Examples of enhancers include the SV40 enhancer, which is located on the late side of the replication origin at bp 100 to 270, the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

A polynucleotide of the present invention, encoding a heterologous structural sequence of a CSG polypeptide of the

WO 01/92528

PCT/US01/17583

- 48 -

present invention, generally will be inserted into the vector using standard techniques so that it is operably linked to the promoter for expression. The polynucleotide will be positioned so that the transcription start site is located 5 appropriately 5' to a ribosome binding site. The ribosome binding site will be 5' to the AUG that initiates translation of the polypeptide to be expressed. Generally, there will be no other open reading frames that begin with an initiation codon, usually AUG, lying between the ribosome binding site 10 and the initiating AUG. Also, generally, there will be a translation stop codon at the end of the polypeptide and there will be a polyadenylation signal and a transcription termination signal appropriately disposed at the 3' end of the transcribed region.

15 Appropriate secretion signals may be incorporated into the expressed polypeptide for secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment. The signals may be endogenous to the polypeptide or they may be 20 heterologous signals.

The polypeptide may be expressed in a modified form, such as a fusion protein, and may include not only secretion signals but also additional heterologous functional regions. Thus, for instance, a region of additional amino acids, 25 particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence in the host cell during purification or during subsequent handling and storage. A region also may be added to the polypeptide to facilitate purification. Such regions 30 may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to polypeptides to engender secretion or excretion, to improve stability and to facilitate purification, among others, are familiar and routine techniques in the art.

WO 01/92528

PCT/US01/17583

- 49 -

Suitable prokaryotic hosts for propagation, maintenance or expression of CSG polynucleotides and polypeptides in accordance with the invention include *Escherichia coli*, *Bacillus subtilis* and *Salmonella typhimurium*. Various species of *Pseudomonas*, *Streptomyces*, and *Staphylococcus* are suitable hosts in this regard. Many other hosts also known to those of skill may also be employed in this regard.

As a representative, but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322. Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM1 (Promega Biotec, Madison, Wis., USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, where the selected promoter is inducible it is induced by appropriate means (e.g., temperature shift or exposure to chemical inducer) and cells are cultured for an additional period. Cells typically then are harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents, such methods are well known to those skilled in the art.

Various mammalian cell culture systems can be employed for expression, as well. An exemplary mammalian expression system is the COS-7 line of monkey kidney fibroblasts described in Gluzman et al., *Cell* 23: 175 (1981). Other mammalian cell lines capable of expressing a compatible vector include for example, the C127, 3T3, CHO, HeLa, human kidney

WO 01/92528

PCT/US01/17583

- 50 -

293 and BHK cell lines. Mammalian expression vectors comprise an origin of replication, a suitable promoter and enhancer, and any ribosome binding sites, polyadenylation sites, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking non-transcribed sequences that are necessary for expression. In certain preferred embodiments in this regard DNA sequences derived from the SV40 splice sites, and the SV40 polyadenylation sites are used for required non-transcribed genetic elements of these types.

CSG polypeptides can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during isolation and or purification.

CSG polypeptides of the present invention include naturally purified products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. Depending upon the host employed in a recombinant production procedure, the CSG polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, CSG polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes.

CSG polynucleotides and polypeptides may be used in accordance with the present invention for a variety of applications, particularly those that make use of the chemical

WO 01/92528

PCT/US01/17583

- 51 -

and biological properties of the CSGs. Additional applications relate to diagnosis and to treatment of disorders of cells, tissues and organisms. These aspects of the invention are illustrated further by the following discussion.

5 **Polynucleotide assays**

As discussed in some detail supra, this invention is also related to the use of CSG polynucleotides to detect complementary polynucleotides such as, for example, as a diagnostic reagent. Detection of a mutated form of CSG
10 associated with a dysfunction will provide a diagnostic tool that can add to or define a diagnosis of a disease or susceptibility to a disease which results from under-expression, over-expression or altered expression of a CSG, such as, for example, a susceptibility to inherited colon
15 cancer.

Individuals carrying mutations in a human CSG gene may be detected at the DNA level by a variety of techniques. Nucleic acids for diagnosis may be obtained from a patient's cells, such as from blood, urine, saliva, tissue biopsy and
20 autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically using PCR prior to analysis (Saiki et al., Nature, 324: 163-166 (1986)). RNA or cDNA may also be used in a similar manner. As an example, PCR primers complementary to a CSG polynucleotide of SEQ ID
25 NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22 can be used to identify and analyze CSG expression and mutations. For example, deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point
30 mutations can be identified by hybridizing amplified DNA to radiolabeled CSG RNA or alternatively, radiolabeled CSG antisense DNA sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase A digestion or by differences in melting temperatures.

WO 01/92528

PCT/US01/17583

- 52 -

Sequence differences between a reference gene and genes having mutations also may be revealed by direct DNA sequencing. In addition, cloned DNA segments may be employed as probes to detect specific DNA segments. The sensitivity of such methods can be greatly enhanced by appropriate use of PCR or another amplification method. For example, a sequencing primer is used with double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The sequence determination is performed by conventional procedures with radiolabeled nucleotide or by automatic sequencing procedures with fluorescent-tags.

Genetic testing based on DNA sequence differences may be achieved by detection of alterations in electrophoretic mobility of DNA fragments in gels, with or without denaturing agents. Small sequence deletions and insertions can be visualized by high resolution gel electrophoresis. DNA fragments of different sequences may be distinguished on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific melting or partial melting temperatures (see, e.g., Myers et al., Science, 230: 1242 (1985)).

Sequence changes at specific locations also may be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (e.g., Cotton et al., Proc. Natl. Acad. Sci., USA, 85: 4397-4401 (1988)).

Thus, the detection of a specific DNA sequence may be achieved by methods such as hybridization, RNase protection, chemical cleavage, direct DNA sequencing or the use of restriction enzymes, (e.g., restriction fragment length polymorphisms ("RFLP") and Southern blotting of genomic DNA. In addition to more conventional gel-electrophoresis and DNA sequencing, mutations also can be detected by *in situ* analysis.

WO 01/92528

PCT/US01/17583

- 53 -

Chromosome assays

The CSG sequences of the present invention are also valuable for chromosome identification. There is a need for identifying particular sites on the chromosome and few chromosome marking reagents based on actual sequence data (repeat polymorphisms) are presently available for marking chromosomal location. Each CSG sequence of the present invention is specifically targeted to and can hybridize with a particular location on an individual human chromosome. Thus, the CSGs can be used in the mapping of DNAs to chromosomes, an important first step in correlating sequences with genes associated with disease.

In certain preferred embodiments in this regard, the cDNA herein disclosed is used to clone genomic DNA of a CSG of the present invention. This can be accomplished using a variety of well known techniques and libraries, which generally are available commercially. The genomic DNA is used for *in situ* chromosome mapping using well known techniques for this purpose.

In some cases, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the cDNA. Computer analysis of the 3' untranslated region of the gene is used to rapidly select primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the primer will yield an amplified fragment.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular DNA to a particular chromosome. Using the present invention with the same oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes or pools of large genomic clones in an analogous manner. Other mapping strategies that can similarly be used to map to its chromosome include *in situ*

WO 01/92528

PCT/US01/17583

- 54 -

hybridization, prescreening with labeled flow-sorted chromosomes and preselection by hybridization to construct chromosome specific-cDNA libraries.

Fluorescence in situ hybridization ("FISH") of a cDNA
5 clone to a metaphase chromosomal spread can be used to provide a precise chromosomal location in one step. This technique can be used with cDNA as short as 50 or 60 bp. This technique is described by Verma et al. (HUMAN CHROMOSOMES: A MANUAL OF
BASIC TECHNIQUES, Pergamon Press, New York (1988)).

10 Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, for example, in V. McKusick, MENDELIAN INHERITANCE
IN MAN, available on line through Johns Hopkins University,
15 Welch Medical Library. The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes).

Next, it is necessary to determine the differences in
20 the cDNA or genomic sequence between affected and unaffected individuals. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative agent of the disease.

25 With current resolution of physical mapping and genetic mapping techniques, a cDNA precisely localized to a chromosomal region associated with the disease could be one of between 50 and 500 potential causative genes. (This assumes 1 megabase mapping resolution and one gene per 20 kb).

30 **Polypeptide assays**

As described in some detail supra, the present invention also relates to diagnostic assays such as quantitative and diagnostic assays for detecting levels of CSG polypeptide in cells and tissues, and biological fluids such as blood and
35 urine, including determination of normal and abnormal levels.

WO 01/92528

PCT/US01/17583

- 55 -

Thus, for instance, a diagnostic assay in accordance with the present invention for detecting over-expression or under-expression of a CSG polypeptide compared to normal control tissue samples may be used to detect the presence of neoplasia. Assay techniques that can be used to determine levels of a protein, such as a CSG polypeptide of the present invention, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays. Among these ELISAs frequently are preferred.

For example, antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 $\mu\text{g/ml}$. The antibodies are either monoclonal or polyclonal and are produced by methods as described herein. The wells are blocked so that non-specific binding of the polypeptide to the well is reduced. The coated wells are then incubated for > 2 hours at room temperature with a sample containing the CSG polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbound polypeptide. Next, 50 μl of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbound conjugate. 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution (75 μl) is then added to each well and the plate is incubated 1 hour at room temperature. The reaction is measured by a microtiter plate reader. A standard curve is prepared using serial dilutions of a control sample, and polypeptide concentration is plotted on the X-axis (log scale) while fluorescence or absorbance is plotted on the Y-

WO 01/92528

PCT/US01/17583

- 56 -

axis (linear scale). The concentration of the CSG polypeptide in the sample is interpolated using the standard curve.

Antibodies

As discussed in some detail supra, CSG polypeptides, 5 their fragments or other derivatives, or analogs thereof, or cells expressing them can be used as an immunogen to produce antibodies thereto. These antibodies can be polyclonal or monoclonal antibodies. The present invention also includes 10 chimeric, single chain, and humanized antibodies, as well as Fab fragments, or the product of an Fab expression library. Various procedures known in the art may be used for the production of such antibodies and fragments.

A variety of methods for antibody production are set forth in Current Protocols, Chapter 2.

15 For example, cells expressing a CSG polypeptide of the present invention can be administered to an animal to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of 20 natural contaminants. This preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity. The antibody obtained will bind with the CSG polypeptide itself. In this manner, even a sequence encoding only a fragment of the CSG polypeptide can 25 be used to generate antibodies binding the whole native polypeptide. Such antibodies can then be used to isolate the CSG polypeptide from tissue expressing that CSG polypeptide.

Alternatively, monoclonal antibodies can be prepared. 30 Examples of techniques for production of monoclonal antibodies include, but are not limited to, the hybridoma technique (Kohler, G. and Milstein, C., Nature 256: 495-497 (1975), the trioma technique, the human B-cell hybridoma technique (Kozbor et al., Immunology Today 4: 72 (1983) and (Cole et al., pg. 77-96 in MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R.

WO 01/92528

PCT/US01/17583

- 57 -

Liss, Inc. (1985). The EBV-hybridoma technique is useful in production of human monoclonal antibodies.

Hybridoma technologies have also been described by Kihler et al. (Eur. J. Immunol. 6: 511 (1976)) Kihler et al. (Eur. J. Immunol. 6: 292 (1976)) and Hammerling et al. (in: Monoclonal Antibodies and T-Cell Hybridomas, Elsevier, N. Y., pp. 563-681 (1981)). In general, such procedures involve immunizing an animal (preferably a mouse) with CSG polypeptide or, more preferably, with a secreted CSG polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 µg/ml of streptomycin. The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP20), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands et al. (Gastroenterology 80: 225-232 (1981)). The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened

WO 01/92528

PCT/US01/17583

- 58 -

to identify clones which produce an antibody whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein specific antibody and can be used 5 to immunize an animal to induce formation of further protein-specific antibodies.

Techniques described for the production of single chain antibodies (U.S. Patent 4,946,778) can also be adapted to produce single chain antibodies to immunogenic polypeptide 10 products of this invention. Also, transgenic mice, as well as other nonhuman transgenic animals, may be used to express humanized antibodies to immunogenic polypeptide products of this invention.

It will be appreciated that Fab, F(ab')₂ and other 15 fragments of the antibodies of the present invention may also be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')₂ fragments). Alternatively, secreted 20 protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. 25 Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art (See, for review, Morrison, Science 229: 1202 (1985); Oi et al., BioTechniques 4: 214 30 (1986); Cabilly et al., U. S. Patent 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312: 643 (1984); Neuberger et al., Nature 314: 268 (1985).)

WO 01/92528

PCT/US01/17583

- 59 -

The above-described antibodies may be employed to isolate or to identify clones expressing CSG polypeptides or purify CSG polypeptides of the present invention by attachment of the antibody to a solid support for isolation and/or purification by affinity chromatography. As discussed in more detail *supra*, antibodies specific against a CSG may also be used to image tumors, particularly cancer of the colon, in patients suffering from cancer. Such antibodies may also be used therapeutically to target tumors expressing a CSG.

10 *CSG binding molecules and assays*

This invention also provides a method for identification of molecules, such as receptor molecules, that bind CSGs. Genes encoding proteins that bind CSGs, such as receptor proteins, can be identified by numerous methods known to those of skill in the art. Examples include, but are not limited to, ligand panning and FACS sorting. Such methods are described in many laboratory manuals such as, for instance, Coligan et al., *Current Protocols in Immunology* 1(2): Chapter 5 (1991).

20 Expression cloning may also be employed for this purpose. To this end, polyadenylated RNA is prepared from a cell responsive to a CSG of the present invention. A cDNA library is created from this RNA and the library is divided into pools. The pools are then transfected individually into cells that are not responsive to a CSG of the present invention. The transfected cells then are exposed to labeled CSG. CSG polypeptides can be labeled by a variety of well-known techniques including, but not limited to, standard methods of radio-iodination or inclusion of a recognition site for a site-specific protein kinase. Following exposure, the cells are fixed and binding of labeled CSG is determined. These procedures conveniently are carried out on glass slides. Pools containing labeled CSG are identified as containing cDNA that produced CSG-binding cells. Sub-pools are then prepared from these positives, transfected into host cells and

WO 01/92528

PCT/US01/17583

- 60 -

screened as described above. Using an iterative sub-pooling and re-screening process, one or more single clones that encode the putative binding molecule, such as a receptor molecule, can be isolated.

5 Alternatively a labeled ligand can be photoaffinity linked to a cell extract, such as a membrane or a membrane extract, prepared from cells that express a molecule that it binds, such as a receptor molecule. Cross-linked material is resolved by polyacrylamide gel electrophoresis ("PAGE") and
10 exposed to X-ray film. The labeled complex containing the ligand-receptor can be excised, resolved into peptide fragments, and subjected to protein microsequencing. The amino acid sequence obtained from microsequencing can be used
15 to design unique or degenerate oligonucleotide probes to screen cDNA libraries to identify genes encoding the putative receptor molecule.

Polypeptides of the invention also can be used to assess CSG binding capacity of CSG binding molecules, such as receptor molecules, in cells or in cell-free preparations.

20 ***Agonists and antagonists - assays and molecules***

The invention also provides a method of screening compounds to identify those which enhance or block the action of a CSG on cells. By "compound", as used herein, it is meant to be inclusive of small organic molecules, peptides,
25 polypeptides and antibodies as well as any other candidate molecules which have the potential to enhance or agonize or block or antagonize the action of CSG on cells. As used herein, an agonist is a compound which increases the natural biological functions of a CSG or which functions in a manner
30 similar to a CSG, while an antagonist, as used herein, is a compound which decreases or eliminates such functions. Various known methods for screening for agonists and/or antagonists can be adapted for use in identifying CSG agonist or antagonists.

WO 01/92528

PCT/US01/17583

- 61 -

For example, a cellular compartment, such as a membrane or a preparation thereof, such as a membrane-preparation, may be prepared from a cell that expresses a molecule that binds a CSG, such as a molecule of a signaling or regulatory pathway modulated by CSG. The preparation is incubated with labeled CSG in the absence or the presence of a compound which may be a CSG agonist or antagonist. The ability of the compound to bind the binding molecule is reflected in decreased binding of the labeled ligand. Compounds which bind gratuitously, i.e., without inducing the effects of a CSG upon binding to the CSG binding molecule are most likely to be good antagonists. Compounds that bind well and elicit effects that are the same as or closely related to CSG are agonists. CSG-like effects of potential agonists and antagonists may be measured, for instance, by determining activity of a second messenger system following interaction of the candidate molecule with a cell or appropriate cell preparation, and comparing the effect with that of CSG or molecules that elicit the same effects as CSG. Second messenger systems that may be useful in this regard include, but are not limited to, AMP guanylate cyclase, ion channel or phosphoinositide hydrolysis second messenger systems.

Another example of an assay for CSG antagonists is a competitive assay that combines CSG and a potential antagonist with membrane-bound CSG receptor molecules or recombinant CSG receptor molecules under appropriate conditions for a competitive inhibition assay. CSG can be labeled, such as by radioactivity, such that the number of CSG molecules bound to a receptor molecule can be determined accurately to assess the effectiveness of the potential antagonist.

Potential antagonists include small organic molecules, peptides, polypeptides and antibodies that bind to a CSG polypeptide of the invention and thereby inhibit or extinguish its activity. Potential antagonists also may be small organic molecules, a peptide, a polypeptide such as a closely related

WO 01/92528

PCT/US01/17583

- 62 -

protein or antibody that binds the same sites on a binding molecule, such as a receptor molecule, without inducing CSG-induced activities, thereby preventing the action of CSG by excluding CSG from binding.

5 Potential antagonists include small molecules which bind to and occupy the binding site of the CSG polypeptide thereby preventing binding to cellular binding molecules, such as receptor molecules, such that normal biological activity is prevented. Examples of small molecules include but are not
10 limited to small organic molecules, peptides or peptide-like molecules.

Other potential antagonists include antisense molecules. Antisense technology can be used to control gene expression through antisense DNA or RNA or through triple-helix
15 formation. Antisense techniques are discussed, for example, in Okano, J. Neurochem. 56: 560 (1991); OLIGODEOXYNUCLEOTIDES AS ANTISENSE INHIBITORS OF GENE EXPRESSION, CRC Press, Boca Raton, Fla. (1988). Triple helix formation is discussed in, for instance Lee et al., Nucleic Acids Research 6: 3073
20 (1979); Cooney et al., Science 241: 456 (1988); and Dervan et al., Science 251: 1360 (1991). The methods are based on binding of a polynucleotide to a complementary DNA or RNA. For example, the 5' coding portion of a polynucleotide that encodes a mature CSG polypeptide of the present invention may
25 be used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription thereby preventing transcription and the production of a CSG polypeptide. The antisense RNA
30 oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of the mRNA molecule into a CSG polypeptide. The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed *in vivo* to inhibit production of a CSG.

35 **Compositions**

WO 01/92528

PCT/US01/17583

- 63 -

The present invention also relates to compositions comprising a CSG polynucleotide or a CSG polypeptide or an agonist or antagonist thereof.

For example, a CSG polynucleotide, polypeptide or an agonist or antagonist thereof of the present invention may be employed in combination with a non-sterile or sterile carrier or carriers for use with cells, tissues or organisms, such as a pharmaceutical carrier suitable for administration to a subject. Such compositions comprise, for instance, a media additive or a therapeutically effective amount of a polypeptide of the invention and a pharmaceutically acceptable carrier or excipient. Such carriers may include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol and combinations thereof. The formulation should suit the mode of administration.

Compositions of the present invention will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the polypeptide or other compound alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1 $\mu\text{g}/\text{kg}/\text{day}$ to 10 $\text{mg}/\text{kg}/\text{day}$ of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 $\text{mg}/\text{kg}/\text{day}$, and most preferably for humans between about 0.01 and 1 $\text{mg}/\text{kg}/\text{day}$ for the hormone. If given continuously, the polypeptide or other compound is typically administered at a dose rate of about 1 $\mu\text{g}/\text{kg}/\text{hour}$ to about 50 $\text{mg}/\text{kg}/\text{hour}$, either by 1-4 injections per day or by continuous subcutaneous

WO 01/92528

PCT/US01/17583

- 54 -

infusion, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention are administered orally, rectally, parenterally, intracisternally, intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), buccally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The polypeptide or other compound is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semipermeable polymer matrices in the form of shaped articles, e. g., films, or microcapsules. Sustained-release matrices include polylactides (U.S. Patent 3,773,919 and EP 58481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., *Biopolymers* 22: 547-556 (1983)), poly (2-hydroxyethyl methacrylate) (R. Langer et al., *J. Biomed. Mater. Res.* 15: 167-277 (1981), and R. Langer, *Chem. Tech.* 12: 98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) and poly-D- (-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also include liposomally entrapped polypeptides. Liposomes containing the polypeptide or other compound are prepared by well known methods (Epstein et al., *Proc. Natl. Acad. Sci. USA* 82: 3688-3692 (1985); Kwang et al., *Proc. Natl. Acad. Sci. USA* 77: 4030-4034 (1980); EP 52322; EP 36676; EP 88046; EP 143949; EP 142641; Japanese Pat. Appl. 83-118008;

WO 01/92528

PCT/US01/17583

- 65 -

U.S. Patent 4,485,045 and 4,544,545; and EP 102324). Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal therapy.

For parenteral administration, in one embodiment, the polypeptide or other compound is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation.

For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to the polypeptide or other compound.

Generally, the formulations are prepared by contacting the polypeptide or other compound uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e. g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino

WO 01/92528

PCT/US01/17583

- 66 -

acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, mannose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The polypeptide or other compound is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts or salts of the other compounds.

Any polypeptide to be used for therapeutic administration should be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e. g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

Kits

The invention further relates to pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of

WO 01/92528

PCT/US01/17583

- 67 -

the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, reflecting approval by the agency of the manufacture, use or sale of the product for human administration.

Administration

CSG polypeptides or polynucleotides or other compounds, preferably agonists or antagonists thereof of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

The pharmaceutical compositions may be administered in any effective, convenient manner including, for instance, administration by topical, oral, anal, vaginal, intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal or intradermal routes among others.

The pharmaceutical compositions generally are administered in an amount effective for treatment or prophylaxis of a specific indication or indications. In general, the compositions are administered in an amount of at least about 10 $\mu\text{g}/\text{kg}$ body weight. However, it will be appreciated that optimum dosage will be determined by standard methods for each treatment modality and indication, taking into account the indication, its severity, route of administration, complicating conditions and the like.

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a CSG polypeptide in an individual can be treated by administering the CSG polypeptide of the present invention, preferably in the secreted form, or an agonist thereof. Thus, the invention also provides a method of treatment of an individual in need of an increased level of a CSG polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the CSG polypeptide or an agonist thereof to increase the activity level of the CSG

WO 01/92528

PCT/US01/17583

- 68 -

polypeptide in such an individual. For example, a patient with decreased levels of a CSG polypeptide may receive a daily dose 0.1-100 $\mu\text{g}/\text{kg}$ of a CSG polypeptide or agonist thereof for six consecutive days. Preferably, if a CSG polypeptide is administered it is in the secreted form.

Compositions of the present invention can also be administered to treating increased levels of a CSG polypeptide. For example, antisense technology can be used to inhibit production of a CSG polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer. A patient diagnosed with abnormally increased levels of a polypeptide can be administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is preferably repeated after a 7-day rest period if the treatment was well tolerated. Compositions comprising an antagonist of a CSG polypeptide can also be administered to decrease levels of CSG in a patient.

20 Gene therapy

The CSG polynucleotides, polypeptides, agonists and antagonists that are polypeptides may be employed in accordance with the present invention by expression of such polypeptides *in vivo*, in treatment modalities often referred to as "gene therapy."

Thus, for example, cells from a patient may be engineered with a polynucleotide, such as a DNA or RNA, encoding a polypeptide *ex vivo*, and the engineered cells then can be provided to a patient to be treated with the polypeptide. For example, cells may be engineered *ex vivo* by the use of a retroviral plasmid vector containing RNA encoding a polypeptide of the present invention. Such methods are well-known in the art and their use in the present invention will be apparent from the teachings herein.

WO 01/92528

PCT/US01/17583

- 69 -

Similarly, cells may be engineered *in vivo* for expression of a polypeptide *in vivo* by procedures known in the art. For example, a polynucleotide of the invention may be engineered for expression in a replication defective retroviral vector, as discussed *supra*. The retroviral expression construct then may be isolated and introduced into a packaging cell transduced with a retroviral plasmid vector containing RNA encoding a polypeptide of the present invention such that the packaging cell now produces infectious viral particles containing the gene of interest. These producer cells may be administered to a patient for engineering cells *in vivo* and expression of the polypeptide *in vivo*. These and other methods for administering a polypeptide of the present invention would be apparent to those skilled in the art upon reading the instant application.

Retroviruses from which the retroviral plasmid vectors herein above mentioned may be derived include, but are not limited to, Moloney Murine Leukemia Virus, spleen necrosis virus, retroviruses such as Rous Sarcoma Virus, Harvey Sarcoma Virus, avian leukosis virus, gibbon ape leukemia virus, human immunodeficiency virus, adenovirus, Myeloproliferative Sarcoma Virus, and mammary tumor virus. In one embodiment, the retroviral plasmid vector is derived from Moloney Murine Leukemia Virus.

Such vectors will include one or more promoters for expressing the polypeptide. The selection of a suitable promoter will be apparent to those skilled in the art from the teachings contained herein. However, examples of suitable promoters which may be employed include, but are not limited to, the retroviral LTR, the SV40 promoter, the human cytomegalovirus (CMV) promoter described in Miller et al., *Biotechniques* 7: 980-990 (1989), and eukaryotic cellular promoters such as the histone, RNA polymerase III, and beta-actin promoters. Other viral promoters which may be employed include, but are not limited to, adenovirus promoters,

WO 01/92528

PCT/US01/17583

- 70 -

thymidine kinase (TK) promoters, and B19 parvovirus promoters. Additional promoters which may be used include respiratory syncytial virus (RSV) promoter, inducible promoters such as the MMT promoter, the metallothionein promoter, heat shock promoters, the albumin promoter, the ApoAI promoter, human globin promoters, viral thymidine kinase promoters such as the Herpes Simplex thymidine kinase promoter, retroviral LTRs, the beta-actin promoter, and human growth hormone promoters. The promoter also may be the native promoter which controls the gene encoding the polypeptide.

The nucleic acid sequence encoding the polypeptide of the present invention will be placed under the control of a suitable promoter.

In one embodiment, the retroviral plasmid vector is employed to transduce packaging cell lines to form producer cell lines. Examples of packaging cells which may be transfected include, but are not limited to, the PES01, PA317, Y-2, Y-AM, PA12, T19-14X, VT-19-17-H2, YCRE, YCRIP, GP+E-86, GP+envAml2, and DAN cell lines as described in Miller, A., Human Gene Therapy 1: 5-14 (1990). The vector may be transduced into the packaging cells through any means known in the art. Such means include, but are not limited to, electroporation, the use of liposomes, and CaPO₄ precipitation. Alternatively, the retroviral plasmid vector may be encapsulated into a liposome, or coupled to a lipid, and then administered to a host. The producer cell line will generate infectious retroviral vector particles which are inclusive of the nucleic acid sequence(s) encoding the polypeptides. Such retroviral vector particles then may be employed to transduce eukaryotic cells, either *in vitro* or *in vivo*. The transduced eukaryotic cells will express the nucleic acid sequence(s) encoding the polypeptide. Eukaryotic cells which may be transduced include, but are not limited to, embryonic stem cells, embryonic carcinoma cells, as well as hematopoietic stem cells, hepatocytes, fibroblasts, myoblasts,

WO 01/92528

PCT/US01/17583

- 71 -

keratinocytes, endothelial cells, and bronchial epithelial cells.

An exemplary method of gene therapy involves transplantation of fibroblasts which are capable of expressing
5 a CSG polypeptide or an agonist or antagonist thereof onto a patient. Generally fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture
10 flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e. g., Ham's F12
15 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week. At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is
20 trypsinized and scaled into larger flasks. pMV-7 (Kirschmeier, P. T. et al., DNA, 7: 219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector
25 is fractionated on agarose gel and purified, using glass beads. The cDNA encoding a CSG polypeptide of the present invention or an agonist or antagonist thereof can be amplified using PCR primers which correspond to their 5' and 3' end sequences respectively. Preferably, the 5' primer contains
30 an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for
35 ligation of the two fragments. The ligation mixture is then

WO 01/92528

PCT/US01/17583

- 72 -

used to transform bacteria HB 101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted. Amphotropic pA317 or GP+aml2 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells). Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced. The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

Alternatively, *in vivo* gene therapy methods can be used to treat CSG related disorders, diseases and conditions. Gene therapy methods relate to the introduction of naked nucleic acid (DNA, RNA, and antisense DNA or RNA) sequences into an animal to increase or decrease the expression of the polypeptide.

WO 01/92528

PCT/US01/17583

- 73 -

For example, a CSG polynucleotide of the present invention or a nucleic acid sequence encoding an agonist or antagonist thereto may be operatively linked to a promoter or any other genetic elements necessary for the expression of the polypeptide by the target tissue. Such gene therapy and delivery techniques and methods are known in the art, see, for example, WO 90/11092, WO 98/11779; U.S. Patents 5,693,622, 5,705,151, and 5,580,859; Tabata H. et al. (1997) *Cardiovasc. Res.* 35 (3): 470-479, Chao J et al. (1997) *Pharmacol. Res.* 35 (6): 517-522, Wolff J. A. (1997) *Neuromuscul. Disord.* 7 (5): 314-318, Schwartz B. et al. (1996) *Gene Ther.* 3 (5): 405-411, Tsurumi Y. et al. (1996) *Circulation* 94 (12): 3281-3290 (incorporated herein by reference). The polynucleotide constructs may be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, lung, liver, intestine and the like). The polynucleotide constructs can be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

The term "naked" polynucleotide, DNA or RNA, refers to sequences that are free from any delivery vehicle that acts to assist, promote, or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, polynucleotides may also be delivered in liposome formulations (such as those taught in Felgner P. L. et al. (1995) *Ann. NY Acad. Sci.* 772: 126-139 and Abdallah B. et al. (1995) *Biol. Cell* 85 (1): 1-7) which can be prepared by methods well known to those skilled in the art.

The polynucleotide vector constructs used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain sequences that allow for replication. Any strong promoter known to those skilled in the art can be used for driving the expression of DNA. Unlike other gene therapies techniques,

WO 01/92528

PCT/US01/17583

- 74 -

one major advantage of introducing naked nucleic acid sequences into target cells is the transitory nature of the polynucleotide synthesis in the cells. Studies have shown that non-replicating DNA sequences can be introduced into 5 cells to provide production of the desired polypeptide for periods of up to six months.

The polynucleotide construct can be delivered to the interstitial space of tissues within the an animal, including of muscle, skin, brain, lung, liver, spleen, bone marrow, 10 thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and connective tissue. Interstitial space of the tissues comprises the intercellular fluid, mucopolysaccharide matrix among the 15 reticular fibers of organ tissues, elastic fibers in the walls of vessels or chambers, collagen fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is similarly the space occupied by the plasma of the circulation and the lymph 20 fluid of the lymphatic channels. Delivery to the interstitial space of muscle tissue is preferred. The polynucleotide construct may be conveniently delivered by injection into the tissues comprising these cells. They are preferably delivered to and expressed in persistent, non-dividing cells which are 25 differentiated, although delivery and expression may be achieved in non-differentiated or less completely differentiated cells, such as, for example, stem cells of blood or skin fibroblasts. In vivo muscle cells are particularly competent in their ability to take up and express 30 polynucleotides.

For the naked polynucleotide injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 $\mu\text{g}/\text{kg}$ body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 35 20 mg/kg and more preferably from about 0.05 mg/kg to about

WO 01/92528

PCT/US01/17583

- 75 -

5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the tissue site of injection. The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration. The preferred route of administration is by the parenteral route of injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to lungs or bronchial tissues, throat or mucous membranes of the nose. In addition, naked polynucleotide constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

The dose response effects of injected polynucleotide in muscle *in vivo* is determined as follows. Suitable template DNA for production of mRNA coding for polypeptide of the present invention is prepared in accordance with a standard recombinant DNA methodology. The template DNA, which may be either circular or linear, is either used as naked DNA or complexed with liposomes. The quadriceps muscles of mice are then injected with various amounts of the template DNA.

Five to six week old female and male Balb/C mice are anesthetized by intraperitoneal injection with 0.3 ml of 2.5% Avertin. A 1.5 cm incision is made on the anterior thigh, and the quadriceps muscle is directly visualized. The template DNA is injected in 0.1 ml of carrier in a 1 cc syringe through a 27 gauge needle over one minute, approximately 0.5 cm from the distal insertion site of the muscle into the knee and about 0.2 cm deep. A suture is placed over the injection site for future localization, and the skin is closed with stainless steel clips.

After an appropriate incubation time (e. g., 7 days) muscle extracts are prepared by excising the entire quadriceps. Every fifth 15 μ m cross-section of the individual quadriceps muscles is histochemically stained for protein

WO 01/92528

PCT/US01/17583

- 76 -

expression. A time course for protein expression may be done in a similar fashion except that quadriceps from different mice are harvested at different times. Persistence of DNA in muscle following injection may be determined by Southern blot analysis after preparing total cellular DNA and HIRT supernatants from injected and control mice.

The results of the above experimentation in mice can be used to extrapolate proper dosages and other treatment parameters in humans and other animals using naked DNA.

Nonhuman Transgenic Animals

The CSG polypeptides of the invention can also be expressed in nonhuman transgenic animals. Nonhuman animals of any species, including, but not limited to, mice, rats, rabbits, hamsters, guinea pigs, pigs, micro-pigs, goats, sheep, cows and non-human primates, e. g., baboons, monkeys, and chimpanzees, may be used to generate transgenic animals. Any technique known in the art may be used to introduce the transgene (i. e., polynucleotides of the invention) into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection (Paterson et al., *Appl. Microbiol. Biotechnol.* 40: 691-698 (1994); Carver et al., *Biotechnology (NY)* 11: 1263-1270 (1993); Wright et al., *Biotechnology (NY)* 9: 830-834 (1991); and Hoppe et al., U.S. Patent 4,873,191); retrovirus mediated gene transfer into germ lines (Van der Putten et al., *Proc. Natl. Acad. Sci., USA* 82: 6148-6152 (1985)), blastocysts or embryos; gene targeting in embryonic stem cells (Thompson et al., *Cell* 56: 313-321 (1989)); electroporation of cells or embryos (Lo, 1983, *Mol. Cell. Biol.* 3: 1803-1814 (1983)); introduction of the polynucleotides of the invention using a gene gun (see, e. g., Ulmer et al., *Science* 259: 1745 (1993)); introducing nucleic acid constructs into embryonic pluripotent stem cells and transferring the stem cells back into the blastocyst; and sperm mediated gene transfer (Lavitranco et

WO 01/92528

PCT/US01/17583

- 77 -

al., Cell 57: 717-723 (1989)). For a review of such techniques, see Gordon, "Transgenic Animals," Intl. Rev. Cytol. 115: 171-229 (1989), which is incorporated by reference herein in its entirety.

5 Any technique known in the art may be used to produce transgenic clones containing polynucleotides of the invention, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal, or adult cells induced to quiescence (Campbell et al., Nature 380: 64-66 (1996);
10 Wilmut et al., Nature 385: 810813 (1997)).

The present invention provides for transgenic animals that carry the transgene in all their cells, as well as animals which carry the transgene in some, but not all their cells, i.e., mosaic or chimeric animals. The transgene may
15 be integrated as a single transgene or as multiple copies such as in concatamers, e. g., head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, for example, the teaching of Lasko et al. (Lasko et al., Proc.
20 Natl. Acad. Sci. USA 89: 6232-6236 (1992)). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. When it is desired that the polynucleotide transgene be integrated into
25 the chromosomal site of the endogenous gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous gene are designed for the purpose of integrating, via homologous recombination with chromosomal
30 sequences, into and disrupting the function of the nucleotide sequence of the endogenous gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene in only that cell type, by following, for example, the teaching of Gu et al. (Science
35 265: 103-106 (1994)). The regulatory sequences required for

WO 01/92528

PCT/US01/17583

- 78 -

such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

Once transgenic animals have been generated, the expression of the recombinant gene may be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to verify that integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot analysis of tissue samples obtained from the animal, *in situ* hybridization analysis, and reverse transcriptase-PCR (rt-PCR). Samples of transgenic gene-expressing tissue may also be evaluated immunocytochemically or immunohistochemically using antibodies specific for the transgene product.

Once the founder animals are produced, they may be bred, inbred, outbred, or crossbred to produce colonies of the particular animal. Examples of such breeding strategies include, but are not limited to: outbreeding of founder animals with more than one integration site in order to establish separate lines; inbreeding of separate lines in order to produce compound transgenics that express the transgene at higher levels because of the effects of additive expression of each transgene; crossing of heterozygous transgenic animals to produce animals homozygous for a given integration site in order to both augment expression and eliminate the need for screening of animals by DNA analysis; crossing of separate homozygous lines to produce compound heterozygous or homozygous lines; and breeding to place the transgene on a distinct background that is appropriate for an experimental model of interest.

Transgenic animals of the invention have uses which include, but are not limited to, animal model systems useful

- 79 -

in elaborating the biological function of CSG polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression of CSGs, and in screening for compounds effective in ameliorating such CSG associated conditions and/or disorders.

Knock-Out Animals

Endogenous gene expression can also be reduced by inactivating or "knocking out" the gene and/or its promoter using targeted homologous recombination (e. g., see Smithies et al., Nature 317: 230-234 (1985); Thomas & Capecchi, Cell 51: 503512 (1987); Thompson et al., Cell 5: 313-321 (1989); each of which is incorporated by reference herein in its entirety). For example, a mutant, non-functional CSG polynucleotide of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous CSG polynucleotide sequence (either the coding regions or regulatory regions of the gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express polypeptides of the invention *in vivo*. In another embodiment, techniques known in the art are used to generate knockouts in cells that contain, but do not express the gene of interest. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the targeted gene. Such approaches are particularly suited in research and agricultural fields where modifications to embryonic stem cells can be used to generate animal offspring with an inactive targeted gene (e. g., see Thomas & Capecchi 1987 and Thompson 1989, *supra*). This approach can also be routinely adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site *in vivo* using appropriate viral vectors that will be apparent to those of skill in the art.

In further embodiments of the invention, cells that are genetically engineered to express the CSG polypeptides of the invention, or alternatively, that are genetically engineered

WO 01/92528

PCT/US01/17583

- 80 -

not to express the CSG polypeptides of the invention (e. g., knockouts) are administered to a patient *in vivo*. Such cells may be obtained from the patient or a MHC compatible donor and can include, but are not limited to, fibroblasts, bone marrow cells, blood cells (e. g., lymphocytes), adipocytes, muscle cells, and endothelial cells. The cells are genetically engineered *in vitro* using recombinant DNA techniques to introduce the coding sequence of polypeptides of the invention into the cells, or alternatively, to disrupt the coding sequence and/or endogenous regulatory sequence associated with the polypeptides of the invention, e. g., by transduction (using viral vectors, and preferably vectors that integrate the transgene into the cell genome) or transfection procedures, including, but not limited to, the use of plasmids, cosmids, YACs, naked DNA, electroporation, liposomes, etc.

The coding sequence of the CSG polypeptides of the invention can be placed under the control of a strong constitutive or inducible promoter or promoter/enhancer to achieve expression, and preferably secretion, of the CSG polypeptides of the invention. The engineered cells which express and preferably secrete the CSG polypeptides of the invention can be introduced into the patient systemically, e.g., in the circulation, or intraperitoneally.

Alternatively, the cells can be incorporated into a matrix and implanted in the body, e.g., genetically engineered fibroblasts can be implanted as part of a skin graft or genetically engineered endothelial cells can be implanted as part of a lymphatic or vascular graft (see, for example, U.S. Patent 5,399,349 and U.S. Patent 5,460,959 each of which is incorporated by reference herein in its entirety).

When the cells to be administered are non-autologous or non-MHC compatible cells, they can be administered using well known techniques which prevent the development of a host immune response against the introduced cells. For example,

WO 01/92528

PCT/US01/17583

- 81 -

the cells may be introduced in an encapsulated form which, while allowing for an exchange of components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

5 Transgenic and "knock-out" animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of CSG polypeptides of the present invention, studying conditions and/or disorders associated with aberrant CSG expression, and
10 in screening for compounds effective in ameliorating such CSG associated conditions and/or disorders.

EXAMPLE

The present invention is further described by the
15 following example. The example is provided solely to illustrate the invention by reference to specific embodiments. This exemplification, while illustrating certain aspects of the invention, does not portray the limitations or circumscribe the scope of the disclosed invention.

20 All examples outlined here were carried out using standard techniques, which are well known and routine to those of skill in the art, except where otherwise described in detail. Routine molecular biology techniques of the following example can be carried out as described in standard laboratory
25 manuals, such as Sambrook et al., MOLECULAR CLONING: A LABORATORY MANUAL, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989).

Identification of CSGs

30 Identification of CSGs (Colon Specific Gene) was carried out by a systematic analysis of data in the LIFESEQ Gold database available from Incyte Pharmaceuticals, Palo Alto, CA using the data mining Cancer Leads Automatic Search Package referred to herein as CLASP.

WO 01/92528

PCT/US01/17583

- 82 -

CLASP performs the following steps. First, highly expressed organ specific genes are selected based on the abundance level of the corresponding EST in the targeted organ versus all the other organs. Next, the expression level of
 5 each highly expressed organ specific gene is analyzed in normal tissue, tumor tissue, and tissue libraries associated with tumor or disease. Candidates are selected based upon demonstration of components of ESTs as well as expression
 10 exclusively or more frequently in tumor tissue or tumor libraries.

Thus, CLASP allows the identification of highly expressed organ and cancer specific genes. A final manual in depth evaluation is then performed to finalize the gene selection.

15 Using the CLASP method, the following Incyte sequences were identified as CSGs.

	SEQ ID NO:	LSGold Gene ID
	1	237623
	2	234891
20	3	262167
	4	246508
	5	203279
	6	983538
	7	206344
25	8	222237
	9	118593
	10	337950
	11	982786
	12	398963
30	13	203640
	14	88875
	15	230552
	16	407124
	17	62662
35	18	230495
	19	470880
	20	898601
	21	29586
40	22	370788

WO 01/92528

PCT/US01/17583

- 83 -

Relative Quantitation of Gene Expression

Real-Time quantitative PCR with fluorescent Taqman probes is a quantitation detection system utilizing the 5'-3' nuclease activity of Taq DNA polymerase. The method uses an internal fluorescent oligonucleotide probe (Taqman) labeled with a 5' reporter dye and a downstream, 3' quencher dye. During PCR, the 5'-3' nuclease activity of Taq DNA polymerase releases the reporter, whose fluorescence can then be detected by the laser detector of the Model 7700 Sequence Detection System (PE Applied Biosystems, Foster City, CA, USA).

Amplification of an endogenous control is used to standardize the amount of sample RNA added to the reaction and normalize for Reverse Transcriptase (RT) efficiency. Either cyclophilin, glyceraldehyde-3-phosphate dehydrogenase (GAPDH) or 18S ribosomal RNA (rRNA) was used as this endogenous control. To calculate relative quantitation between all the samples studied, the target RNA levels for one sample was used as the basis for comparative results (calibrator). Quantitation relative to the "calibrator" can be obtained using the standard curve method or the comparative method (User Bulletin #2: ABI PRISM 7700 Sequence Detection System).

The tissue distribution and the level of the target gene were determined for each sample of normal and cancer tissue. Total RNA was extracted from normal tissues, cancer tissues and from cancers and the corresponding matched adjacent tissues. Subsequently, first strand cDNA was prepared with reverse transcriptase and the polymerase chain reaction was done using primers and Taqman probe specific to each target gene. The results were analyzed using the ABI PRISM 7700 Sequence Detector. The absolute numbers are relative levels of expression of the target gene in a particular tissue compared to the calibrator tissue.

The following primers were used for real-time quantitative PCR:

35

WO 01/92528

PCT/US01/17583

- 84 -

forward primer:
 TGGAATAGATTCAGGGTCAT (SEQ ID NO:23)
 reverse primer:
 CGGGGTACCTCACTGACTTC (SEQ ID NO:24)

5 Q-PCR probe:
 TGTCTCCGAGAGAACCCAGGCTCCG (SEQ ID NO:25)

The absolute numbers depicted in Table 1 are relative levels of expression of Gene ID 203279 (also referred to herein as Cln129 or SEQ ID NO:5) in 24 normal different 10 tissues. All the values were compared to normal liver (calibrator). These RNA samples are commercially available pools, originated by pooling samples of a particular tissue from different individuals.

15 Table 1: Relative Levels of CSG Cln129 Expression in Pooled Samples

	TISSUE	NORMAL
	Adrenal Gland	0
20	Bladder	0
	Brain	0
	Cervix	0
	Colon	0.7
	Endometrium	0.4
25	Esophagus	0
	Heart	0
	Kidney	3.7
	Liver	1
	Lung	0
30	Mammary Gland	0.2
	Muscle	0
	Ovary	0
	Pancreas	0
	Prostate	0
35	Rectum	23
	Small Intestine	1.5
	Spleen	0
	Stomach	0.8
	Testis	0.1
40	Thymus	0.4
	Trachea	0
	Uterus	0

The relative levels of expression in Table 1 show that 45 Cln129 mRNA expression is detected at high levels in the pool

WO 01/92528

PCT/US01/17583

- 85 -

of normal rectum (23), and at a lower levels in kidney (3.7). In contrast, Cln129 is expressed at very low levels in the other 22 normal tissue pools analyzed. Further, the level of expression in rectum is 6 fold higher compared to the expression in kidney. These results demonstrate that Cln129 mRNA expression is highly specific for rectum tissue.

The absolute numbers in Table 1 were obtained analyzing pools of samples of a particular tissue from different individuals. They can not be compared to the absolute numbers originated from RNA obtained from tissue samples of a single individual in Table 2.

The absolute numbers depicted in Table 2 are relative levels of expression of Cln129 in 21 pairs of matching samples. All the values are compared to normal liver (calibrator). A matching pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual.

Table 2: Relative Levels of CSG Cln129 Expression in Individual Samples

Sample ID	Tissue	CANCER	NORMAL
ClnA598	Colon ascending (C)1	383	24
ClnCM67	Colon cecum (B)2	15	8
ClnCXGA	Colon rectum (A)3	85	118
ClnMF38	Colon splenic flexure (D)4	33	18
ClnRC24	Colon rectum (D)5	77	29
ClnRC67	Colon rectum (B)6	0.9	15
ClnRS45	Colon rectosigmoid (C)7	161	25
ClnSG27	Colon sigmoid (C)8	48	13
ClnSG33	Colon sigmoid (C)9	190	100
ClnSG36	Colon sigmoid (B)10	186	93
ClnRC89	Colon rectum (D)11	0	28
Bld32XK	Bladder 1	0	0
CvxKS52	Cervix 1	0	0
Endo8XA	Endometrium 1	0	0.7
Kid106XD	Kidney 1	0	6.7
Liv15XA	Liver 1	1.7	3.2
Lng47XQ	Lung 1	3.4	0
Mam59X	Mammary Gland 1	1.3	0
Pro34B	Prostate 1	0	0

WO 01/92528

PCT/US01/17583

- 86 -

SmInt	Small Intestine 1	5.4	1.7
Utr85XU	Uterus 1	0.9	0

0= Negative

5 Among 42 samples in Table 2 representing 11 different tissues significant expression is seen only in colon, kidney, and small intestine tissues. These results confirm the tissue specificity results obtained with normal samples shown in Table 1. Table 1 and Table 2 represent a combined total of 10 66 samples in 24 human tissue types. Only one small intestine sample, one lung sample, one liver sample, and one kidney sample showed expression of Cln129, out of a total of forty-two samples representing 22 different tissue types different than colon and rectum.

15 Comparisons of the level of mRNA expression in colon cancer samples and the normal adjacent tissue from the same individuals are shown in Table 2. Cln129 is expressed at higher levels in 8 of 11 (73%) cancer samples (colon 1, 2, 4, 5, 7, 8, 9, 10) compared to normal adjacent tissue.

20 Altogether, the high level of tissue specificity, plus the mRNA upregulation in 73% of the colon cancer matching samples tested indicate Cln129 to be a diagnostic marker for colon cancer.

25 It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of 30 the appended claims.

The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other disclosures) in the 35 Background of the Invention, Detailed Description, and

WO 01/92528

PCT/US01/17583

- 87 -

Examples is hereby incorporated herein by reference. Further, the hard copy of the sequence listing submitted herewith and the corresponding computer readable form are both incorporated herein by reference in their entireties.

WO 01/92528

PCT/US01/17583

- 88 -

What is claimed is:

1. An CSG comprising:
 - (a) a polynucleotide of SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22, or a variant thereof;
 - (b) a protein expressed by a polynucleotide of SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22, or a variant thereof; or
 - (c) a polynucleotide which is capable of hybridizing under stringent conditions to the antisense sequence of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, or 22.
2. A method for diagnosing the presence of colon cancer in a patient comprising:
 - (a) determining levels of a CSG of claim 1 in cells, tissues or bodily fluids in a patient; and
 - (b) comparing the determined levels of CSG with levels of CSG in cells, tissues or bodily fluids from a normal human control, wherein a change in determined levels of CSG in said patient versus normal human control is associated with the presence of colon cancer.
3. A method of diagnosing metastases of colon cancer in a patient comprising:
 - (a) identifying a patient having colon cancer that is not known to have metastasized;
 - (b) determining levels of a CSG of claim 1 in a sample of cells, tissues, or bodily fluid from said patient; and
 - (c) comparing the determined CSG levels with levels of CSG in cells, tissue, or bodily fluid of a normal human control, wherein an increase in determined CSG levels in the patient versus the normal human control is associated with a cancer which has metastasized.

WO 01/92528

PCT/US01/17583

- 89 -

4. A method of staging colon cancer in a patient having colon cancer comprising:
- (a) identifying a patient having colon cancer;
 - (b) determining levels of a CSG of claim 1 in a sample of cells, tissue, or bodily fluid from said patient; and
 - (c) comparing determined CSG levels with levels of CSG in cells, tissues, or bodily fluid of a normal human control, wherein an increase in determined CSG levels in said patient versus the normal human control is associated with a cancer which is progressing and a decrease in the determined CSG levels is associated with a cancer which is regressing or in remission.
5. A method of monitoring colon cancer in a patient for the onset of metastasis comprising:
- (a) identifying a patient having colon cancer that is not known to have metastasized;
 - (b) periodically determining levels of a CSG of claim 1 in samples of cells, tissues, or bodily fluid from said patient; and
 - (c) comparing the periodically determined CSG levels with levels of CSG in cells, tissues, or bodily fluid of a normal human control, wherein an increase in any one of the periodically determined CSG levels in the patient versus the normal human control is associated with a cancer which has metastasized.
6. A method of monitoring a change in stage of colon cancer in a patient comprising:
- (a) identifying a patient having colon cancer;
 - (b) periodically determining levels of a CSG of claim 1 in cells, tissues, or bodily fluid from said patient; and
 - (c) comparing the periodically determined CSG levels with levels of CSG in cells, tissues, or bodily fluid of a normal human control, wherein an increase in any one of the

WO 01/92528

PCT/US01/17583

- 90 -

periodically determined CSG levels in the patient versus the normal human control is associated with a cancer which is progressing in stage and a decrease is associated with a cancer which is regressing in stage or in remission.

5

7. A method of identifying potential therapeutic agents for use in imaging and treating colon cancer comprising screening compounds for an ability to bind to or decrease expression of a CSG of claim 1 relative to the CSG in the
10 absence of the compound wherein the ability of the compound to bind to the CSG or decrease expression of the CSG is indicative of the compound being useful in imaging and treating colon cancer.

15 8. An antibody which specifically binds a polypeptide encoded by a CSG of claim 1.

9. A method of imaging colon cancer in a patient comprising administering to the patient an antibody of claim
20 8.

10. The method of claim 9 wherein said antibody is labeled with paramagnetic ions or a radioisotope.

25 11. A method of treating colon cancer in a patient comprising administering to the patient a compound which downregulates expression or activity of a CSG of claim 1.

12. A method of inducing an immune response against a
30 target cell expressing a CSG of claim 1 comprising delivering to a human patient an immunogenically stimulatory amount of a CSG polypeptide so that an immune response is mounted against the target cell.

WO 01/92528

PCT/US01/17583

- 91 -

13. The method of claim 12 wherein the CSG polypeptide is encoded by a polynucleotide of SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, or 22.

5

14. A vaccine for treating colon cancer comprising an CSG of claim 1.

WO 01/92528

PC/T/US01/17583

SEQUENCE LISTING

<110> Macina, Roberto A
Chen, Sei-Yu
Pluta, Jason
Sun, Yongming
Recipon, Herve
diaDexus, Inc.

<130> Method of Diagnosing, Monitoring, Staging, Imaging and
Treating Colon Cancer

<130> DEX-0208

<140>

<141>

<150> 60/207,383

<151> 2000-05-26

<160> 25

<170> PatentIn Ver. 2.1

<210> 1

<211> 911

<212> DNA

<213> Homo sapiens

<400> 1
tttttttttt ttgctgtttt gtttaaatg tttactgtac aaagaacaa aaccagga 60
tagtaoagt attgaacagt agcagagatg gttgtgaaat aagggaaccac ttggagac 120
agttttattg gcttctgttc ttacooaga aagactgttg atttttgaaa acttctact 180
gaaatgtatt tttctgtctt tcccagaggaa gggcaactta cagtgttcct aggttttct 240
gtgagtgagg tgccagtcgt gattcaaaat atccttgcct gcactgcagc tecttaggga 300
gtcttttctt gcccttgagg cctgggcaaga ctctccctct acaccctccc gccctctccc 360
acgacgcagc agaaataaag cacaaacctca gaaagtctca ggcacgaaga actgtctctg 420
ggtggagcat gggaccttta ttctgtaaga catcaggctc cagatatgaa ctttcagcag 480
aagoccttgc cgggagcaaa gggacagaaa agctgagatg aacagtgcct ggcagcaate 540
acagccgggc aagggtgctc ctagcctctc atccccggc cgggggcagc tggaggtgce 600
tcagaaggtg cttctgtctt cctgcagggg ctgaaacac caaggcactc cagggatctc 660
ggagtaaang cagcagcccc ggttgttgca ctctctgggg gtgacatggg ggtagccgca 720
gtccacctg tcttggctg gaaaggcaca ctggtttgca gctgtcccag acaaagccct 780
gtcagctgcc agagcccttg ctgggacagg ccaactact tctcagcag agctggagga 840
cagcaaggcc aggaccagcc ccagcatgca gaggctctg gtagccatga ccaccgtggg 900
ctccgggaag c 911

WO 01/92528

PC/T/US01/17583

<210> 2
<211> 322
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> (244)

<400> 2
gacaagcaac aaaccobtga tgathattca tcaactggat gagtgcccac acagtcaegc 60
tttaaagaaa gtgtttgtgtg aaaaataaga aatccagaaa ttggcagagc agtttgtcct 120
cctcaatctg gtttatgaaa caactgacaa acacobbct cctgatggcc agtatgtccc 180
caggattatg tttgttgacc catctctgac agtttagacc gatatactct gaagatattc 240
aaancctctc tatgcttaag aaocctgnaga taacagctctg ttgcttgaca acatgaagaa 300
agctctcaag ttgctgaaga ct 322

<210> 3
<211> 4569
<212> DNA
<213> Homo sapiens

<400> 3
atggataaat tcttcaaac ataccctctc ccaagactaa accaggaaga agttgaaatc 60
cigaatagac caataacagg ctctgatatt gtggcaataa tcaagagctt accaaccaaa 120
aagagtccag gaccoagatgg atccacagct gaattctacc agaggtacaa gggaggaactg 180
gtaccattcc ctctgaaagt attacaatca atagaaaaag aggcaatcct cctaactcg 240
ttttatgagg ccaacatcat cctgatacca aagccgggca gggcccaaac caaaaaagag 300
aatttttagc caatatcttt gatgaacatt gatgcaaaaa tcttcaataa aatactggca 360
aaccgaatcc agcagcacat caaaaagctt atccaccatg atcaagtggg cticactcct 420
gggataacca aagacaataa cccatgatt atctcaatag atgcagaaaa ggcctttgac 480
aaaaatcaac aacccttcat gctaaaaacc ctcaataaag tagatattga tgggacatat 540
ctcaaaaataa taagagctat ctatggcaaa gccacagcca atatcatact gaatgggcaa 600
aaactggaag cattcccttt gaaaactggc acaagacagg gatgccctct ctccaccctc 660
ctattcaaca tagttttgga agttctggcc agggcaatta ggcaggagaa ggaataaag 720
ggtttcaat taggaaaaga ggaagtcaaa ttgtccctgt ttgcaggtga catgattgta 780
tacctagaaa accccattct ctcaagccaa aatctcctta agctgataag caacttcagc 840
aaagctcagc gatcaaaaat caatgtacaa aaatcccaag cattcctata caccataaac 900
agagaaacag agagccaaat catgaatgaa ctccattca caattgcttc aagagaata 960
aaatcactag gaatccaact tacaagggat gtgaaggacc tcttcaagga gaactacaaa 1020
ccactgctca atgaataaa agaggataca aacaaatgga agaacttcc atgctcatgg 1080
ataggaagaa tcaatatogt gaaaatggcc atactgcca agattatgct agatataaag 1140
ggcattcaat taggaaaaga ggaagtcaaa ttgtccctgt ttgcagatga catgattgta 1200
tatctagaaa accccattgt ctcaagccaa aatctcctta agctgataag caacttcagc 1260
aaagtctcag gatcaaaaat caatgtacaa aaatcccaag cattcttata caccacaac 1320
aganaaacag agagccaaat catgagtcaa ctccattca caattgcttc aagagaata 1380
aaatcactag gaatccaact tacaagggac gtgaaggacc tottcaagga gaactacaaa 1440

WO 01/92528

PC/US01/17583

```

ccactgctca aggaataaaa agaggataca aaccaatgga agaacaattc atgctcatgg 1500
atagggaagaa tcaatatcgt gaaaatggcc atactgcco aagagaaat cacagggaga 1560
tgtacagcaa tggggccatt taagagttct gtgttcatct tgattcttca cctctagaa 1620
ggggccctga gtaattcact cttcagctg acaacaatg gctatgaagg cttgtcgtt 1680
gcaatgcacc coastgtgcc agaagatgaa acactcattc acaaatataa ggggggtac 1740
aogtcaaac atgaggaagg gagagtcaga gagaactct ctctccccc gtcacaata 1800
catacacaca caccacacgc acaagctcgt gtgcacacac acacgccat gcacacacgc 1860
agacatacac gccacacacgc acgtcagaag gacatgtgga ccagggcctc tctgtatctg 1920
cttgaagcta caggaaagcg attttatttc aaaaatgttg ccattttgat tctgaacaa 1980
tgggaagcaa aggtcgacta tctgagacca aaacttgaga cctacaaaaa tctgtatgt 2040
tgtttcacc aagtctactc tccaggtaat gatgaacct acactgagca gatggcaac 2100
tgtggagaga aggggtgaaag gatccacctc actcctgatt tcattgcagg aaaaaagta 2160
gctgaatag gccacaagc tagggcattt gtcctagat gggctcatct acgatggga 2220
gtatttgaog agtacaataa tgatgagaaa ttctacttat ccaatggaag aatacaagca 2280
gtaagatgtt cagcaggtat tactgtgata aatgtagtaa agaagtgtca gggaggcagc 2340
tgtttcacc aagatgac acatcaataa gtaacaggac tctatgaaa aggatgtgag 2400
tttbtctcc aatcccgcga gccggagaag gcttctataa tgtttgcaca acatgttgat 2460
tctatagtt aattctgtac agacaataac cacacaagag aagctccaaa caagcaaat 2520
caaaaatgca atctccgaag cacatgggaa gtgatccgtg attctgagga ctttaagaa 2580
accactccta tgacaacaca gccaccaaat cccaccttct cactgtgca gatggacaa 2640
agaattgtgt gtttqctct tgacaaactt ggaagcctgg cgaactgtaa ccgctcaat 2700
cgaactgcat aagcagccca gctttctctg ctgcagacag ttgagctggg gtcctgggt 2760
gggatgtga catttgacag tctgcccact gtaacaastg aactcataca gataaacagt 2820
ggcagtgada gggacacact gcgcaaaaaga ttacctgcag cagctcagc agggactgct 2880
atctgcagcg ggtctgact ggcatttact gatattgtggc aacatttgcc tgttttccat 2940
gacacacagc agttatgggg agtgcgacaa gaaaatccaa attgggctc tctggcctgc 3000
agcttagtga ttagggaagaa atatcccaact gatggatctg aaattgtgct gctgacgat 3060
ggggaagaca acactataag tgggtgcttt aacgaggtca acaaaagtgg tgcctcctc 3120
cacacagtcg ctttggggcc cctctcagct caagaactag aggagctgct caaaatgaca 3180
ggaggtttac agacatagc ttccagatcaa gttcagaaca atggcctcat tgatgcttt 3240
ggggcccttt cactcagaaa tggagctgtc tctcagcctt ccactccagct tgegagtaag 3300
ggatfaacc tccagaacag ccagtgatg aatggcagc tgatctgga cagcagcgtg 3360
ggaaggaca ctttgttct tatcactgg acaatgcagc ctcccacat cctctctg 3420
gatcccagtg gacagaagca aggtggcttt gtagtgaca aaaaaccaa aatggctac 3480
ctccaaatcc caggcattgc taaggttggc acttggaaat acagctctga agcaagctca 3540
caaaccttga cctgactgt cactcctcgt gcctccaatg ctacctgccc tccaattaca 3600
gtgacttcca aacgacaa ggacacacgc aatctccca gccctctggt agtttatgca 3660
aatattgccc aaggagcctc cccaattctc agggccagtg tcaagcctc gattgatac 3720
gtgaatgaa aaacagttac cttggaacta ctggataatg gagcaggtgc tgatgctact 3780
aaggatgacg gtgtctactc aaggtatttc acaactatg acacgaatgg tagatacagt 3840
gtaaaagtgc gggctctggg aggagttaac gcagccagac ggaagatgat accccagcag 3900
agtgagcacc tgaactacc tggctggatt gagaatgat aatacaatg gaatccacca 3960
agacctgaaa ttaataagga tgatgttcaa cacaagcaag tgtgttccag cagaacatcc 4020
tcggggaggt catttgggcc tctctgatgc ccaaatgctc ccatacctga tctcttccc 4080
cctggccaaa tcaccagcct gaagcggaa attcacgggg ccagbctcat taactgact 4140
tggacagctc ctgggatga ttatgacct ggaacagctc acaagtatat cattgaata 4200
agtaacagta ttcttgatct cagagacaag ttcaatgaat ctctcaagt gaatacact 4260
gctctcctc caaaggagc caactctgag gaagctttt tgtttaaac agaaaacatt 4320

```

WO 01/92528

PCT/US01/17583

```

accttfgaaa atggcacaga tcttttcatt gctattcagg ctgttgataa ggtcgcctcg 4380
aaatcagaaa tctccaacct tgcacagata tctttgttta ttcctccaca gactccgcca 4440
gagacacctc gtccctgatga aacgtctgct ccttgctcta atattcatat caacagcacc 4500
attcctggca ttcacatctt aaaaattatg tgggaagtga taggagaact gcagctgten 4560
atagcctag 4569

```

<210> 4

<211> 3206

<212> DNA

<213> Homo sapiens

<400> 4

```

ttcggctgga gtgtaaaact gccaaaggaaa gtaattacct gtaggagttt gctgagcttg 60
aagagtgaaa actgtttgta atgagcctga tcataaaacg gaccagggcca ttcattattc 120
ctcaagtgtt aatataactga cttatgoagt attcaaacaa aaacattgca ctgagatggtg 180
caagaacage gtaaaatgaa agccatcatt catcttactc ttcttgccgc tcctttctgt 240
aaacaacacc accaaaccaag gcaactcagc tgatgctgta acaaccacag aaactcggac 300
tagtggctct acagttagctg cagctgatac cactgaaact aatttgccct gaaactgcta 360
gcaccacage aaatacaact tctttcccaa cagctacttc acctgctccc ccataatta 420
gtcaacatag ttcctccaca attcctacac ctgctcccc cataattagt acacatagtt 480
cctccacaat tctataacct actgctgcag acagtggctc aaccaaaaat gtaaatctag 540
ttagctacct ctgacataat caacgcttca tctccaaagt atggattaat tcacaatggc 600
tcctctgaaa acaaaaagta acaatgaaat gtcctccacc acagaagaca atcaatctc 660
agtggctccc cactgggacc cgtcttattt ggatgacct gcacgcctaa accgcccagt 720
gtcccagcaa tctctgcaaa agatgatccc cctgtgcaga taattcgta ttgtttgta 780
agcttgctat aatacaagtt ttgctctgtg tttagaaggg tattaactca actctctac 840
atgtaagaaa ggaagggtat tccctggaga agatttcagt gacagatca gaaacattg 900
accagaaga gaaacattcc atggcctatc aagsettgca tagtgaatt actagctgtg 960
ttaaagatgt atttggcaca tctgtttatg gacagactgt aattctact gtaaggcaca 1020
tctctgtcac caagattctg aaatgogtgc ttgatgadaa gttttgttaa tghtaacaata 1080
gtaacaattt tggcaagaaa caaagtgac aatgagaaga ctgtgactgg agaaaattaa 1140
taaagcaatt tataagtgc taaagcaact tttctaaact atgattggac cctgtoggtg 1200
ggattgatt gagggtggtg aaccaagact ggtggatga ctgctcaat gggtttagca 1260
tcgatgtgc aaatgctgac ctgcaaaagg ctaaccaca gacccttcc tggcttgctt 1320
cagctctcag agtctctga tgcctgcaac gcacagcaca agcgaatgct taataaagaa 1380
gagtggtggg gtcccctgca gtgttgcctt gctgcccgg tctaccagga agatgctaat 1440
gggaactgcc aaaagtgtgc atttgggcta cagtggact gactgtaagg acaaatcca 1500
gctgactctc acttattgt gggcaccatc gctggcattg tcaattctcag catgataatt 1560
gcattgattg tcactagcaa gatcaataaa caaaagcga gcatattgaa gaacgagaac 1620
ttgatgacg aagactttca aaatctaaaa ctgctgtcgc acaggcttca ccaatctatg 1680
gagcataacg gagcgtcttc cctcaggtca ggatcaggc ctccaagaga ccgcttagat 1740
gcaaaaatcc cgtagtttca agacacagca geatgcccoc ggcctgacta ttagaatcca 1800
tcagaatgtg gaacccgcca tggcccocaa ccatatgac atatotatta ttctagcagt 1860
gtttagacaa gactgcatgg agaagtgagc accacgtaaa gactctggcc tccgggagtt 1920
tcttcttcca tctagacata ctgcccagtc tcatctgcaa tggcaacglt gtgcaatgtc 1980
ttgcaacga catccagct cacttgctaa aataagaato tatgacatta acatgtagct 2040
cgatgetatt agogctgtgo tdagagaggt gggctttctt caatcagtaa caaagtactg 2100

```

WO 01/92528

PC/US01/17583

```

agacaatgct taggggttgg ttcttaatt ctttccctg gtagggcaac aagaccocat 2160
ttccaaatct agaggaaagc ctcccagca ttgctttgct cctggggcca aaccatgctt 2220
cttgagttaa gttgacctaa ctcccctgg gacgacatac cgcatacaat gtaggagttc 2280
gagggggatg agaaaaggat accccacatc ttccataggg tcaacaageta cactctcttg 2340
acaagtCaga ataggggaca cctgcltcta tccctccaat ggaggagatt ctggccaaac 2400
ccccctttt ttgaaaacca ggcgccaga gcttgccaac cttagcctca cccaagaaga 2460
ctggaaaagg acaatatttt ttoagctttt tcaaggaggcg tgccttggga atccaggaac 2520
gtttttgatg ctaattagaa ggccctggact ataataatgt ccatctatgg ggttttaac 2580
taccagtttt gaacatgcta ggaggcagaa cggggccaga gagtaaaaa acatgacctg 2640
gtagaaggaa gaaggcmeta ggaactggg tggggaggat caattagaga ggaggcact 2700
ggatccacc ttggttoott aggtccctc ctccatgcag caaaggagca ctctctaaag 2760
tcattgcctc ccgaagactg gctggggagaa ggtttaaaaa acaaaaaaac caggagtaa 2820
gagocctagg gtcagttttg aaaattggag acaactttgt cttggcaaaag ggtccaaga 2880
ggaggacttg ttgctcagga gtcccagcog tccagcctcg ggggttaagg tctctgaggt 2940
gtgcacatgg ggctcagcc ttctctggtg acccagagct cagctgtggc caccacaca 3000
caaccacaca cacacacaca cacacacaaa tgggggcaac ccacatccac gtaaccaagc 3060
tttaacaca atgttattag tgtccctttt tatttctaag agccctgtcc tcttaaaagt 3120
tattttattt gtattattta ttgttcttg actgttaatt gtagaatgta atgcaataaa 3180
gtgcctttg tagatggaaa aaaaaa 3206

```

```

<210> 5
<211> 3610
<212> DNA
<213> Homo sapiens

```

```

<400> 5
gatgtgggca cgcctcagag ccagaagttt atggctcca cctgctaat ctgacaggaa 60
gcttctgctc cccagttctc cccagccact gttgtctaca gattccagga aaccatccc 120
cctgtgacct cagggttgcg tctgttctcc accctaggga ccagaaggag ccaggagtaa 180
agaactggct taacttggcg ccaactggaa attctgggta attcagagc cctggaaatt 240
tggaccactc ccctgatagc gtagtgggca gggttctagg gaacacaaga ggggggcca 300
ggtggcttcc ctgtctgctg attcttggct ctctctctct ctctctctct ctctctgtct 360
ctctctctct ctctctctct cagocctgca gcccgtttc cctccctgag ctccagttgt 420
agtgtgactc gattccaggg aaagggaaat cgcctgggct gaggagaccg gactggaccg 480
gctggggagc gacccgtgat gcccgcaacc cccgtccctc ggaaggggtg gtccatgagc 540
tgctctctct taccctctgt cgggggcctc tggaggatgc ggtgaccatt cctctgggac 600
acaacttctg ccggctctgc ctcccgcgc tctcccagat gggggcccaa tctctgtggc 660
aagatcctgc tctgcccctc ctgccaagag gagtagcagg cagagactcc catggccctc 720
gtgcccctgg gcccgctggg agataactta ctgcgaggag caccgcgaga agatctactt 780
cttcttgaga gaacgatgcc gacttctctc gttgttctcg caggaggggt cccacgacc 840
aggcgcacac cgtgggttcc ctggaagagg ccatcagccc ctaccgggat cgtctcagga 900
gtcagctgga agctctgagc acggagagag atgagattgt aggatgtaaa gtttcaagaa 960
gaccagaagc ttcaagtggc gctgactcag atcgaacaag caagaagccg tccgggtgca 1020
cacagctcct tgagaggtcg caagcgggag ctgcagcagc agcagatgct cctgctggcg 1080
caggactgag tggtaocgctc gtagtoacag atttggagg agaggatga atatatcaca 1140
aaggtctctg aggaagtcac ccggttggaa gccccagctc aaggagctcg gaggagaggt 1200
gtcagcagcc agcaagtgag ottctacaag atgtcagagt caagccagag cagggtgtgag 1260

```

WO 01/92528

PC/US01/17583

```

atgaagactt ttgtgagctc tgaggcoatt tctccctgac ctgttcaaga agatccgtga 1320
tttccacagg aaaatactca cccctccaga gatgatgaga atgttctcaa gaaaacttgg 1380
cgatcatctt ggaatatgat tcagggttca tcaactctgga cccctcagacc gccagccgga 1440
gacctgttcc tctcgggaaga cagggaagca gtgaggtaca ccggcagaa gaagagccctg 1500
ccagacagcc cctcgcgctt cgaagcctc ccggcggctc tgggcttccc gggctctctc 1560
tcggggcccc acccgtggca ggttgcctg cagctgggag ccggcggcgg ctgcacggtg 1620
ggggtggccg ggaagggggt gaggaggaca gggagagatg ggaactcagc ccgaggaccg 1680
cgtctggccc gtgatcatct ctgcaccaag cagtctggg ccagcacctc cccggcacc 1740
gacctgtccg ctgagcgaga tcccgcgag gctgagagt ccctctggac tacgaggcgg 1800
ggcaggtgac cctccacaac gcccagagcc caggggccca tccctcacct tcactggctc 1860
ttttctcggg ccaaggtctt cctgtctctt gcccgcctgg acacaaggg toctggcctt 1920
aggctgacac ggggaaatg gggcgcgca agggcgcgca agcggagagc gggctctcc 1980
gggatccagc tcgcgccctg gccagtgatg gcccggggg cccctctgac ccgctggagg 2040
cgagagaaac ccggggaact gagtctgaa cagcggttgt ttttaactta tttatctag 2100
gccctcagct cctcagctc ctgagcctcc ctgtgacgct ctggccttct ctgcacctca 2160
gagtgcagaa ccacagacgg ctctggctgt gccatgggca acagcaacc taggaacctg 2220
ccggccttcc ggggaaaaac taagaagga gactctaaa atgtaatgtt taactgttt 2280
cagataaatt atctgggaa aaactcaggtt ttgtctggac ttgactaat ttgtacagt 2340
aaactctgac ttgacacac accctgaagat gccctcacct ttgtaggct tagggcctt 2400
ttatcagccc tgggtgagcc ccaggccccc ttccttccc tccctctctg gtcatctctc 2460
tggacttgta gagaatgtcc taagaagtg tgaactcacg acctctggat tccatgtctc 2520
caattagccc tgatgggact ggaagaagg taaatccaa tgggatctg cctgtgttg 2580
caatttaggg ccagatggc tcgaggag 2610

```

- <210> 6
- <211> 1627
- <212> DNA
- <213> Homo sapiens

```

<400> 6
ttttatttcc tagagtgata tatattttt ggcttttcc ttttttttc ttccaaaaca 60
aacaaataga gctttaggcc cctcgcctc cccacacca ccgcagaacc ctccatata 120
atcgacaact gaaaacaagc gagacaatca cccccaaag gatcacgaa cagagccaca 180
agtctcacag acagccaccg acaaaacaaa aaaacttgct actaggaatg tccgccttgc 240
atgatcatgt agaagcagga gcaagagtct acaaatgaa tggggacctg attagtatg 300
gggttagcagg gggatggctc ggaatcagaa ggttaagctc tccatgctga tggcttaggt 360
gccatttgc cctttcctg ttgcacggcg ggtactgttt tcccagaagc gcgcgcagc 420
acctggccac gcagatctgc agtccatgag cctgtgtagt caggatgccc atagcccgtt 480
ccctggggcg ggtctccttt ggcctgggg ctagagccc caagcccggg gcttctctgc 540
gtgggtcgag aagccagcgg gatctggagg aagcctgcag agcgttctgc cactggggcc 600
gttgcactct cctcgtccca tgtaccactt gtaccgggaa gggagtcaat gggaaatcgag 660
tgcccaata aattctcatt tggactctcc tggcctggct tccctgtcta cagtggggtt 720
gacctagcg gtggaacgga aggtggaggg attttctac aagggcgggc ttgacttgcg 780
ggtgcaaggt ggatacgacc aagagagatt gatctcagag ctaggagagg tgcggaagaa 840
tgcagtcccg gtcgaagagc aagagaagct acagtctgct aagtgtgca cagatgaaca 900
ggaggcaaac atgtcgaagg ctcatagcag ccacagtgct acctatatt gttggaagga 960
tgagggaaac atcctgctgg taaatataac atttctgca acaataatgt atataatgt 1020

```

WO 01/92528

PC/US01/17583

```

gggagggtggg gaggtagctcc acctaaagata ccttcataaa accaactgctt gcttttttt 1090
gfacttttcta gcccccggg tggggggcta ggthtgetcc atcttcccc tggcccttgg 1140
cctgagaata gtgggccact coactgggaat ggtatggcca tgcctgagcc tttgggctgc 1200
aacctcctac ttagggagctc gctctagac atctccctgg tgggtatttg cacttaggggt 1260
agaacccggg cttgcctgac agtctgaggg ctgtttttgc caatttgggt tggcattggtc 1320
tgcaactggg agtgcacct cacttgacty aatgggtggt gtgagctcc cccattactg 1380
tgtgtgaatg tctgtgagc tgtgtagagt tggagtgtcc ctgggtgact tttgggtggg 1440
tgttagaag aaacgggcaa gctggaagt aggggctagg acttcccaga aaaattacag 1500
ggcatactag gacttgact ggggtctctc tttcctgtg gccatcaca ttcttaggaa 1560
ccaactattt ctatcttcta aatcaacaaa actttctct gacacctaga gacctgaga 1620
agccatg 1627

```

```

<210> 7
<211> 929
<212> DNA
<213> Homo sapiens

```

```

<400> 7
catgtatgca ataaaaata aaagatacat acacaaaatt ctttaaatgt cccacacaca 60
agacaaatac gbtccaat acatcagctt ctgaaagctc tgcaccactc tacacgctgc 120
tcctctctgac tagtactgoc ctctgcccc tctgtccac gtgtcaact cccatcacc 180
ctttaaaaoc agattggaat attttgctc tgtgaagctt tccctgacta tcccgggat 240
agaataatgt ttccaactgt gttttgtcat ttactgcta taataagaat aegaagaac 300
atgtattttt gaaaagtatc tgtgatctct aatgagcttg taacaatctt gaggaataga 360
gactaagttt tgcctctttg ttccccaaa gagaactta ttaataacat ttaccatctc 420
tttagagaga gggtttttcc catctctgtg aaaaagctcc agaacttaca accaggaata 480
agtgtaatg ggatagaacc aatgtagaga acagcatatg atatgtgaaa tgtactttat 540
tattaatacg aattcagttg gctccacaga tgaacctttt tgcacaaactg ggggaaagc 600
atcttctgta aaggtatctt tagaaaaata tgtataattt gaaaaatggt tatccaaatt 660
taacatttgt catataaag gctcataaaa cgtgtgtggc tgtgtttctc aaatgtgg 720
ggcraattgg tcacattatg cctagacatt ctggttttgt tcttgggt taataatggt 780
tgtgtctta tacagaaaag gaaatctgga catcttgcct ctgttattaa tacactgtc 840
attactaata aaagtggttt gttgatatgc taaaatagggt gaaaaagctg tccctttgca 900
tgaaattaac tagggaatac ttctttata 929

```

```

<210> 8
<211> 2302
<212> DNA
<213> Homo sapiens

```

```

<400> 8
gagagggaagc agcatcagga caocttacca ccaactgcgc tgcctcagca tccacccgc 60
agcccacgtg tggcaaacgg ggaaggggt ggaagcaac gccggagacc acgtggaaga 120
aggggcctgt ttggcccttc catctgggtg cggggagccc ctaggccctc cggccatggc 180
cgacagcggc gatgctggca gctccggccc ctgggtgaaa tgcctcacc aacagagaaa 240
gaaaaacaag gaagcgcag tgggggtgco gctcctccgc cagcccctc cggggagcc 300

```

WO 01/92528

PC/US01/17583

```

cacgccacct ggcgcgccca gcccgactg gaccagcagc tcccgggaga acccagcccc 360
cacaatctect cggggggccc ggcgagcccc ccaaaccaga caagttatac ggggacaaat 420
cggcgagcag ccgcccgaat ttgaagatct cggctccgg ccgctttaag gagaacagga 480
aagtgcgggc cacgctgctc ccggaggcgg gcaggtcttc ggaggaggca ggtttctctg 540
gtgaccccca cgaggacaag cagtagcccc aatagcctgc ggcctccagg actgctctacc 600
cagcactacc ccaaaccccc agttccaaac ccgagacttc aggcocgccc ccttacgctg 660
tgcctcaatc caccaaatcc agaataatta cacaatgctc tcatgattaa atttttctgg 720
aacttgaggt gtcaattggg ttctcaagat ttcatgacc caaggatgcc ttgaatattt 780
atthtgggta agagaagata cctgcgcggc agtagggctg cataattatt tttttctctc 840
agtgcacagg ttttaaatag ccacactaaa ataggctgta cacttttgta gtttaacatc 900
tcaaaqcaat cctgccttat gttlaaaatg cttctaacta agaatgcttc tgcctctccc 960
gcactccggt cacttacagg tataagtcta cccctagaag tgcatttctc accgcaatta 1020
aaaactagca ctgtgatbtt ctttctctaca gactcctgaa abaaactagcc accttctctg 1080
catttgatga ggtctactaga gttccaagct ctagctctgt actaggagca cagggggcca 1140
ggggcccaag aatacgtctt cttagaagaa aaaactaatt atgccacctc tottccgctg 1200
caggtatctc tetettaaca caaataaata tttcaaatgc atccttgagg gtcacgaaat 1260
attgagaacc caabaagaca ctacaatttc cagaaaaata aatcatgaa ggcattgctg 1320
taaatattct gcaatttggg ggaatgagaa caacgcgtaa gggggcggac ctgaagctctc 1380
gthtttgtaa ctgggggttt agaggtagtg ctgggtaggc ctgctcggag cggcaggtct 1440
atthgggcta ctgctgtctc tctgggggtt caccaggaaa gcctgcctcc tccgaggacc 1500
tgcccgcctc cgggagcagc gtagcgcgca ctttctctct ctcttaaaag cggcggagc 1560
gcgagatctt caacattgctg cgggctgctg cgggatgtgt ccccgataaa ctgtctctgt 1620
ttgggggctt cgcgcgcgcc cccgaggaga cttcgggggt ctggttctcc cgggagctgc 1680
tggtcaactc cgggtggggc ggcgaagtg cgtgggctc cccggagcgg ggtggggcgg 1740
gaggcggcac cccaactgct gcttctctgc ttttctctct gctgttgggt agcgatttcc 1800
accaggggcc ctagctgcca gcatcgcgct tctcggccat ggcggagggg cctaggggct 1860
cccggcaccc agatgggaagg gccaaagcgg ccccttctc cacgtgctct cgggcgcttc 1920
actccacccc ttcccggct tgcacacgt ggggctgctg ggtggatgct gaggcagcgg 1980
cctgtgctgg gaggagggcc ctgggaacca agtgcactct ctctacaggt gaacggtatt 2040
aattaagctc atggtcaaac aagtcacgaa atttccctcc aaagatttgc ccccatcgac 2100
ttctgtccca ggaagctctt tcatgagat acttaggaga attttatctc ccagttagga 2160
agagaaggac aagcttatga taattgggtt tgggttctct ttaaaattct ggttttgac 2220
caattctgct ttgtgacttt caaagaagca tgcctagact taactttccc ttgaaaaaac 2280
gcatcctaaa tcttcccttt act 2303

```

```

<210> 9
<211> 1769
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> unsure
<222> (578)..(948)

```

```

<400> 9
attctccagt cacttctcat agacttctgg cttcctgtca ggcataaac aagcttgaaa 60
ttttcactg gtttctaacc cttaagtaaaa agctgaacaa actcaaaagt caacaacttg 120

```

WO 01/92528

PC/US01/17583

```

ttaaaatccc tcagagatgg ctgggcactc catctctgag tggactcttg accccatcct 180
cactcatgac gccatctcca acctgctgtg gcgctcatal cctccagtgg atcctgggac 240
ctccccccag tggagctggo caggcaggtg ctgtctgata ggtttgctgc ccatccaca 300
tacacctgtg tctcatgat gatgcattg tcataagggtg ggttccctg gactgagaag 360
tgaaccagcc actggcgctc cacttagact ctaccagttg acaaaaactt aaactctagt 420
tgtgttttct gaggttgata gtagaggaag aaaaactttc acatgcctgt tttgaggctt 480
ctcctctttt tgcctaacto tgcacaggaa ctaggggcag ggagcgcttt ctaaatctac 540
taacatcaca cacattgctt ctctaaactt ggcctcattt ctccccttat gtaactgaca 600
cacacctaaq agttcctctc tgaccggttc tgcctcttta acaggtctca catcctctc 660
totgttccag gactcactga ttccaacca ctctcagcat ctgctcttag agcataatgt 720
gatcaatttg gaattccagc cagacctaaa ccttagcata atattaaaat gaatactac 780
tctctagaaa attagataat tagatcttta ggaccaatga taagaattgt ccacctatg 840
gaaaagactt taeggtgttc ccccAAatgt cttcacnnn nnnnnnnnn nnnnnnnnn 900
nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn tacagattga 960
gatccccaaa tccgaaaatc caaaaatcca aaatgracca aaactctgaa atgctcccaa 1020
aatcaaaac ttttgagtgc caacataaca attaaaaca aaatgctcac tggagcattt 1080
cggatttggg attggatttt ggattttcag attagggatg ctacagctggg tctcagatgc 1140
ctgatacatt caattcctgg tttcttaka cctactcca cgtctgggag atttatgtag 1200
tggaaatttg tgttggcatt gtaagtgtta acagatttgt agagactccc cttttcaat 1260
tgtatggag cactagtacc tctcagctgc agaaatcaat ttcaaaaat ggaatggaac 1320
aaataaaatt gaaacatacc tatgatggag gctgtcctgt ggcctcctg ctccccccag 1380
aagggttagg ctctcatgtg agggagtttg ggaaccagg tggagatagc catgtacaca 1440
gcccotggaa agggatgtgt ctagtccgaa tgaagcagga agcccgagat gggagatca 1500
tgtgtgtgat catagtctat tttatgtggg aggatgttca gcaagcgggc agagtcatgg 1560
ggtgggtctg tggctctcgt gacttcaaga atgaagcgc agactctca agcaagtgtt 1620
accagctctt aaaggctgtg cggaccocaa gagtggcag cagcaagatt tatggtgaag 1680
accgaaagaa caaagcttcc acagtgtgga agggggacct gagcgggttg cactgctg 1740
ctaggsgoaa agttctcct gtggactga 1769

```

<210> 10

<211> 2159

<212> DNA

<213> Homo sapiens

<400> 10

```

cactagcaga gaagctgttg tecttccacc accagocccg gccaccctgc tccaagacca 60
gctctctygg gggaccagge acccggcctt cactggcacc cagggaagcc tctcagcag 120
cgtcaacatg tcaaggccca gcagcagaga cattaacttg caccggaag agtactccca 180
gaacctcacc tcagagccca cctctctgca gcacaggttg gagcaottga tgcactgcaa 240
gcaagggagt cagagagctc aggggoccca ggatgccttg cagaagctgt tcgagatgga 300
tgcacagggc cgggtgtgga gccaaagact gatctctcag gtcaggagcg gctggtgca 360
gctcctggac attgagaoca aggaggaact ggaactctac cgcctagaca gcattccaggc 420
catgaabgtg gogctcaaca catgctccta caactccatc ctgtccatca cgtgagagga 480
gcccggcttg ccaggcaeta gcaactctgt ctccagctgc caggagctgg gggcagagcg 540
actgaagacc agctgacaga aggctctgga ggaagagctg gagcaaaagc ctccacttgg 600
aggcctcag ccaggcaggg acagatggag ggggocctgt atggaaagcg cgtcctctat 660
ggagcaggca cgttatcttg agcgggggat cctccagaa cagccccacc agaggacctt 720

```

WO 01/92528

PC/T/US01/17583

```

agagcagcagc ctccaccat ccccaaggcc cctgcccgc cacaccagt cccgagaacc 780
aagtgccctt actctgcctc ctccaaggcg gteectctcc cccgaggacc cagcagggga 840
cgagggaagtg ctgaaccatg tccaaagga cattgagctg ttoatggaa agctggagaa 900
ggcccaggca aagaccagca ggaagaaga atttgggaaa gaagagaaca aggccaggg 960
aggtctcacc caggccacgt acagttgact gottccagaa gatcagcac agctcaacc 1020
tcctgggaag gctggcccacc tggctgaagg agacaagtgc cctgagctc gtacacatcc 1080
tcctcaagtc cctgaacttc atoctggcca ggtgcccctga gctggccta gcagcccag 1140
tgatctcacc cctcctcacc cctaaagcta tcaacctgt acagtctgt ctaagctcac 1200
ctgagagtaa cctttggatg gggttggccc cagcctggac cactagccgg gccgactgga 1260
caggcgatga gcccttgccc taccaccca cattctcaga tgactggcaa ctccagagc 1320
ctccacgcca agcaacctta ggataccagg acctgttctc ccttggggcc tccagtccc 1380
aaaacctgcc agccagtcct tgaaaatgca agtcttctac gagtttgaag ctggaatcc 1440
cacgggaaac tgcctgtggt ccaggtagag aagctggagg ttctggacca cagcaagcgg 1500
tggtggctgg tgaagaatga ggcgggacgg agcggctaca tccaagcaa catcctggag 1560
ccctcagacc cggggacccc tgggacccag ggcagctcac cctctcagg ttccaatgct 1620
tcgacttagc tcgagggcctg aagaggtcac agactggctg caggcagaga acttctcac 1680
tgccaccggtg agsacactg ggtccctgac gggggagccc agctacttg cattagacc 1740
tggggagcta ccaggtgct atgcccaca gaggcccctc accaaatcct gtcccggctg 1800
gaggctgta gaaggatgct tgggataag ccttaggca ccagctaga caectcaag 1860
aacccagccc cgtctgatga agatggcaga totgataccc attagagccc cggaaatcc 1920
tcttctggal cccagtttgc agcaaaccc acacctccag cgtcacacag caaaaacat 1980
ggacagggcc agaggtgaa gcaaacagt tccctctggt ctgtgttgg gctcccagc 2040
taaccaccta ttattttac ctcttccca acctggagc atttatgct aggtctgca 2100
agaatctggt cagtcctct cctctcaat aaaaacatct tcaagcttga aaaaaaaaa 2159

```

```

<210> 11
<211> 3872
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> (2663)..(2664)

```

```

<400> 11
gaaaccgaca caaatacctg aaatacacag ccacagcag acacacacgg aagcactcta 60
tgacaaaac actcacacag tacacaccat gctgcacata cctgaccca aacagctaa 120
caagccctga gggctctccg ggtgcccctg gggctattgc ccacccctcc caccgtccc 180
gctagggtga gatggtgttc cccagggaa0 agaagtctcc agtcccatct taagctctgc 240
cggatcccgc gtgacatcag ctgaccccct cggcgtctcc gggagctgtg agctctgtg 300
tggggccagg ccggcaccag gcacagaccc ttaggccctt gttgggagaa cagcagaggg 360
ctctctgttc caatgctctg ctctgggttc aactgctggt tctctagag gctctctctc 420
agactcctag gtatgtggga ccagggaggg cgggtctctg ccaaggggcc actggggtca 480
gcccaggaga gggctgtggca gtgtgttggg ccgtttgtag gacacacac gctctggcatt 540
ggctaggggc aggtctgcct tcccttagcag ttctgcagct tgctcttaag gcttggcagg 600
gctgggcttc tcagggagag ctgggtggg ggtatctctc agttcccctt cactttctct 660
gttcccaga aggccatgag gttggtgccc ccaggacccc ccttctgtaa gataggaaat 720

```

WO 01/92528

PC/T/US01/17583

```

ctctactcag agaggctggg ctgcagccca ggccccacag tgggccaaga ctaaggctctt 780
gagatgcgeg gcaactgggc tttcaggtga gatctctgct cttcagcctt ttccaagcaa 840
ggatgagact ttggggcccc aagcaatctg tttgcagggc ctgggcaacc tggccccctc 900
tcccttgcag ggtggaagca aggaagcac tatctctggc cacatagatc agctggctac 960
acctctggtt gtttggcccc gaatagabat tggccagctc tgggtctctc tgtgtcccca 1020
gccccaggct tccagggcag ctgctttcc tgaggcattg ggcagaatbc cttgtggcaa 1080
ggagatcgtc gcaacagagcc cagctgggac tggcacaagt aattcagggc tgccttctt 1140
cctctatggg agtccggaga gcccaagcctg tgettcacaa ggtatgtgg cctcaagaa 1200
gtctcttttt aggccacagg ccttccatct gtgaaatggg ggatgggttc agactttatg 1260
cctcgaagaa atccttccag ccttggccat cttggacttc tgggctacc ctggcttaca 1320
ggggtcttgt tgccctgggt gtcccagtt cttgaaagx atcagctgg gaggggccac 1380
accctgaoca tccccttta tccctctga gatgtttggt aggaagctg ggtccaggg 1440
atctcatttc ttgttccatc catgcagggg ttgcttacct cggtaggaa accctcaggc 1500
gggtgcaggt gcacaggtag gggaggatgg agagggcagt ggtgcctgaa gccctggatg 1560
ggcggagctg accccccaac accaactcta tcatgctcgc tctcctctgt cccccagag 1620
ctgctgato attgctaacg aatgaactct agcccagctg gtgaccccaa tctccacgc 1680
cctccaggg gccaaatggg aacatcaacc tgggtgctct tccagcaacc caaatgccc 1740
gccccaggac ttgcacttcc tcaaatgcat cggcagaagg gaactcagtg gaagtctct 1800
actgtcccaa gcgcaagtct gatggggcgt tctatgcagt gaatggctact acagaagaa 1860
gtctacttta aatgaagaaa gacagatgc cacatcatgg cagagcgcag tgtctctct 1920
aagaaagtgc ggcaaccctt cctctggggc ctgcttact cctccagac accctgaga 1980
ctctacttct gtctcagct atgtcaagg gggaggagct ctctctccac ctgcagcgt 2040
tgggcctggt tcttggagcc cctgggccc atgtctacgt gctgaggtg ccagccgcc 2100
ttggctacct gaactccctc aacatcatt acagggatct gaaaacagga gaacattct 2160
cttggactgc cagcccctgc cctccctcat tctcagggac cctggtctgt gacggattt 2220
ggcctctgca aggaaggtgt agagctgaa gaccccacat ccactctctg tggtaacct 2280
gagtatctgt cccctgaaag tgcctctgga aagagcctta tgcctgagca gctgactgt 2340
ggtgcttggg ggcagctctc tacgagatgc tccatggct gcccccctc tacagccaag 2400
atgtatccca gatgtatgg aacatctgc accagcgcct acagatcccc ggtgcccga 2460
cagtggccgc ctgtgacctc ctgcaagcc tctccacaa ggcacagagg cagcggctgg 2520
gtcccaagc agactttctt tgagattaag aaaccatgta ttcttcagcc ccataaactg 2580
ggatgactgt taccacaaga ggtcaactcc acccttcaac ccaaatgtga caggacctgg 2640
ctgacttggg agcatttttt gamcccaga gttcccccag gaagctgtgt ccaagtcct 2700
tggctgtacc cctgacact gtggccagca gctctggggc ctaagctgc atttctctgg 2760
atttcttat gcgocagagg atgatgaact cttggattgc tagaagagaa ggcctgtga 2820
aaactactgag gccagctggt attagtaagg aattaccttc agctgctagg aagagcact 2880
caactaaca atgcttctat ccgagttagt caggtttatt gttattgcca gcatcata 2940
aagatgagaa tatagtctc tacgaggtg ccatggatct ggcaggatca ggtcctcag 3000
actacctca cagagactgt atctctgccc tgcacaact gacaaatggc ttccaaatgt 3060
ttaggtttc ttacaaagat ggttaactggg agctctaagg ctgcttatt ttgggtttt 3120
taggaaaggg aaaaatggag gaaagggag aagagcaag ggcctttt aaagacttt 3180
ccctaaagc tccatccaat gagctttctg ctctcctc acttaaccac ccacccctac 3240
ctgggaatgg agcctggga gatgtgctt atttctggg tadgtgact tccctataa 3300
caagggggt ctgacctaa gacattaggg gagaatgctt ggtaggcagc cagcactct 3360
ttaccagagg gcctcctggt gtttggatt tgatctcaat gtgaaacat gacagagatg 3420
taacaagctc atagggatc aatatctct attgtctat gttgatgata ttgtctctg 3480
ttgtgggtaa taactgacat tttgttatt ggtctgggt gccctggtta tctgaaccc 3540
cttctgtct ccaagagaacc cctatttta tgagacttca tggggggcca ataactact 3600

```

WO 01/92528

PC/US01/17583

```

ccacettaaga gtacotgasa atgtatgaca ctgactttcc cagcctcccc tttagctaggg 3660
ccaggcatgg ggaccaggyca taacacctgtg ccacattttg actcagggaa gggatogggg 3720
gagctctttt gtgtggtaac tgtgataaca gtaccoccaa aattgagttc ctggtgtaga 3780
agtacaagg atgcaaacctg tagcagttgg tgctcagtgg cagcaacgcc atcagaccag 3840
ccctgcaatg tcatctctgg aagcctcaag tg 3872

```

<210> 12

<211> 4728

<212> DNA

<213> Homo sapiens

<400> 12

```

atggccagcc agcgggtaag ctccagcac gaggtgtacc cagcggagcc agccacagcc 60
cctggggccc ccagccagga gctggaggag cagaccgtgt cccgtcaggt gttcatcgtg 120
caggagctgg aggtccgaga ccggctcgcc tctccccaga tcaacaagtt cctgtacctg 180
cacaccagtg agcggatgcc ccgactgccc cactctaaca tgcaccaaat caaagcgtcg 240
catgtggccc ccaactacaa cctgggtggg cctgagtgct gtctccgctg ctccgtgatg 300
ccctgcgggc tcaatgtgga ccaggatgcc ctctctctcc tcaaggactt cttaactagt 360
ctgggtggcc gcaatcaacc cgtgggtccc ggggagacct ccctgagggc tggccccagc 420
actcagagcc agccacagag ccccccggaa gggcagcccg aagcgtaga gaccctgggt 480
ctcagggagg cccagggagg tggacacagc cctcccctc ctgaccagca gcccatctac 540
ttcagagagt tccgcttca cctgtgaggt cccatctggc tggattacca tggcaagcac 600
gtcacgatgg accagtgagg cacttttgcg ggctctctca tggcctggc ccaactcaac 660
tgctccagag tgaagctaaa ccggctctgt tgcagggcac ggcctcctgg tgtggacaag 720
gtgctgggct atgacctcaa ctagtggtg caggacatcc gcaagaacca gctgcccgcc 780
ctgctgggag gcctggggccc catgcactcg gttgtccagc tcttccaagg gttccgggac 840
ctgctgtggc tggccattga ccaatcagc aaggatggcc gctcatgcg ggggctgca 900
caggggctgg cctcccttgg ctcaaccaca gcctctgccc ccttggaaat cagcaacggg 960
ttggtagagg ctatccagcc cacagctgag accgtgtatg acatcctgtc cccggcagcc 1020
cccgctctcc gctccctgca ggataagcgc tctggcggga ggtgcgcag gggccagcag 1080
cctgccagcc tggcgggggg tgtggccaag gctcagcaca cagtgcgaga gggcatctg 1140
gatacagctc agaccatctg tgcgktggca tggcggggcc atgagtagaa ggggctgacg 1200
ggcgcctggt ggggctgat ccgccaagctg ccccagactg tggtagagcc gctcatctg 1260
gcccagggag ccaagtcagc cctgctcggg ggcctgagca accagattgt ccccgagccc 1320
cacaaaggacc acgcccctcaa gactggcaac tgtcacggga acctgtctgg gaggagcag 1380
aacacgcttt gcaagagaaa gctctgctc acagaccctt gggctcactc agggaccctg 1440
gccagcagct gcttctctc cccacagcgg agagagacc aagggtcca gggcggatg 1500
ttccaccagc gccagccag cgtgcagggt ggcctcccc ccaactctt tcttagtctc 1560
atcttcagct tcccatcaga ggcacacct atgaaatcag gcactgggag gtcctctggg 1620
actgacaagt gccagctgtc ccttgcctgc tctctgccc atggtgcaag caggagggga 1680
aggagtctgt gcagcacacg gggcggcagg tgtgggccc ggtgataag aagcctcgg 1740
gaaaagacca tggacctggg gccacgaaga ctggggagcc cagcaactcc atgtggaagt 1800
gccactcaggt tccagtgggg ctgctgttat ctggggcgag gcccagtaac caccgaaga 1860
gagagggcag taagcttoca gcaaggggtg taccagcgg agccagccc agggcctgag 1920
gcccacagcc aggagctgga ggaagcagcc ctgtcccgtc aggtgtcat cgtgcaggag 1980
ctggaggtcc gagaccggct gcctcctcc cagatcaaca agttcctgta cctacacag 2040
agtgaagcga tggcggagc tggccactct aacatgtcca ccatcaaac cctgcatgtg 2100

```

WO 01/92528

PC/T/US01/17583

```

gccccacta ccaacctggg tgggocfag tgetgtctcc gogbtctgct gatgccccg 2160
cggctcaatg tggaccagga tgcctcttc tctctcaagg acttctctcac tagtctggtg 2220
gocggatca acccctggtt ccaagggag acctccctg aggtctgccc cgagactcga 2280
gcccagccca ccagccccc ggaagggcag gccgaaggcg tagagaccac tgglttcag 2340
gagggccdag gaggtagaca cagcccctcc cctctgacc agcagccat ctacttcaga 2400
gagttccgct tcaagcttga ggtccccatc tggttgatt accatggcaa gcaagtcaag 2460
atggaccag tgggcaactt tgcctggctc ctcatcgccc tggcccaact caactgctcc 2520
gagctgaagc taaagcggct ctggtgcagg cacgggctcc tgggttgga caaggtgctg 2580
ggctatgccc tcaacgagt gctgcaggac atccgcaaga accagctgcc cggcctgctg 2640
ggggcgtgg gccccatgca ctggttgtc cagctcttcc aagggttccg ggacctgctg 2700
tggctgocca ttgagcagta caggagggat ggcgcctca tggggggct gcaggaggg 2760
gctggctctc ttggctcaco cacagcctct gccgcccctg aactcagcaa ccggttgga 2820
caggctatcc agggccacgc tgagaccctg tatgacatcc tgtcccggc agcccctgc 2880
tcccctccc tgcaggataa ggcctctgog cggaggctgc gcaggggcca gcagctgccc 2940
gacctcggg aggggtggtc caagcctac gacacagtgc gagagggcat ctgtgataca 3000
gctocagcca tctgtgagct ggcctcggg ggcctagac agaagggct gcagggccc 3060
gtggggggg tgatccgcca gctgcccog actgtggtga agccctcat cctggccag 3120
gagggcaegt ccagcctgct cggggccatg cgaaccaga ttgtcccga cggcccaag 3180
gaccacgccc tcaagactgg caactgtca cgaacctgt cgggagggg cggagaacag 3240
ctttgcaaga ggaagctctg cctcaagag cctcgggctc actcaggac cctggccagc 3300
agctgctcc tctcccaca gccgagagag acccaaggtt ccagggcgg atgcttcca 3360
ccaggccagc ccagcgtgca gggtagctc cccccacac ttctctctag tctcatctc 3420
aagttccat acagggcat cctcatgaaa tcaggcactg gtaggtccct gggactgac 3480
aagtgcacag tgtccctgct tgtctctctg ccccatggct gcagcagga gggagggagt 3540
gctggcagca ccggggcgc caggtgtgg ccccgatga taagaagct cggtaaaag 3600
accatggacc tggggccacg aagactggg agcccagca ctccatgtgg aagtccca 3660
tggttccagt gggctgctg ttatctggg cagggccag taccacgaa gaaggaggg 3720
caggtctgct ccagcagacc agccaggact accgtggca cgtcccagg ccagatggtg 3780
goggttagtg gaggctgtc tggtaggctg ccgagccga gtycacagg ctctgacct 3840
tgaattgaca gccagtctc tggctctccc ttggtctgac aatccatag gtcacagga 3900
tgtctgctc aatgcccagc accaggaact gcaggatag gggagggccg ggggtgtcca 3960
cagctcagca gagatcctg gaccccagt cagcactcat ggtcccact cctctgtct 4020
cattccccgt gaatgagcct gaacagctc agtctctgcc ctgcccctg tgcctgtgg 4080
cacctctatg ctttgccat gctgtctctc tgggctgcaa tactctctct agcttattg 4140
ccaggctcac tcttactaac ccttcaagc tctgtccaag catttgctgc ctccagaag 4200
ccttatgaa gcttotaagt ccccaactgg gcaccccac acagtgtgc cgcagagca 4260
tgcctctcg gaccccggg tgcctgttcc tgcctatgct tgcactctc ttcccactc 4320
gtgagctcag tcccagccc aaggcgcgt cccaataaa tgtttgctga accaatctg 4380
agcctctgtc ttgcaacctg aggaagcaac ccaccgaca atgcagtggt gccaaaggg 4440
ggctgagtg ctatggccca gtrttgtgc ttggagccc ccccccagg atggggccct 4500
gagccagcct ccccatctgc ttctactct cccctccttt gccagtctca tctccctgga 4560
gacagccct gtagttggtg gacagctct tccagccct aggatctca agagggcca 4620
ggaccccagc tgetggtaga ggaagagcag ccaaccagg acaggacagc tgaccccac 4680
cctgtcccgc ccccacaac agcctcattt ccaactattt ctttctgg 4728

```

<210> 13
<211> 6650

WO 01/92528

PC/T/US01/17583

<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> (4298)

<220>
<221> unsure
<222> (4307)

<220>
<221> unsure
<222> (4311)

<220>
<221> unsure
<222> (4313)

<220>
<221> unsure
<222> (4315)

<220>
<221> unsure
<222> (4327)

```

<400> 13
tctctccacat accggctcag ctctctccagg acgcagcccg ccagacacgc tctggaagct 60
gggaaccggc ccttgttttg ttcabgaaca ttgggtttag tgccctggcaa cttgatgcat 120
atggaagagc aatgccaagt gatctgacat aatacaaat cacgaagtga cattcaatca 180
caagcaaatg tggaaattcc aaagagaagt ggtgagatct ttaactagta cagtgaagat 240
ggggaaaaat gacatacctg cagcagatgt gggctgaaaa taccctcttc tctgccaat 300
caggaaatgt acctgttttt gggataaac ttttagaaaa ggaaggcca aaactaccg 360
ttggctttct gaaaagggaag cataaatgtt ctttccctcc atttgcctgg atctgagaac 420
ctgcaatttg tattagctag tggaaagcagt atgtatggtt gaagtgcatt gctgcagctg 480
gtagctgag tggtagccac cagctgcagc tggctgccct ctggccctgg ctgctgatgg 540
ctaccctgca ggcaggcttt ggcgcgcacg gactggtact ggcagcagcg gtcgagctcg 600
aaagatcagc agaacagaaa gotattatca gactgatccc cttgaaaatg gaccccacg 660
gaaaactgaa tctcaacttg gaaggtgtgt ttgctggtgt tgcgaaata actccagcag 720
aaggaaaatt aatgcagtc caccocgtgt acctgtgcaa tgcagtgat gacgacaatc 780
tggagctcgg attcacagc atcgtcaagc tggagagtc cgcagggcc ccccgccct 840
gectgtaact gctcagcaag gctcggatgg cgggtgagcg aggagccagt gctgtcctt 900
ttgacatcac tgggatcga gctgctgctg agcagctgca gcagccctg gggctgacct 960
ggcagtggt gttgatctgg ggtaatgacg ctgagaagct gatggagttt tctgtacaat 1020
gaaccgaaaa ggcctatglt gaggattgac gctgagagga gcccccggtc gtcgccaagca 1080
ttatgatgt gtcgatccta actgacatgt ggtgggcacc atctttgtga tcatcctggc 1140
ttcgggtgctg cgcatacggg. gccgccccg ccacagcagg ccggtatcgc ttcagcagag 1200

```

WO 01/92528

PC/US01/17583

```

aacagcctgg gccatcagcc agctggccac caggaggtac caggccagct gacggcagge 1260
cgggggtgag tggccagact caggagagag ctgcagctca gccctgtgt gtgcatctg 1320
tcggggagag ttctctgagg ggcaggagct acgggtcatt tccctgctcc atgagttcca 1380
tcgtaactgt gtggaccctt ggttacatca gcctgggact tggccctct gctgttcaa 1440
catcacagag ggagattcat ttccccagtc cctgggaccc tctgatctt accaagaacc 1500
aggtcagaag ctccacctca ttccccagca tcccgcccat gccactacc accctcctgc 1560
tgcctacctg ttgggcccct ccaggagtcg agtggctcgg cccccaagac ctggtccctt 1620
cctgccatcc caggagccag gcactggccc tggcatcac cgttcccc gaagtccaca 1680
tccccgggtc ccaggagagc agcagcgcct ggcaggagcc cagcacccct atgcaacaag 1740
ctggggaatg agccacctcc aatccacctc acagcaccct gctgcttgc cagtgcctct 1800
ggctgggccc aggcccctg acagcagtg atctggaga agctattgca cagaacgag 1860
tgggtacctg gcagatgggc cagccagtga ctccagctca gggccctgtc atgctcttc 1920
cagtgaactt gtgtaoact gcaaggacat cagccctacg ggggtccatg gcagcagttc 1980
tactttctgc agctccctaa gcagtgaatt tgaccctcta gtgtaactga gcccaaaag 2040
ggatccccag cgagtggaca tgcagcctag tgtgacctct cggcctcgtt ccttggactc 2100
ggctgtgccc acaggggaaa ccagggttcc cagccatgtc cactaccacc gccaccgga 2160
ccaccctacc aaaaagcggg ttccagtgca tggcaggaa cctggcccag aaaccggagt 2220
ccccagtcct aggcctccta ttccctggag acagccccag ccagagcccc ctctctctga 2280
tcagcagagc accggatcca actcagcagc cccctcgggg cggctctcta acccaagtg 2340
ccccaggccc ctccctgagc cagcccttgg cccagtlgac gcctccagca tctgcccag 2400
taccagcagt ctgttcaagt tgcaacagaat ccacgctct tctgcccga cacctcaac 2460
gagaaaaggg acggggcggg tccctctga gcccccctt gggccctcgg ccaccaagga 2520
tgcacatgt gcccccagta ctgcccagat ttttccccat tacaccacca gtgtgcag 2580
atccttggtc cccagaggca cacccttga actgtggacc tccaggctg gaacaagag 2640
ctgtaccag aaaaaccagc gccctgtta ctcaaatca acagccagtg tggctgtgcc 2700
tgactctcg accagcccct ggaaccact ccacctgggg agggccttc tgcaatggag 2760
ttctgaccc gccaggggca ggcctatgcc ttatccgac tgcaggtgc tgtggccca 2820
gectggctca gaggaggaac tccaggagct gtgtgaacag gactgtgtga gatgtccag 2880
cctagctcca accaagagtg tgcctcagga tgtttttgg cccctacctg gcacagagtc 2940
ctgctcctgt gtgaaatgga atggaccaca gcaaacacca ttcttttgg cgtacttct 3000
aggaagcact gggaaagagg ctggatgat gtgggagggt gagaggtgc cgtttctgc 3060
tccagctcca gaccttgcct tgaacaaaa catctgcaga tgcagcaac atccatgtcc 3120
agccaggaca accagctgct gccctgtggc tgtgtggctt ggtcccttg aagctctgct 3180
ttttgaagg cagaagctca gotatggta gccagggtt tccaaagtg ctgctcttc 3240
tccaaccct acttggtttc cctacacccc aatgctcat gttcatacca gccaaagtgg 3300
ttcagagaa accaatgaca cctttatcac ctccctctct tgggtagagc tctgtgagca 3360
ccagcgttt gcccccctca cagtaaggct gctccatcag gggcaacctt ggtctatca 3420
tttctcttt ttgctaaag gaccagtag cataggtgag cctgagcac taaaaggag 3480
gggtccctgg aagctttccc agctatagtg cgggagttct gttccctgga ggggtgggta 3540
cagcagcctt tggttctctt gggggttgag aataagaat agtgggtag ggaaaaactc 3600
ctctttgag atttctctgc tcagagtcct tgagtagtta gaaagygaga atttctgctg 3660
ggcctttatt ctggggcagg aggaagagat gggaaatta ggtagaanga ggcaaaaatt 3720
tccgtttag cgggggccc aaaaaagttt tttttttgg aaaaagttt tttcttagaa 3780
caaggatggo aaaaagggtg daccagccat aggaagagat caaacgtgtg aaccttggg 3840
gtttgggaca ggcctatgag gccccagctc cctagtata agccatacag gtccaaagga 3900
tctctcagat gagagtgagc tttagagcag aagtgctggc gctgcgatct gagtgcgacc 3960
aagagctga tagggcctag atgcaggta gacaatctca gcgcccagg ccagctctga 4020
cccactcttt gcccctcag cgcactatc ccactttgga aabgtgaatt gtgggggca 4080

```

WO 01/92528

PC/T/US01/17583

```

aaegtgggg caagaggacc cccaactggg aaacttttcc cctccagggt tagttgggga 4140
actagacccc tcaggtaaac caccactggc gtaatttata tctgaaccca gaccgagcgc 4200
tttgaatcag gcaactaaact ccagaatat atttatttgc taatatattt atccccaaat 4260
tggtgtggt cctgtgggtt tgttctgtcg tggagctngt ccagctngca ngngngtga 4320
gcaagcngtc catgcttctg ttgtctgaca tctaaagaaa gtaaatlact batgtttaca 4380
gaggtaggc tccgattcat gaaatggata ggttagagta gaggggcttg gccaatlaag 4440
aactggtttg taagccocca aaagctgggc ttaagtgaag atcagggaaa ggaagaaga 4500
catgaactgg aatccttaac tgtgcttcca gtctattatt attatactgt tcacttcaca 4560
cattatocat acttcagggt gactcagacc tggggcaaat actctgtggc ctgctctttt 4620
cagtcctata aatgggccta cttaattgct gtttagcaggc ctatacatga gataatagag 4680
tgtagaazga tatgttccaa aagtggaaaa gtttattcca agtgatagaa gaacatccaa 4740
acotgtcaca agaagcccat ctgaaaacaa gcatgggacc gccaaocaga agaaagccc 4800
cccggaagca gctcaatcaa ggaggtcggg ctggaatgac agccagcgg ggccgaaac 4860
tatttatatc ccaagctcc tctcagata acacaaatga ctgctctctg cctgcactcg 4920
ggctatttgc aggaacagaa gctgtgtctc cactggcctg aagttccag gggccagaaa 4980
ggggcctttg tgcctctctc acaaggacaa agtctccct ctgctctccc gagaaggtt 5040
tgggtagggg gtgggtggtt tagtgcctat agaacaagc atttccctc ctgacoggt 5100
aaatgaaggg gaaaaaagg acacctaact tcttcaaat ggtctttagt aaaggaacc 5160
tgtctaaagg ctaagaactg ccgaagtat aaattatcag ccggaacgag caaacagacg 5220
gagttttaa agataaatc gcaatcttct ccgctctag tcccaggcca gcaatctctg 5280
gggaagcaag tggaaacct atagcctct ccagcttagg aaggaggggt gggctgtcc 5340
ctggattct tctcgtctc tgcagagaca ataccaggg gagagcaatg gatctactgc 5400
cccgaatgct tctaaaocgg ggggcaaaa caaaaaaaa caaacgtcc ggttaccatc 5460
ggggaacagg adcgagccc agggccacca gccagatca aacagcccgc gtctcggcgc 5520
tggctctcag ccgacacac tcccgccaa gcgcagccgc cccccccc cggggcccg 5580
ctgactaccc cacacagcct ccgcccgc ctccggggc tcaagtgctc ggcagcgcct 5640
ccggcccaag tggggccgg ccgcccagc tcccgcctg ctggcgggag aaacatctc 5700
ctctggcggg ggtaggggcg gagctgscgt ccgcccac ccgaaagga agtctaagc 5760
ccggaagtga tggcattct gggtaacyag ctatttactt cctgoggtg cacaggctgt 5820
ggtctctat ctccctgtt ttcttccat ccgcaagat ggcctggag accgtccga 5880
aggaactgag gcaatgagg gactgtttgc tgttttctg ggtcaagtg taagtgggg 5940
acctggttgt agggcccat ggggacaaag gtccgggaaa gaggggcga tgggctcgt 6000
aggatcggg acaggtctt gactgaggg caggggggt cttaactgc ttgcaatcc 6060
cagctcttag acgttcggg aacttaagtt ggaagcgaag gacactggga gtcagaggcg 6120
ggtggggatc cgtctctgag tgaatagtc gaaaggaag ctgaacctga gtgactcac 6180
agaaactggt ctlttctgc ttcaggatc gtgaggagc tgaagaatcg aggagtgccc 6240
tcaotgggtc agcatgaga tcaagcaga tccagcttga gtagttca tcaagttctc 6300
tagctgctg gctgctctc ctctctcgg cccagagtc agaacttga ggtgaacggg 6360
atgaatocaa gctggtctgc agggcagtc tcaactgaga gtctcttcc aactctcac 6420
acottttcca gctggtctgc ggaatgaggg aatcctgttg ggggcaggag gctggcagg 6480
ggaatagat agctcttgc ccttcttcc cagacaagat aaggggagaa tctactaga 6540
gcaatctca gccacctgc ctctctgca ttttggagg tctgctctg agccagctga 6600
gaagatacca tggctgctg ggggtggg aggatttga acactctgt 6650

```

<310> 14
 <211> 1206
 <212> DNA

WO 01/92528

PC/T/US01/17583

<213> Homo sapiens

```

<400> 14
gcagtgccag gaacctccc ggaggcggg cagagcaga gcttctggc cctgtccga 60
gcccaggcct gcacccctaa ggcaggcact gctccgtgat ccaggaacca cctctctcta 120
cagctgggag tgagcagtca gggagggaga cagccttgcc cgggtgctacc cagcaagcta 180
gtcacccagt gggcagaggg aggagcggcc ctccaccgat gtcaggcagc ctgggtcccc 240
agtcaccgctc tgcctgtccc tcgcaataac gcctcagtga cgaaccttg tgagccatct 300
ctctgtctca ggcacgggtc tacatgcaa cgaaacctgc tcccatgaa cctggcccag 360
ccagtgaaag aagggttgag cctggggagt gccactttac agacaggggc accaaggggc 420
agggtggcag gaggcccccc ggaacttccc catgaagtac cagtcaccag atcccaacc 480
agcaggcaac acgtggccc gcagcctccc tgcacagcag cctggcttc cggcctcggg 540
acttgatctg ctccctcttc cggacactgg ggtctctgcc aagtctctgg ctgggcaaga 600
actgtctaac attctaagaa atccctccca gggttttctc aggagccagg gtggggcagg 660
aagtcaccag gggctgagg gacogtggc gacgggtggc cccagagcag cactctctctg 720
gggccaccgg tgttgggcca gaggcaggac tctgaggcct agtctagggc ctctctccag 780
tggccggcac ctacttgttg ggtcgggggt tccccacga ggttgggct cccacctgac 840
acactcacag acctgtgccc ttggagagcc agtgttcccg gggccacata gctatgcgc 900
ccaggggctg ggcctgtccc agctctggtc ccccgcccc aggtcctgga cgtggctccg 960
cgcagcagca ggcgctccc ggagacacag atgtgaatgg ctgacagtac gtcgactca 1020
gatgagtctg gcgccgatcg acctgctgcc gagtctctcc ggcacggcac aggcaggag 1080
tgaaaattat ctacccttt ttatttctta ataactgaat gaaaataaac attggtggtt 1140
tgacaataaa ctacatattt tcaaacccag ccagtcaccg ggatgcagtt tccaggtgag 1200
ttatgc 1206

```

<210> 15

<211> 1443

<212> DNA

<213> Homo sapiens

```

<400> 15
gccttttacc actgacccaa agcgaaaagc accaggttta actctgttcc cctgtgcta 60
ggterccaca ggttttgta tccctgtacc ttccttactc ctagcagcta ctctgatcga 120
ttttctctca coctcagagc agacttgtgg cctgttttgg ggaagcactg gaattttgaa 180
ccccagcctc atttgggtca atgttttggc aagagtgtcc gcttoatgat gctgggtgatg 240
gcattgcacct cgtcaactgt gcaaggctag gcttctgacg gtggcctcta ttaaccaaac 300
actgagggga agcccccttg tgcctttgga gagatgccag gtgcttagtt tacatttttg 360
cctgcttggg gagctaacag cttgaagtaa accaatccat cggggactcc tgaggttttc 420
accagccagg accacccaat cgtgctgtaa gaacttctga ctccctggac attgccatgg 480
actcaacctg tcaactcagg acctgttttt gaactaacaa agctagactt ctgattctct 540
cttgctctca cctacctgta cattcogaac acatggtaga gactctacaa aatgcttaat 600
atgtgatcta tggacgggtc cccctgaaat tataaatgct gccakcttca tccctctggt 660
tttcccagc taitaacctc atccatttgt ctgtgttata caacgtaact atccaggcct 720
cogtctcgga actgtgtgaa gctcttttgt ctagggacca aaggcaggaa ttatttagtg 780
atcagacaaat aagaataaac tgaaagagat gatttgcctt tgatggatgt aaaaaacta 840
aaaatttatt ttcaatttat ggttaabgta cttagccatt ttctctdaaa caccactgga 900
gaatttataa aacatgaagc atatacaaaa tgcactctagg ggttaatgag gcttctcttt 960

```

WO 01/92528

PC/US01/17583

```

catcaattc tgccttttag gatttgeccc aatattgtac ttggaggtaa atattaaaac 1020
tccattgagg actggfataa agttgtaaaq tgaacaaaac ccagtagaaa gctattgata 1080
aagaatctat ttataaaaat aagttttata caataaaaac tactctgtaa ttccttttc 1140
aaagtatatt tctaaaaatag ctatataatg cttctgtacc aaatttteta aataagggat 1200
tatgttoaca cttctocagt cctccttoca gctcttoaac ctactatccc aataagggtc 1260
ataagactga ggcagtttca acagctcctg ctaagggtaa agaaagatac ggggaagcat 1320
catgaaaaga taggaacttc cctatcctaa gtatgtttat acatacotta tatatggagg 1380
ctantaagtt tcccttaagt atatcaataa ttaagatctg tactaagtga ccactataag 1440
tgt 1493

```

```

<210> 16
<211> 1957
<212> DNA
<213> Homo sapiens

```

```

<400> 16
ggggccgccc agctccgccc ggggcaaac ccgccgccc gccatgccc gaggtaagt 60
atctgocctgt cgcgccaggg cgtgggaagg cgcgccccc ctctctcttc caggatgaa 120
ggaaaagaa aatgccgcaa tgaanaaccc tctgccctcc caanaacaca tcttgccgt 180
gtgcccgtg ctcctgcage tctgtgcaec caeggagtg gctctcact gttgagtgga 240
gtggggcag aagcctgccc tgcctccagg agagcccgg ctgcctggg gctcctgga 300
gtctcgggg agcgggacgg ggtggtgga cgaactggcg gtgacccga gaacgccca 360
ctccacccc ccactttcca aagacggct tcccgggga gcccccaca taacgccag 420
cgaactgccc ctccgtgaaa gtcttagcca gaaactttcc cgttttgtc gccagtgca 480
cagagagtcg tgtgctctg ggcggcgct gctggtcaa gagccagcct ggcgttctt 540
gccctaccg tccccttctc aggcagttc tcaattgcc ctgagagccc attcccggct 600
cggtagaaaa ggcactatat ccatccctgc atcgtctcca agactcattc cctctaaac 660
ttcaagttcc atggaatg ggagaccacc tgatcctgca gactgggccc gtagggatg 720
cgtcaattat ttccgaaccg tgggatgtga ggagcaagc agtgcttttc agaacagga 780
aattgatgga aaatccctgc tatfgatgac aagaaatgat gtgttgacag gacttcagtt 840
aaaattgggg ctgctctga aaactcaaga atatctgta aaactctgc agacaaagca 900
tttaagaac aactctctat agtaoagta aattggggtc ttcgaactca aaaaaaac 960
ataatgacat aattcagtt catgtaatga aactttgaa acagaataca tacatgtgt 1020
tatgtaaaga atttcaatca aatgaaagct tatcctattg gatagactag gcaattcac 1080
agctcacctg aaatccgcca ggagagada ggacaagatg cgcacagggt ggttttctc 1140
atggatlttg tcaaatgat gatctttgac acgattagac actcctcccc ccaaaagcct 1200
tgaatcata aggattttcc tcaatccttt atagctttcc caaaatcttt taaaaaaga 1260
atttaattaa atgacagctc tttggttaca gactaggat gagtaaaaac aagaaaattt 1320
ggggaggggg agaaagaga aaggatgct tgtctccott gaattcctct gttccttaga 1380
gottgtgtta ttggacgga abtgcaaca ccocttttta tagagggttc tccactgac 1440
cttattagg ttttattggg atatgctgca gtgtttgaaa tgaactgca tcatggccc 1500
ttcaggagca gaatcatagc tctgaaaaga gaagctccgt tgtgtactga ggatattcat 1560
coatlattca ctagctttca aatgggggtt aatgatattt tctgactaga tttctttta 1620
aattggttct tgtttctga agaaagaatt tttttaact tcatggtrtt atttataata 1680
atltgtttct gaagaaattt gccgagagtt acaggtcaaa aagccttgtt actagtacag 1740
aatattttta tatatattcc ttcctgatgg tgaatttttt ttaattgtc ctatgctttg 1800
tcggctctcc gggtaagta cttgttttta agagcttggg aaaagtgggc ttgctacac 1860

```

WO 01/92528

PC/US01/17583

```

tcgtttcaaa gagacatttg ttcaatctct gtgtgtcaac gecttttga attggtgctt 1920
tgcggtagca ataaagcatt gcttcaagttt ataaaaa 1957

```

```

<210> 17
<211> 2074
<212> DNA
<213> Homo sapiens

```

```

<400> 17
tgcagctatt ttaggttctc taacttcac gtagtttata gggtlaagtaa agggaagggg 60
aaagtgalg gtgtggtgtg ctcccataag aactgatttt ttctactga agcatgtata 120
aagttthata atgacttttt atatttgttt aataaaaaatt ttacaggaac taaatttgat 180
tatcaataty agtttttct ttaatttcag atttcaacta ttgcagaaag tgaagattca 240
caggagtcag tggatagtg aactgatcc caaaagcgaa gggaaattct tcaaggagg 300
cttctctaca gggagaagtc tgaagaggag actcagcac ctgcatrac cactgtaacg 360
gtgccaacct caatttaca aactagcagt ggacagtata ttgccattac ccaggaggga 420
gcaatacagc tggctaaca tggtaaccgat ggggtacagg gcttcaaac attaacatg 480
accaatgcaag cagccactca gccgggtact accattctac agtatgcaca gaccactgat 540
ggacagcaga tottagtgcc cagcaacca gttgtgttcc aaggtactca aaaattgtaa 600
agcaggatgt cagtgaattt gaattctgaa cgtcagtttg aagatggtaa catgtttagt 660
atataaatct ttccactca aaccatacat tttaattgat attaataatt aatataaact 720
aattttataa agacctcaa attttttaa graaccatag gttcttatt aggagagcat 780
attattaccg tgtttttaga agcagtttga caaatagtga ttgtgtttgt ttttacaact 840
ggtgaatcag ttagaanaaa anaacttcag ttattttagc cattatcatt tacattaaaa 900
caatagtgtt ttcaataaat ataattggca tcaagtata cacltttca tacttttagt 960
ttgttttaa ttcaaaattt ataatagttg accataatgc ttatcttct ttttcaattt 1020
gctcatttta tgaanaaaca tggctgttt ttatgtctgt ggcagagtc taactgatat 1080
ttgtttaata tgaattttac caatacaaaa ggtatagtag tactgaggaa ctatactcta 1140
tctaggtaag atcaccatca gtctgtgccc catctgtacc ttttagaccg taegcgtgcc 1200
tctggagagc tacaactacta taccagtatt cgtactagc taacctacta gctactattg 1260
gccctggag ttgttatggc atctccctct agctacttcc tacacagcct gctcgaagat 1320
agcagctacg tataagttag gaggctcgtc taatgaagat acaggaagc tagttctaga 1380
gtgtcgtaga aagaagttaa gaatatgtga aatgtttaga aaacagagtg gctagtogt 1440
tgaaatcaa taactagaca ttgattgagg agcttaagc acttaaggac ctttactgce 1500
acaaatcaga ttaatttggg atttaatttt tcacctgta aggtggaaaa tggactggct 1560
tggcccaaac ctgaaagaca aaataaacat ttatthttct aaacatttct tttttctat 1620
gctcaaaact gctgaaagc aactacagaa ttccattcat ttgtgctttt gcattaact 1680
gtgaatgttc cagcaactgc ctccactctt cccctcaaga catttcaac gccaggaaac 1740
atgaagagac ttctgctttt caaccccacc ctctcaaga agtaataatt tgttacttg 1800
taaattgatg ggagacatga ggaanaagaa atctthttaa aattgatctc aaggtttgtg 1860
ctgagctcct tgatggcctt agggacagaa ttaccccagc ctcttgagct gaagtaatgt 1920
gtggccgca tgcataaagt aagtaaggtg caatgaagaa gtgttgattg ccaaatgac 1980
atgtgtgca atctcattg tgaattatgt aaagtgtta agagacatac cctcaaaaa 2040
agaactttag catggctatg aggaactaga aatg 2074

```

<210> 18

WO 01/92528

PC/T/US01/17583

<220>
<221> unsure
<222> (74)

<220>
<221> unsure
<222> (92)

<220>
<221> unsure
<222> (126)

<220>
<221> unsure
<222> (135)

<220>
<221> unsure
<222> (113)

<400> 20
ctcaaccaac atctgacatc ttccocngg agcaactcc tgcctcccg gaaagaggcc 60
gaaggattta ccontggacc cataagtctg anctcctgc tgaagtccc tencattgc 120
tccttnaagc caaanotaca ctttctggt tcdtctccc tctgagaaag gggatagaaa 180
gtctcttccot ctatgtcctc ccctcgagat ctgttctggg gatggagett ccaacttcc 240
cttgccagcg gaaagaatgc tgcctaccct tctgtcttgc agagtgggat tgtgggaggg 300
attggoagcc ttcttctcca ccacctgtcc agcttcttcc tggctcaggg tgggacccc 360
aggaatatta tgttgcc 377

<210> 21
<211> 709
<212> DNA
<213> Homo sapiens

<400> 21
tctgaatgtt ttggtgaata aatctgttct tcagcaacc taactgcttc tccaaactgc 60
ctaagagat ccagtactga tgacgctggt ctccactct tactccctgg aaactaacc 120
cgttctcttc gtttccctca ccacgcacca ggagctcaga gatcaaaagc gcttccate 180
ttgttctccc agccccagga cactgactct gtacaggatg gggccgtcct cttgcccctc 240
ttctctcct aatccccct ctccagctga tcaaccggg gactactcag tgttccctag 300
actccgttat gataaagaag atcaaggatg ttctcaacag tctagagtac agtccctctc 360
ctataagca gaaactctcg tgtgctagt tcaaaagcca aggcagacc tctcactgc 420
ctgctgggg atggctgta ctgctgtgc ttgtggctat gctgtggtt cgtgggatgt 480
tcagctggaa accacctgcc actgccagt cagtggtggt gactggacca ctgcccctg 540
ctgcaactg aactgacagg gaggaaggct gagaactcag ttctgtgacc atgacagtaa 600
tgaaaccagg gtcccaacca agaaactctaa ctcaaacgtc ccacttcatt tgttccattc 660

WO 01/92528

PC/T/US01/17583

ctgattcttg ggtataaag acaaaccttg tacctotcaa aaaaaaaaa

709

<210> 22
<211> 3195
<212> DNA
<213> Homo sapiens

<400> 22
gocaggaata actagagaggg aacaatgggg ttattdagag gttttgtrtt ccterragtt 60
ctgtgctgce tgcaccagtc aaatacticc ttcattagc tgaataata tggctttgaa 120
gataattgca ttgttataga tcttagtgty ccagagatg aaaaaataa tgaacaaata 180
gaggetatgg tgaactacagc ttctacgtac ctgtttgaa ccacagaaaa aagattttt 240
ttcaaaaatg tactatatt aattcctgag aatgggaagg aaatcctca gtaaaaaagg 300
ccaaaacatg aaaaacctaa acatgctgat gttatagttg caccacctac actccagggt 360
agagatgnac catacaccaa gcagttcaca gaatgtggag agaaaggcga atacattcac 420
ttcacocctg accttctact tggaaaaaa ncaaatgaa batggacac caggcaaac 480
gtttgtccat gagtgggctc acctccggtg gggagtgctt gatgagaca atgaagatca 540
gccttctcac cgtgctaagt caaaaaaat cgaagcaaca aggtgttccg caggtatctc 600
tggtagaagt agagtttata agtgcgaag aggcagctgt cttagtagag catgcgaat 660
tgattctaca acaaaactgt atggaaaaga ttgtcaatc ttctcgata aagtaacaac 720
agaaaaagca tccataatgt ttaatgnaag tattgattct gtgttgaat ttgtkaaga 780
aaaaacocat aatcaagaag ctccaagcct acaaacata aagtcaatt ttgaagtae 840
atggggaggty attagaatt ctgaggattt taaaaacac ataccatgg tgaacacac 900
tctccacct gtcttctcat tgcagaagt cagtcacaaga atgtgtgct tagttcttga 960
taagtctgga agcatggggg gbaaggaacc cctaaatcga atgaatcaag cagcaaaaa 1020
tttctcctg cagactgttg aaatggatc ctgggtgggg atggttcaat ttgatagtae 1080
tgccactatt gtaataaagc taatccaat aaaaagcagt gatgaagaa acacactcat 1140
ggcaggatta cctacatata ctctgggagg aactccatc tgccttggaa ttaaatatgc 1200
atctcagtyg atggagagc tacattccca actcgtatga tccgaagtae tgcctgtgac 1260
tgatggggag gataacactg caagttcttg tattgatgae gtgaaacaaa gtagggocat 1320
tgttcatttt atgtctttgg gaagagctgc tgatgaagca gtaatagaga tgaacaagat 1380
aacaggagga agtcattttt atgtttcaga tgaagctcag aacatggcc bcattgatgc 1440
ttttggggct cttacatcag gaaatactga tctctcccag aagtccttcc agctcgaag 1500
taagggattâ acactgaata gtaatgcctg gatgaacgac actgtoataa ttgatagtae 1560
agtgggaaag gacacgttct tctctatcac atggacagt ctgctccca gtatttctct 1620
ctgggtccc agtggaaaca taatggaaaa ttccacagt gatgcaact ccaaaatggc 1680
ctatctcagt atccaggaa ctgcaaaagt gggcacttgg gaatacaatc ttcaagcaca 1740
agogaaccca gaaacattaa ctattacagt aacttctcga gcagcaaat ctctctgtgc 1800
tccaatcaca gtgaatgcta aatgaataa ggaagttaac agtttccca gcccaatgat 1860
tgtttacgca gaaattctac aaggatagt acctgttctt ggagccaatg tgactgcttt 1920
cattgataca cagaatggac atacagaagt ttggaaact ttgataatg gtgcaaggcg 1980
tgattcttcc asgaatgatg gactctactc cagggtatttt acagcatata cagaaatgg 2040
cagatatact taaaagttcg ggcctatgga gggcacaaca ctgcaaggct aaaaatcag 2100
ctccactga ategagccgc gacatacca ggcctggtag tgaacgggga aattgaagca 2160
aacccgccaa gacctgaat tgatgaggat actcagacca ccttggagga ttteagccga 2220
acagcatcdg gaggtgcatt tgtggtatca caagtcccaa gcttccctt gctgaccaa 2280
taccaccaa gtcnaatcac agacctgat gccacagttc atgaggataa gattattctt 2340

WO 01/92528

PC/T/US01/17583

```

acatggacag caccaggaga taatbttgat gttgaaaaag ttcaacgtta tateataaga 2400
ataagtgcac gtattcttga tctaagagac agttttgatg atgctcttca agtaataact 2460
actgatctgt caccnaagga ggccaactcc aaggaaagct ttgcattta accaanaaat 2520
atccagaaag aaaatgcaac ccacataatt atgcccatta aaagtataga taaaagcaat 2580
ttgacatcaa agtatccaaa cattycaaaa gtsactttgt ttatccctca egcaaatcct 2640
gatgacattg atcctacacc tactctact cctactccta ctctcgataa agtcataat 2700
ctcggagtta atattttctac gctgggtattg ctctgtgattg ggtctgtctgt aattgttaac 2760
tttattttaa gtaccaccat ttgaacctta acgaagaaaa aatcttcaag tagacctaga 2820
agagagtttt aaaaaaacaa aacaatgtaa gtaaggata tttctgaaac ttaaaattoa 2880
tcccattgtg gatcataaac tcataaaaat aattttaaga tgcggaaaa ggatactttg 2940
atcaataaaa acaactcatg gatatgtaaa aactgtcaag attaaattt aatagtttca 3000
ttcatttght attttatttg taagaatag tgatgaacaa agatcccttt tcatactgat 3060
acctggttgt atatttattg atgcaacagt tttctgaaat gatatttcaa attgcatcaa 3120
gaaattaaaa tcactctatct gagtatgcaa aatacaagta aaggagagca aataaacac 3180
atttgaaaa aaatg 3195

```

```

<210> 23
<211> 22
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence: Synthetic

```

```

<400> 23
tggaaataga ttcagggggtc at 22

```

```

<210> 24
<211> 21
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence: Synthetic

```

```

<400> 24
egggtgtacc tcactgactt c 21

```

```

<210> 25
<211> 25
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence: Synthetic

```

WO 01/92528

PCT/US01/17583

<400> 25
tgtcttcga gagaaccagg ctccg

25

WO 01/92528

PCT/US01/17583

<211> 933

<212> DNA

<213> Homo sapiens

<400> 18

```

atggcggagg ctgtactgag ggtgcccagg oggcagctga gccagcggcg cgagctcttcg 60
agctcccacg ctccctgggc agatgttoga gctgtgagc tgcacctca cgracctgct 120
gggtgacaga gactcccagg acgctgtctc gatcgaccca gtccctggaa cagcgcctcg 180
ggatgtccag ctgatcaagg agctggggct gccgctgctc tatgtctgta ataccactg 240
ccacgggaaa ccaccattaca ggcctggggc tgcctccgtc cctccctccct ggcctccagt 300
ctgcactctc ccgccttagt ggggcccagg ctgacttaca cattgaggat gggagactcc 360
atccgcttcg ggcgcttagt tacagcccaa ctccctggctg ctttcaaggg ctggtgtgga 420
gtatctgtgg cttttccagg cacatggtgc aagctctcgg tggatctaac actctgggtt 480
ctggaggggc atggccctct tctcacagct ccactagggg cagtgcocca gtgggaactc 540
tctgcgttgg agaccagggc cagccctggc cacaccocag gctgtgtcac cttcgtctgt 600
aatgaccaca gcatggcctt cactggagat gccctgttga tccgtgggtg tggcgggacn 660
gactccagc aaggctgtgc caagaccttg taccactcgg tccatgaaaa gatcttaca 720
cttccaggag actgtctgat ctaccctgct cacgattacc atgggttcc acgttccacc 780
gtggaggggg agaggactct gaacctcgg ctccacctca gctgtgagga gtttgtcaaa 840
atcatggcca acctgaactt gcttaacct cagcagatag actttgctgt tccagcdaac 900
atgcgctgtg ggggtcagac acccaactgc tga 933

```

<210> 19

<211> 525

<212> DNA

<213> Homo sapiens

<400> 19

```

gccatggggt cccctccagc ctgtccatac agagtgtgca ttccttggca ggggctcctg 60
ctccacagct cgccttctaac cttctggaac ctgccaaca gtgccagac caatattgat 120
gggtgctcgt tcaatgtcgc agaagggaag gaggctcttc tagtagtoca taatgagtc 180
cagaatcttt atggctacaa ctggtacaaa gggcaagggt tgcctgccc ctatcgaatt 240
ataggatag taaaaaatat aagtcaagaa aatgcccag gcccgcaca caacggtcga 300
gagacaatat accccaatgg aacctgctg atccagaacg taccaccaca tgacgcagga 360
atctataccc tacacgttat aaagaaaaat cttgtgaaat saggagtaac cagacaattc 420
tacgtattct atgagtcagt acaagcaagt taccctgacc tctcagctgg gaccgctgtc 480
agcatcatga ttgggtact ggtgggatg gctctgatat agcag 525

```

<210> 20

<211> 377

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> (28)

【国際公開パンフレット(コレクトバージョン)】

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
6 December 2001 (06.12.2001)

PCT

(10) International Publication Number
WO 01/092528 A3

- (61) International Patent Classification: C12N 15/12, C07K 14/17, C12Q 1/68, G01N 33/574, C07K 16/18, A51K 51/00, 39/395, 38/17, A61P 35/00
- (21) International Application Number: PCT/US01/17583
- (22) International Filing Date: 29 May 2001 (29.05.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 60/207,383 26 May 2000 (26.05.2000) US
- (71) Applicant (for all designated States except US): BI-ADENUS, INC. [US/US]; 365 Oyster Point Boulevard, South San Francisco, CA 94080 (US)
- (72) Inventors; and
(75) Inventors/Applicants (for US only): MACINA, Roberto, A. [AR/US]; 4,118 Crescentia Avenue, San Jose, CA 95136 (US); CHEN, Sei-Yu [US]; 160 Mira Street, Redwood City, CA 94062 (US); PLUTA, Jason [US/US]; Apartment 15, 1260 Dale Avenue, Mountain View, CA 94040 (US); SUN, Yongming [US/US]; Apartment 260, 869 S. Winchester Boulevard, San Jose, CA 95128 (US); RECIPON, Hervé [FR/FR]; 85 Tartana Avenue, San Francisco, CA 94115 (US)
- (74) Agents: LICATA, Jane, Massey et al.; Licata & Tyrrell P.C., 65 E. Main Street, Madison, NJ 08053 (US)
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GF, GR, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW
- (84) Designated States (regional): ARIPO patent (GI, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FR, GB, GR, IE, IT, LI, LU, MC, NL, PT, SI, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG)
- Published:
with international search report
- (88) Date of publication of the international search report: 3 January 2003
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 01/092528 A3

(54) Title: METHOD OF DIAGNOSING, MONITORING, STAGING, IMAGING AND TREATING COLON CANCER

(57) Abstract: The invention relates to C/Sg polypeptides, polymucleotides encoding the polypeptides, methods for producing the polypeptides, in particular by expressing the polymucleotides, and agonists and antagonists of the polypeptides. The invention further relates to methods for utilizing such polymucleotides, polypeptides, agonists and antagonists for applications, which relate, in part, to research, diagnostic and clinical arts.

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		Date of Application No.
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/12 C07K14/47 C12Q1/68 G01N33/571 C07K16/18 A61K51/00 A61K39/395 A61K38/17 A61P35/00		PC1/US 01/17583
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Main search system used (classification system followed by classification symbols) IPC 7 GC1N Documentation searched other than main search system (date of search) (date of search) (date of search) EPO-Internal, SEQUENCE SEARCH, WPI Data, PAJ, BIOSIS, EMBL		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim(s)
X	WO 96 39419 A (HUMAN GENOME SCIENCES INC; ROSEN CRAIG A (US); YU GUO LIANG (US)) 12 December 1996 (1996-12-12) see in particular Fig. 8 relevant to invention 1 the whole document	1-14
X	WO 00 07632 A (MACINA ROBERTO A; SUN YONGMING (CN); RECIFON HERVE (FR); DIADEXUS) 17 February 2000 (2000-02-17) relevant to invention 1 the whole document	1-14
-/-		
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex		
* Special categories of cited documents: ** Earlier document published after the international filing date to priority date and not in conflict with the applicant but cited to disclose the principle or theory underlying the invention *** 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 **** 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 *5* Document member of the same patent family		
Date of the actual completion of the international search 26 June 2002		Date of mailing of the international search report 16.07.02
Name and mailing address of the ISA European Patent Office, P.O. Box 5016, Patentstr. 2 8001 Zurich, Switzerland Tel. (+41-76) 340 2040, Tx. 21 651 0000 Fax: (+41-76) 340 2070		Authorized officer: Morawetz, R

Form PCT/ISA/210 (second sheet) (July 1999)

INTERNATIONAL SEARCH REPORT

Int: Application No
 PCT/US 01/17583

C (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Description of document, with indication of the relevant passages	Relevant to claim No.
X	<p>DATABASE EMBL 'Online! 18 April 1997 (1997-04-18) ADAMS, M.D., ET AL.: "EST187535 Colon carcinoma (HCC) cell line II Homo sapiens cDNA 5' end similar to similar to trefol1 factor, intestinal." Database accession no. AA315762 XP002194184 relevant to invention 1 the whole document</p>	1-14
X	<p>WO 96 06861 A (THIM LARS ;NOVONORDISK AS (DK); NIELSEN PER FRANKLIN (DK); WOELDIK) 7 March 1996 (1996-03-07) relevant to invention 1 the whole document</p>	1
X	<p>WO 98 37093 A (CORIXA CORP) 27 August 1998 (1998-08-27) relevant to invention 1; see SEQ ID NO: 179</p>	1
X	<p>DATABASE EMBL 'Online! 18 November 1999 (1999-11-18) ABOLA, A.P. ET AL.: "Homo sapiens chromosome 21 clone RP11-113F1, complete sequence." Database accession no. AC015555 XP002194185 relevant to invention 1 the whole document</p>	1
X	<p>DATABASE SWALL 'Online! 1 October 1994 (1994-10-01) HAUSER, F., ET AL.: "INTESTINAL TREFOIL FACTOR PRECURSOR (HPI.B)." Database accession no. Q07654 XP002194186 relevant to invention 1 the whole document</p>	1
X	<p>DATABASE EMBL 'Online! 24 November 1998 (1998-11-24) "gk58e09.x1 NCL_CGAP_C08 Homo sapiens cDNA clone IMAGE:1873192 3', mRNA sequence" Database accession no. AI281212 XP002202713 relevant to invention 5 the whole document</p>	1-14

Form PCT/ISA/210 (continued from previous sheet) (July 1999)

INTERNATIONAL SEARCH REPORT

Inventor: Applicant: Name No:
PCI/US 01/17583

C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication where appropriate, of the relevant paragraph	Relevant claim No.
X	DATABASE EMBL 'Online! 31 September 1997 (1997-09-11) "nm73f01.s1 NCI_CGAP_Co9 Homo sapiens cDNA clone IMAGE:1073881 3', mRNA sequence" Database accession no. AA576913 XP002202714 relevant to invention 5 the whole document	1-14
X	DATABASE EMBL 'Online! 3 August 1999 (1999-08-03) "Homo sapiens genomic DNA, chromosome 6p21.3, HLA class I region, clone:679622." Database accession no. AB023054 XP002202715 relevant to invention 5 the whole document	1
X	DATABASE EMBL 'Online! 18 April 1997 (1997-04-18) "EST32671 Embryo, 12 week I Homo sapiens cDNA 5' end." Database accession no. AA328866 XP002202716 relevant to invention 5 the whole document	1
X	DATABASE EMBL 'Online! 5 February 1999 (1999-02-05) "tg24b01.x1 NCI_CGAP_CLL1 Homo sapiens cDNA clone IMAGE:2109673 3' similar to TR:000635 000635 similar to the 52 kD RO/SSA Protein and RFP: mRNA sequence" Database accession no. AI394677 XP002202717 relevant to invention 5 the whole document	1

From PCI/US 01/17583 (Continuation of Invention No. 01/17583)

INTERNATIONAL SEARCH REPORT

Int. No. of Application No.
PCT/US 01/17583

C (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of its relevant passages	Relevant Class. No.
X	GOEI VITA L ET AL: "Isolation of Novel Non-ILA Gene Fragments from the Hemochromatosis Region (6p21.3) by cDNA Hybridization Selection." AMERICAN JOURNAL OF HUMAN GENETICS, vo. 54, no. 2, 1994, pages 244-251, XP001085146 ISSN: 0002-9297 relevant to invention 5 the whole document	1
X	-& DATABASE EMBL 'Online! 6 January 1999 (1999-01-06) GOEI, V.L. ET AL.: "Human putative zinc finger protein (ZNF87) mRNA, complete cds" Database accession no. U34249 XP002202718 the whole document	1
X	-& DATABASE SMALL 'Online! 1 May 1999 (1999-05-01) GOEI, V.L., ET AL.: "ZINC FINGER PROTEIN." Database accession no. 0956C4 XP002202719 the whole document	1
X	DATABASE EMBL 'Online! 8 January 1997 (1997-01-08) "Homo sapiens RFB30 gene for RING finger protein" Database accession no. Y07829 XP002202720 relevant to invention 5 the whole document	1
X	DATABASE EMBL 'Online! 7 January 1999 (1999-01-07) "tb95a12.x1 NCI_C6AP_Co16 Homo sapiens cDNA clone IMAGE:2062078 3', mRNA sequence" Database accession no. AI343354 XP002202721 relevant to invention 6 the whole document	1-14
X	DATABASE EMBL 'Online! 30 June 1999 (1999-06-30) "wi60a08.x1 NCI_C6AP_Co16 Homo sapiens cDNA clone IMAGE:2594614 3', mRNA sequence" Database accession no. AI761312 XP002202722 relevant to invention 6 the whole document	1-14

Form PCT/US 99/10 (Instruction for use of form) (July 1999)

INTERNATIONAL SEARCH REPORT

C: (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		Inter- national Application No.
Category	Class of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, X	<p>DATABASE WPI Section Ch, Week 200124 Derwent Publications Ltd., London, GB; Class B04, AN 2001-235357 XP002202724 & WO 01 22920 A (HUMAN GENOME SCI INC), 5 April 2001 (2001-04-05) relevant to invention 6 the whole document -& DATABASE: 6SP (Online! 3 September 2001 (2001-09-03) "Human colon cancer antigen protein SEQ ID NO: 5315" Database accession no. AAG74551 XP002202723 The whole document: -----</p>	1-14

Form PCT/ISA/210 (continuation of form PCT/ISA/210) July 1999

INTERNATIONAL SEARCH REPORT		international application No. PCT/JP 01/17583
Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)		
This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:		
1.	<input checked="" type="checkbox"/> Claims Nos. 1-14 because they relate to subject matter not required to be searched by this Authority, namely: see FURTHER INFORMATION sheet PCT/ISA/210	
2.	<input type="checkbox"/> Claims Nos. 1-14 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:	
3.	<input type="checkbox"/> Claims Nos. 1-14 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:		
see additional sheet		
1.	<input type="checkbox"/> As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.	
2.	<input type="checkbox"/> As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.	
3.	<input checked="" type="checkbox"/> 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 claim Nos. 1-14 (all partially) insofar related to SEQ ID NO: 1, 5 or 6	
4.	<input type="checkbox"/> 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. 1-14	
Remark on Protest		
<input type="checkbox"/> The additional search fees were accompanied by the applicant's protest.		
<input checked="" type="checkbox"/> No protest accompanied the payment of additional search fees.		

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US 01 17583

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this International application, as follows:

Invention 1: Claims 1-14 (all partially)

A colon specific gene (CSG) comprising a polynucleotide of SEQ ID NO: 1, and subject-matter related thereto.

Inventions 2-22: Claims 1-14 (all partially)

A colon specific gene (CSG) comprising a polynucleotide of SEQ ID NO: 2 - 22, and subject-matter related thereto.

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US 01 47583

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although claims 2-6, 9 and 10 are directed to a diagnostic method practised on the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Although claims 7 and 11-13 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No.

PC1/US 01/17583

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9639419	A	12-12-1996	CA 2221798 A1 12-12-1996
			US 5733748 A 31-03-1998
			WO 9639419 A1 12-12-1996
			AU 2820595 A 24-12-1996
			EP 0847398 A1 17-06-1998
			JP 11506342 T 08-06-1999
			US 6337195 B1 08-01-2002
WO 0007632	A	17-02-2000	EP 1107798 A1 20-06-2001
			WO 0007632 A1 17-02-2000
WO 9606861	A	07-03-1996	AU 694796 B2 30-07-1998
			AU 3341595 A 22-03-1998
			BR 9506772 A 30-12-1997
			CA 2196876 A1 07-03-1996
			CN 1156459 A, B 06-08-1997
			CZ 9702548 A3 13-08-1997
			CZ 289864 B6 17-04-2002
			WO 9606861 A1 07-03-1996
			EP 0777627 A1 11-06-1997
			FI 970783 A 25-02-1997
			HU 77917 A2 28-10-1998
			JP 1050472C T 12-05-1998
			KR 255911 B1 01-05-2000
			NO 970844 A 28-04-1997
			PL 31878E A1 07-07-1997
			RU 2162857 C2 10-02-2001
WO 9837093	A	27-08-1998	US 6261562 B1 17-07-2001
			AU 731840 B2 05-04-2001
			AU 6181898 A 09-09-1998
			BR 9808881 A 11-09-2001
			CN 1252837 T 10-05-2000
			CZ 9903016 A3 13-03-2002
			EP 1005546 A2 07-06-2000
			FI 0007095 A2 28-10-2000
			NO 994059 A 22-10-1999
			NZ 337446 A 23-02-2001
			PL 335348 A1 25-04-2000
			TR 9902053 T2 21-04-2000
			US 6262245 B1 17-07-2001
			WO 9837093 A2 27-08-1998
			US 6395278 B1 28-05-2002
			US 6329505 B1 11-12-2001
			US 2002022248 A1 21-02-2002
US 2002051977 A1 02-05-2002			
ZA 9801585 A 04-09-1998			
WO 0122920	A	05-04-2001	AU 7721500 A 30-04-2001
			WO 0122920 A2 05-04-2001

フロントページの続き

(51) Int.Cl. ⁷	F I	テーマコード(参考)
A 6 1 P 35/00	C 0 7 K 16/18	4 H 0 4 5
C 0 7 K 16/18	C 1 2 Q 1/02	
C 1 2 Q 1/02	C 1 2 Q 1/68	A
C 1 2 Q 1/68	G 0 1 N 33/53	D
G 0 1 N 33/53	G 0 1 N 33/53	M
G 0 1 N 33/58	G 0 1 N 33/58	A
	A 6 1 K 49/02	Z

(81) 指定国 AP(GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, 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, 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, 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, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW

(72) 発明者 プルータ, ジェイソン
 アメリカ合衆国 カリフォルニア州 9 4 0 4 0、マウンテン ビュー -、デール アベニュー 1
 2 4 0、アパートメント 1 5

(72) 発明者 スン, ヨンミン
 アメリカ合衆国 カリフォルニア州 9 5 1 2 8、サン ホセ、エス・ウィンチェスター ブルバ
 ード 8 6 9、アパートメント 2 6 0

(72) 発明者 レシボン, エルヴェ
 アメリカ合衆国 カリフォルニア州 9 4 1 1 5、サン フランシスコ、フォーチュナ アベニュー
 - 8 5

F ターム(参考) 2G045 AA26 FB08
 4B024 AA01 AA11 BA41 CA04 CA09 CA12 HA12
 4B063 QA01 QA18 QA19 QQ02 QQ03 QQ08 QQ20 QQ79 QR31 QR56
 QR62 QS25 QS34 QX01
 4C084 AA13 AA17 DC50 NA14 ZB261
 4C085 AA03 AA13 CC02 CC04 DD33 DD61
 4H045 AA10 AA11 AA30 BA10 CA40 DA75 DA86 EA51

专利名称(译)	结直肠癌的诊断，监测，分期，成像和治疗方法		
公开(公告)号	JP2004510408A	公开(公告)日	2004-04-08
申请号	JP2002500720	申请日	2001-05-29
申请(专利权)人(译)	Diadokusasu公司		
[标]发明人	マシーナロベルトエー チェンセイユ プルータジェイソン スンヨンミン レシボンエルヴェ		
发明人	マシーナ,ロベルト エー. チェン,セイ-ユ プルータ,ジェイソン スン,ヨンミン レシボン,エルヴェ		
IPC分类号	G01N33/53 A61K39/00 A61K45/00 A61K48/00 A61K49/00 A61K51/00 A61P35/00 C07K16/18 C12N15/09 C12N15/12 C12Q1/02 C12Q1/68 C12Q1/6886 G01N33/574 G01N33/58		
CPC分类号	A61K49/0008 A61P35/00 A61P35/04 C12Q1/6886 C12Q2600/136 G01N33/57419		
FI分类号	C12N15/00.ZNA.A A61K39/00.H A61K45/00 A61K48/00 A61P35/00 C07K16/18 C12Q1/02 C12Q1/68.A G01N33/53.D G01N33/53.M G01N33/58.A A61K49/02.Z		
F-TERM分类号	2G045/AA26 2G045/FB08 4B024/AA01 4B024/AA11 4B024/BA41 4B024/CA04 4B024/CA09 4B024/CA12 4B024/HA12 4B063/QA01 4B063/QA18 4B063/QA19 4B063/QQ02 4B063/QQ03 4B063/QQ08 4B063/QQ20 4B063/QQ79 4B063/QR31 4B063/QR56 4B063/QR62 4B063/QS25 4B063/QS34 4B063/QX01 4C084/AA13 4C084/AA17 4C084/DC50 4C084/NA14 4C084/ZB261 4C085/AA03 4C085/AA13 4C085/CC02 4C085/CC04 4C085/DD33 4C085/DD61 4H045/AA10 4H045/AA11 4H045/AA30 4H045/BA10 4H045/CA40 4H045/DA75 4H045/DA86 4H045/EA51		
优先权	60/207383 2000-05-26 US		
外部链接	Espacenet		

摘要(译)

本发明涉及CSG多肽，编码多肽的多核苷酸，产生多肽的方法，特别是通过表达多核苷酸，以及多肽的激动剂和拮抗剂。本发明还涉及利用这些多核苷酸，多肽，激动剂和拮抗剂的应用的方法，其部分涉及研究，诊断和临床领域。

組織	正常
副腎	0
膀胱	0
脳	0
子宮頸部	0
大腸	0.7
子宮内膜	0.4
食道	0
心臓	0
腎臓	3.7
肝臓	1
肺	0
乳腺	0.2
筋	0
睪丸	0
脾臓	0
前立腺	0
直腸	23
小腸	1.5
脾臓	0
胃	0.8
精巣	0.1
胸腺	0.4
気管	0
子宮	0