

(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2004-500819

(P2004-500819A)

(43) 公表日 平成16年1月15日(2004.1.15)

(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	2 G O 4 5
A 6 1 K	38/00	A 6 1 K	45/00		4 B O 2 4
A 6 1 K	45/00	A 6 1 P	3/00		4 B O 6 3
A 6 1 P	3/00	A 6 1 P	5/00		4 B O 6 4
A 6 1 P	5/00	A 6 1 P	15/00		4 B O 6 5
		審査請求	未請求	予備審査請求	有 (全 186 頁) 最終頁に続く

(21) 出願番号	特願2001-554421 (P2001-554421)	(71) 出願人	301005050
(86) (22) 出願日	平成13年1月26日 (2001. 1. 26)		インサイト・ゲノミックス・インコーポレ
(85) 翻訳文提出日	平成14年7月29日 (2002. 7. 29)		イテッド
(86) 国際出願番号	PCT/US2001/002741		アメリカ合衆国カリフォルニア州9430
(87) 国際公開番号	W02001/055392		4・パロアルト・ポータードライブ 31
(87) 国際公開日	平成13年8月2日 (2001. 8. 2)		60
(31) 優先権主張番号	60/178, 578	(74) 代理人	100089266
(32) 優先日	平成12年1月28日 (2000. 1. 28)		弁理士 大島 陽一
(33) 優先権主張国	米国 (US)	(72) 発明者	バーフォード、ニール
(31) 優先権主張番号	60/185, 079		アメリカ合衆国コネチカット州06422
(32) 優先日	平成12年2月25日 (2000. 2. 25)		・ダラム・ワイルドウッドサークル 10
(33) 優先権主張国	米国 (US)		5
(31) 優先権主張番号	60/189, 999		
(32) 優先日	平成12年3月17日 (2000. 3. 17)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 核ホルモン受容体

(57) 【要約】

本発明は、ヒト核ホルモン受容体 (NHREC) と、NHREC を同定及びコードするポリヌクレオチドとを提供する。本発明はまた、発現ベクター、宿主細胞、抗体、アゴニスト、およびアンタゴニストを提供する。更に、本発明は、NHREC の異常発現に関連する疾患の診断・治療・予防方法を提供する。

【特許請求の範囲】

【請求項 1】

単離されたポリペプチドであって、

(a) SEQ ID NO: 1 乃至 SEQ ID NO: 3 (SEQ ID NO: 1-3) からなる群から選択されたアミノ酸配列と、

(b) SEQ ID NO: 1-3 からなる群から選択されたアミノ酸配列と少なくとも 90% の配列同一性を有する天然のアミノ酸配列と、

(c) SEQ ID NO: 1-3 からなる群から選択されたアミノ酸配列の生物学的に活性な断片と、

(d) SEQ ID NO: 1-3 からなる群から選択されたアミノ酸配列の免疫原性断片とで構成される群から選択されたアミノ酸配列を含むことを特徴とする単離されたポリペプチド。 10

【請求項 2】

SEQ ID NO: 1-3 からなる群から選択された請求項 1 の単離されたポリペプチド。

【請求項 3】

請求項 1 のポリペプチドをコードする単離されたポリヌクレオチド。

【請求項 4】

請求項 2 のポリペプチドをコードする単離されたポリヌクレオチド。

【請求項 5】

SEQ ID NO: 4-6 からなる群から選択された請求項 4 の単離されたポリヌクレオチド。 20

【請求項 6】

請求項 3 のポリヌクレオチドに機能的に結合されたプロモーター配列を含む組換えポリヌクレオチド。

【請求項 7】

請求項 6 の組換えポリヌクレオチドで形質転換された細胞。

【請求項 8】

請求項 6 の組換えポリヌクレオチドを含む遺伝子組換え生物。

【請求項 9】

請求項 1 のポリペプチドを生産する方法であって、

(a) 前記ポリペプチドの発現に好適な条件下で、請求項 1 のポリペプチドをコードするポリヌクレオチドに機能的に結合されたプロモーター配列を含む組換えポリヌクレオチドで形質転換された細胞を培養するステップと、

(b) そのように発現したポリペプチドを回収するステップとを含むことを特徴とする請求項 1 のポリペプチドの生産方法。 30

【請求項 10】

請求項 1 のポリペプチドに特異的に結合する単離された抗体。

【請求項 11】

単離されたポリヌクレオチドであって、 40

(a) SEQ ID NO: 4-6 からなる群から選択されたポリヌクレオチド配列と、

(b) SEQ ID NO: 4-6 からなる群から選択されたポリヌクレオチド配列と少なくとも 90% の配列同一性を有する天然のポリヌクレオチド配列と、

(c) 前記 (a) に相補的なポリヌクレオチド配列と、

(d) 前記 (b) に相補的なポリヌクレオチド配列と、

(e) 前記 (a) 乃至 (d) の RNA 等価物とで構成される群から選択されたポリヌクレオチド配列を含む単離されたポリヌクレオチド。

【請求項 12】

請求項 11 のポリヌクレオチドの少なくとも 60 個の連続するヌクレオチドを含む単離されたポリヌクレオチド。 50

【請求項 1 3】

サンプルにおいて、請求項 1 1 に記載のポリヌクレオチド配列を有する標的ポリヌクレオチドを検出する方法であって、

(a) 前記サンプルをプローブでハイブリダイズするステップであって、前記プローブが、前記サンプル内の前記標的ポリヌクレオチドと相補的な配列を含む少なくとも 20 個の連続するヌクレオチドを含み、前記プローブと前記標的ポリヌクレオチドまたはその断片との間でハイブリダイゼーション複合体が形成される条件下で、前記プローブが前記標的ポリヌクレオチドと特異的にハイブリダイズする、該ステップと、

(b) 前記ハイブリダイゼーション複合体が存在するか否かを検出し、存在する場合には随意選択でその収量を測定するステップとを含むことを特徴とする標的ポリヌクレオチドの検出方法。 10

【請求項 1 4】

前記プローブが少なくとも 60 個の連続するヌクレオチドを含むことを特徴とする請求項 1 3 に記載の方法。

【請求項 1 5】

サンプルにおいて、請求項 1 1 のポリヌクレオチド配列を有する標的ポリヌクレオチドを検出する方法であって、

(a) ポリメラーゼ連鎖反応増幅を用いて、前記標的ポリヌクレオチドまたはその断片を増幅するステップと、

(b) 増幅された前記標的ポリヌクレオチドまたはその断片が存在するか否かを検出し、存在する場合には随意選択でその収量を測定するステップとを含むことを特徴とする標的ポリヌクレオチドの検出方法。 20

【請求項 1 6】

有効量の請求項 1 のポリペプチド及び医薬的に容認できる賦形剤を含む組成物。

【請求項 1 7】

前記ポリペプチドが、SEQ ID NO: 1-3 からなる群から選択されたアミノ酸配列を含むことを特徴とする請求項 1 6 の組成物。

【請求項 1 8】

機能的な NHRHC (新規の核ホルモン受容体) の発現の低下に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項 1 6 の組成物を投与することを含むことを特徴とする治療方法。 30

【請求項 1 9】

請求項 1 のポリペプチドのアゴニストとして効果的な化合物をスクリーニングする方法であって、

(a) 請求項 1 のポリペプチドを含むサンプルを化合物に曝露するステップと、

(b) 前記サンプルにおいてアゴニスト活性を検出するステップとを含むことを特徴とするスクリーニング方法。

【請求項 2 0】

請求項 1 9 のスクリーニング方法によって同定されたアゴニスト化合物及び医薬的に容認できる賦形剤を含む組成物。 40

【請求項 2 1】

機能的な NHRHC の発現の低下に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項 2 0 の組成物を投与することを含むことを特徴とする治療方法。

【請求項 2 2】

請求項 1 のポリペプチドのアンタゴニストとして効果的な化合物をスクリーニングする方法であって、

(a) 請求項 1 のポリペプチドを含むサンプルを化合物に曝露するステップと、

(b) 前記サンプルにおいてアンタゴニスト活性を検出するステップとを含むことを特徴とするスクリーニング方法。 50

【請求項 2 3】

請求項 2 2 のスクリーニング方法によって同定されたアンタゴニスト化合物及び医薬的に容認できる賦形剤を含む組成物。

【請求項 2 4】

機能的な N H R E C の過剰な発現に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項 2 3 の組成物を投与することを含むことを特徴とする治療方法。

【請求項 2 5】

請求項 1 のポリペプチドに特異的に結合する化合物をスクリーニングする方法であって、
(a) 請求項 1 のポリペプチドを好適な条件下で少なくとも 1 つの試験化合物と結合させるステップと、
(b) 請求項 1 のポリペプチドと前記試験化合物との結合を検出して、請求項 1 のポリペプチドと特異的に結合する化合物を同定するステップとを含むことを特徴とするスクリーニング方法。 10

【請求項 2 6】

請求項 1 のポリペプチドの活性を変化させる化合物をスクリーニングする方法であって、
(a) 請求項 1 のポリペプチドを、その活性が許容される条件下で少なくとも 1 つの試験化合物と結合させるステップと、
(b) 前記試験化合物の存在下での請求項 1 のポリペプチドの活性を評価するステップと、
(c) 前記試験化合物の存在下での請求項 1 のポリペプチドの活性と、前記試験化合物の不在下での請求項 1 のポリペプチドの活性とを比較するステップとを含み、
前記試験化合物の存在下での請求項 1 のポリペプチドの活性の変化が、請求項 1 のポリペプチドの活性を変化させる化合物の存在を示唆すること特徴とするスクリーニング方法。 20

【請求項 2 7】

請求項 5 の配列を含む標的ポリヌクレオチドの発現を変化させるのに効果的な化合物をスクリーニングする方法であって、
(a) 前記標的ポリヌクレオチドの発現に好適な条件下で、前記標的ポリヌクレオチドを含むサンプルを化合物に曝露するステップと、
(b) 前記標的ポリヌクレオチドの発現の変化を検出するステップと、
(c) 様々な量の前記化合物の存在下での前記標的ポリヌクレオチドの発現と、前記化合物の不在下での前記標的ポリヌクレオチドの発現とを比較するステップとを含むことを特徴とするスクリーニング方法。 30

【請求項 2 8】

試験化合物の毒性を評価する方法であって、
(a) 核酸を含む生体サンプルを前記試験化合物で処理するステップと、
(b) 処理した前記生体サンプルの核酸と、請求項 1 1 のポリヌクレオチドの少なくとも 2 0 の連続するヌクレオチドを含むプローブをハイブリダイズさせるステップであって、このハイブリダイゼーションゼーションが、前記プローブと前記生体サンプルの標的ポリヌクレオチドとの間で特異的なハイブリダイゼーション複合体が形成される条件下で行われ、前記標的ポリヌクレオチドが、請求項 1 1 のポリヌクレオチドのポリヌクレオチド配列またはその断片を含むポリヌクレオチドである、前記ステップと、
(c) ハイブリダイゼーション複合体の収量を定量するステップと、
(d) 前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量を、未処理の生体サンプルにおけるハイブリダイゼーション複合体の収量と比較するステップとを含み、
前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量の差が試験化合物の毒性を示唆すること特徴とする試験化合物の毒性評価方法。 40

【発明の詳細な説明】

【0 0 0 1】

(技術分野)

本発明は、核ホルモン受容体に関する。本発明はまた、これらの配列を利用した代謝障害、内分泌障害、生殖障害、発生または発達障害、免疫疾患、細胞増殖異常の診断・治療・予防に関する。本発明はさらに、核ホルモン受容体の核酸配列及びアミノ酸配列の発現における外来性化合物の影響についての評価にも関する。

【0002】

(発明の背景)

核受容体又は細胞内受容体とも呼ばれる核ホルモン受容体は、タンパク質スーパーファミリーを構成する。このスーパーファミリーのメンバーは、受容体及び転写調節因子の両方である。核ホルモン受容体は、その受容体機能及び転写調節機能の両方によって、発生、恒常性、細胞増殖、及び細胞分化を含む広範な生体プロセスに影響を与える (Mangelsdorf, D. J. 他 (1995) Cell 83 : 835-840 ; Wen, D. X. 及び D. P. McDonnell (1995) Curr. Opin. Biotechnol. 6 : 582-589 ; Perlmann, T. and R. M. Evans (1997) Cell 90 : 391-397 ; Tenbaum, S. 及び A. Baniahmad (1997) Int. J. Biochem. Cell Biol. 29 : 1325-1341 ; Moras, D. 及び H. Gronemeyer (1998) Curr. Opin. Cell Biol. 10 : 384-391 ; Willy, P. J. 及び D. J. Mangelsdorf (1998) in : Hormones and Signaling (ed : B. W. O'Malley) vol. 1, Academic Press, San Diego CA, pp. 307-358 ; Weatherman, R. V. 他 (1999) Annu. Rev. Biochem. 68 : 559-581.)。 10 20

【0003】

受容体としての核ホルモン受容体

一般に、受容体とは、他の分子を特異的に認識するタンパク質を指す。核ホルモン受容体は受容体として、それらの同種リガンドを特異的に認識してそれらに結合する。核ホルモン受容体は細胞内に局在するが、多くの受容体は、細胞外リガンドと結合する細胞外の細胞表面タンパク質である。このような細胞外受容体は、成長、分化、エンドサイトーシス、及び免疫反応に影響を与える細胞応答を引き起こす。他の受容体は、小胞体からのタンパク質の選択的な輸送を促進し、細胞の特定の領域に酵素を局在化させる。核ホルモン受容体による転写調節、細胞外受容体による細胞内へのシグナル伝達、及び他の受容体によるタンパク質の輸送及び局在化は全て、受容体と様々な細胞成分との間の特異的な相互作用に依存する。多くの場合、受容体が結合する同種リガンドが同定されていない。このような受容体はオーファン受容体と呼ばれる。リガンドと結合しないで転写調節を行う核ホルモン受容体もオーファン受容体である。 30

【0004】

転写調節因子としての核ホルモン受容体

多細胞生物は、構造及び機能が著しく異なる多様な細胞型から構成されている。細胞型は遺伝子発現パターンの特徴によって決定される。異なった細胞型は発生の間、重複するが固有の遺伝子群として発現する。遺伝子発現の空間的及び一時的な調節は細胞増殖、細胞分化、アポトーシス、及び生物の発達に必要なその他のプロセスにとって極めて重要である。核ホルモン受容体は転写調節因子として、これらの基本的な生体プロセスの調節において重要な役割を果たしている。他の転写調節因子は、細胞-細胞伝達の仲介及び様々な細胞型の作用の調節を行なう細胞外シグナルに応答して遺伝子発現に影響を与える。 40

【0005】

一般に、核ホルモン受容体等の転写調節因子は、配列特異的な要領でプロモーター、エンハンサー、及び遺伝子上流の調節領域と結合して、遺伝子転写の開始、活性化、抑制、または停止を行う。しかしながら、或る種の転写調節因子は、遺伝子コード領域内或いはその下流の調節エレメントと結合する。転写調節タンパク質が、単体で或いは他のアクセ 50

サリーファクターと複合体を形成してDNAの特定の領域に結合し得る (Lewin, B. (1990) in: Genes IV, Oxford University Press, New York NY, and Cell Press, Cambridge MA, pp. 554-570.)。

【0006】

核ホルモン受容体機能のメカニズム

核ホルモン受容体は、リガンドと結合していない状態では、hsp90等の熱ショックタンパク質及びhsp56等のイムノフィリンを含むシャペロンの多タンパク質複合体と結合して存在する。これらのシャペロンは、リガンド非結合受容体を遊離リガンドの結合を素直に受け入れる不活性な状態に維持し、リガンド非結合受容体が核に転移するのを防止する。同種リガンドによって活性化されると、核ホルモン受容体はホモ二量体又はヘテロ二量体を形成して核に転移したり、特定のDNA配列に結合したり、転写調節機能を発揮したりし得る。核ホルモン受容体の調節の役割を効果的に果たすために、活性化した核ホルモン受容体はヒストンデアセチラーゼ含有コリプレッサー複合体から解離し、ヒストンアセチルトランスフェラーゼ含有コアクティベーター複合体と結合する (Xu, L. 他 (1999) *Curr. Opin. Genet. Dev.* 9 : 140-147)。活性化した受容体がコアクティベータタンパク質と結合してクロマチンが再構築されることにより、転写的に開いた活性な状態になり、活性化した核受容体の転写調節エレメントに到達できるようになる (Lemon, B. D. 及び L. P. Freedman (1999) *Curr. Opin. Genet. Dev.* 9 : 499-504)。

【0007】

核ホルモン受容体の構造

核ホルモン受容体は、ホルモンシグナルを転写反応に変換してシグナルトランスデューサとして作用する。一般に、核ホルモン受容体は、可変アミノ末端ドメイン、高度に保存されたDNA結合ドメイン、及び保存されたC末端リガンド結合ドメインからなる。

【0008】

ステロイド結合核ホルモン受容体では、アミノ末端ドメインがAF-1と呼ばれるトランス活性化エレメントを匿っている。或る種の核ホルモン受容体はまた、AF-2と呼ばれるリガンド結合ドメインにトランス活性化エレメントを含む。核ホルモン受容体のDNA結合ドメイン及びリガンド結合ドメインは二量体化エレメントを含み得る。また、DNA結合ドメインは核局在化シグナルを含み得る (Weatherman, R. V. 他 (1999) *Annu. Rev. Biochem.* 68 : 559-581)。

【0009】

核ホルモン受容体のDNA結合ドメインは、特定のDNA配列の認識を仲介する2つのZnフィンガーモチーフからなる。Znフィンガーモチーフは、 Zn^{+2} を配位結合する周期的に離間したヒスチジン残基及びシステイン残基を含む。この配列パターンの例には、C2H2型、C4型、及びC3HC4型 (「RING」フィンガー) のZnフィンガー、及びPHDドメインが含まれる (Lewin, 前出 ; Aasland, R. 他 (1995) *Trends Biochem. Sci.* 20 : 56-59)。Znフィンガーモチーフは、構造及び隣接状態が亜鉛イオンによって維持されている1つの逆平行βシート及び1つのαヘリックスを含む。DNAとの接触は、αヘリックスの前のアルギニン残基によって、並びにαヘリックスの第2の残基、第3の残基、及び第6の残基によってなされる。タンパク質の各ZnフィンガーのαヘリックスがDNA二重螺旋の主溝と接触できるように、Znフィンガーモチーフはそのタンパク質内にタンデム型にリピートされ得る。このタンパク質とDNAとの間の接触の繰り返しによって、強力且つ特異的なDNA-タンパク質相互作用が起こる。この相互作用の強度及び特異性は、タンパク質内のZnフィンガーモチーフの数によって調節され得る。Znフィンガーは、初めはDNAと直接相互作用する領域としてDNA結合タンパク質において見出されたが、その後DNAと結合しないタンパク質からも見出された (例えばLodish, H. 他 (1995) 40 50

Molecular Cell Biology, Scientific American Books, New York NY, pp. 447-451. 参照)。

【0010】

核ホルモン受容体のリガンド結合ドメインは、リガンド、コアクティベータタンパク質、及びコリプレッサータンパク質との結合に必要である。このドメインは三層の α ヘリックスからなり、中間の層は多数の疎水性測鎖を含む2つのヘリックスからなる (Moras, D. 及び H. Gronemeyer (1998) *Curr. Opin. Cell Biol.* 10 : 384-391)。従って、これらの2つの中間層の α ヘリックスが、リガンド結合部位である疎水性ポケットを形成している。この疎水性リガンド結合部位のリガンド結合は、受容体タンパク質の内部に完全に埋もれ、溶液に露出していない。このことから、リガンドの結合にはリガンド結合ドメインにおける大きな構造の変化を伴うことが示唆される。リガンド結合ドメインの1つの α ヘリックスは、核受容体の二量体に多くのインターサブユニット接触を提供する。この α ヘリックスはリガンド結合ポケットにおいて結合している時はそのリガンドに接触することから、リガンド結合が受容体の二量体の形成に影響を与え得ることが示唆される (Weatherman, R. V. 他 (1999) *Annu. Rev. Biochem.* 68 : 559-581)。

【0011】

核ホルモン受容体のクラス

核ホルモン受容体は、ステロイド受容体、RXR-ヘテロ二量体受容体、及びオーファン核ホルモン受容体の3つのクラスに大きく分類することができる。ステロイド受容体はステロイドホルモンに結合し、このクラスにはアンドロゲン受容体、ミネラルコルチコイド受容体、エストロゲン受容体、糖質コルチコイド受容体、及びプロゲステロン受容体が含まれる。RXR-ヘテロ二量体受容体は非ステロイドリガンドと結合し、このクラスには、甲状腺ホルモン受容体、レチノイン酸受容体、ビタミンD受容体、エクジソン受容体、及びペルオキシソーム増殖因子活性化受容体が含まれる。オーファン核ホルモン受容体には、ステロイド産生因子1、神経成長因子誘導性受容体、及びX染色体オーファン受容体DAX-1が含まれる。

【0012】

ステロイドホルモン受容体は、特定のステロイドホルモンと結合すると活性化する。リガンドの結合によって起こる構造の変化により、熱ショックタンパク質から受容体が解離し、ホルモン応答エレメント (HRE) と呼ばれる特定のパ lindロームDNA配列を認識する受容体ホモ二量体が形成される。ステロイドホルモン受容体ホモ二量体はHREに結合すると標的遺伝子の転写を調節し得る。

【0013】

例えばプロゲステロン受容体 (PR) は、女性の生殖周期を変化させる重要な要素であるコレステロール由来のプロゲステロン即ち4-プレグネン-3, 20-ジオンによって活性化されるステロイドホルモン受容体である。これらの変動は、月経及び妊娠を含む解剖学的及び形態学的変化と相関する。プロゲステロンの作用はPRによって仲介される。細胞質では、PRは、幾つかの他のタンパク質及びPRヘテロ複合体として知られる因子と結合している。このヘテロ複合体には、hsp70、hsp90、hsp27、p59 (hsp56)、p48、及びp23等のイムノフィリン及び熱ショックタンパク質が含まれる (Johnson, J. L. 他 (1994) *Mol. Cell. Biol.* 14 : 1956-1963)。プロゲステロンの結合で活性化されたPRが核に転移し、正準DNA転写エレメントに結合して分化及び細胞周期に関係するプロゲステロン調節性遺伝子を調節する (Moutsatsou, P. 及び C. E. Sekeris (1997) *Ann. N. Y. Acad. Sci.* 816 : 99-115)。妊娠を終了させるのに用いられ得るPRアンタゴニストRU486は、ステロイドホルモン受容体を標的とした市販の治療薬の一例である。

【0014】

RXRヘテロ二量体核受容体は、そのグループのメンバーがレチノイドX受容体(RXR)とヘテロ二量体を形成するとその標的DNA配列と結合することから、ステロイドホルモン受容体と区別される(Mangelsdorf, D. J. 及び R. M. Evans (1995) Cell 83 : 841-850)。RXRの3種類のアイソフォームが同定された(Minucci, S. 及び K. Ozato (1996) Curr. Opin. Genet. Dev. 6 : 567-574)。レチノイン酸受容体(RAR)はRXRヘテロ二量体核受容体の例である。レチノイン酸(RA)は、主に魚の肝油、肝臓、卵黄、バター、及びクリームに存在する脂溶性ビタミンであるビタミンA(レチノール)の生物学的に活性な代謝産物である。9-シス-RAはRAR/RXRに結合するが、オールトランス-RAはRXRのみと結合する。RAR/RXRヘテロ二量体は、レチノイン酸応答エレメント(RARE)として知られる特定のDNA配列と高親和的に結合するため、RA依存性転写の制御因子として作用する。 10

【0015】

ペルオキシソーム増殖因子活性化受容体(PPAR)は、脂肪酸及びエイコサノイドによって誘導される治療にとって重要なRXRヘテロ二量体核受容体である。それぞれが特定の発現パターンを有するPPARの3種類の既知のイソ型が存在し、これらはグルコース及び脂質の全身の恒常性に関与する遺伝子の調節に関与する(Kliwer, S. A. 及び T. M. Willson (1998) Curr. Opin. Genet. Dev. 8 : 576-581 ; Michalik, L. 及び W. Wahli (1999) Curr. Opin. Biotechnol. 10 : 564-570)。従って、PPARは糖尿病、異脂肪血症、及び肥満症等の疾患の治療標的である(Smith, S. A. (1996) Pharmacol. Rev. Commun. 8 : 57-64 ; Willson, T. M. 及び W. Wahli (1997) Curr. Opin. Chem. Biol. 1 : 235-241 ; Barroso, I. 他 (1999) Nature 402 : 880-883)。 20

【0016】

オーファン核受容体は、活性化するリガンドを有していない核受容体、或いはリガンド結合してもその影響がなく転写調節活性を発揮し得る核受容体である。例えば、線虫において、X染色体によってコードされる核受容体相同体SEX-1は、性決定遺伝子xol-1の転写を調節する(Carmi, I. 他 (1998) Nature 396 : 168-173)。SEX-1はリガンド結合に依存するのではなく、X染色体カウント機構による性分化調節において、用量応答的に転写調節因子として影響を与える。 30

【0017】

或る種の核ホルモン受容体は、通常は核ホルモン受容体ファミリーと結合する通常のDNA結合ドメインを含まない。DAX-1は、通常のDNA結合ドメインを含まない核ホルモン受容体の一つであって、DAX-1における突然変異は先天性X染色体副腎形成不全を引き起こすことが分かっている(Zanaria, E. F. 他 (1994) Nature 372 : 635-641)。DAX-1はオーファン核受容体であって、ステロイド産生因子1(SF-1)と直接相互作用する(Ito, M. 他 (1997) Mol. Cell. Biol. 17 : 1476-1483)。DAX-1はまた、性特異的遺伝子発現におけるSF-1の作用を調節する能力がある(Nachtigal, M. W. 他 (1998) Cell 445-454)。SF-1は、副腎及び生殖腺の幾つかのステロイド産生酵素遺伝子に対する転写調節因子として作用し(Lala, D. S. 他 (1992) Mol. Endocrinol. 6 : 1249-1258 ; Lynch, J. P. 他 (1993) Mol. Endocrinol. 7 : 776-786 ; Clemens, J. W. 他 (1994) Endocrinology 134 : 1499-1508)、下垂体性腺刺激細胞で発現される幾つかの遺伝子も調節し得る(Barnhart, K. M. 及び P. L. Mellon (1994) Mol. Endocrinol. 8 : 878-885 ; Ingraham, H. A. 他 (1994) Genes Dev. 8 : 2302-2312 40

; Halvorson, L. M. 他 (1996) J. Biol. Chem. 271 : 6645-6650 ; Keri, R. A. 及び J. H. Nilsson (1996) J. Biol. Chem. 271 : 10782-10785)。

【0018】

SF-1はまた小ヘテロ二量体 (SHP ; short heterodimer partner) パートナーの強力なトランス活性化因子として作用する (Lee, Y. K. 他 (1999) J. Biol. Chem. 274 : 20869-20873)。SHPもまた、通常のDNA結合ドメインを含まない核ホルモン受容体である (Seol, W. 他 (1996) Science 272 : 1336-1339 ; Lee, H.-K. 他 (1998) J. Biol. Chem. 273 : 14398-14402)。SHPは核ホルモン受容体ファミリーの多くのメンバーと相互作用する。その中には、レチノイド受容体、エストロゲン受容体、甲状腺ホルモン受容体、及びオーファン受容体CARが含まれる。SHPは、エストロゲン受容体への結合においてコアクティベーターと競合することで、エストロゲン受容体仲介性転写活性化のインヒビターとして作用する (Johansson, L. 他 (1999) J. Biol. Chem. 274 : 345-353)。SHPはまた、オーファン受容体肝細胞核因子4及びレチノイドX受容体によってトランス活性化を阻害する (Lee, Y. K. 他 (2000) Mol. Cell. Biol. 20 : 187-195)。

【0019】

不完全な転写調節による結果

ヒトにおける腫瘍疾患の多くは不適切な遺伝子発現によって起こり得る。悪性の細胞成長は、腫瘍促進遺伝子の過剰な発現又は腫瘍抑制遺伝子の不十分な発現によって起こり得る (Cleary, M. L. (1992) Cancer Surv. 15 : 89-104)。染色体転座によって、或る遺伝子のコード配列を別の関連のない遺伝子の調節領域と結合させるキメラ座 (chimeric loci) が生成され得る。このような配列は不適切な遺伝子転写を引き起こし、悪性腫瘍を引き起こす可能性もある。

【0020】

更に、免疫系は、細胞防御機構の漸進的な選択、増幅、及び動員を調整する現象のカスケードを活性化することにより感染や外傷に応答する。遺伝子の活性化及び抑制の複雑かつバランスの取れたプログラムが、このプロセスに関与している。しかしながら、遺伝子発現の不適切或いは不十分な調節によって起こる免疫系の過度の活性化は、組織や器官に相当な損傷を与え得る。この損傷は、関節炎、アレルギー、心臓発症、発作、及び感染に関連する免疫応答についての文献に詳しく記載されている (例えば Isselbacher 他 (1996) Harrison's Principles of Internal Medicine, 13/e, McGraw Hill, Inc. and Tetra Data Systems Software 参照)。

【0021】

更に、多細胞生物の成長は発生の適当な段階における細胞分化の誘導及び調整に基づいている。このプロセスの中心は、体中の細胞及び組織に固有性を与える差次的な遺伝子発現である。発生中に遺伝子発現の調節に失敗すると発生障害が起こり得る。

【0022】

新規の核ホルモン受容体、およびそれらをコードするポリヌクレオチドの発見により、新規の組成物を提供することで当分野の要望に応えることができる。この新規の組成物は、代謝障害、内分泌障害、生殖障害、発生または発達障害、免疫疾患、細胞増殖異常の診断・治療・予防において有用であり、また、核ホルモン受容体の核酸配列及びアミノ酸配列の発現における外来性化合物の影響についての評価にも有用である。

【0023】

(発明の要約)

本発明は、総称して「NHREC」、個別にはそれぞれ「NHREC-1」、「NHREC-2」、「NHREC-3」と呼ぶ核ホルモン受容体である精製されたポリペプ

チドを提供する。本発明の一実施態様では、(a) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列と、(b) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列と、(c) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) SEQ ID NO: 1-3 とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含む単離されたポリペプチドを提供する。別法では、SEQ ID NO: 1-3 のアミノ酸配列を含む単離されたポリペプチドを提供する。

【0024】

更に本発明は、(a) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列と、(b) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列と、(c) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含むポリペプチドをコードする単離されたポリヌクレオチドを提供する。別法では、このポリヌクレオチドは、SEQ ID NO: 1-3 からなる一群から選択されたポリペプチドをコードする。別法では、このポリヌクレオチドは、SEQ ID NO: 4-6 からなる一群から選択される。

10

【0025】

更に、本発明は、(a) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列と、(b) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列と、(c) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含むポリペプチドをコードするポリヌクレオチドと機能的に結合されたプロモーター配列を含む組換えポリヌクレオチドを提供する。別法では、本発明は、この組換えポリヌクレオチドで形質転換された細胞を提供する。更なる別法では、本発明は、この組換えポリヌクレオチドを含む遺伝子組換え生物を提供する。

20

【0026】

また、本発明は、(a) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列と、(b) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列と、(c) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) SEQ ID NO: 1-3 とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含むポリペプチドの生産方法を提供する。この方法は、(a) このポリペプチドの発現に好適な条件下で、このポリペプチドをコードするポリヌクレオチドと機能的に結合されたプロモーター配列を含む組換えポリヌクレオチドで形質転換された細胞を培養するステップと、(b) このように発現したポリペプチドを回収するステップとを含む。

30

【0027】

更に、本発明は、(a) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列と、(b) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列と、(c) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) SEQ ID NO: 1-3 とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含むポリペプチドに特異的に結合する単離された抗体を提供する。

40

【0028】

更に、本発明は、(a) SEQ ID NO: 4-6 からなる一群から選択されたポリヌクレオチド配列と、(b) SEQ ID NO: 4-6 からなる一群から選択されたポリヌク

50

レオチド配列と90%以上の配列同一性を有する天然のポリヌクレオチド配列と、(c)前記(a)に相補的なポリヌクレオチド配列と、(d)前記(b)に相補的なポリヌクレオチド配列と、(e)前記(a)乃至(d)のRNA等化物とで構成される一群から選択されたポリヌクレオチド配列を含む単離されたポリヌクレオチドを提供する。別法では、このポリヌクレオチドは、少なくとも60個の連続するヌクレオチドを含む。

【0029】

更に本発明は、(a)SEQ ID NO: 4-6からなる一群から選択されたポリヌクレオチド配列と、(b)SEQ ID NO: 4-6からなる一群から選択されたポリヌクレオチド配列と90%以上の配列同一性を有する天然のポリヌクレオチド配列と、(c)前記(a)に相補的なポリヌクレオチド配列と、(d)前記(b)に相補的なポリヌクレオチド配列と、(e)前記(a)乃至(d)のRNA等化物とで構成される一群から選択されたポリヌクレオチド配列を含むポリヌクレオチド配列を有する標的ポリヌクレオチドをサンプルにおいて検出する方法を提供する。この方法は、(a)前記サンプル内の標的ポリヌクレオチドと相補的な配列を構成する少なくとも20個の連続するヌクレオチドを含むプローブと前記サンプルをハイブリダイズさせるステップであって、前記プローブと前記標的ポリヌクレオチドまたはその断片とでハイブリダイゼーション複合体が形成される条件下で、前記プローブが前記標的ポリヌクレオチドに特異的にハイブリダイズする、該ステップと、(b)前記ハイブリダイゼーション複合体の存在するか否かを検出し、存在する場合には随意選択でその収量を測定するステップとを含む。別法では、前記プローブは、少なくとも60個の連続するヌクレオチドを含む。

10

20

【0030】

更に本発明は、(a)SEQ ID NO: 4-6からなる一群から選択されたポリヌクレオチド配列と、(b)SEQ ID NO: 4-6からなる一群から選択されたポリヌクレオチド配列と90%以上の配列同一性を有する天然のポリヌクレオチド配列と、(c)前記(a)に相補的なポリヌクレオチド配列と、(d)前記(b)に相補的なポリヌクレオチド配列と、(e)前記(a)乃至(d)のRNA等化物とで構成される一群から選択されたポリヌクレオチド配列を含むポリヌクレオチド配列を有する標的ポリヌクレオチドをサンプルにおいて検出する方法を提供する。この方法は、(a)ポリメラーゼ連鎖反応増幅を用いて、前記標的ポリヌクレオチドまたはその断片を増幅するステップと、(b)増幅された前記標的ポリヌクレオチドまたはその断片が存在するか否かを検出し、存在する場合には随意選択でその収量を測定するステップとを含む。

30

【0031】

更に本発明は、(a)SEQ ID NO: 1-3からなる一群から選択されたアミノ酸配列と、(b)SEQ ID NO: 1-3からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列と、(c)SEQ ID NO: 1-3からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d)SEQ ID NO: 1-3とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含む効果的な量のポリペプチド及び好適な医薬用賦形剤を含む組成物を提供する。一実施例では、SEQ ID NO: 1-3からなる一群から選択されたアミノ酸配列を含む組成物を提供する。更に、本発明は、患者にこの組成物を投与することを含む、機能的NHRECの発現の低下に関連した疾患やその症状の治療方法を提供する。

40

【0032】

更に本発明は、(a)SEQ ID NO: 1-3からなる一群から選択されたアミノ酸配列と、(b)SEQ ID NO: 1-3からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列と、(c)SEQ ID NO: 1-3からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d)SEQ ID NO: 1-3とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含むポリペプチドのアゴニストとして効果的な化合物をスクリーニングする方法を提供する。この方法は、(a)このポリペプチドを含むサン

50

ブルを化合物に曝露するステップと、(b) このサンプルのアゴニスト活性を検出するステップとを含む。別法では、本発明は、この方法によって同定されたアゴニスト化合物と好適な医薬用賦形剤とを含む組成物を提供する。更なる別法では、本発明は、この組成物の患者への投与を含む、機能的NHRECの発現の低下に関連した疾患やその症状の治療方法を提供する。

【0033】

更に、本発明は、(a) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列と、(b) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列と、(c) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) SEQ ID NO: 1-3 とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含むポリペプチドのアンタゴニストとして効果的な化合物をスクリーニングする方法を提供する。この方法は、(a) このポリペプチドを含むサンプルを化合物に曝露するステップと、(b) このサンプルのアンタゴニスト活性を検出するステップとを含む。別法では、本発明は、この方法によって同定されたアンタゴニスト化合物と好適な医薬用賦形剤とを含む組成物を提供する。更なる別法では、本発明は、この組成物の患者への投与を含む、機能的NHRECの過剰な発現に関連した疾患やその症状の治療方法を提供する。

10

【0034】

更に本発明は、(a) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列と、(b) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列と、(c) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) SEQ ID NO: 1-3 とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含むポリペプチドに特異的に結合する化合物をスクリーニングする方法を提供する。この方法は、(a) このポリペプチドを好適な条件下で少なくとも1つの化合物と結合させるステップと、(b) このポリペプチドとこの試験化合物との結合を検出して、このポリペプチドと特異的に結合する化合物を同定するステップとを含む。

20

【0035】

更に本発明は、(a) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列と、(b) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列と、(c) SEQ ID NO: 1-3 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) SEQ ID NO: 1-3 とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含むポリペプチドの活性を調節する化合物をスクリーニングする方法を提供する。このスクリーニング方法は、(a) このポリペプチドを、その活性が許容される条件下で少なくとも1つの化合物と結合させるステップと、(b) この試験化合物の存在下でのこのポリペプチドの活性を評価するステップと、(c) この試験化合物の存在下でのこのポリペプチドの活性と、この試験化合物の不在下でのこのポリペプチドの活性とを比較するステップとを含み、この試験化合物の存在下でのこのポリペプチドの活性の変化が、このポリペプチドの活性を調節する化合物の存在を示唆するという特徴を有する。

30

40

【0036】

更に本発明は、SEQ ID NO: 4-6 からなる一群から選択された配列を含む標的ポリヌクレオチドの発現を変化させるのに効果的な化合物をスクリーニングする方法であって、(a) この標的ポリヌクレオチドを含むサンプルを化合物に曝露するステップと、(b) この標的ポリヌクレオチドの発現の変化を検出するステップとを含む、該スクリーニング方法を提供する。

【0037】

50

本発明はさらに、試験化合物の毒性を評価する方法を提供する。この方法は、(a) 核酸を含む生体サンプルを前記試験化合物で処理するステップと、(b) 処理した前記生体サンプルの核酸をプローブとハイブリダイズするステップと、(c) ハイブリダイゼーション複合体の収量を測定するステップと、(d) 前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量を、未処理の生体サンプルにおけるハイブリダイゼーション複合体の収量とを比較するステップとを含み、前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量の差異が試験化合物の毒性を示唆する。この方法における前記プローブは、(1) SEQ ID NO: 4-6 からなる一群から選択されたポリヌクレオチド配列と、(2) SEQ ID NO: 4-6 からなる一群から選択されたポリヌクレオチド配列と少なくとも90%の配列同一性を有する天然のポリヌクレオチド配列と、(3) 前記(1)に相補的なポリヌクレオチド配列と、(4) 前記(2)に相補的なポリヌクレオチド配列と、(5) 前記(1)乃至(4)のRNA等価物とで構成される一群から選択されたポリヌクレオチド配列を含むポリヌクレオチドの連続する少なくとも20個のヌクレオチドを含む。また、前記ハイブリダイゼーションは、前記プローブと前記生体サンプルの標的ポリヌクレオチドとの間で特異的なハイブリダイゼーション複合体が形成される条件下で行われる。また、前記標的ポリヌクレオチドが、(1) SEQ ID NO: 4-6 からなる一群から選択されたポリヌクレオチド配列と、(2) SEQ ID NO: 4-6 からなる一群から選択されたポリヌクレオチド配列と少なくとも90%の配列同一性を有する天然のポリヌクレオチド配列と、(3) 前記(1)に相補的なポリヌクレオチド配列と、(4) 前記(2)に相補的なポリヌクレオチド配列と、(5) 前記(1)乃至(5)のRNA等価物とを含む。代替的に前記標的ポリヌクレオチドは前記ポリヌクレオチド配列の断片である。

【0038】

(本発明の記載について)

本発明のタンパク質及び核酸配列、方法について説明する前に、本発明は、ここに開示した特定の装置及び材料、方法に限定されず、その実施形態を変更できることを理解されたい。また、ここで用いられる用語は、特定の実施例のみを説明する目的で用いられたものであり、後述の請求の範囲によってのみ限定され、本発明の範囲を限定することを意図したものではないということも理解されたい。

【0039】

本明細書及び請求の範囲において単数形を表す「或る」、「その(この等)」は、文脈で明確に示していない場合は複数形を含むことに注意されたい。従って、例えば「或る宿主細胞」は複数の宿主細胞を含み、その「抗体」は複数の抗体は含まれ、当業者には周知の等価物なども含まれる。

【0040】

本明細書で用いた全ての科学技術用語は、別の方法で定義されていない限り、本発明の属する技術分野の一般的な技術者が普通に解釈する意味と同じである。本明細書で記述したものと類似、或いは同等の全ての装置及び材料、方法は本発明の実施及びテストに使用できるが、好適な装置及び材料、方法をここに記す。本明細書に記載の全ての文献は、本発明に関連して使用する可能性のある文献に記載された細胞系、プロトコル、試薬、ベクターを記述し開示するために引用した。従来の特許を引用したからと言って、本発明の新規性が損なわれると解釈されるものではない。

【0041】

(定義)

用語「NHREC」は、天然、合成、半合成或いは組換え体など全ての種(特にウシ、ヒツジ、ブタ、マウス、ウマ及びヒトを含む哺乳動物)から得られる実質的に精製されたNHRECのアミノ酸配列を指す。

【0042】

用語「アゴニスト」は、NHRECの生物学的活性を強める、或いは模倣する分子を指す。このアゴニストは、NHRECに直接相互作用するか、或いはNHRECが関与する生

物学的経路の成分と作用して、N H R E Cの活性を調節するタンパク質、核酸、糖質、小分子、任意の他の化合物や組成物を含み得る。

【0043】

用語「アレル変異配列」は、N H R E Cをコードする遺伝子の別の形を指す。アレル変異配列は、核酸配列における少なくとも1つの変異によって生じ、変異m R N A若しくは変異ポリペプチドになり、これらの構造や機能は変わる場合もあれば変わらない場合もある。ある遺伝子は、天然型のアレル変異配列が存在しないもの、1つ或いは多数存在するものがある。一般にアレル変異配列を生じる変異は、ヌクレオチドの自然な欠失、付加、或いは置換による。これらの各変異は、単独或いは他の変異と同時に起こり、所定の配列内で一回或いはそれ以上生じる。

10

【0044】

N H R E Cをコードする「変異」核酸配列は、様々なヌクレオチドの欠失、挿入、或いは置換が起こっても、N H R E Cと同じポリペプチド或いはN H R E Cの機能特性の少なくとも1つを備えるポリペプチドを指す。この定義には、N H R E Cをコードするポリヌクレオチド配列の正常な染色体の遺伝子座ではない位置でのアレル変異配列との不適合或いは予期しないハイブリダイゼーション、並びにN H R E Cをコードするポリヌクレオチドの特定のオリゴヌクレオチドプローブを用いて容易に検出可能な或いは検出困難な多形性を含む。コードされたタンパク質も変異され得り、サイレント変化を生じN H R E Cと機能的に等価となるアミノ酸残基の欠失、挿入、或いは置換を含み得る。意図的なアミノ酸置換は、生物学的或いは免疫学的にN H R E Cの活性が保持される範囲で、残基の極性、電荷、溶解度、疎水性、親水性、及び／または両親媒性についての類似性に基づいて成され得る。例えば、負に荷電したアミノ酸にはアスパラギン酸及びグルタミン酸が含まれ、正に荷電したアミノ酸にはリシン及びアルギニンが含まれ得る。類似の親水性の値をもち極性非荷電側鎖を有するアミノ酸には、アスパラギン、グルタミン、セリン、トレオニンが含まれ得る。類似の親水性の値をもち非荷電側鎖を有するアミノ酸には、ロイシン、イソロイシン、バリン、グリシン、アラニン、フェニルアラニン及びチロシンが含まれ得る。

20

【0045】

用語「アミノ酸」及び「アミノ酸配列」は、オリゴペプチド、ペプチド、ポリペプチド、タンパク質配列、或いはそれらの任意の断片を指し、天然の分子及び合成分子を含む。「アミノ酸配列」が天然のタンパク質分子の配列を指す場合、「アミノ酸配列」及び類似の用語は、アミノ酸配列を記載したタンパク質分子に関連する完全で元のままのアミノ酸配列に限定するものではない。

30

【0046】

用語「増幅」は、核酸配列の複製物を作製することに関連する。一般に増幅は、この技術分野で周知のポリメラーゼ連鎖反応（P C R）技術によって行われる。

【0047】

用語「アンタゴニスト」は、N H R E Cの生物学的活性を阻害或いは減弱する分子である。アンタゴニストは、N H R E Cに直接相互作用するか、或いはN H R E Cが関与する生物学的経路の成分と作用して、N H R E Cの活性を調節する抗体、核酸、糖質、小分子、任意の他の化合物や組成物などのタンパク質を含み得る。

40

【0048】

用語「抗体」は、抗原決定基と結合可能なF a b及びF（a b'）₂、及びそれらの断片、F v断片などの無傷の分子を指す。N H R E Cポリペプチドと結合する抗体は、抗体を免疫する小ペプチドを含む無傷の分子またはその断片を用いて作製可能である。動物（例えば、マウス、ラット、若しくはウサギ）を免疫化するのに使用されるポリペプチド或いはオリゴペプチドは、R N Aの翻訳から、或いは化学的に合成可能であり、必要に応じて担体タンパク質と結合させることも可能である。ペプチドと化学的に結合した一般に用いられる担体は、ウシ血清アルブミン、チログロブリン、及びキーホールリンペットヘモニアン（K L H）を含む。次ぎに、この結合したペプチドを用いて動物を免疫化する。

50

【0049】

用語「抗原決定基」は、特定の抗体と接触する分子の領域（即ちエピトープ）を指す。タンパク質或いはタンパク質の断片が、宿主動物を免疫化するのに用いられるとき、このタンパク質の種々の領域は、抗原決定基（タンパク質上の特定の領域或いは三次元構造体）に特異的に結合する抗体の産生を誘発し得る。抗原決定基は、抗体と結合するために無傷の抗原（即ち、免疫応答を引き出すために用いられる免疫原）と競合し得る。

【0050】

本明細書において「アンチセンス」は、特定の核酸配列のセンス（コーディング）鎖と塩基対を形成し得る任意の組成物を指す。アンチセンス成分には、DNAと、RNAと、ペプチド核酸（PNA）と、ホスホロチオネートやメチルホスホネート、ベンジルホスホネート（benzylphosphonate）などの修飾された骨格（backbone linkage）を有するオリゴヌクレオチドと、2'-メトキシエチル糖または2'-メトキシエトキシ糖などの修飾された糖を有するオリゴヌクレオチドと、5-メチルシトシンまたは2'-deoxyuracil、7-deaza-2'-deoxyguanosineなどの修飾された塩基を有するオリゴヌクレオチドを含み得る。アンチセンス分子は、化学合成や転写を含む任意の方法で作ることができる。相補的アンチセンス分子は、一度細胞に導入されると、細胞によって作られた天然の核酸配列と塩基対となって二重鎖を形成し、転写や翻訳を阻害する。「負」または「マイナス」という表現はアンチセンス鎖であり、「正」または「プラス」という表現はセンス鎖である。

10

【0051】

用語「生物学的に活性」は、天然分子の構造的、調節的、或いは生化学的な機能を有するタンパク質を指す。同様に、用語「免疫学的に活性」または「免疫原性」は、天然或いは組換え体のNHREC、合成のNHRECまたはそれらの任意のオリゴペプチドが、適当な動物或いは細胞の特定の免疫応答を誘発して特定の抗体と結合する能力を指す。

20

【0052】

用語「相補的」は、塩基対合によってアニールする2つの一本鎖核酸配列間の関係を指す。例えば、配列「5' AGT 3'」が相補的な配列「3' TCA 5'」と対をなす。

【0053】

「所定のポリヌクレオチド配列を含む組成物」または「所定のアミノ酸配列を含む組成物」は広い意味で、所定のヌクレオチド配列若しくはアミノ酸配列を含む任意の組成物を指す。この組成物は、乾燥した製剤或いは水溶液を含み得る。NHREC若しくはNHRECの断片をコードするポリヌクレオチド配列を含む組成物は、ハイブリダイゼーションプローブとして使用され得る。このプローブは、凍結乾燥状態で保存可能であり、糖質などの安定化剤と結合させることが可能である。ハイブリダイゼーションにおいて、プローブは、塩（例えば、NaCl）及び界面活性剤（例えば、SDS：ドデシル硫酸ナトリウム）、その他の物質（例えば、デンハート液、乾燥ミルク、サケ精子DNAなど）を含む水溶液に展開され得る。

30

【0054】

「コンセンサス配列」は、不要な塩基を分離するためにDNA配列の解析を繰り返し行い、XL-PCRキット（PE Biosystems, Foster City CA）を用いて5'及び／または3'の方向に伸長され、再度シーケンシングされた核酸配列、またはGELVIEW断片構築システム（GCG, Madison, WI）またはPhrap（University of Washington, Seattle WA）等の断片構築用のコンピュータプログラムを用いて1つ或いはそれ以上の重複するcDNAやEST、またはゲノムDNA断片から構築された核酸配列を指す。伸長及び重複の両方によって構築されるコンセンサス配列もある。

40

【0055】

用語「保存的なアミノ酸置換」は、元のタンパク質の特性を殆ど変えない置換を指す。即ち、置換によってそのタンパク質の構造や機能が大きくは変わらず、そのタンパク質の構造、特にその機能が保存される。以下に、あるタンパク質の元のアミノ酸が別のアミノ酸

50

に置換される保存的なアミノ酸置換を示す。

元の残基

保存的な置換

A l a	G l y, S e t
A r g	H i s, L y s
A s n	A s p, G l n, H i s
A s p	A s n, G l u
C y s	A l a, S e r
G l n	A s n, G l u, H i s
G l u	A s p, G l n, H i s
G l y	A l a
H i s	A s n, A r g, G l n, G l u
I l e	L e u, V a l
L e u	I l e, V a l
L y s	A r g, G l n, G l u
M e t	L e u, I l e
P h e	H i s, M e t, L e u, T r p, T y r
S e r	C y s, T h r
T h r	S e r, V a l
T r p	P h e, T y r
T y r	H i s, P h e, T r p
V a l	I l e, L e u, T h r

10

20

一般に、保存されたアミノ酸置換の場合は、a) 置換された領域のポリペプチドの骨格構造、例えば、 β シートや α ヘリックス高次構造、b) 置換された部位の分子の電荷または疎水性、及び／または、c) 側鎖の大半が維持される。

【0056】

用語「欠失」は、1個以上のアミノ酸残基が欠如するアミノ酸配列の変化、或いは1個以上のヌクレオチドが欠如する核酸配列の変化を指す。

【0057】

用語「誘導体」は、化学修飾されたポリヌクレオチドまたはポリペプチドを指す。ポリヌクレオチド配列の化学修飾には、例えば、アルキル基、アシル基、ヒドロキシル基、或いはアミノ基による水素の置換がある。誘導体ポリヌクレオチドは、自然分子（未修飾の分子）の生物学的或いは免疫学的機能の少なくとも1つを維持するポリペプチドをコードする。誘導体ポリペプチドとは、もとのポリペプチドの生物学的機能、或いは免疫学的機能の少なくとも1つを維持する、グリコシル化、ポリエチレングリコール化、或いは任意の同様のプロセスによって修飾されたポリペプチドのことである。

30

【0058】

「検出可能な標識」は、測定可能な信号を生成し得る、ポリヌクレオチドやポリペプチドに共有結合或いは非共有結合するレポーター分子や酵素を指す。

【0059】

用語「断片」は、NHRECまたはNHRECをコードするポリヌクレオチドの固有の部分であって、その親配列（parent sequence）と同一であるがその配列より長さが短いものを指す。「断片」の最大の長さは、親配列から1つのヌクレオチド／アミノ酸残基を差し引いた長さである。例えば、ある断片は、5～1000個の連続するヌクレオチド或いはアミノ酸残基を含む。プローブ、プライマー、抗原、治療用分子、またはその他の目的に用いる断片は、少なくとも5、10、15、16、20、25、30、40、50、60、75、100、150、250若しくは500個の連続するヌクレオチド或いはアミノ酸残基の長さである。断片は、優先的に分子の特定の領域から選択される場合もある。例えば、ポリペプチド断片は、所定の配列に示された最初の250若しくは500のアミノ酸（或いは、ポリペプチドの最初の25%または50%）から選択された連続するアミノ酸の所定の長さを含み得る。これらの長さは一例であり、配列表及び表

40

50

、図面を含む明細書に記載の任意の長さが、本発明の実施例に含まれ得る。

【0060】

SEQ ID NO: 4-6の断片は、例えば、この断片を得たゲノム内の他の配列とは異なる、SEQ ID NO: 4-6を明確に同定する固有のポリヌクレオチド配列の領域を含む。SEQ ID NO: 4-6のある断片は、例えば、ハイブリダイゼーションや増幅技術、またはSEQ ID NO: 4-6に関連ポリヌクレオチド配列から区別する類似の方法に有用である。ある断片と一致するSEQ ID NO: 4-6の正確な断片の長さや領域は、その断片の目的に基づいて当分野で一般的な技術によって日常的に測定できる。

【0061】

SEQ ID NO: 1-3のある断片は、SEQ ID NO: 4-6のある断片によってコードされる。SEQ ID NO: 1-3のある断片は、SEQ ID NO: 1-3を特異的に同定する固有のアミノ酸配列の領域を含む。例えば、SEQ ID NO: 1-3のある断片は、SEQ ID NO: 1-3を特異的に認識する抗体の開発における免疫原性ペプチドとして有用である。ある断片と一致するSEQ ID NO: 1-12の正確な断片の長さや領域は、その断片の目的に基づいて当分野で一般的な技術によって日常的に決定できる。

【0062】

「完全長」ポリヌクレオチド配列とは、少なくとも1つの翻訳開始コドン（例えばメチオニン）、それに続くオープンリーディングフレーム及び翻訳終止コドンを有する配列である。「完全長」ポリヌクレオチド配列は、「完全長」ポリペプチド配列をコードする。

【0063】

「相同性」は、2つ以上のポリヌクレオチド配列間または2つ以上のポリペプチド配列間の配列類似性である。この配列類似性は配列同一性と言い換えることができる。

【0064】

ポリヌクレオチド配列についての用語「パーセントの同一性」または「%の同一性」とは、標準化されたアルゴリズムを用いてアラインメントされる、2つ以上のポリヌクレオチド配列間の一致する残基の百分率のことである。このようなアルゴリズムは、標準化され再現できる方法で、2つの配列間のアラインメントを最適化するべく、配列にギャップを挿入して、より意味をもつ2つの配列間の比較を行うことができる。

【0065】

ポリヌクレオチド配列間の同一性のパーセントは、MEGALIGN version 3.12e配列アラインメントプログラムに組込まれるCLUSTAL Vアルゴリズムのデフォルトパラメータを用いて決定可能である。このプログラムはLASERGENEソフトウェアパッケージの一部であり、分子生物学分析プログラム一式（DNASTAR, Madison WI）である。このCLUSTAL Vは、Higgins, D. G. 及び P. M. Sharp (1989) CABIOS 5:151-153、Higgins, D. G. 他 (1992) CABIOS 8:189-191に記載されている。ポリヌクレオチド配列の対のアライメントの場合、デフォルトパラメータは、Ktuple=2、gap penalty=5、window=4、「diagonals saved」=4と設定する。「重み付けされた」残基重み付け表が、デフォルトとして選択された。同一性のパーセントは、アラインメントされたポリヌクレオチド配列間の「類似性のパーセント」としてCLUSTAL Vによって報告される。

【0066】

別法では、National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S. F. 他 (1990) J. Mol. Biol. 215:403-410) が提供する、広く用いられている無料の配列比較アルゴリズム一式が、NCBI (Bethesda, MD) を含む幾つかのソース及びインターネット (<http://WWW.ncbi.nlm.nih.gov/BLAST/>) で入手可能である。このBLASTソフトウェア一式には、既知のポリヌク

レオチド配列と様々なデータベースの別のポリヌクレオチド配列とのアラインメントに用いられる「blastn」を含む、様々な配列分析プログラムが含まれる。「BLAST 2 Sequences」と呼ばれるツールが入手可能であり、2つのヌクレオチド配列の対を直接比較するために用いられる。「BLAST 2 Sequences」は、<http://WWW.ncbi.nlm.nih.gov/gorf/bl2.html>にアクセスして、対話形式で利用できる。「BLAST 2 Sequences」ツールは、blastn 及び blastp（以下に記載）の両方に用いることができる。BLASTプログラムは、一般的には、デフォルトを設定するギャップ及び他のパラメータと共に用いられる。例えば、2つのヌクレオチド配列を比較する場合、ある者は、デフォルトパラメータに設定された「BLAST 2 Sequences」ツールVersion 2.0.12（April-21-2000）でblastnを使用するであろう。そのようなデフォルトパラメータは、例えば、以下のようにする。

【0067】

Matrix: BLOSUM62
 Reward for match: 1
 Penalty for mismatch: -2
 Open Gap: 5 及び Extension Gap: 2 penalties
 Gap x drop-off: 50
 Expect: 10
 Word Size: 11
 Filter: on

同一性のパーセントは、例えば、特定の配列番号で決められた、所定の配列の全長に対して測定してもよいし、それより短い長さに対して、例えば、ある大きな所定の配列から得られた断片、例えば、連続する少なくとも、20または30、40、50、70、100、200のヌクレオチドの断片の長さに対して測定してもよい。このような長さは単なる例であり、配列表及び表、図面を含む明細書に記載の配列の任意の長さの断片を用いて、同一性のパーセントが測定される長さを示すことができる。

【0068】

高い同一性を示さない核酸配列でも、遺伝子コードの縮重によって類似のアミノ酸配列をコードし得る。縮重を利用して核酸配列を変え、それぞれが実質的に同じタンパク質をコードする様々な核酸配列を作製できることを理解されたい。

【0069】

ポリペプチド配列に用いられる用語「パーセントの同一性」または「%の同一性」とは、標準化されたアルゴリズムを用いてアラインメントされる2つ以上のポリペプチド配列間の一致する残基の百分率のことである。ポリペプチド配列アラインメントの方法は周知である。アラインメント方法の中には、保存的なアミノ酸置換を考慮したものもある。詳細に上述したこのような保存的な置換は、一般に、置換部位の電荷や疎水性が保存され、ポリペプチドの構造（従って機能も）が保存される。

【0070】

ポリペプチド配列間の同一性のパーセントは、MEGALIGN バージョン3.12e 配列アラインメントプログラム（上記）に組込まれるCLUSTAL Vアルゴリズムのデフォルトパラメータを用いて決定可能である。CLUSTAL Vを用いる対方式のポリペプチド配列のアラインメントの場合、デフォルトパラメータは、Ktuple=1、gap penalty=3、window=5、及び「diagonals saved」=5と設定する。PAM250マトリクスが、デフォルトの残基重み付け表として選択される。ポリヌクレオチドアラインメントと同様に、アラインメントされたポリペプチド配列の対の同一性のパーセントは、「類似性のパーセント」としてCLUSTAL Vによって報告される。

【0071】

別法では、NCBI BLASTソフトウェア一式が用いられる。例えば、2つのポリペ

プチド配列を対で比較をする場合、ある者は、デフォルトパラメータで設定された「BLAST 2 Sequences」ツール Version 2.0.12 (Apr-21-2000) で blastp を使用するであろう。そのようなデフォルトパラメータは、例えば、以下のようにする。

【0072】

Matrix: BLOSUM62

Open Gap: 11 及び Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

10

同一性のパーセントは、例えば、特定の配列番号で決められた、所定のポリペプチド配列の全長に対して測定してもよいし、それより短い長さに対して、例えば、ある大きな所定のポリペプチド配列から得られた断片、例えば、連続する少なくとも15、20または30、40、50、70、150の残基の断片の長さに対して測定してもよい。このような長さは単なる例であり、配列表及び表、図面を含む明細書に記載の配列の任意の長さの断片を用いて、同一性のパーセントが測定される長さを示すことができる。

【0073】

「ヒト人工染色体 (HAC)」は、約6 kb (キロベース) ~ 10 Mb のサイズの DNA 配列を含み得る、安定した有糸分裂染色体の分離及び維持に必要な全てのエレメントを含む直鎖状の小染色体である。

20

【0074】

用語「ヒト化抗体」は、もとの結合能力を保持しつつよりヒトの抗体に似せるために、非抗原結合領域のアミノ酸配列が変えられた抗体分子を指す。

【0075】

「ハイブリダイゼーション」とは、所定のハイブリダイゼーション条件下で、ある一本鎖ポリヌクレオチドがある相補的な一本鎖と塩基対を形成するアニーリングのプロセスである。特異的なハイブリダイゼーションとは、2つの核酸配列が高い相同性を有することを意味する。アニーリングが許容される条件下で、特異的なハイブリダイゼーション複合体が形成され、洗浄過程の後にもハイブリダイズしたままである。洗浄過程は、ハイブリダイゼーションプロセスの厳密性即ちストリンジェンシー (stringency) の決定において特に重要であり、よりストリンジェントな条件では、非特異的な結合、即ち完全には一致しない核酸鎖間の対の結合が減少する。核酸配列間のアニーリングが許容される条件は、当業者によって日常的に決定され、ハイブリダイゼーションの間は一定であるが、洗浄過程は、目的のストリンジェンシーにするためにその最中に条件の変更が可能であり、ハイブリダイゼーション特異性が得られる。アニーリングが許容される条件は、例えば、温度が68℃で、約6×SSC、約1% (w/v) のSDS、並びに約100 µg/ml のせん断して変性したサケ精子DNAが含まれる。

30

【0076】

一般に、ハイブリダイゼーションのストリンジェンシーは、洗浄過程を行う際の温度によっても左右される。この洗浄温度は通常、所定のイオン強度とpHにおける特定の配列の熱融点 (Tm) より約5~20℃低く選択される。このTmは、(所定のイオン強度とpHの下) 標的の配列の50%が完全に一致するプローブとハイブリダイズする温度である。Tmを計算する式及び核酸のハイブリダイゼーションの条件は、周知であり、Sambrook, J. 他による, 1989, Molecular Cloning: A Laboratory Manual, 第2版の1-3巻, Cold Spring Harbor Press, Plainview NY; 特に2巻の9章に記載されている。

40

【0077】

本発明のポリヌクレオチド間の高いストリンジェンシーのハイブリダイゼーションでは、約0.2×SSC及び約1%のSDSの存在の下、約68℃で1時間の洗浄過程を含む

50

。別法では、65℃、60℃、55℃、42℃の温度で行う。SSCの濃度は、約0.1%のSDSが存在の下、約0.1~2x SSCの範囲である。通常は、ブロッキング試薬を用いて非特異的なハイブリダイゼーションを阻止する。このようなブロッキング試薬には、例えば、約100~200 µg/mlの切断され変性したサケ精子DNAが含まれる。約35~50% v/vの濃度のホルムアミドなどの有機溶剤が、例えば、RNAとDNAのハイブリダイゼーションなどの特定の場合に用いることができる。これらの洗浄条件の有用な改変は、当業者には周知である。特に高いストリンジェントな条件でのハイブリダイゼーションは、ヌクレオチド間の進化における類似性を示唆し得る。このような類似性は、それらのヌクレオチド及びコードされたポリペプチドが類似の役割を果たしていることを強く示唆する。

10

【0078】

用語「ハイブリダイゼーション複合体」は、相補的な塩基間の水素結合によって、形成された2つの核酸配列の複合体を指す。ハイブリダイゼーション複合体は溶液中（例えば、C₀tまたはR₀t分析）で形成されるか、或いは溶液中の1つの核酸配列と固体の支持物（例えば、紙、膜、フィルタ、チップ、ピン、或いはスライドガラス、または細胞及びその核酸を固定する任意の適当な基板）に固定されたもう一つの核酸配列とで形成され得る。

【0079】

用語「挿入」或いは「付加」は、1個以上のアミノ酸残基或いはヌクレオチドがそれぞれ追加されるアミノ酸配列或いは核酸配列の変化を指す。

20

【0080】

「免疫応答」は、炎症性疾患及び外傷、免疫異常、感染症、遺伝病などに関連する症状を指す。これらの症状は、細胞系及び全身防衛系に影響を及ぼすサイトカイン及びケモカイン、別の情報伝達分子などの様々な因子の発現という特徴をもつ。

【0081】

用語「免疫原性断片」は、例えば哺乳動物などの生きている動物に導入すると、免疫反応を引き起こすGVREDのポリペプチド断片またはオリゴペプチド断片を指す。用語「免疫原性断片」はまた、本明細書で開示するまたは当分野で周知のあらゆる抗体生産方法に有用なNHRECのポリペプチド断片またはオリゴペプチド断片を含む。

【0082】

用語「マイクロアレイ」は、基質上の複数のポリヌクレオチド、ポリペプチドまたはその他の化合物の構成を指す。

30

【0083】

用語「エレメント」または「アレイエレメント」は、マイクロアレイ上に固有の指定された位置を有する、ポリヌクレオチド、ポリペプチドまたはその他の化合物を指す。

【0084】

用語「調節」は、NHRECの活性の変化を指す。例えば、調節によって、NHRECのタンパク質活性、或いは結合特性、またはその他の生物学的特性、機能的特性或いは免疫学的特性の変化が起こる。

【0085】

用語「核酸」及び「核酸配列」は、ヌクレオチド、オリゴヌクレオチド、ポリヌクレオチド、或いはそれらの断片を指し、一本鎖若しくは二本鎖であって、センス鎖或いはアンチセンス鎖であるゲノム起源若しくは合成起源のDNA或いはRNA、ペプチド核酸（PNA）、任意のDNA様物質、及びRNA様物質である。

40

【0086】

「機能的に結合した」は、第1の核酸配列と第2の核酸配列が機能的な関係にある状態を指す。例えば、プロモーターがコード配列の転写または発現に影響を与える場合、そのプロモーターはそのコード配列に機能的に結合している。一般に、機能的に結合したDNA配列は、同じ読み枠内で2つのタンパク質をコードする領域が結合する必要がある場合は、非常に近接或いは連続する。

50

【0087】

「ペプチド核酸（PNA）」は、末端がリシンで終わるアミノ酸残基のペプチド骨格に結合した、少なくとも約5ヌクレオチドの長さのオリゴヌクレオチドを含む、アンチセンス分子または抗遺伝子剤を指す。この末端のリシンにより、この組成物が溶解性となる。PNAは、相補的な一本鎖DNAやRNAに優先的に結合して転写物の伸長を止め、ポリエチレングリコール化して細胞における寿命を延ばし得る。

【0088】

NHRECの「翻訳後修飾」には、脂質化、グリコシル化、リン酸化、アセチル化、ラセミ化、蛋白分解性切断及びその他の当分野で既知の修飾を含まれ得る。これらのプロセスは、合成或いは生化学的に生じ得る。生化学的修飾は、NHRECの酵素環境に依存し、細胞の種類によって異なり得る。 10

【0089】

「プローブ」とは、同一配列或いはアレル核酸配列、関連する核酸配列の検出に用いる、NHRECやそれらの相補配列、またはそれらの断片をコードする核酸配列のことである。プローブは、検出可能な標識またはレポーター分子が結合され単離されたオリゴヌクレオチドやポリヌクレオチドである。典型的な標識には、放射性アイソトープ及びリガンド、化学発光試薬、酵素がある。「プライマー」とは、相補的な塩基対を形成して標的のポリヌクレオチドにアニーリング可能な、通常はDNAオリゴヌクレオチドである短い核酸である。プライマーがポリヌクレオチドにアニーリングした後、あるDNAポリメラーゼ酵素によって、標的のDNA一本鎖に沿って伸長される。プライマーの組は、例えば、PCR法における核酸配列の増幅（及び同定）に用いることができる。 20

【0090】

本発明に用いられるプローブ及びプライマーは、既知の配列の少なくとも15の連続するヌクレオチドを含む。特異性を高めるために、より長いプローブ及びプライマーが用いることも可能である。例えば、開示した核酸配列の連続する少なくとも20または25、30、40、50、60、70、80、90、100、150のヌクレオチドを含む。プローブ及びプライマーは、上記した例より相当長いものも用いることができ、本明細書の表及び図面、配列表に示された任意の長さのヌクレオチドも用いることができることを理解されたい。

【0091】

プローブ及びプライマーの準備及び使用方法については、例えば、Sambrook, J. 他による、1989年、名称「Molecular Cloning: A Laboratory Manual」、第2版の1-3巻（Cold Spring Harbor Press, Plainview NY）、またはAusubel, F. M. 他による、1987年、名称「Current Protocols in Molecular Biology」（Greene Publi. Assoc. & Wiley-Intersciences, New York NY）、並びにInnis他による、1990年、名称「PCR Protocols, A Guide to Methods and Applications」（Academic Press, San Diego CA.）を参照されたい。PCR用のプライマーの組は、例えば、Primer（Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA）などのそのような目的のためのコンピュータプログラムを用いて、ある既知の配列から引き出すことができる。 40

【0092】

プライマーとして用いるオリゴヌクレオチドは、当分野で周知のプライマー選択用のコンピュータプログラムで選択される。例えば、OLIGO 4.06ソフトウェアは、それぞれが最大100ヌクレオチドまでのPCR用のプライマーの対の選択、及び32,000塩基までの入力ポリヌクレオチド配列から最大5,000ヌクレオチドまでの大きなポリヌクレオチド及びオリゴヌクレオチドの分析に有用である。類似のプライマー選択用プログラムには、能力を拡大する追加の機能が含まれている。例えば、PrimOUプライ 50

マー選択プログラム (Genome Center at University of Texas South West Medical Center, Dallas TXより入手可能) は、メガベース配列から特定のプライマーを選択できるため、ゲノムワイドスコープ (genome-wide scope) におけるプライマーの設計に有用である。Primer3プライマー選択プログラム (Whitehead Institute/MIT Center for Genome Research, Cambridge MAより入手可能) によって、ユーザーは、プライマー結合部位として避けたい配列を指定できる「非プライミングライブラリ (mispriming library)」を入力できる。また、Primer3は、特にマイクロアレイのオリゴヌクレオチドの選択に有用である (後の方の2つのプライマー選択プログラムのソースコードは、それぞれのソースから得ることができ、ユーザーのニーズを満たすように変更することもできる)。PrimerGenプログラム (UK Human Genome Mapping Project Resource Centre, Cambridge UKより入手可能) は、多数の配列アラインメントに基づいてプライマーを設計するため、アラインメントされた核酸配列の最も保存された領域或いは最も保存されていない領域のどちらかとハイブリダイズするプライマーを選択することができる。従って、このプログラムは、固有及び保存されたオリゴヌクレオチドやポリヌクレオチドの断片の同定に有用である。上記した任意の選択方法で同定されたオリゴヌクレオチドやポリヌクレオチドの断片は、例えば、PCR法やシーケンシングプライマー、マイクロアレイエレメント、或いはサンプルの核酸の完全或いは部分的に相補的なポリヌクレオチドを同定する特定のプローブなどの、ハイブリダイゼーション技術に有用である。オリゴヌクレオチドの選択方法は、上記した方法に制限されるものではない。

10

20

30

40

50

【0093】

本明細書における「組換え核酸」は天然の配列ではなく、2つ以上の配列の離れたセグメントを人工的に組み合わせた配列である。この人工の組み合わせは、化学合成によって作られる場合も多いが、前出の Sambrook に記載されたような遺伝子工学の技術を用いて核酸の離れたセグメントを人工的に操作する方がより一般的である。この「組換え核酸」には、単に核酸の一部の追加または置換、欠失によって変更された核酸も含む。組換え核酸は、あるプロモーター配列に機能的に結合した核酸配列を含む場合もある。このような組換え核酸は、例えば、ある細胞を形質転換するのに用いられるベクターの一部であり得る。

【0094】

別法では、このような組換え核酸は、この組換え核酸を発現する哺乳動物のワクチン接種に用いると、その哺乳動物の防衛的な免疫応答を誘発する、ワクシニアウイルスに基づいたウイルスベクターの一部であり得る。

【0095】

「調節エレメント」は、通常は遺伝子の非翻訳領域に由来する核酸配列であり、エンハンサー、プロモーター、イントロン及び5'及び3'の非翻訳領域 (UTR) を含む。調節エレメントは、転写や翻訳、またはRNAの安定性を調節する宿主またはウイルスタンパク質と相互作用する。

【0096】

「レポーター分子」は、核酸、アミノ酸または抗体の標識に用いられる化学的または生化学的な部分である。レポーター分子には、放射性核種、酵素、蛍光剤、化学発光剤、発色剤、基質、補助因子、インヒビター、磁気粒子及びその他の当分野で既知の成分が含まれる。

【0097】

本明細書において、DNA配列に対する「RNA等価物」とは、基準となるDNA配列と同じ直鎖の核酸配列から構成されるが、窒素性塩基のチミンがウラシルに置換され、糖鎖の背骨がデオキシリボースではなくリボースからなる。

【0098】

用語「サンプル」は、その最も広い意味で用いられている。N H R E C、N H R E Cをコードする核酸、またはその断片を含むと推定されるサンプルは、体液と、細胞からの抽出物や細胞から単離された染色体や細胞内小器官、膜と、細胞と、溶液中に存在するまたは基板に固定されたゲノムDNA、RNA、cDNAと、組織と、組織プリント等を含み得る。

【0099】

用語「特異的結合」及び「特異的に結合する」は、タンパク質若しくはペプチドと、アゴニスト、抗体、アンタゴニスト、小分子、若しくは任意の天然若しくは合成の結合組成物との間の相互作用を指す。この相互作用は、結合する分子によって認識される、例えば、抗原決定基つまりエピトープなどのタンパク質の特定の構造の存在によって左右される。例えば、抗体がエピトープ「A」に対して特異的である場合、結合していない標識した「A」及び抗体を含む反応液に、エピトープAを含むポリペプチド或いは結合していない無標識の「A」が存在すると、抗体と結合する標識Aの量が減少する。

10

【0100】

用語「実質的に精製された」は、自然の環境から取り除かれてから、単離或いは分離された核酸配列或いはアミノ酸配列であって、自然に結合している組成物が少なくとも約60%除去されたものであり、好ましくは約75%以上の除去、最も好ましくは90%以上除去されたものを指す。

【0101】

「置換」とは、一つ以上のアミノ酸またはヌクレオチドをそれぞれ別のアミノ酸またはヌクレオチドに置き換えることである。

20

【0102】

用語「基板」は、任意の好適な固体或いは半固体の支持物を指し、膜及びフィルタ、チップ、スライド、ウェハ、ファイバー、磁気または非磁気ビード、ゲル、チューブ、プレート、ポリマー、微小粒子、毛細管が含まれる。この基板には、壁または塹壕、ピン、チャンネル、細孔などの様々な表面形態があり、そこにポリヌクレオチドやポリペプチドが結合する。

【0103】

「転写イメージ」は、所定条件下での所定時間における特定の細胞の種類または組織による集合的遺伝子発現のパターンを指す。

30

【0104】

「形質転換」とは、外来DNAが受容細胞に導入されるプロセスのことである。形質転換は、当分野で周知の種々の方法により、自然或いは人工の条件下で起こり、原核宿主細胞若しくは真核宿主細胞の中に外来核酸配列を挿入する任意の周知の方法によって行うことができる。この形質転換の方法は、形質転換される宿主細胞のタイプによって選択される。この方法には、バクテリオファージまたはウイルス感染、電気穿孔法（エレクトロポレーション）、リポフェクション、及び微粒子照射が含まれるが、これらに限定されるものではない。「形質転換された」細胞には、導入されたDNAが自律的に複製するプラスミドとして或いは宿主染色体の一部として複製可能である安定的に形質転換された細胞が含まれる。さらに、限られた時間に一時的に導入DNA若しくは導入RNAを発現する細胞も含まれる。

40

【0105】

本明細書における「遺伝子組換え生物」とは、当分野で周知の遺伝子組換え技術などを用いて、人間が生物の1つ以上の細胞に異種の核酸を導入した任意の生物であり、動物及び植物を含むが、それらに限定されるものではない。微量注入や組換えウイルスに感染させるなどの慎重な遺伝子操作によって、細胞の前駆体に直接或いは間接的に異種核酸を細胞に導入する。「遺伝子操作」とは、典型的な交雑育種や*in vitro*での受精ではなく、組換えDNA分子を導入することである。本発明に従った遺伝子組換え生物には、細菌及びラン藻類、菌類、植物、動物が含まれる。本発明の単離されたDNAは、当分野で周知の、例えば、感染、形質移入、形質転換、トランス接合（*trans conjugation*）

50

tion)などの方法によって、宿主に導入することができる。本発明のDNAをそのような生物に導入する技術は周知であり、前出のSambrook他(1989)に記載されている。

【0106】

特定の核酸配列の「変異配列」とは、デフォルトパラメータ設定の「BLAST 2 Sequences」ツールVersion 2.0.9 (May-07-1999)を用いるblastnによって、ある核酸配列のある長さに対する該特定の核酸配列の同一性が、少なくとも40%と決定された核酸配列のことである。このような核酸の対は、ある長さにおいて、例えば、少なくとも50%または60%、70%、80%、85%、90%、95%、98%、或いはそれ以上の同一性を示し得る。ある変異配列は、例えば、「アレル」変異配列(上述)または「スプライス」変異配列、「種」変異配列、「多型」変異配列と表すことができる。スプライス変異配列は基準分子と同一性が極めて高い可能性があるが、mRNAプロセッシング中のエキソンの択一的スプライシングによってポリヌクレオチドの数が多くなったり、少なくなったりする。対応するポリペプチドは、基準分子に存在する追加の機能ドメインを有したり、基準分子に存在するドメインが欠落したりし得る。種変異配列は、種によって異なるポリヌクレオチド配列である。得られるポリペプチドは、互いに高いアミノ酸同一性を有する。多型変異配列は、所定の種と種における特定の遺伝子のポリヌクレオチド配列が異なる。多型変異配列はまた、ポリヌクレオチド配列の1つのヌクレオチドが異なる「1ヌクレオチド多型」(SNP)も含み得る。SNPの存在は、例えば、或る集団、病態、病態の性向を示唆し得る。

10

20

【0107】

特定のポリペプチド配列の「変異体」とは、デフォルトパラメータ設定の「BLAST 2 Sequences」ツールVersion 2.0.9 (May-07-1999)を用いるblastpによって、あるポリペプチド配列のある長さに対する該特定のポリペプチド配列の同一性が、少なくとも40%と決定されたポリペプチド配列のことである。このようなポリペプチドの対は、ある長さにおいて、例えば、少なくとも50%または60%、70%、80%、85%、90%、95%、98%、或いはそれ以上の同一性を示し得る。

【0108】

(発明)

本発明は、新規のヒト核ホルモン受容体(NHREC)及びNHRECをコードするポリヌクレオチドの発見に基づき、これらの組成物を利用した代謝障害、内分泌障害、生殖障害、発生または発達障害、免疫疾患、細胞増殖異常の診断、治療、及び予防に関する。

30

【0109】

表1は、本発明のポリヌクレオチド配列及びポリペプチド配列の識別番号を示す。各ポリヌクレオチドおよびそれに対応するポリペプチドは、1つのインサイトプロジェクト識別番号(Incyte Project ID)に相関する。各ポリペプチド配列は、記載されているようにポリペプチド配列識別番号(Polypeptide SEQ ID NO:)およびインサイトポリペプチド配列番号(Incyte Polypeptide ID)の両方によって示されている。各ポリヌクレオチド配列は、記載されているようにポリヌクレオチド配列識別番号(Polypeptide SEQ ID NO:)およびインサイトポリヌクレオチドコンセンサス配列番号(Incyte Polypeptide ID)の両方によって示されている。

40

【0110】

表2は、GenBankタンパク質(genept)データベースにおいてBLAST解析により同定された本発明のポリペプチドに相同性を有する配列を示す。列1および列2はそれぞれ、本発明の各ポリペプチドに対するポリペプチド配列識別番号(Polypeptide SEQ ID NO:)およびそれに対応するインサイトポリペプチド配列番号(Incyte Polypeptide ID)を示す。列3は、GenBankの最も近い相同体のGenBankの識別番号(Genbank ID NO:)を示す。列

50

4 は、各ポリペプチドとその G e n B a n k 相同体との間の一致を表す確率スコアを示す。列 5 は、G w n B a n k 相同体のアノテーションを示し、更に該当箇所には適当な引用文も示す。これらを引用することを以って本明細書の一部とする。

【0111】

表 3 は、本発明の各ポリペプチドの様々な構造的特徴を示す。列 1 および列 2 はそれぞれ、本発明の各ポリペプチドのポリペプチド配列識別番号 (S E Q I D N O :) およびそれに対応するインサイトポリペプチド配列番号 (I n c y t e P o l y p e p t i d e I D) を示す。列 3 は、各ポリペプチドのアミノ酸残基数を示す。列 4 および列 5 はそれぞれ、G C G 配列分析ソフトウェアパッケージの M O T I F S プログラム (G e n e t i c s C o m p u t e r G r o u p , M a d i s o n W I) によって決定された、潜在的なリン酸化部位および潜在的なグリコシル化部位を示す。列 6 は、シグネチャ (s i g n a t u r e) 配列、ドメイン、およびモチーフを含むアミノ酸残基を示す。列 7 は、タンパク質の構造／機能の分析のための分析方法を示し、該当箇所にはさらに分析方法に利用した検索可能なデータベースを示す。

10

【0112】

表 2 および表 3 は共に、本発明の各ポリペプチドの特性を要約したものであって、これらの特性は請求するポリヌクレオチドが核ホルモン受容体であることを立証するものである。例えば、S E Q I D N O : 3 はドブネズミ f a r n e s o i d X 活性化受容体 (G e n B a n k I D g 8 6 8 0 3 2) と 9 1 % の同一性を有することが B a s i c L o c a l A l i g n m e n t S e a r c h T o o l (B L A S T) によって示された (表 2 を参照) 。 B L A S の確率スコアは 3.8×10^{-236} であり、探しているポリペプチド配列アラインメントが偶然の一致により得られる確率を示す。S E Q I D N O : 3 はまた、核ホルモン受容体リガンド結合ドメインを有する。これは、隠れマルコフモデル (H M M) を基にした保存されたタンパク質ファミリードメインの P F A M データベースにおいて、統計的に有意な一致を検索して決定された (表 3 を参照) 。 B L I M P S 、 M O T I F S 、 および P R O F I L E S C A N 解析から得られたデータによって、S E Q I D N O : 3 がステロイドホルモン受容体であることが裏付けられた。S E Q I D N O : 1 および S E Q I D N O : 2 は同様の方法で解析してアノテーションを付けた。S E Q I D N O : 1 - 3 の解析のためのアルゴリズムおよびパラメータを表 7 に記載する。

20

【0113】

表 4 に示されているように、本発明の完全長ポリヌクレオチド配列は、c D N A 配列、またはゲノム D N A 由来のコード (エキソン) 配列、或いはこれらの 2 種類の配列のあらゆる組み合わせを用いて組み立てた。列 1 および列 2 はそれぞれ、本発明の各ポリヌクレオチドのポリヌクレオチド配列識別番号 (P o l y n u c l e o t i d e S E Q I D N O :) およびそれに対応するインサイトポリヌクレオチドコンセンサス配列番号 (I n c y t e P o l y n u c l e o t i d e I D) を示す。列 3 は、塩基対における各ポリヌクレオチド配列の長さを示す。列 4 は、例えば、S E Q I D N O : 4 - 6 を同定するため、或いは S E Q I D N O : 4 - 6 と関連するポリヌクレオチド配列とを区別するためのハイブリダイゼーションまたは増幅技術に有用なポリヌクレオチド配列の断片を示す。列 5 は、c D N A 配列、ゲノム D N A から推定されるコード配列 (エキソン) 、および / または c D N A およびゲノム D N A の両方からなる群に対応する識別番号を示す。これらの配列を用いて本発明の完全長ポリヌクレオチド配列を組み立てた。表 4 の列 6 および列 7 はそれぞれ、列 5 の配列に対応する c D N A 配列およびゲノム配列の開始ヌクレオチド (5 ') 位置および終了ヌクレオチド (3 ') 位置を示す。

30

40

【0114】

表 4 の列 5 に示されている識別番号は、具体的には、例えばインサイト c D N A およびそれらに対応する c D N A ライブラリの識別番号を示す。例えば、6 8 1 2 3 2 9 H 1 はインサイト c D N A 配列の識別番号であり、A D R E T U R 0 1 はそれが由来する c D N A ライブラリの識別番号である。c D N A ライブラリが示されていないインサイト c D N A は、プールされている c D N A ライブラリ (例えば、7 1 2 1 8 7 2 4 V 1) に由来する

50

。または、列5の識別番号は、ポリヌクレオチド配列の組み立てに用いたGenBankのcDNAすなわちEST（例えば、g2368743）の識別番号の場合もある。または、列5の識別番号は、Genscan分析によって推定されるゲノムDNAのコード領域の場合もある。例えば、g4096054. v113. gs_2. nt. は、Genscan推定コード配列の識別番号であって、g4096054がGenscan分析によって得られたGenBankの配列の識別番号である。このGenscan推定コード配列は、配列を組み立てる前に編集する場合がある（実施例4を参照）。または、列5の識別番号は、“exon-stitching”アルゴリズムによってcDNAおよびGenscan推定エキソンの両方からなる群の場合もある（実施例5を参照）。または、列5の識別番号は、“exon-stretching”アルゴリズムによってcDNAおよびGenscan推定エキソンの両方からなる群の場合もある（実施例5を参照）。場合によっては、列5に示されている配列の範囲と重複するインサイトcDNAの範囲が得られ、最終的なコンセンサス配列が決定されるが、それに相当するインサイトcDNAの識別番号は示されていない。

【0115】

表5は、インサイトcDNA配列を用いて組み立てられたこれらの完全長ポリヌクレオチド配列が由来する代表的なcDNAライブラリを示す。代表的なcDNAライブラリとは、上記ポリヌクレオチド配列の組み立ておよび決定に用いられたインサイトcDNA配列を最も多く含むインサイトcDNAライブラリのことである。表5に示されているcDNAライブラリを作製するために用いた組織およびベクターが表6に示されている。

【0116】

本発明はまた、NHRECの変異体も含む。好適なNHRECの変異体は、NHRECの機能的或いは構造的特徴の少なくともどちらか一方を有し、かつNHRECアミノ酸配列に対して少なくとも約80%のアミノ酸配列同一性、或いは少なくとも約90%のアミノ酸配列同一性、更には少なくとも約95%のアミノ酸配列同一性を有する。

【0117】

本発明はまた、NHRECをコードするポリヌクレオチドを提供する。特定の実施例において、本発明は、NHRECをコードするSEQ ID NO: 4-6からなる一群から選択された配列を含むポリヌクレオチド配列を提供する。配列表に示したSEQ ID NO: 4-6のポリヌクレオチド配列は、窒素系塩基のチミンがウラシルに置換され、糖鎖の背骨がデオキシリボースではなくリボースからなる等価RNA配列を含む。

【0118】

本発明はまた、NHRECをコードするポリヌクレオチド配列の変異配列を含む。詳細には、このようなポリヌクレオチド配列の変異配列は、NHRECをコードするポリヌクレオチド配列と少なくとも70%のポリヌクレオチド配列同一性、或いは少なくとも85%のポリヌクレオチド配列同一性、更には少なくとも95%のポリヌクレオチド配列同一性を有する。本発明の特定の実施形態は、SEQ ID NO: 4-6からなる一群から選択された核酸配列と少なくとも70%のポリヌクレオチド配列同一性、或いは少なくとも85%のポリヌクレオチド配列同一性、更には少なくとも95%のポリヌクレオチド配列同一性を有するSEQ ID NO: 4-6からなる一群から選択された配列を含むポリヌクレオチド配列の変異配列を提供する。上記したポリヌクレオチド変異配列は何れも、NHRECの機能的或いは構造的特徴の少なくとも1つを有するアミノ酸配列をコードする。

【0119】

遺伝暗号の縮重により作り出され得るNHRECをコードする種々のポリヌクレオチド配列には、既知の自然発生する任意の遺伝子のポリヌクレオチド配列と最小の類似性しか有しないものも含まれることを、当業者は理解するであろう。したがって本発明には、可能なコドン選択に基づいた組み合わせの選択によって作り出され得る可能なポリヌクレオチド配列の変異の全てが含まれ得る。これらの組み合わせは、天然のNHRECのポリヌクレオチド配列に適用される標準的なトリプレット遺伝暗号を基に作られ、全ての変異が明

確に開示されていると考慮する。

【0120】

NHRECをコードするヌクレオチド配列及びその変異配列は一般に、好適に選択されたストリンジェントな条件下で、天然のNHRECのヌクレオチド配列とハイブリダイズ可能であるが、非天然のコドンを含めるなどの実質的に異なった使い方のコドンを有するNHREC或いはその誘導体をコードするヌクレオチド配列を作るとは有利となり得る。特定のコドンが宿主によって利用される頻度に基づいてコドンを選択して、ペプチドの発現が特定の真核細胞または原核宿主に発生する割合を高めることが可能である。コードされたアミノ酸配列を変えないで、NHREC及びその誘導体をコードするヌクレオチド配列を実質的に変更する別の理由は、天然の配列から作られる転写物より例えば長い半減期など好ましい特性を備えるRNA転写物を作ることにある。

【0121】

本発明はまた、NHREC及びその誘導体をコードするDNA配列またはそれらの断片を完全に合成化学によって作り出すことも含む。作製後にこの合成配列を、当分野で良く知られた試薬を用いて、種々の入手可能な発現ベクター及び細胞系の何れの中にも挿入可能である。更に、合成化学を用いて、NHRECまたはその任意の断片をコードする配列の中に突然変異を導入することも可能である。

【0122】

更に本発明には、種々のストリンジェントな条件下で、請求項に記載されたポリヌクレオチド配列、特に、SEQ ID NO: 4-6及びそれらの断片とハイブリダイズ可能なポリヌクレオチド配列が含まれる（例えば、Wahl, G. M. 及び S. L. Berger (1987) *Methods Enzymol.* 152: 399-407; and Kimmel, A. R. (1987) *Methods Enzymol.* 152: 507-511. を参照）。アニーリング及び洗浄条件を含むハイブリダイゼーションの条件は、「定義」に記載されている。

【0123】

当分野で周知のDNAのシーケンシング方法を用いて、本発明の何れの実施例も実行可能である。この方法には、例えばDNAポリメラーゼIのクレノウ断片、SEQUENASE (US Biochemical, Cleveland OH)、Taqポリメラーゼ (Applied Biosystems)、熱安定性T7ポリメラーゼ (Amersham, Pharmacia Biotech Piscataway NJ)、或いはELONGASE増幅システム (Life Technologies, Gaithersburg MD) にみられるような校正エキソヌクレアーゼとポリメラーゼとの組み合わせなどの酵素が用いられる。好ましくは、MICROLAB 2200液体転移システム (Hamilton, Reno, NV)、PTC 200 Thermal Cycler 200 (MJ Research, Watertown MA) 及びABI CATALYST 800 (PE Biosystems) などの装置を用いて配列の準備を自動化する。次に、ABI 373或いは377 DNAシーケンシングシステム (PE Biosystems)、MEGABACE 1000 DNAシーケンシングシステム (Molecular Dynamics, Sunnyvale CA) または当分野で周知の他の方法を用いてシーケンシングを行う。得られた配列を当分野で周知の様々なアルゴリズムを用いて分析する（例えば、Ausubel, F. M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R. A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853. を参照）。

【0124】

当分野で周知のPCR法をベースにした種々の方法で、部分的なヌクレオチド配列を利用して、NHRECをコードする核酸配列を伸長し、プロモーターや調節エレメントなどの上流にある配列を検出する。例えば制限部位PCR法を利用する1つの方法では、一般的

なプライマー及び入れ子プライマー (nested primer) を用いてクローニングベクター内のゲノムDNAから未知の配列を増幅する (例えば、Sarkar, G. (1993) PCR Methods Applic 2: 318-322を参照)。逆PCR法を用いる別法では、広範な方向に伸長して環状化した鋳型から未知の配列を増幅するプライマーを用いる。この鋳型は、既知のゲノム遺伝子座及びその周辺の配列を含む制限断片に由来する (例えば、Triglia, T. ら (1988) Nucleic Acids Res 16: 8186を参照)。キャプチャPCR法を用いる第3の方法は、ヒト及び酵母菌人工染色体DNAの既知の配列に隣接するDNA断片のPCR増幅を含む (例えば、Lagerstrom, M. 他 (1991) PCR Methods Applic 1: 111-119を参照)。この方法では、多数の制限酵素による消化及びライゲーションを用いて、PCRを行う前に未知の配列の領域の中に組換え二本鎖配列を挿入することが可能である。また、当分野で周知の別の方法を用いて未知の配列を得ることも可能である。 (例えば、Parker, J. D. 他 (1991) Nucleic Acids Res. 19: 3055-3060を参照)。更に、PCR、ネスト化プライマー、PROMOTERFINDERライブラリ (Clontech, Palo Alto CA) を用いれば、ゲノムDNA内の歩行が可能である。この方法ではライブラリをスクリーニングする必要がなく、イントロン/エキソン接合部を探すのに有用である。全てのPCR法をベースにした方法では、プライマーは、市販のOLIGO 4.06 Primer Analysis software (National Bioscience, Plymouth MN) 或いは別の好適なプログラムなどを用いて、長さが22~30ヌクレオチド、GC含量が50%以上、約68℃~72℃の温度で鋳型に対してアニーリングするよう設計される。

10

20

【0125】

完全長のcDNAをスクリーニングする場合は、大きなcDNAを含むようにサイズが選択されたライブラリを用いるのが好ましい。更に、オリゴd (T) ライブラリが完全な長さのcDNAを産生できない場合は、遺伝子の5'領域を有する配列を含むものが多いランダムに初回抗原刺激を受けたライブラリが有用である。ゲノムライブラリは、5'非転写調節領域への配列の伸長に有用であろう。

【0126】

市販のキャピラリー電気泳動システムを用いて、シークエンシングまたはPCR産物のヌクレオチド配列のサイズの分析、または確認が可能である。詳しくは、キャピラリーシークエンシングには、電気泳動による分離のための流動性ポリマー、及び4つの異なったヌクレオチドに特異的なレーザーで活性化される蛍光色素、放出された波長の検出に利用するCCDカメラを使用することが可能である。出力/光強度は、適切なソフトウェア (例えば、GENOTYPER及びSEQUENCE NAVIGATOR、PE Biosystems) を用いて電気信号に変換され、サンプルのローディングからコンピュータ分析までのプロセス及び電子データ表示がコンピュータ制御可能である。キャピラリー電気泳動法は、特定のサンプルに少量しか存在しない場合もあるDNAの小片のシークエンシングに特に適している。

30

【0127】

本発明の別の実施例では、NHRECをコードするポリヌクレオチド配列またはその断片を組換えDNA分子にクローニングして、適切な宿主細胞内にNHREC、その断片または機能的等価物を発現させることが可能である。遺伝暗号固有の縮重により、実質的に同じ或いは機能的に等価のアミノ酸配列をコードする別のDNA配列が作られ得り、これらの配列をNHRECのクローン化及び発現に利用可能である。

40

【0128】

種々の目的でNHRECをコードする配列を変えるために、当分野で一般的に知られている方法を用いて、本発明のヌクレオチド配列を組換えることができる。この目的には、遺伝子産物のクローン化、プロセッシング及び/または発現の調節が含まれるが、これらに限定されるものではない。ランダムな断片によるDNAの混合や遺伝子断片と合成オリゴ

50

ヌクレオチドのPCR再組み立てを用いて、ヌクレオチド配列の組換えが可能である。例えば、オリゴヌクレオチド媒介性定方向突然変異誘発を利用して、新しい制限部位を生成する突然変異の導入、グリコシル化パターンの変更、コドン選択の変更、スプライスバリエーションの作製等が可能である。

【0129】

本発明のヌクレオチドを、MOLECULAR BREEDING (Maxygen Inc., Santa Clara CA; 米国特許第5,837,458号; Chang, C.-C. 他 (1999) Nat. Biotechnol. 17:793-797; Christians, F. C. 他 (1999) Nat. Biotechnol. 17:259-264; Cramer, A. 他 (1996) Nat. Biotechnol. 14:315-319)などのDNAシャフリング技術を用いてシャフリングして、NHRECの生物学的または酵素的な活性、或いは他の分子や化合物と結合する能力などのNHRECの生物学的特性を変更或いは改良することができる。DNAシャフリングは、PCR法による遺伝子断片の組換えで遺伝子変異体のライブラリを作製するプロセスである。次に、このライブラリを、目的の特性を有する遺伝子変異体を同定するために選択或いはスクリーニングする。これらの好ましい変異体をプールし、DNAシャフリング及び選択／スクリーニングを繰り返す。従って、人工的な育種及び急速な分子の進化によって多様な遺伝子が作られる。例えば、ランダムな位置に変異がある1つの遺伝子の断片を、目的の特性が最適化するまで、組換え及びスクリーニング、シャフリングを実施することもできる。別法では、所定の遺伝子の断片を、同じ或いは異なった種の同じ遺伝子ファミリーの相同な遺伝子の断片で組換え、それによってプロトコルに従った調節可能な方法で、多数の天然遺伝子の遺伝子多様性を最大にすることができる。

【0130】

別の実施例によれば、NHRECをコードする配列は、当分野で周知の化学的方法を用いて、全体或いは一部が合成可能である（例えば、Caruthers, M. H. ら (1980) Nucl. Acids Res. Symp. Ser 7:215-223; 及びHorn, T. 他 (1980) Nucl. Acids Res. Symp. Ser. 225-232を参照）。別法として、化学的方法を用いてNHREC自体またはその断片を合成することが可能である。例えば、ペプチド合成は種々の固相技術を用いて実行可能である（例えば、Creighton, T. (1984) Proteins. Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; Roberge, J. Y. ら (1995) Science 269:202-204を参照）。また、合成の自動化は例えばABI 431Aペプチドシンセサイザー (PE Biosystems) を用いて達成し得る。更にNHRECのアミノ酸配列または任意のその一部は、直接的な合成の際の変更、及び／または化学的方法を用いた他のタンパク質または任意のその一部からの配列との組み合わせにより、天然のポリペプチド配列を有するポリペプチドまたは変異体ポリペプチドを作製することが可能である。

【0131】

このペプチドは、分離用高速液体クロマトグラフィー（例えば、Chiez, R. M. 及び F. Z. Regnier (1990) Methods Enzymol. 182:392-421を参照）を用いて実質的に精製可能である。合成されたペプチドの組成は、アミノ酸分析或いはシーケンシングにより確認することができる（例えば、Creighton、前出、pp 28-53を参照）。

【0132】

生物学的に活性なNHRECを発現させるために、NHRECをコードするヌクレオチド配列またはその誘導体を好適な発現ベクターに挿入する。この発現ベクターは、好適な宿主に挿入されたコーディング配列の転写及び翻訳の調節に必要なエレメントを含む。これらのエレメントには、ベクター及びNHRECをコードするポリヌクレオチド配列におけるエンハンサー、構成型及び発現誘導型のプロモーター、5'及び3'の非翻訳領域など

の調節配列が含まれる。このようなエレメントは、その長さ及び特異性が様々である。特定の開始シグナルによって、NHRECをコードする配列のより効果的な翻訳を達成することが可能である。このようなシグナルには、ATG開始コドン及びコザック配列などの近傍の配列が含まれる。NHRECをコードする配列及びその開始コドン、上流の調節配列が好適な発現ベクターに挿入された場合は、更なる転写調節シグナルや翻訳調節シグナルは必要なくなるであろう。しかしながら、コーディング配列或いはその断片のみが挿入された場合は、インフレームのATG開始コドンを含む外来性の翻訳調節シグナルが発現ベクターに含まれなければならない。外来性の翻訳要素及び開始コドンは、自然及び合成の様々なものから得ることが可能である。用いられる特定の宿主細胞系に好適なエンハンサーを含めることで発現の効率を高めることが可能である。（例えば、Scharf, D. 他 (1994) *Results Probl. Cell Differ.* 201-18-162. を参照）。

10

【0133】

当業者に周知の方法を用いて、NHRECをコードする配列、好適な転写及び翻訳調節エレメントを含む発現ベクターを作製することが可能である。これらの方法には、in vitro組換えDNA技術、合成技術、及びin vivo遺伝子組換え技術が含まれる。（例えば、Sambrook, J. 他. (1989) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, 4章及び8章, 及び16-17章; 及び Ausubel, F. M. 他. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9章及び13章1-4章を参照）。

20

【0134】

種々の発現ベクター／宿主系を利用して、NHRECをコードする配列の保持及び発現が可能である。これらには、限定するものではないが、組換えバクテリオファージ、プラスミド、またはコスミドDNA発現ベクターで形質転換された細菌などの微生物や、酵母菌発現ベクターで形質転換された酵母菌や、ウイルス発現ベクター（例えば、バキュロウイルス）に感染した昆虫細胞系や、ウイルス発現ベクター（例えば、カリフラワーモザイクウイルス、CaMV; タバコモザイクウイルス、TMV）または細菌発現ベクター（例えば、TiまたはpBR322プラスミド）で形質転換された植物細胞系や、動物細胞系などが含まれる（例えば、前出のSambrook、前出のAusubel、Van Heeke, G. および S. M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509、Engelhard, E. K. 他 (1994) *Proc. Natl. Acad. Sci. USA* 91:3224-3227、Sandig, V. 他 (1996) *Hum. Gene Ther.* 7:1937-1945、Takamatsu, N. (1987) *EMBOJ.* 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196、Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659、Harrington, J. J. 他 (1997) *Nat. Genet.* 15:345-355を参照）。レトロウイルス、アデノウイルス、ヘルペスウイルスまたはワクシニアウイルス由来の発現ベクター、または種々の細菌性プラスミド由来の発現ベクターを用いて、ヌクレオチド配列を標的器官、組織または細胞集団へ輸送することができる（Di Nicola, M. 他 (1998) *Cancer Gen. Ther.* 5(6):350-356、Yu, M. 他 (1993) *Proc. Natl. Acad. Sci. USA* 90(13):6340-6344、Buller, R. M. 他 (1985) *Nature* 317(6040):813-815; McGregor, D. P. 他 (1994) *Mol. Immunol.* 31(3):219-226、Verma, I. M. and N. Somia (1997) *Nature* 389:239-242

30

40

50

等を参照)。本発明は使用される宿主細胞によって限定されるものではない。

【0135】

細菌系では、多数のクローニングベクター及び発現ベクターが、NHRECをコードするポリヌクレオチド配列の使用目的に応じて選択可能である。例えば、NHRECをコードするポリヌクレオチド配列の日常的なクローニング、サブクローニング、増殖には、PBLUESCRIPT (Stratagene, La Jolla CA) または pSPORT1 プラスミド (GIBCO BRL) などの多機能の大腸菌ベクターを用いることができる。ベクターの多数のクローニング部位にNHRECをコードする配列をライゲーションすると lacZ 遺伝子が破壊され、組換え分子を含む形質転換された細菌の同定のための比色スクリーニング法が可能となる。更に、これらのベクターを用いて、クローニングされた配列の *in vitro* での転写、ジデオキシススクリーニング、ヘルパーファージによる一本鎖の救出、入れ子状態の欠失を作り出すことが可能である (例えば、Van Heeke, G. 及び S. M. Schuster (1989) J. Biol. Chem. 264: 5503-5509. を参照)。例えば、抗体の産生のためなどに多量のNHRECが必要な場合は、NHRECの発現をハイレベルで誘導するベクターが使用できる。例えば、強力に発現を誘発する T5 または T7 バクテリオファージプロモーターを含むベクターを使用できる。

10

【0136】

NHRECの発現に酵母の発現系の使用が可能である。 α 因子やアルコールオキシダーゼやPGHプロモーターなどの構成型或いは誘導型のプロモーターを含む多種のベクターが、酵母菌サッカロミセス・セレビジエまたは *Pichia pastoris* に使用可能である。更に、このようなベクターは、発現したタンパク質の分泌か細胞内への保持のどちらかを誘導し、安定した増殖のために宿主ゲノムの中に外来配列を組み込む。(例えば、Ausubel, 1995, 前出、Bitter, G. A. ら (1987) Methods Enzymol. 153: 516-544、及び Scorer, C. A. ら (1994) Bio/Technology 12: 181-184. を参照)。

20

【0137】

植物系もNHRECの発現に使用可能である。NHRECをコードする配列の転写は、例えば、CaMV由来の35S及び19Sプロモーターなどのウイルスプロモーターが単独で、或いはTMV (例えば、Coruzzi, G. ら (1984) EMBO J. 3: 1671-1680; Broglie, R. ら (1984) Science 224: 838-843; および Winter, J. ら (1991) Results Probl. Cell Differ. 17: 85-105 を参照) 由来のオメガリーダー配列と組み合わせて促進される。これらの作製物は、直接のDNA形質転換或いは病原体を介したトランスフェクションによって、植物細胞の中に導入可能である。(例えば、The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill NY, pp. 191-196 を参照)。

30

【0138】

哺乳動物細胞では、多種のウイルスベースの発現系が利用され得る。アデノウイルスが発現ベクターとして用いられる場合、後発プロモーター及び3連リーダー配列からなるアデノウイルス転写物／翻訳複合体にNHRECをコードする配列を結合し得る。ウイルスのゲノムの非必須のE1またはE3領域への挿入により、感染した宿主細胞にNHRECを発現する生ウイルスを得ることが可能である (Logan, J. 及び Shenk, T. (1984) Proc. Natl. Acad. Sci. 81: 3655-3659 を参照)。さらに、ラウス肉腫ウイルス (RSV) エンハンサーなどの転写エンハンサーを用いて、哺乳動物宿主細胞における発現を増大させることが可能である。タンパク質を高レベルで発現させるために、SV40 または EBV を基にしたベクターを用いることが可能である。

40

【0139】

50

ヒト人工染色体 (HAC) を用いて、プラスミドで発現しそれに含まれているものより大きな DNA の断片を供給可能である。治療のために約 6 kb ~ 10 Mb の HACs を作製し、従来の輸送方法 (リポソーム、ポリカチオンアミノポリマー、またはベシクル) で供給する。(例えば、Harrington, J. J. 他 (1997) Nat Genet. 15: 345-355. を参照)。

【0140】

哺乳動物系の組換えタンパク質の長期にわたる産生のためには、株化細胞における NHREC の安定した発現が望ましい。例えば、発現ベクターを用いて、NHREC をコードする配列を株化細胞に形質転換することが可能である。このような発現ベクターは、ウイルス起源の複製及び／または内在性の発現要素や、同じ或いは別のベクターの上の選択マーカー遺伝子を含む。ベクターの導入の後、細胞を選択培地に移す前に、強化培地で約 1 ~ 2 日の間増殖させる。選択マーカーの目的は選択的な媒介物に対する抵抗性を与えるとともに、その存在により導入された配列を確実に発現する細胞の増殖及び回収が可能となる。安定的に形質転換された細胞の耐性クローンは、その細胞型に好適な組織培養技術を用いて増殖可能である。

10

【0141】

任意の数の選択系を用いて、形質転換された細胞系を回収することが可能である。選択系には、以下のものに限定はしないが、単純ヘルペスウイルスチミジンキナーゼ遺伝子及びアデニンホスホリボシルトランスフェラーゼ遺伝子が含まれ、それぞれ tk⁻ または apr⁻ 細胞において使用される。(例えば、Wigler, M. 他 (1977) Cell 11: 223-232; 及び Lowy, I. 他 (1980) Cell 22: 817-823 を参照)。また代謝拮抗物質、抗生物質或いは除草剤への耐性を選択のベースとして用いることができる。例えば dhfr はメトトレキセートに対する耐性を与え、neo はアミノグリコシッドネオマイシン及び G-418 に対する耐性を与え、als 或いは pat はクロルスルフロン (chlorosulfuron)、ホスフィノトリシンアセチルトランスフェラーゼ (phosphinotricin acetyltransferase) に対する耐性を与える (例えば、Wigler, M. 他. (1980) Proc. Natl. Acad. Sci. 77: 3567-3570; Colbere-Garapin, F. 他 (1981) J. Mol. Biol. 150: 1-14 を参照)。さらに選択に利用できる遺伝子、例えば、代謝のために細胞が必要なものを変える trpB 及び hisD が文献に記載されている (例えば、Hartman, S. C. 及び R. C. Mulligan (1988) Proc. Natl. Acad. Sci. 85: 8047-51 を参照)。アニトシアニン、緑色蛍光タンパク質 (GFP; Clontech)、β グルクロニダーゼ及びその基質 GUS、ルシフェラーゼ及びその基質ルシフェリンなどの可視マーカーが用いられる。緑色蛍光タンパク質 (GFP) (Clontech, Palo Alto, CA) も使用できる。これらのマーカーを用いて、トランスフォーマントを特定するだけでなく、特定のベクター系に起因する一過性或いは安定したタンパク質発現を定量することが可能である (例えば、Rhodes, C. A. 他 (1995) Methods Mol. Biol. 55: 121-131 を参照)。

20

30

40

【0142】

マーカー遺伝子の発現の存在／不在によって目的の遺伝子の存在が示されても、その遺伝子の存在及び発現の確認が必要な場合もある。例えば、NHREC をコードする配列がマーカー遺伝子配列の中に挿入された場合、NHREC をコードする配列を含む形質転換された細胞は、マーカー遺伝子機能の欠落により特定可能である。または、1 つのプロモーターの制御下でマーカー遺伝子が NHREC をコードする配列と一列に配置することも可能である。誘導または選択に応答したマーカー遺伝子の発現は、通常タンデム遺伝子の発現も示す。

【0143】

一般に、NHREC をコードする核酸配列を含み、NHREC を発現する宿主細胞は、当

50

業者に周知の種々の方法を用いて特定することが可能である。これらの方法には、DNA-DNA 或いは DNA-RNA ハイブリダイゼーションや、PCR 法、核酸或いはタンパク質の検出及び／または数量化のための膜系、溶液ベース、或いはチップベースの技術を含むタンパク質生物学的試験法または免疫学的アッセイが含まれるが、これらに限定されるものではない。

【0144】

特異的なポリクローナル抗体またはモノクローナル抗体のどちらかを用いる NHREC の発現の検出及び計測のための免疫学的方法は、当分野で周知である。このような技法には、酵素に結合したイムノソルベントアッセイ (ELISA)、ラジオイムノアッセイ (RIA)、蛍光標示式細胞分取器 (FACS) などがある。NHREC 上の 2 つの非干渉エピトープに反応するモノクローナル抗体を用いた、2 部位のモノクローナルベースイムノアッセイ (two-site, monoclonal-based immunoassay) が好ましいが、競合の結合アッセイも用いることもできる。これらのアッセイ及びその他のアッセイは、当分野では十分に知られている。(例えば、Hampton, R. 他. (1990) Serological Methods, a Laboratory Manual, APS Press, St Paul, MN, Sect. IV; Coligan, J. E. 他 Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York, NY; 及び Pound, J. D. (1990) Immunochemical Protocols, Humana Press, Totowa NJ)。

【0145】

種々の標識技術及び結合技術が当業者には周知であり、様々な核酸アッセイおよびアミノ酸アッセイに用いられ得る。NHREC をコードするポリヌクレオチドに関連する配列を検出するための、標識されたハイブリダイゼーションプローブ或いは PCR プローブを生成する方法には、オリゴ標識化、ニックトランスレーション、末端標識化、または標識されたヌクレオチドを用いる PCR 増幅が含まれる。別法として、NHREC をコードする配列、またはその任意の断片を mRNA プローブを生成するためのベクターにクローニングすることも可能である。当分野では周知であり市販されているこのようなベクターを、T7, T3, または SP6 などの好適な RNA ポリメラーゼ及び標識されたヌクレオチドの追加によって、in vitro で RNA プローブの合成に用いることができる。これらの方法は、例えば、Amersham Pharmacia Biotech 及び Promega (Madison WI)、U. S. Biochemical Corp (Cleveland OH) が市販する種々のキットを用いて行うことができる。容易な検出のために用い得る好適なレポーター分子或いは標識には、基質、コファクター、インヒビター、磁気粒子、及び放射性核種、酵素、蛍光剤、化学発光剤、色素産生剤などが含まれる。

【0146】

NHREC をコードするヌクレオチド配列で形質転換された宿主細胞は、細胞培地でのこのタンパク質の発現及び回収に好適な条件下で培養される。形質転換された細胞から産生されたタンパク質が分泌されるか細胞内に留まるかは、使用されるその配列及び／またはそのベクターによる。NHREC をコードするポリヌクレオチドを含む発現ベクターは、原核細胞膜及び真核細胞膜を透過する NHREC の分泌を誘導するシグナル配列を含むように設計できることは、当業者には理解されよう。

【0147】

更に、挿入した配列の発現調節能力または発現したタンパク質を所望の形にプロセッシングする能力によって宿主細胞株が選択される。このようなポリペプチドの修飾には、アセチル化、カルボキシル化、グリコシル化、リン酸化、脂質化 (lipidation)、及びアシル化が含まれるが、これらに限定されるものではない。タンパク質の「プレプロ」または「プロ」形を切断する翻訳後のプロセッシングを利用して、標的タンパク質、折りた

たみ及び／または活性を特定することが可能である。翻訳後の活性のための特定の細胞装置及び特徴のある機構をもつ種類の宿主細胞（例えば、CHO、HeLa、MDCK、MEK293、WI38）がAmerican Type Culture Collection（ATCC；Bethesda，MD）より入手可能であり、外来のタンパク質の正しい修飾及びプロセッシングを確実にするために選択される。

【0148】

本発明の別の実施例では、NHRECをコードする自然或いは変更された、または組換えの核酸配列を上記した任意の宿主系の融合タンパク質の翻訳となる異種配列に結合させる。例えば、市販の抗体によって認識できる異種部分を含むキメラNHRECタンパク質が、NHRECの活性のインヒビターに対するペプチドライブラリのスクリーニングを促進し得る。また、異種タンパク質部分及び異種ペプチド部分が、市販の親和性基質を用いて融合タンパク質の精製を促進し得る。このような部分には、グルタチオンSトランスフェラーゼ（GST）、マルトース結合タンパク質（MBP）、チオレドキシン（Trx）、カルモジュリン結合ペプチド（CBP）、6-His、FLAG、c-mc、赤血球凝集素（HA）が含まれるが、これらに限定されるものではない。GST及びMBP、Trx、CBP、6-Hisによって、固定されたグルタチオン、マルトース、フェニルアルシン酸化物（phenylarsine oxide）、カルモジュリン、金属キレート樹脂のそれぞれで同族の融合タンパク質の精製が可能となる。FLAG、c-mc、及び赤血球凝集素（HA）によって、これらのエピトープ標識を特異的に認識する市販のモノクローナル抗体及びポリクローナル抗体を用いた融合タンパク質の免疫親和性の精製ができる。また、NHRECをコードする配列と異種タンパク質配列との間にあるタンパク質分解切断部位を融合タンパク質が含むように遺伝子操作すると、NHRECが精製の後に異種部分から切断され得る。融合タンパク質の発現と精製の方法は、Ausubel.（1995、前出ch 10）に記載されている。市販されている様々なキットを用いて、融合タンパク質の発現及び精製を促進できる。

10

20

30

【0149】

本発明の別の実施例では、TNTウサギ網状赤血球可溶化液またはコムギ胚芽抽出系（Promega）を用いて*in vitro*で放射能標識したNHRECの合成が可能である。これらの系は、T7またはT3、SP6プロモーターと機能的に結合したタンパク質をコードする配列の転写と翻訳をつなげる。翻訳は、例えば、³⁵Sメチオニンである放射能標識されたアミノ酸前駆体の存在の下で起こる。

【0150】

本発明のNHRECまたはその断片を用いて、NHRECに特異結合する化合物をスクリーニングすることができる。少なくとも1つまたは複数の試験化合物を用いて、NHRECへの特異的な結合をスクリーニングすることが可能である。試験化合物の例には、抗体、オリゴヌクレオチド、タンパク質（例えば受容体）または小分子が挙げられる。

【0151】

一実施例では、このように同定された化合物は、例えばリガンドやその断片などのNHRECの天然のリガンド、または天然の基質、構造的または機能的な擬態性または自然結合パートナーに密接に関連している（Coligan, J. E. 他（1991）Current Protocols in Immunology 1（2）の5章等を参照）。同様に、化合物は、NHRECが結合する天然受容体、或いは例えばリガンド結合部位などの少なくとも受容体のある断片に密接に関連し得る。何れの場合も、既知の技術を用いてこの化合物を合理的に設計することができる。一実施例では、このような化合物に対するスクリーニングには、分泌タンパク質或いは細胞膜上のタンパク質の何れか一方としてNHRECを発現する好適な細胞の作製が含まれる。好適な細胞には、哺乳動物、酵母、大腸菌からの細胞が含まれる。NHRECを発現する細胞またはNHRECを含有する細胞膜断片を試験化合物と接触させて、NHRECまたは化合物の何れかの結合、刺激または阻害を分析する。

40

【0152】

50

あるアッセイは、単に試験化合物をポリペプチドに実験的に結合させ、結合を、フルオロフォア、放射性同位体、酵素抱合体またはその他の検出可能な標識により検出することができる。例えば、このアッセイは、少なくとも1つの試験化合物を、溶液中の或いは固体支持物に固定されたNHRECと結合させるステップと、NHRECとこの化合物との結合を検出するステップを含み得る。別法では、標識された競合物の存在下での試験化合物の結合の検出及び測定を行うことができる。更にこのアッセイでは、細胞遊離剤、化学ライブラリまたは天然の生成混合物を用いて実施することができ、試験化合物は、溶液中で遊離させるか固体支持体に固定させる。

【0153】

本発明のNHRECまたはその断片を用いて、NHRECの活性を調整する化合物をスクリーニングすることが可能である。このような化合物には、アゴニスト、アンタゴニスト、或るいは部分的または逆アゴニスト等が含まれる。一実施例では、NHRECが少なくとも1つの試験化合物と結合する、NHRECの活性が許容される条件下でアッセイを実施し、試験化合物の存在下でのNHRECの活性が試験化合物不在下でのNHRECの活性と比較する。試験化合物の存在下でのNHRECの活性の変化は、NHRECの活性を調整する化合物の存在を示唆する。別法では、試験化合物をNHRECの活性に適した条件下でNHRECを含む*in vitro*または細胞遊離系と結合させてアッセイを実施する。これらアッセイの何れかにおいて、NHRECの活性を調整する試験化合物は間接的に結合することが可能であり、試験化合物と直接接触する必要がない。少なくとも1つから複数の試験化合物をスクリーニングすることができる。

10

20

【0154】

別の実施例では、胚性幹細胞（ES細胞）における相同組換えを用いて動物モデル系内で、NHRECまたはその哺乳動物相同体をコードするポリヌクレオチドを「ノックアウト」する。このような技術は当技術分野において周知であり、ヒト疾患動物モデルの作製に有用である（米国特許第5,175,383号及び第5,767,337号等を参照）。例えば129/SvJ細胞株等のマウスES細胞は初期のマウス胚に由来し、培地で増殖させることができる。このES細胞は、ネオマイシンホスホトランスフェラーゼ遺伝子（neo:Capecchi, M. R. (1989) Science 244:1288-1292）等のマーカー遺伝子で破壊した目的の遺伝子を含むベクターで形質転換する。このベクターは、相同組換えにより宿主ゲノムの対応する領域に組み込まれる。別法では、Cre-loxP系を用いて相同組換えを行い、組織特異的または発生段階特異的に目的遺伝子をノックアウトする（Marth, J. D. (1996) Clin. Invest. 97:1999-2002; Wagner, K. U. 他 (1997) Nucleic Acids Res. 25:4323-4330）。形質転換したES細胞を同定し、例えばC57BL/6マウス系等から採取したマウス細胞胚盤胞に微量注入する。胚盤胞を偽妊娠メスに外科的に導入し、得られるキメラ子孫の遺伝形質を決め、これを交配させてヘテロ接合性系またはホモ接合性系を作製する。このようにして作製した遺伝子組換え動物は、潜在的な治療薬や毒性薬剤で検査することができる。

30

【0155】

NHRECをコードするポリヌクレオチドを*in vitro*でヒト胚盤胞由来のES細胞において操作することが可能である。ヒトES細胞は、内胚葉、中胚葉及び外胚葉の細胞の種類を含む少なくとも8つの別々の細胞系統に分化する可能性を有する。これらの細胞系統は、例えば神経細胞、造血系統及び心筋細胞に分化する（Thomson, J. A. 他 (1998) Science 282:1145-1147）。

40

【0156】

NHRECをコードするポリヌクレオチドを用いて、ヒト疾患をモデルとした「ノックイン」ヒト化動物（ブタ）または遺伝子組換え動物（マウスまたはラット）を作製することが可能である。ノックイン技術を用いて、NHRECをコードするポリヌクレオチドの或る領域を動物ES細胞に注入し、注入した配列を動物細胞ゲノムに組み込ませる。形質転換細胞を胞胚に注入し、胞胚を上記のように移植する。遺伝子組換え子孫または近交系に

50

ついて研究し、潜在的な医薬品を用いて処理し、ヒトの疾患の治療に関する情報を得る。別法では、例えばNHRECを乳汁内に分泌するなどNHRECを過剰に発現する哺乳動物近交系は、便利なタンパク質源となり得る（J a n n e , J . 他（1998）B i o t e c h n o l . A n n u . R e v . 4 : 5 5 - 7 4）。

【0157】

（治療）

NHRECのある領域と核ホルモン受容体のある領域との間に、例えば配列及びモチーフの文脈における化学的及び構造的類似性が存在する。更に、NHRECの発現は、消化管疾患および胃組織に密接に関連する。従って、NHRECは、代謝障害、内分泌障害、生殖障害、発生または発達障害、免疫疾患、細胞増殖異常においてある役割を果たすと考えられる。NHRECの発現若しくは活性の増大に関連する疾患の治療においては、NHRECの発現または活性を低下させることが望ましい。また、NHRECの発現または活性の低下に関連する疾患の治療においては、NHRECの発現または活性を増大させることが望ましい。

10

【0158】

従って、一実施例において、NHRECの発現または活性の低下に関連した疾患の治療または予防のために、患者にNHRECまたはその断片や誘導体を投与することが可能である。限定するものではないが、このような疾患には代謝障害、内分泌障害、生殖障害、発生または発達障害、免疫疾患、細胞増殖異常がふくまれ、代謝障害の中には、副腎機能不全、脳髄黄色腫症、副腎皮質過形成、クマリン耐性、嚢胞性線維症、糖尿病、脂肪性肝硬変、果糖-1, 6-ジホスファターゼ欠損症、ガラクトース血症、甲状腺腫、グルカゴノーマ、糖原病、遺伝性果糖不耐症、アドレナリン過剰症、腎臓不全症、上皮小体亢進症、副甲状腺低下症、高コレステロール血症、甲状腺亢進症、低血糖症、甲状腺低下症、高脂血症、脂質ミオパシー（*lipid myopathies*）、脂肪異常栄養症、リソソーム蓄積症、マンノシドーシス、ノイラミニダーゼ欠損症、肥満症、ペントースフェニルケトン尿症、プソイドビタミンD欠損症が含まれ、内分泌障害の中には原発脳腫瘍及び腺腫、妊娠性梗塞、下垂体切除、動脈瘤、血管奇形、血栓症、感染症、免疫異常、頭部外傷による合併症などの病変から起こる視床下部及び下垂体の障害と、性機能低下及びシーハン症候群、尿崩症、カルマン病、ハンド-シュラー-クリスチャン病、レトラ-シヴェ病、サルコイドーシス、エンブティセラ症候群、小人症を含む下垂体低下に関連した障害と、良性線種によって発生しやすい不适当抗利尿ホルモン（ADH）分泌症候群（SIADH）及び先端巨大症、巨人症を含む下垂体亢進に関連した障害と、甲状腺腫及び粘液水腫、細菌感染性急性甲状腺炎、ウイルス感染性亜急性甲状腺炎、自己免疫性甲状腺炎（橋本病）、クレチン病を含む甲状腺機能低下症に関連した障害と、甲状腺中毒症及びその様々な型、グレーブス病、前脛骨粘液水腫、中毒性多結節性甲状腺腫、甲状腺癌、プランマー病を含む甲状腺機能亢進症と、Conn病（*chronic hypercalcemia*）を含む副甲状腺機能亢進症と、I型及びII型糖尿病及び合併症などの膵臓疾患と、過形成及び副腎皮質の癌腫や腺腫、アルカローシスに関連した高血圧、アミロイド症、低カリウム血、クッシング病、リドル症候群、Arnold-Healy-Gordon症候群、褐色細胞腫瘍、副腎機能不全などの副腎に関連した障害と、女性の異常プロラクチン産生及び不妊症、子宮内膜症、月経周期の摂動、多嚢胞性卵巣疾患、高プロラクチン血症、選択的性腺刺激ホルモン不全（*isolated gonadotropin deficiency*）、無月経、乳汁漏出症、半陰陽、多毛症及び男性化、乳癌、閉経期後の骨粗鬆症、男性のライジッヒ細胞過形成、男性更年期、生殖細胞無形成症、ライジッヒ細胞腫瘍に関連した性機能亢進、アンドロゲン受容体の欠如に関連したアンドロゲン耐性、5 α -還元酵素症候群、女性乳房症などの生殖腺ステロイドホルモンに関連した疾患とが含まれ、生殖障害の中には、プロラクチンの産生異常と、卵管病及び排卵異常、子宮内膜症、発情期異常、月経周期異常、多嚢胞卵巣症候群、卵巣過剰刺激症候群、子宮内膜癌及び卵巣癌、子宮筋腫、自己免疫異常、異所性妊娠、奇形発生を含む不妊症と、乳癌、線維嚢胞性乳腺症、乳漏症と、精子形成異常、生理学上の精子異常、精巣癌、前立腺癌、良性の

20

30

40

50

前立腺過形成、前立腺炎、ペーロニー病、インポテンス、男性乳房及び女性化乳房癌とが含まれ、発生または発達障害の中には尿細管性アシドーシス、貧血、クッシング症候群、軟骨形成不全性小人症、デュシェンヌ-ベッカー型筋ジストロフィー、癲癇、性腺形成異常、WAGR症候群（ウィルムス腫瘍、無虹彩症、尿生殖器異常、精神薄弱）、スミス-マジェニス症候群（Smith-Magenis syndrome）、脊髄形成異常症候群、遺伝性粘膜上皮異形成、遺伝性角皮症、シャルコー-マリー-トゥース病及び神経線維腫症などの遺伝性神経病、甲状腺機能低下症、水頭症、Syndenham 舞蹈病（Syndenham's chorea）及び脳性小児麻痺などの発作障害、脊髄二分裂、無脳症、頭蓋脊椎披裂、先天性緑内障、白内障、感覚神経性聴力損失が含まれ、免疫疾患の中には、炎症及び日光性角化症、後天性免疫不全症候群（AIDS）及び副腎機能不全、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化症、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫性多腺性内分泌カンジダ性外胚葉ジストロフィー（APCED）、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球毒素性一時性リンパ球減少症、赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性大腸症候群、多発性硬化症、重症筋無力症、心筋または心膜炎、骨関節炎、骨粗しょう症、脾炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性硬化症、原発性血小板血症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症、外傷が含まれ、細胞増殖異常の中には日光性角化症及びアテローム性動脈硬化、滑液包炎、硬変、肝炎、混合型結合組織病（MCTD）、骨髄線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症、並びに腺癌及び白血病、リンパ腫、黒色腫、骨髄腫、肉腫、及び奇形癌、具体的には、副腎、膀胱、骨、骨髄、脳、乳房、頸部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、脾臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、甲状腺、子宮の癌が含まれる。

10

20

30

40

50

【0159】

別の実施例では、限定するものではないが上に列記した疾患を含むNHRECの発現または活性の低下に関連した疾患の治療または予防のために、NHRECまたはその断片や誘導体を発現し得るベクターを患者に投与することも可能である。

【0160】

更に別の実施例では、限定するものではないが上に列記した疾患を含むNHRECの発現または活性の低下に関連した疾患の治療または予防のために、実質的に精製されたNHRECを含む組成物を好適な医薬用担体と共に患者に投与することも可能である。

【0161】

更に別の実施例では、限定するものではないが上に列記した疾患を含むNHRECの発現または活性の低下に関連した疾患の治療または予防のために、NHRECの活性を調節するアゴニストを患者に投与することも可能である。

【0162】

更なる実施例では、NHRECの発現または活性の増大に関連した疾患の治療または予防のために、患者にNHRECのアンタゴニストを投与することが可能である。限定するものではないが、このような疾患の例には、上記した代謝障害、内分泌障害、生殖障害、発生または発達障害、免疫疾患、細胞増殖異常が含まれる。一実施態様では、NHRECと特異的に結合する抗体が直接アンタゴニストとして、或いはNHRECを発現する細胞または組織に薬剤を運ぶターゲティング或いは運搬機構として間接的に用いられ得る。

【0163】

別の実施例では、限定するものではないが上に列記した疾患を含むNHRECの発現または活性の増大に関連した疾患の治療または予防のために、NHRECをコードするポリヌクレオチドの相補配列を発現するベクターを患者に投与することも可能である。

【0164】

別の実施例では、本発明の任意のタンパク質、アンタゴニスト、抗体、アゴニスト、相補的な配列、ベクターを別の好適な治療薬と組み合わせて投与することもできる。当業者は、従来の医薬原理にしたがって併用療法で用いる好適な治療薬を選択可能である。治療薬との組み合わせにより、上に列記した種々の疾患の治療または予防に相乗効果をもたらし得る。この方法を用いて少ない量の各薬剤で医薬効果をあげることが可能であり、広範囲な副作用の可能性を低減し得る。

【0165】

NHRECのアンタゴニストは、当分野で一般的な方法を用いて製造することが可能である。詳しくは、精製されたNHRECを用いて抗体を作ったり、治療薬のライブラリをスクリーニングしてNHRECと特異的に結合するものを同定が可能である。NHRECの抗体も、当分野で一般的な方法を用いて製造することが可能である。このような抗体には、ポリクローナル抗体、モノクローナル抗体、キメラ抗体、一本鎖、Fabフラグメント、及びFab発現ライブラリによって作られたフラグメントが含まれる。但し、これらに限定されるものではない。治療用には、中和抗体（即ち、二量体の形成を阻害するもの）が特に好ましい。

10

【0166】

抗体の産生のためには、ヤギ、ウサギ、ラット、マウス、ヒト及びその他のものを含む種々の宿主が、NHRECまたは任意の断片、または免疫原性の特性を備えるそのオリゴペプチドの注入によって免疫化され得る。宿主の種に応じて、種々のアジュバントを用いて免疫応答を高めることもできる。このようなアジュバントにはフロイントアジュバント、水酸化アルミニウムなどのミネラルゲルアジュバント、リゾレシチン、プルロニックポリオール、ポリアニオン、ペプチド、油性乳剤、キーホールリンペットヘモシニアン、及びジニトロフェノールなどの界面活性剤が含まれるが、これらに限定されるものではない。ヒトに用いられるアジュバントの中では、BCG (*bacilli Calmette-Guerin*) 及び *Corynebacterium parvum* が特に好ましい。

20

【0167】

NHRECに対する抗体を誘発するために用いられるオリゴペプチド、ペプチド、または断片は、少なくとも約5個のアミノ酸からなり、一般的には約10個以上のアミノ酸からなるものが好ましい。これらのオリゴペプチド或いはペプチド、またはそれらの断片は、天然のタンパク質のアミノ酸配列の一部と同一であることが望ましい。NHRECアミノ酸の短いストレッチは、KLHなどの別のタンパク質の配列と融合し、キメラ分子に対する抗体が産生され得る。

30

【0168】

NHRECに対するモノクローナル抗体は、培地内の連続した細胞株によって、抗体分子を産生する任意の技術を用いて作製することが可能である。これらの技術には、ハイブリドーマ技術、ヒトB細胞ハイブリドーマ技術、及びEBV-ハイブリドーマ技術が含まれるが、これらに限定されるものではない（例えば、Kohler, G. ら. (1975) *Nature* 256:495-497; Kozbor, D. ら. (1985) *J. Immunol. Methods* 81-8-42; Cote, R. J. ら. (1983) *Proc. Natl. Acad. Sci.* 80:2026-2030; Cole, S. P. ら. (1984) *Mol. Cell Biol.* 62:109-120を参照）。

40

【0169】

更に、「キメラ抗体」作製のために発達したヒト抗体遺伝子にマウス抗体遺伝子をスプライシングするなどの技術が、好適な抗原特異性及び生物学的活性を備える分子を得るために用いられる（例えば、Morrison, S. L. 他. (1984) *Proc. Natl. Acad. Sci.* 81-4851-4855; Neuberger, M. S. 他. (1984) *Nature* 312:604-608; Takeda, S. ら. (1985) *Nature* 314:452, 454を参照）。別法では、当分

50

野で周知の方法を用いて、一本鎖抗体の産生のための記載された技術を適用して、NHRE C 特異性一本鎖抗体を生成する。関連する特異性を備えるが別のイディオタイプの組成の抗体は、ランダムな組み合わせの免疫グロブリンライブラリから鎖混合によって生成することもできる（例えば、Burton D. R. (1991) Proc. Natl. Acad. Sci. 88:11120-3を参照）。

【0170】

抗体は、リンパ球集団の中の *in vivo* 産生を誘発することによって、または免疫グロブリンライブラリのスクリーニングまたは文献に示されているような、高度に特異的な結合試薬のパネルをスクリーニングすることによって、得ることもできる（例えば、Orlandi, R. 他. (1989) Proc. Natl. Acad. Sci. 86:3833-3837; Winter, G. 他. (1991) Nature 349:293-299を参照）。

【0171】

NHRE C に対する特異的な結合部位を含む抗体も得ることができる。例えば、このような断片には、抗体分子のペプシン消化によって生成される F(ab') に断片と、F(ab') に断片のジスルフィド架橋を減じることによって生成される Fab 断片が含まれるが、これらに限定されるものではない。別法では、Fab 発現ライブラリを作製することによって、所望の特異性とモノクローナル Fab 断片の迅速且つ容易な同定が可能となる（例えば、Huse, W. D. 他. (1989) Science 254:1275-1281を参照）。

【0172】

種々のイムノアッセイを用いてスクリーニングし、所望の特異性を有する抗体を同定する。隔離された特異性を有するポリクローナル抗体またはモノクローナル抗体の何れかを用いる競合的な結合、または免疫放射線活性のための数々のプロトコルが、当分野では周知である。通常このようなイムノアッセイには、NHRE C とその特異性抗体との間の複合体調整の計測が含まれる。二つの非干渉性 NHRE C エピトープに対して反応性のモノクローナル抗体を用いる、2 部位モノクローナルベースのイムノアッセイが一般に利用されるが、競合的結合アッセイも利用することができる（Pound、前出）。

【0173】

ラジオイムノアッセイ技術と共に Scatchard 分析などの様々な方法を用いて、NHRE C に対する抗体の親和性を評価する。親和性を結合定数 K_a で表すが、この K_a は、平衡状態の下で NHRE C 抗体複合体のモル濃度を遊離抗体と遊離抗原のモル濃度で除して得られる値である。多数の NHRE C エピトープに対して親和性が不均一なポリクローナル抗体医薬の K_a は、NHRE C に対する抗体の平均親和性または結合活性を表す。特定の NHRE C エピトープに単一特異的なモノクローナル抗体医薬の K_a は、親和性の真の測定値を表す。 K_a 値が $10^9 \sim 10^{12} \text{ L/mol}$ の高親和性抗体医薬は、NHRE C 抗体複合体が激しい操作に耐えなければならないイムノアッセイに用いるのが好ましい。 K_a 値が $10^6 \sim 10^7 \text{ L/mol}$ の低親和性抗体医薬は、NHRE C が抗体から最終的に活性化状態で解離する必要がある免疫精製 (immunopurification) 及び類似の処理に用いるのが好ましい。（Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington, DC; Liddell, J. E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY）。

【0174】

ある下流での適用におけるこのような医薬品の品質及び適性を調べるために、ポリクローナル抗体医薬の抗体価及び結合活性を更に評価する。例えば、少なくとも $1 \sim 2 \text{ mg/ml}$ の特異的な抗体、好ましくは $5 \sim 10 \text{ mg/ml}$ の特異的な抗体を含むポリクローナル抗体医薬は一般に、NHRE C 抗体複合体を沈殿させなければならない処理に用いられる

。様々な適用例における抗体の特異性及び抗体価、結合活性、抗体の品質や使用法の指針は一般に入手可能である。（例えば、C a t t y, 前出, 及びC o l l i g a n 他、前出を参照）。

【0175】

本発明の別の実施例では、NHRECをコードするポリヌクレオチド、またはその任意の断片や相補配列が、治療目的で 사용할 ことができる。ある実施態様では、NHRECをコードする遺伝子のコーディング領域や調節領域に相補的な配列やアンチセンス分子（DNA及びRNA、修飾ヌクレオチド）を設計して遺伝子発現を変更することができる。このような技術は当分野では周知であり、センスまたはアンチセンスオリゴヌクレオチドまたは大きな断片が、NHRECをコードする配列の制御領域から、またはコード領域に沿ったさまざまな位置から設計可能である。 10

【0176】

治療に用いる場合、アンチセンス配列を好適な標的細胞に導入するのに好適な任意の遺伝子送達系を用いることができる。アンチセンス配列は、転写時に標的タンパク質をコードする細胞配列の少なくとも一部に相補的な配列を発現する発現プラスミドの形で細胞内に送達することができる（例えば、S l a t e r, J. E. 他（1998）J. A l l e r g y C l i n. I m m u n o l. 102（3）：469-475; and S c a n l o n, K. J. 他（1995）9（13）：1288-1296. を参照）。また、アンチセンス配列は、例えばレトロウイルスやアデノ関連ウイルスベクター等のウイルスベクターを用いて細胞内に導入することもできる（例えば、M i l l e r, A. D.（1990）B l o o d 76：271; A u s u b e l, 前出; U c k e r t, W. and W. W a l t h e r（1994）P h a r m a c o l. T h e r. 63（3）：323-347を参照）。その他の遺伝子送達機構には、リポソーム系、人工的なウイルスエンベロップ、及び当分野で周知のその他の系が含まれる（R o s s i, J. J.（1995）B r. M e d. B u l l. 51（1）：217-225; B o a d o, R. J. 他（1998）J. P h a r m. S c i. 87（11）：1308-1315; and M o r r i s, M. C. 他（1997）N u c l e i c A c i d s R e s. 25（14）：2730-2736. を参照）。 20

【0177】

本発明の別の実施例では、NHRECをコードするポリヌクレオチドを、体細胞若しくは生殖細胞の遺伝子治療に用いることが可能である。遺伝子治療は、（i）遺伝子欠損症（例えば、X染色体連鎖遺伝（C a v a z z a n a-C a l v o, M. 他（2000）S c i e n c e 288：669-672）によって特徴づけられる重度の複合型免疫欠損（SCID）-X1）、遺伝性アデノシンデアミナーゼ（ADA）欠損症（B l a e s e, R. M. 他（1995）S c i e n c e 270：475-480; B o r d i g n o n, C. 他（1995）S c i e n c e 270：470-475）に関連する重度の複合型免疫欠損、嚢胞性繊維症（Z a b n e r, J. 他（1993）C e l l 75：207-216; C r y s t a l, R. G. 他（1995）H u m. G e n e T h e r a p y 6：643-666; C r y s t a l, R. G. 他（1995）H u m. G e n e T h e r a p y 6：667-703）、サラセミア（t h a l a s s a m i a）、家族性高コレステロール血症、第V I I I 因子若しくは第I X 因子欠損による血友病（C r y s t a l, R. G.（1995）S c i e n c e 270：404-410; V e r m a, I. M. and S o m i a, N.（1997）N a t u r e 389：239-242）を治療したり、（ii）条件的致死性遺伝子産物（例えば、細胞増殖の制御不能による癌の場合）を発現させたり、及び（i i i）細胞内の寄生虫（例えば、ヒト免疫不全ウイルス（HIV）（B a l t i m o r e, D.（1988）N a t u r e 335：395-396; P o e s c h l a, E. 他（1996）P r o c. N a t l. A c a d. S c i. U S A. 93：11395-11399）や、B型若しくはC型肝炎ウイルス（HBV、HCV）、C a n d i d a a l b i c a n s及びP a r a c o c c i d i o i d e s b r a s i l i e n s i 30 40 50

s等の真菌寄生虫、Plasmodium falciparum及びTrypanosoma cruzi等の原虫寄生体)に対する防御機能を有するタンパク質を発現させて行うことができる。NHRECの発現若しくは調節に必要な遺伝子の欠損が疾患を引き起こす場合、導入した細胞の好適な集団からNHRECを発現させて、遺伝子欠損によって起こる症状の発現を緩和することが可能である。

【0178】

本発明の更なる実施例では、NHRECの欠損による疾患や異常症は、NHRECをコードする哺乳動物発現ベクターを作製して、これらのベクターを機械的手段によってNHREC欠損細胞に導入することによって治療する。in vivo或いはex vitroの細胞に用いる機械的な導入技術には、(i)個々の細胞内へのDNAのマイクロインジェクション、(ii)金粒子の打ち込み、(iii)リポソーム仲介性トランスフェクション、(iv)受容体仲介性遺伝子導入、及び(v)DNAトランスポソン(Morgan, R. A. and W. F. Anderson (1993) Annu. Rev. Biochem. 62:191-217; Ivics, Z. (1997) Cell 91:501-510; Boulay, J-L. and H. Recipon (1998) Curr. Opin. Biotechnol. 9:445-450)の使用が含まれる。

【0179】

NHRECの発現に影響を及ぼし得る発現ベクターには、限定するものではないが、PCDNA 3.1、EPITAG、PRCCMV2、PREP、PVAXベクター(Invitrogen, Carlsbad CA)、PCMV-SCRIPT、PCMV-TAG、PEGSH/PERV(Stratagene, La Jolla CA)、PTE-T-OFF、PTET-ON、PTRE2、PTRE2-LUC、PTK-HYG(Clontech, Palo Alto CA)が含まれる。NHRECを発現させるために、(i)恒常的に活性なプロモーター(例えば、サイトメガロウイルス(CMV)、ラウス肉腫ウイルス(RSV)、SV40ウイルス、チミジンキナーゼ(TK)、若しくはβ-アクトチン遺伝子等)、(ii)誘導性プロモーター(例えば、市販されているT-R-EXプラスミド(Invitrogen)に含まれている、テトラサイクリン調節性プロモーター(Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. U. S. A. 89:5547-5551; Gossen, M. 他 (1995) Science 268:1766-1769; Rossi, F. M. V. and H. M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456))、エクジソン誘導性プロモーター(市販されているプラスミドPVGRXR及びPINDに含まれている:Invitrogen)、FK506/ラパマイシン誘導性プロモーター、またはRU486/ミフェプリストン誘導性プロモーター(Rossi, F. M. V. and H. M. Blau, 前出)、または(iii)正常な個体に由来するNHRECをコードする内在性遺伝子の天然のプロモーター若しくは組織特異的プロモーターを用いることが可能である。

【0180】

市販のリポソーム形質転換キット(例えば、Invitrogenが販売しているPERFECT LIPID及びTRANSFECTION KIT)を用いれば、当業者は経験にそれほど頼らないでもポリヌクレオチドを培養中の標的細胞に導入することが可能である。別法では、リン酸カルシウム法(Graham, F. L. and A. J. Eb (1973) Virology 52:456-467)若しくは電気穿孔法(Neumann, B. 他 (1982) EMBO J. 1:841-845)を用いて形質転換を行う。初代細胞にDNAを導入するためには、これらの標準的な哺乳動物トランスフェクションプロトコルを変更する必要がある。

【0181】

本発明の別の実施例では、NHRECの発現に関連する遺伝子欠損によって起こる疾患や異常症は、(i)レトロウイルス末端反復配列(LTR)プロモーター若しくは独立した

プロモーターのコントロール下でNHRECをコードするポリヌクレオチドと、(ii) 好適なRNAパッケージングシグナルと、(iii) 追加のレトロウイルス・シス作用性RNA配列及び効率的なベクターの増殖に必要なコーディング配列を伴うRev応答性エレメント(RRE)とからなるレトロウイルスベクターを作製して治療することができる。レトロウイルスベクター(例えば、PFB及びPFBNEO)はStratagene社から入手可能であり、公表データ(Riviere, I. 他. (1995) Proc. Natl. Acad. Sci. U. S. A. 92:6733-6737)に基づいている。上記データを引用することをもって本明細書の一部とする。このベクターは、VSVg(Armentano, D. 他 (1987) J. Virol. 61:1647-1650; Bender, M. A. 他 (1987) J. Virol. 61:1639-1646; Adam, M. A. and A. D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. 他 (1998) J. Virol. 72:8463-8471; Zufferey, R. 他 (1998) J. Virol. 72:9873-9880)等の乱交雑エンベロープタンパク質若しくは標的細胞上の受容体に対する親和性を有するエンベロープ遺伝子を発現する好適なベクター産生細胞系(VPCL)において増殖される。RIGGに付与された米国特許第5,910,434号(「Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant」)において、レトロウイルスパッケージング細胞系を得るための方法が開示されており、引用することをもって本明細書の一部とする。レトロウイルスベクターの増殖、ある細胞集団(例えば、CD4⁺T細胞)の形質導入、並びに形質導入した細胞を患者に戻す方法は、遺伝子治療の分野では周知であり、多数の文献に記載されている(Ranga, U. 他. (1997) J. Virol. 71:7020-7029; Bauer, G. 他 (1997) Blood 89:2259-2267; Bonyhadi, M. L. (1997) J. Virol. 71:4707-4716; Ranga, U. 他 (1998) Proc. Natl. Acad. Sci. U. S. A. 95:1201-1206; Su, L. (1997) Blood 89:2283-2290)。

【0182】

別法では、アデノウイルス系遺伝子治療の送達系を用いて、NHRECの発現に関連する1或いは複数の遺伝子異常を有する細胞にNHRECをコードするポリヌクレオチドを送達する。アデノウイルス系ベクターの作製及びパッケージングは当分野では周知である。複製欠損型アデノウイルスベクターは、免疫調節タンパク質をコードする遺伝子を脾臓の無損傷の脾島の中に導入するために可変性であることが証明された(Csete, M. E. 他. (1995) Transplantation 27:263-268)。使用できる可能性のあるアデノウイルスベクターが、米国特許第5,707,618号(「Adenovirus vectors for gene therapy」)に記載されており、引用することをもって本明細書の一部とする。アデノウイルスベクターについてはまた、Antinuzzi, P. A. 他 (1999) Annu. Rev. Nutr. 19:511-544; and Verma, I. M. and N. Somia (1997) Nature 18:389:239-242を参照し、引用することをもって本明細書の一部とする。

【0183】

別法では、ヘルペス系遺伝子治療の送達系を用いて、NHRECの発現に関連する1或いは複数の遺伝子異常を有する標的細胞にNHRECをコードするポリヌクレオチドを送達する。単純疱疹ウイルス(HSV)系のベクターは、HSV親和性の中枢神経細胞にNHRECを導入する際に特に重要である。ヘルペス系ベクターの作製及びパッケージングは当分野では周知である。複製適格性の単純疱疹ウイルス(HSV)I型系のベクターは、霊長類の眼にレポーター遺伝子を送達するために用いられてきた(Liu, X. 他 (

1999) Exp. Eye Res. 169:385-395)。HSV-1ウイルスベクターの作製は、De Lucaに付与された米国特許第5,804,413号(Herpes simplex virus swains for gene transfer)に記載されており、引用することをもって本明細書の一部とする。米国特許第5,804,413号には、ヒト遺伝子治療を含む目的のために、好適なプロモーターのコントロールの下で、細胞に導入される少なくとも1つの内在性遺伝子を含むゲノムからなる組換えHSV d92についての記載がある。また上記特許には、ICP4、ICP27及びICP22のために除去される組換えHSV株の作製及び使用方法が開示されている。HSVベクターについては、Goins, W. F. 他(1999) J. Virol. 73:519-532 and Xu, H. 他(1994) Dev. Biol. 163:152-161を参照し、引用することをもって本明細書の一部とする。クローニングされたヘルペスウイルス配列の操作や、巨大ヘルペスウイルスのゲノムの異なった部分を含む多数のプラスミドをトランスフェクトした後の組換えウイルスの継代、ヘルペスウイルスの成長及び増殖、並びにヘルペスウイルスの細胞への感染は当分野で周知の技術である。

【0184】

別法では、αウイルス(正の一本鎖RNAウイルス)ベクターを用いてNHRECをコードするポリヌクレオチドを標的細胞に送達する。プロトタイプのαウイルスであるセムリキ森林熱ウイルス(Semliki Forest Virus, SFV)の生物学的な研究が広範に行われ、遺伝子伝達ベクター(gene transfer vector)がSFVゲノムに基づいていることが分かった(Garoff, H. and K.-J. Li(1998) Cunn. Opin. Biotech. 9:464-469)。αウイルスRNAの複製中に、通常はウイルスカプシドタンパク質をコードするサブゲノムRNAが作り出される。このサブゲノムRNAが完全長のゲノムRNAより高いレベルで複製されるため、酵素活性(例えばプロテアーゼ及びポリメラーゼ)を有するウイルスタンパク質に対してカプシドタンパク質が過剰に産生される。同様に、NHRECをコードする配列をαウイルスゲノムのカプシドをコードする領域に導入することによって、ベクター導入細胞において多数のNHRECをコードするRNAが産生され、高いレベルでNHRECが合成される。通常はαウイルス感染は2〜3日以内の細胞溶解に関係するが、シンドビスウイルス(SIN)の変異体を有するハムスターの正常な腎細胞(BHK-21)の持続的な感染を確立する能力は、αウイルスの溶解性の複製が遺伝子治療に適用できるように好適に変更することが可能であることを示唆している(Dryga, S. A. 他.(1997) Virology 228:74-83)。様々な宿主にαウイルスを導入できることから、様々なタイプの細胞にNHRECを導入することできる。ある集団における細胞のサブセットの特定の形質導入には、形質導入する前に細胞のソーティングを必要とする場合がある。αウイルスの感染性cDNAクローンの操作、αウイルスcDNA及びRNAのトランスフェクション、並びにαウイルスの感染方法は当分野で周知である。

【0185】

例えば開始部位から約-10から約+10までの転写開始部位に由来するオリゴヌクレオチドを用いて、遺伝子の発現を阻害することが可能である。同様に、三重らせん塩基対合法を用いて阻害することができ。三重らせん構造は、二重らせんがポリメラーゼ、転写因子、または調節分子の結合のために十分に広がるのを阻止するため有用である。三重式DNAを用いる最近の治療の進歩は文献に記載されている(例えば、Gee, J. E. ら.(1994) In: Huber, B. E. 及び B. I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, NYを参照)。相補的な配列またはアンチセンス分子もまた、転写物がリボソームに結合するのを阻止することによってmRNAの翻訳を阻止するように設計できる。

【0186】

酵素性RNA分子であるリボザイムは、RNAの特異的切断を触媒するために用いることができる。リボザイム作用の機構には、相補的な標的RNAへのリボザイム分子の配列特異性ハイブリダイゼーションが含まれ、ヌクレオチド鎖切断が続く。例えば、NHRECをコードする配列のヌクレオチド鎖切断を、特異的且つ効果的に触媒する組換え型のハンマーヘッド型リボザイム分子が含まれる。

【0187】

任意の潜在的RNA標的内の特異的なリボザイム切断部位が、後続の配列GUA、GUU、及びGUCを含むリボザイム切断部位に対して、標的分子をスキニングすることによって初めに同定される。一度同定されると、切断部位を含む標的遺伝子の領域に対応する15個から20個のリボヌクレオチドの短いRNA配列を、オリゴヌクレオチドの機能を不全にする二次的な構造の特徴について評価することが可能である。候補標的の適合性も、リボヌクレアーゼ保護アッセイを用いて、相補的なオリゴヌクレオチドとのハイブリダイゼーションの容易性をテストすることによって評価することが可能である。

10

【0188】

本発明の相補的なリボ核酸分子及びリボザイムは、当分野で周知の方法を用いて、核酸分子の合成のために作製することができる。これらの方法には、固相ホスホラミダイト化合物などのオリゴヌクレオチドを化学的に合成する方法が含まれる。別法では、RNA分子が*in vitro*及び*in vivo*でNHRECをコードするDNA配列の転写によって生成され得る、このようなDNA配列はT7またはSP6等の好適なRNAポリメラーゼプロモータを用いて、種々のベクターの中に組み入れることが可能である。別法では、相補的なRNAを構成的または誘導的に合成するこれらのcDNA作製物は、細胞株、細胞、または組織の中に導入することができる。

20

【0189】

RNA分子を修飾することによって、細胞内の安定性を高め、半減期を長くすることができる。可能な修飾には、分子の5'及び/または3'端部でのフランキング配列の追加、または分子のバックボーン内のホスホジエステル結合の代わりにホスホロチオネートまたは2' Oメチルを用いる修飾が含まれるが、これらに限定されるものではない。PNAの生成に固有のこの概念は、内在性のエンドヌクレアーゼによって容易には認識されないアデニン、シチジン、グアニン、チミン、及びウリジンのアセチル、メチル、チオール、及び同様の修飾形態だけでなく、イノシン、キューエオシン (queosine)、及びワイブトシン (wybutosine) などの従来のものでない塩基を含めることによって、これらの分子の全体に拡大することができる。

30

【0190】

本発明の更なる実施例は、NHRECをコードするポリヌクレオチドの発現の変化に有効な化合物をスクリーニングする方法を含む。特定のポリヌクレオチドの発現の変化に有効な化合物には、限定するものではないが、特定のポリヌクレオチド配列と相互作用可能な非高分子化学物質、オリゴヌクレオチド、アンチセンスオリゴヌクレオチド、三重らせん形成オリゴヌクレオチド、転写因子やその他のポリペプチド転写調節因子が含まれる。有効な化合物は、ポリヌクレオチド発現のインヒビター或いはエンハンサーとして作用し、ポリヌクレオチドの発現を変化させ得る。従って、NHRECの発現または活性の増加に関連する疾患の治療においては、NHRECをコードするポリヌクレオチドの発現を特異的に阻害する化合物が治療上有用であり、NHRECの発現または活性の低下に関連する疾患の治療においては、NHRECをコードするポリヌクレオチドの発現を特異的に促進する化合物が治療上有用であり得る。

40

【0191】

特定のポリヌクレオチドの発現の変化の有効性を調べるために、少なくとも1個から複数個の試験化合物をスクリーニングすることができる。試験化合物は、有効な化合物の化学修飾を含む当分野で周知の任意の方法で得ることができる。このような方法は、ポリヌクレオチドの発現を変化させる場合、一般に市販されている或いは専売の天然または非天然の化合物ライブラリから選択する場合、標的ポリヌクレオチドの化学的及び/または構造

50

的特性に基づいて化合物を合理的にデザインする場合、更に組合せ的にまたは無作為に生成した化合物のライブラリから選択する場合に有効である。NHRECをコードするポリヌクレオチドを含むサンプルは、少なくとも1つの試験化合物に曝露して得る。サンプルには、例えば無傷細胞、透過化処理した細胞、in vitro細胞遊離系または再構成生化学系が含まれ得る。NHRECをコードするポリヌクレオチドの発現における変化は、当分野で周知の任意の方法でアッセイする。通常、NHRECをコードするポリヌクレオチドの配列に相補的なヌクレオチド配列を有するプローブを用いたハイブリダイゼーションにより、特定のヌクレオチドの発現を検出する。ハイブリダイゼーションの収量を定量し、その値が1或いは複数の試験化合物に曝露される及び曝露されないポリヌクレオチドの発現の比較における基準となり得る。試験化合物に曝露されるポリヌクレオチドの発現の変化が検出される場合は、ポリヌクレオチドの発現の変化に試験化合物が有効であることを示している。特定のポリヌクレオチドの発現の変化に有効な化合物を調べるために、例えば Schizosaccharomyces pombe 遺伝子発現系 (Atkins, D. 他 (1999) 米国特許第5,932,435号、Arndt, G. M. 他 (2000) *Nucleic Acids Res.* 28:E15) または HeLa 細胞等のヒト細胞株 (Clarke, M. L. 他 (2000) *Biochem. Biophys. Res. Commun.* 268:8-13) を用いてスクリーニングする。本発明の特定の実施例は、特異的ポリヌクレオチド配列に対するアンチセンス活性を調べるための、各オリゴヌクレオチド (デオキシリボヌクレオチド、リボヌクレオチド、ペプチド核酸、及び修飾オリゴヌクレオチド) の組み合わせライブラリのスクリーニングを含む (Bruice, T. W. 他 (1997) 米国特許第5,686,242号、Bruice, T. W. 他 (2000) 米国特許第6,022,691号)。

【0192】

ベクターを細胞または組織に導入する多数の方法が利用でき、in vivo、in vitro、及び ex vivoでの使用に等しく適している。ex vivoでの治療の場合、患者から採取された肝細胞の中にベクターを導入して、自家移植で同じ患者に戻すためにクローニング増殖される。トランスフェクション、リボソーム注入またはポリカチオンアミノポリマーによる運搬は、当分野で周知の方法を用いて実行することができる (例えば、Goldman, C. K. 他. (1997) *Nature Biotechnology* 15:462-66:を参照)。

【0193】

上記したいかなる治療方法も、例えば、ヒト、イヌ、ネコ、ウシ、ウマ、ウサギ及びサルなどの哺乳動物を含む、治療が必要な全ての被験者に適用できる。

【0194】

本発明の別の実施例は、上記した全ての治療効果のために、医学上認められる担体と共に医薬品或いは無菌組成物の投与に関連する。このような組成物は、NHREC、NHRECの抗体、擬態、アゴニスト、アンタゴニスト、またはNHRECのインヒビターなどからなる。この組成物は、単体で、或いは安定剤などの1種類以上の別の薬剤と共に、無菌の生体適合性医薬品担体に投与することができる。このような医薬品担体には、生理食塩水、緩衝食塩水、ブドウ糖、及び水などが含まれるがこれらに限定されるものではない。この組成物は、単独或いは薬物またはホルモンなどの別の薬剤と共に投与することができる。

【0195】

本発明に用いられる組成物は、様々な経路を用いて投与するが可能である。この経路には、経口、静脈内、筋肉内、動脈内、骨髓内、クモ膜下、心室内、経皮、皮下、腹腔内、鼻腔内、腸内、局所、舌下、または直腸が含まれるがこれらに限定されるものではない。

【0196】

肺投与用の組成物は、液状または乾燥粉末状に調製することができる。このような組成物は通常、患者が吸入する直前にエアロゾル化する。小分子 (例えば、従来の低分子量有機薬剤) の場合には、速効製剤のエアロゾル輸送が当分野で周知である。高分子 (例えばよ

り大きなペプチドやタンパク質)の場合には、肺の肺胞領域を介する肺輸送の技術が近年向上したため、インスリン等の薬剤を実際に血中に輸送することが可能となった(Patterson, J. S. 他, 米国特許第5,997,848号等を参照)。肺輸送は、針注射を用いなくて投与できるという点で優れており、潜在的に有毒な浸透エンハンサーが必要でなくなる。

【0197】

本発明に用いる好適な組成物には、目的を達成するため、効果的な量の活性処方成分を含む組成物が含まれる。当業者は、十分に自身の能力で効果的な服用量を定めることができる。

【0198】

NHRECまたはその断片を含む高分子を直接細胞内に輸送するべく、特殊な形態に組成物が調製されるのが好ましい。例えば、細胞不透過性高分子を含むリポソーム製剤は、細胞融合及び高分子の細胞内輸送を促進し得る。別法では、NHRECまたはその断片をHIV Tat-1タンパク質の陽イオンN末端部に結合することもできる。このようにして作製された融合タンパク質は、マウスモデル系の脳を含む全ての組織の細胞に形質導入されることが確認されている(Schwarze, S. R. 他 (1999) Science 285:1569-1572)。

【0199】

どのような組成物であっても、治療に効果的な薬用量は、初めは、例えば腫瘍細胞の腫瘍細胞アッセイで、或いは動物モデルのどちらかで推定することができる。通常、動物モデルには、マウス、ウサギ、イヌ、サル、またはブタなどが用いられる。動物モデルはまた、好適な濃縮範囲及び投与の経路を決めるのに用いることができる。このような治療をもとに、ヒトへの有益な薬用量及び投与経路を決定することができる。

【0200】

医学的に効果的な薬用量は、症状や容態を回復させる、たとえばNHRECまたはその断片、NHRECの抗体、NHRECのアゴニストまたはアンタゴニスト、インヒビターなどの活性処方成分の量に関連する。薬用有効度及び毒性は、たとえば、ED₅₀ (服用に対して集団の50%に医薬的效果がある用量)またはLD₅₀ (服用に対して集団の50%に致命的である用量)統計を計算するなど、細胞培養または動物実験における標準的な薬剤手法によって決定することができる。毒性効果と治療効果との薬用量比は治療指数であり、LD₅₀/ED₅₀と示すことができる。高い治療指数を示す組成物が望ましい。細胞培養アッセイ及び動物実験から得られたデータが、ヒトへの適用のために、薬用量の範囲を調剤するのに用いられる。このような組成物が含まれる薬用量は、毒性を殆ど或いは全く含まず、ED₅₀を含む血中濃度の範囲であることが望ましい。薬用量は、用いられる投与形態及び患者の感受性、投与の経路によって、この範囲内で様々である。

【0201】

正確な薬用量は、治療が必要な患者に関する要素を考慮して、実務者によって決められるであろう。薬用量及び投与は、効果的なレベルの活性成分を与えるため或いは所望の効果を維持するために調節される。薬用量の要素として考慮されるものには、疾患の重症度、患者の一般的な健康状態、年齢、体重、及び患者の性別、投与の時間及び頻度、併用する薬剤、反応感受性、及び治療に対する応答が含まれる。作用期間が長い組成物は、三日か四日に一度、一週間に一度、二週間に一度、特定の製剤の半減期及びクリアランス率によって左右され、投与され得る。

【0202】

通常、薬用量は投与の経路によって異なるが、約0.1~100,000 µgまでの最大約1グラムまでである。特定の薬用量及び運搬の方法に関するガイダンスは文献に記載されており、一般に当分野の実務者はそれを利用することができる。当業者は、タンパク質またはインヒビターとは異なったヌクレオチドの製剤を利用するであろう。同様に、ポリヌクレオチドまたはポリペプチドの運搬は、特定の細胞、状態、位置などに対して特異的であろう。

10

20

30

40

50

【0203】

(診断)

別の実施例では、NHRECに特異的に結合する抗体が、NHRECの発現によって特徴付けられる疾患の診断、またはNHRECやNHRECのアゴニストまたはアンタゴニスト、インヒビターで治療を受けている患者をモニターするためのアッセイに用いられる。診断に有用な抗体は、治療のところで記載した方法と同じ方法で製剤される。NHRECの診断アッセイには、抗体及び標識を用いてヒトの体液或いは細胞や組織から採取されたものからNHRECを検出する方法が含まれる。これらの抗体は、修飾をして或いはしないで使用され、レポーター分子の共有結合性或いは非共有結合性の接着によって標識化され得る。当分野で周知の種々のレポーター分子が用いられるが、その内の幾つかは上記した。

【0204】

NHRECを測定するためのELISA, RIA, 及びFACSを含む種々のプロトコルは、当分野では周知であり、変わった或いは異常なレベルのNHRECの発現を診断する元となるものを提供する。正常或いは標準的なNHRECの発現の値は、複合体の形成に適した条件の下、正常な哺乳動物、例えばヒトなどの被験者から採取した体液または細胞とNHRECに対する抗体とを結合させることによって決定する。標準的な複合体形成の量は、測光法 (photometric) などの種々の方法で定量され得る。被験者のNHRECの発現の量、制御及び疾患、生検組織からのサンプルが基準値と比較される。基準値と被験者との間の偏差が、診断の指標となる。

【0205】

本発明の別の実施例によれば、NHRECをコードするポリヌクレオチドを診断のために用いることもできる。用いられるポリヌクレオチドには、オリゴヌクレオチド配列、相補的なRNA及びDNA分子、及びPNAが含まれる。このポリヌクレオチドを用いて、疾患と相関し得るNHRECを発現する生検組織における遺伝子の発現を検出し定量する。この診断アッセイを用いて、NHRECの存在の有無、更に過剰な発現を調べ、治療中のNHREC値の調節を監視する。

【0206】

一実施形態では、NHRECまたは近縁の分子をコードする遺伝子配列を含むポリヌクレオチド配列を検出可能なPCRプローブを用いたハイブリダイゼーションによって、NHRECをコードする核酸配列を同定することが可能である。例えば5'調節領域である高度に特異的な領域か、例えば保存されたモチーフであるやや特異性の低い領域から作られているかのプローブの特異性と、ハイブリダイゼーション或いは増幅のストリンジェントは、プローブがNHRECをコードする自然界の配列のみを同定するかどうか、或いはアレルや関連配列コードする自然界の配列のみを同定するかどうかによって決まるであろう。

【0207】

プローブはまた、関連する配列の検出に利用され、NHRECをコードする任意の配列と少なくとも50%の配列同一性を有し得る。目的の本発明のハイブリダイゼーションプローブには、DNAあるいはRNAが可能であり、SEQ ID NO: 4-6の配列、或いはNHREC遺伝子のプロモーター、エンハンサー、イントロンを含むゲノム配列に由来し得る。

【0208】

NHRECをコードするDNAに対して特異的なハイブリダイゼーションプローブの作製方法には、NHREC及びNHREC誘導体をコードするポリヌクレオチド配列をmRNAプローブの作製のためのベクターにクローニングする方法がある。このようなベクターは市販されており、当業者には周知であり、好適なRNAポリメラーゼ及び好適な標識されたヌクレオチドを加えることによって、in vitroでRNAプローブを合成するために用いられる。ハイブリダイゼーションプローブは、例えば³²P或いは³⁵Sなどの放射性核種、或いはアビジン/ビオチン (biotin) 結合系によってプローブに結

合されたアルカリホスファターゼなどの酵素標識等の種々のレポーターの集団によって標識され得る。

【0209】

NHRECをコードするポリヌクレオチド配列を用いて、NHRECの発現に関連する疾患を診断することが可能である。限定するものではないが、このような疾患には代謝障害、内分泌障害、生殖障害、発生または発達障害、免疫疾患、細胞増殖異常がふくまれ、代謝障害の中には、副腎機能不全、脳腱黄色腫症、副腎皮質過形成、クマリン耐性、嚢胞性線維症、糖尿病、脂肪性肝硬変、果糖-1, 6-ジホスファターゼ欠損症、ガラクトース血症、甲状腺腫、グルカゴノーマ、糖原病、遺伝性果糖不耐症、アドレナリン過剰症、腎臓不全症、上皮小体亢進症、副甲状腺低下症、高コレステロール血症、甲状腺亢進症、低血糖症、甲状腺低下症、高脂血症、脂質ミオパシー (lipid myopathies)、脂肪異栄養症、リソソーム蓄積症、マンノシドーシス、ノイラミニダーゼ欠損症、肥満症、ペントースフェニルケトン尿症、プソイドビタミンD欠損症が含まれ、内分泌障害の中には原発脳腫瘍及び腺腫、妊娠性梗塞、下垂体切除、動脈瘤、血管奇形、血栓症、感染症、免疫異常、頭部外傷による合併症などの病変から起こる視床下部及び下垂体の障害と、性機能低下及びシーハン症候群、尿崩症、カルマン病、ハンド-シュラー-クリスチャン病、レトラ-シヴェ病、サルコイドーシス、エンプティセラ症候群、小人症を含む下垂体低下に関連した障害と、良性線種によって発生しやすい不适当抗利尿ホルモン (ADH) 分泌症候群 (SIADH) 及び先端巨大症、巨人症を含む下垂体亢進に関連した障害と、甲状腺腫及び粘液水腫、細菌感染性急性甲状腺炎、ウイルス感染性亜急性甲状腺炎、自己免疫性甲状腺炎 (橋本病)、クレチン病を含む甲状腺機能低下症に関連した障害と、甲状腺中毒症及びその様々な型、グレーブス病、前脛骨粘液水腫、中毒性多結節性甲状腺腫、甲状腺癌、プランマー病を含む甲状腺機能亢進症と、Conn病 (chronic hypercalcemia) を含む副甲状腺機能亢進症と、I型及びII型糖尿病及び合併症などの膵臓疾患と、過形成及び副腎皮質の癌腫や腺腫、アルカローシスに関連した高血圧、アミロイド症、低カリウム血、クッシング病、リドル症候群、Arnold-Hearley-Gordon症候群、褐色細胞腫瘍、副腎機能不全などの副腎に関連した障害と、女性の異常プロラクチン産生及び不妊症、子宮内膜症、月経周期の摂動、多嚢胞性卵巣疾患、高プロラクチン血症、選択的性腺刺激ホルモン不全 (isolated gonadotropin deficiency)、無月経、乳汁漏出症、半陰陽、多毛症及び男性化、乳癌、閉経期後の骨粗鬆症、男性のライジッヒ細胞過形成、男性更年期、生殖細胞無形成症、ライジッヒ細胞腫瘍に関連した性機能亢進、アンドロゲン受容体の欠如に関連したアンドロゲン耐性、5 α -還元酵素症候群、女性乳房症などの生殖腺ステロイドホルモンに関連した疾患とが含まれ、生殖障害の中には、プロラクチンの産生異常と、卵管病及び排卵異常、子宮内膜症、発情期異常、月経周期異常、多嚢胞卵巣症候群、卵巣過剰刺激症候群、子宮内膜癌及び卵巣癌、子宮筋腫、自己免疫異常、異所性妊娠、奇形発生を含む不妊症と、乳癌、線維嚢胞性乳腺症、乳漏症と、精子形成異常、生理学上の精子異常、精巣癌、前立腺癌、良性の前立腺過形成、前立腺炎、ペーロニー病、インポテンス、男性乳房及び女性化乳房癌とが含まれ、発生または発達障害の中には尿細管性アシドーシス、貧血、クッシング症候群、軟骨形成不全性小人症、デュシェンヌ-ベッカー型筋ジストロフィー、癲癇、性腺形成異常、WAGR症候群 (ウィルムス腫瘍、無虹彩症、尿生殖器異常、精神薄弱)、スミス-マジェニス症候群 (Smith-Magenis syndrome)、脊髄形成異常症候群、遺伝性粘膜上皮異形成、遺伝性角皮症、シャルコー-マリー-トゥス病及び神経線維腫症などの遺伝性神経病、甲状腺機能低下症、水頭症、Sydenham舞踏病 (Sydenham's chorea) 及び脳性小児麻痺などの発作障害、脊髄二分裂、無脳症、頭蓋脊椎披裂、先天性緑内障、白内障、感覚神経性聴力損失が含まれ、免疫疾患の中には、炎症及び日光性角化症、後天性免疫不全症候群 (AIDS) 及び副腎機能不全、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化症、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫性多腺性内分泌カンジダ性外胚葉ジストロフィー (APCED)、気

管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球毒素性一時性リンパ球減少症、赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性大腸症候群、多発性硬化症、重症筋無力症、心筋または心膜炎、骨関節炎、骨粗しょう症、瘰癧、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性硬化症、原発性血小板血症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症、外傷が含まれ、細胞増殖異常の中には日光性角化症及びアテローム性動脈硬化、滑液包炎、硬変、肝炎、混合型結合組織病（MCTD）、骨髄線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症、並びに腺癌及び白血病、リンパ腫、黒色腫、骨髄腫、肉腫、及び奇形癌、具体的には、副腎、膀胱、骨、骨髄、脳、乳房、頸部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、脾臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、甲状腺、子宮の癌が含まれる。MADをコードするポリヌクレオチド配列は、サザン法やノーザン法、ドットブロット法、或いはその他の膜系の技術、PCR法、ディップスティック（dipstick）、ピン（pin）、ELISA式アッセイ、及び変異NHRECの発現を検出するために患者から採取した体液或いは組織を利用するマイクロアレイに使用することが可能である。このような質的或いは量的方法は、当分野では周知である。

10

20

30

40

50

【0210】

ある実施態様では、NHRECをコードするヌクレオチド配列は、関連する疾患、特に上記した疾患を検出するアッセイにおいて有用であろう。NHRECをコードするヌクレオチド配列は、標準的な方法で標識化され、ハイブリダイゼーション複合体の形成に好適な条件の下、患者から採取した体液或いは組織のサンプルに加えることができるであろう。好適な培養期間の後、サンプルを洗浄し、シグナルを定量して基準値と比較する。患者のサンプルのシグナルの量が、制御サンプルと較べて著しく変わっている場合は、サンプル内のNHRECをコードするヌクレオチド配列の変異レベルにより、関連する疾患の存在が明らかになる。このようなアッセイを用いて、動物実験、臨床試験、或いは個人の患者の治療を監視における、特定の治療効果を推定することが可能である。

【0211】

NHRECの発現に関連する疾患の診断の基準となるものを提供するために、発現の正常すなわち標準的なプロファイルが確立される。これは、ハイブリダイゼーション或いは増幅に好適な条件の下、動物或いはヒトの何れかの正常な被験者から抽出された体液或いは細胞と、NHRECをコードする配列或いはその断片とを結合させることにより達成され得る。標準的なハイブリダイゼーションは、正常な被験者から得た値と周知の量の実質的に精製されたポリヌクレオチドが用いられる実験からの値とを比較することによって定量可能である。正常なサンプルから得た標準的な値を、疾患の症状を示す被験者から得た値と比較可能である。基準値と被験者の値との偏差を用いて罹患しているかどうかを決定する。

【0212】

疾患の存在が確定され、治療プロトコルが開始されると、ハイブリダイゼーションアッセイを通常ベースで繰り返して、被験者における発現のレベルが正常な患者に示される値に近づき始めたかどうかを推定することが可能である。繰り返し行ったアッセイの結果を、数日から数ヶ月の期間の治療の効果を見るのに用いることができる。

【0213】

癌では、個体からの生体組織における異常な量の転写物が、疾患の発生の素因を示し、また実際に臨床的症状が出る前に疾患を検出する方法を提供することが可能である。この種のより明確な診断により、医療の専門家が予防方法或いは積極的な治療法を早くから利用して、癌の発生または進行を防ぐことが可能となる。

【0214】

NHRECをコードする配列から設計されたオリゴヌクレオチドのさらなる診断への利用には、PCRの利用が含まれ得る。このようなオリゴマーは、化学的な合成、酵素を用いた生成、或いは*in vitro*で生成され得る。オリゴマーは、好ましくはNHRECをコードするポリヌクレオチドの断片、或いはNHRECをコードするポリヌクレオチドと相補的なポリヌクレオチドの断片を含み、最適な条件の下、特定の遺伝子や条件を識別するために利用される。また、オリゴマーは、やや緩いストリンジェントな条件の下、近縁のDNA或いはRNA配列の検出及び／または定量のため用いることが可能である。

【0215】

或る実施態様において、NHRECをコードするポリヌクレオチド配列由来のオリゴヌクレオチドプライマーを用いて、一塩基多型 (SNP) を検出し得る。SNPは、ヒトの先天性または後天性遺伝病の原因となる場合が多いヌクレオチドの置換、挿入及び欠失である。限定するものではないが、SNPの検出方法には、一本鎖立体構造多型 (SSCP) 及び蛍光SSCP (fSSCP) 法が含まれる。SSCPでは、NHRECをコードするポリヌクレオチド配列由来のオリゴヌクレオチドプライマーを用いたポリメラーゼ連鎖反応 (PCR) でDNAを増幅する。このDNAは、例えば病変或いは正常な組織、生検サンプル、体液等に由来し得る。このDNA内のSNPは、一本鎖形状のPCR産物の2次及び3次構造に差異を生じさせる。この差異は非変性ゲル中でのゲル電気泳動法を用いて検出可能である。fSSCPでは、オリゴヌクレオチドプライマーを蛍光標識することによって、DNAシーケンシング装置などのハイスループット機器でアンプリマー (amplimer) の検出をすることが可能になる。更に、インシリコSNP (insilico SNP: isSNP) と呼ばれる配列データベース分析法は、共通のコンセンサス配列の構築に用いられる個々の重複するDNA断片の配列を比較することによって、多型を同定することができる。これらのコンピュータベースの方法は、DNA配列クロマトグラムの自動分析及び統計モデルを用いたシーケンシングエラーや研究室でのDNAの調整に起因する配列のばらつきを排除する。別法では、例えばハイスループットのMASSARRAYシステム (Sequenom, Inc., San Diego CA) を用いた質量分析によりSNPを検出し、特徴付ける。

【0216】

NHRECの発現を定量するために用いられ得る方法には、ヌクレオチドの放射標識或いはビオチン標識、調節核酸の相互増幅 (coamplification)、及び標準的な曲線に結果が加えられたものが含まれる (例えば、Melby, P. C. ら (1993) J. Immunol. Methods, 159:235-44; Duplaa, C. ら (1993) Anal. Biochem. 229-236を参照)。多数のサンプルの定量速度は、ハイスループット型のアッセイを用いることで速くなるであろう。このアッセイでは、目的のオリゴマーやポリヌクレオチドが様々な希釈液中に含まれ、分光光度法或いは非色応答によって定量が迅速である。

【0217】

更に別の実施例では、本明細書で記載した任意のポリヌクレオチド配列に由来するオリゴヌクレオチドまたはより長い断片を、マイクロアレイにおける標的として用いることができる。マイクロアレイを、上記したように多数の遺伝子の相対的な発現レベルを同時にモニタリングする転写イメージング技術に用いることができる。マイクロアレイはまた、遺伝子変異、突然変異及び多型の同定に用いることができる。この情報を用いて、遺伝子機能を決定し、疾患の遺伝的根拠を解明し、疾患を診断し、遺伝子発現に関連する疾病の進行／後退をモニタリングし、疾患の治療における治療薬の開発や活性のモニタリングを行うことができる。特に、患者にとって最適かつ有効な治療法を選択するために、この情報を用いて患者の薬理ゲノムプロファイルを作成することができる。例えば、患者の薬理ゲノムプロファイルに基づいて、患者に対して極めて効果的でありながら副作用を殆ど示さない治療薬を選択することができる。

【0218】

別の実施例では、NHREC、NHRECの断片、NHRECに特異的な抗体をマイクロ

アレイ上のエレメントとして用いることができる。マイクロアレイを用いて、上記のようにタンパク質間相互作用、薬剤－標的相互作用及び遺伝子発現プロファイルをモニタリング及び測定することが可能である。

【0219】

特定の実施例は、或る組織または細胞型の転写イメージを生成する本発明のポリヌクレオチドの使用に関連する。転写イメージは、特定の組織または細胞型により遺伝子発現の包括的パターンを表す。包括的遺伝子発現パターンは、所定の条件下で所定の時間に発現した遺伝子の数及び相対存在量を定量することにより分析される（Seilliamer 他、米国特許第5,840,484号の”Comparative Gene Transcript Analysis”を参照。この特許に言及することを以って本明細書の一部とする）。従って、特定の組織または細胞型の転写物または逆転写物の全てに本発明のポリヌクレオチドまたはその相補配列をハイブリダイズすることにより、転写イメージが生成され得る。或る実施例では、本発明のポリヌクレオチドまたはその相補配列がマイクロアレイ上に複数のエレメントのサブセットを構成するハイスループット型でハイブリダイゼーションさせる。結果として得られる転写イメージは、遺伝子活性のプロファイルとなり得る。

10

【0220】

転写イメージは、組織、細胞株、生検サンプル、またはその他の生体サンプルから単離した転写物を用いて生成し得る。従って、転写イメージは、組織または生検サンプルの場合には in vivo、または細胞株の場合には in vitro における遺伝子発現を反映する。

20

【0221】

本発明のポリヌクレオチドの発現プロファイルを示す転写イメージはまた、合成化合物または天然化合物の毒性試験のみならず、in vitroモデル系及び薬剤の前臨床評価に関連して使用され得る。全ての化合物は、作用及び毒性の機構を示唆する、頻繁に分子フィンガープリント若しくは毒性シグネチャ（signature）と称されるような特徴的な遺伝子発現パターンを引き起こす（Nuwaysir, E. F. 他（1999）Mol. Carcinog. 24:153-159、Steiner, S. and N. L. Anderson（2000）Toxicol. Lett. 112-113:467-471、また言及することを以って本明細書の一部とする）。試験化合物が、毒性を有する既知の化合物のシグネチャと同一のシグネチャを有する場合には、毒性特性を共有している可能性が高い。フィンガープリンまたはシグネチャが、より多くの遺伝子及び遺伝子ファミリーからの発現情報を含んでいれば、より有用かつ正確になる。理想としては、発現のゲノム全域にわたって測定し、最高品質のシグネチャを提供することである。任意の試験化合物によっても発現が変化しない遺伝子も同様に重要である。それは、これらの遺伝子の発現レベルを用いて残りの発現データを標準化することができるためである。標準化処理は、異なる化合物で処理した後の発現データの比較に有用である。毒性シグネチャのエレメントへの遺伝子機能を割り当てることは毒性機構の解明に役立つが、毒性の予測につながるシグネチャの統計的な一致には遺伝子機能の知識は必要ではない（例えば2000年2月29日にNational Institute of Environmental Health Sciencesより発行されたPress Release 00-02を参照されたい。これについては<http://www.niehs.nih.gov/oc/news/toxchip.htm>で入手可能である）。従って、毒性シグネチャを用いる毒性スクリーニングにおいて、全ての発現した遺伝子配列を含めることは重要でありまた望ましいことである。

30

40

【0222】

一実施例では、試験化合物の毒性は、核酸を含有する生体サンプルをその試験化合物で処理して評価する。処理した生体サンプル中で発現した核酸は、本発明のポリヌクレオチドに特異的な1若しくは複数のプローブでハイブリダイズさせ、それによって本発明のポリヌクレオチドに対応する転写レベルを定量することができる。処理した生体サンプル中の

50

転写レベルを、非処理生体サンプル中のレベルと比較する。両サンプルの転写レベルの差が、処理されたサンプル中で試験化合物が引き起こす毒性反応を示唆する。

【0223】

別の実施例は、本発明のポリペプチド配列を用いて組織または細胞型のプロテオームを分析することに関連する。「プロテオーム」という用語は、或る特定の組織または細胞型におけるタンパク質発現の包括的パターンを指す。プロテオームを構成する各タンパク質は、個々に更なる分析をすることができる。プロテオーム発現パターン即ちプロファイルは、所定の条件下で所定の時間に発現したタンパク質の数及びそれらの相対的な存在量を定量することにより分析する。従って、ある細胞のプロテオームのプロファイルは、特定の組織または細胞型のポリペプチドを分離及び分析することにより作成し得る。或る実施例では、このような分離は2次元ゲル電気泳動によって行う。この2次元ゲル電気泳動法では、まず、1次元の等電点電気泳動によりサンプルからタンパク質を分離し、次に、2次元のドデシル硫酸ナトリウムスラブゲル電気泳動により分子量に従って分離する（前出の Steiner and Anderson）。これらのタンパク質は、通常クーマシーブルーまたはシルバーまたは蛍光染色などの染色剤を用いてゲルを染色して、分散した個別の位置にあるスポットとしてゲル中で可視化される。各タンパク質スポットの光学密度は、通常サンプル中のタンパク質レベルに比例する。異なるサンプル、例えば試験化合物または治療薬で処理済みまたは未処理のいずれかの生体サンプルから得られる等位置にあるタンパク質スポットの光学密度を比較し、処理に関連するタンパク質スポット密度の変化を調べる。スポット内のタンパク質は、例えば化学的または酵素的に切断した後、質量分析する標準的な方法を用いて部分的にシーケンシングする。スポット内のタンパク質の同一性は、好適には少なくとも5個の連続するアミノ酸残基であるその部分的な配列を、本発明のポリペプチド配列と比較することにより決定し得る。場合によっては、決定的なタンパク質同定のための更なる配列が得られる。

10

20

【0224】

プロテオームのプロファイルは、NHRECに特異的な抗体を用いてNHREC発現レベルを定量することによっても作成可能である。或る実施例では、マイクロアレイ上のエレメントとして抗体を用い、マイクロアレイをサンプルに曝露して各アレイエレメントへのタンパク質結合レベルを検出することによりタンパク質発現レベルを定量する（Lueking, A. ら. (1999) Anal. Biochem. 270:103-111, Mendoz, L. G. ら. (1999) Biotechniques 27:778-788）。検出は当分野で既知の様々な方法で行うことができ、例えば、チオール反応性またはアミノ反応性蛍光化合物を用いてサンプル中のタンパク質を反応させ、各アレイエレメントにおける蛍光結合の量を検出し得る。

30

【0225】

プロテオームレベルでの毒性シグネチャも中毒学的スクリーニングに有用であり、転写レベルでの毒性シグネチャと並行して分析するべきである。或る組織における或るタンパク質では、転写物の存在量とタンパク質の存在量との相関性が低いことがあるため（Anderson, N. L. and J. Seilhamer (1997) Electrophoresis 18:533-537）、プロテオーム毒性シグネチャは、転写イメージにはそれ程影響しないがプロテオームのプロファイルを変化させる化合物の分析において有用たり得る。更に、体液中での転写の分析は、mRNAが急速に分解するため困難である。したがって、このような場合にはプロテオームのプロファイル作成はより信頼でき、情報価値がある。

40

【0226】

別の実施例では、試験化合物の毒性は、タンパク質を含む生体サンプルをその試験化合物で処理して評価する。処理された生体サンプル中で発現したタンパク質を分離して、各タンパク質の量が定量できるようにする。各タンパク質の量を、未処理生体サンプル中の対応するタンパク質の量と比較する。両サンプル中のタンパク質の量の差は、処理されたサンプル中の試験化合物に対する毒性反応を示唆する。個々のタンパク質は、それらのアミ

50

ノ酸残基をシーケンシングし、これらの部分配列を本発明のポリペプチドと比較することで同定する。

【0227】

別の実施例では、試験化合物の毒性は、タンパク質を含む生体サンプルをその試験化合物で処理することにより評価する。生体サンプルから得たタンパク質を、本発明のポリペプチドに特異的な抗体と共にインキュベートする。その抗体により認識されたタンパク質の量を定量する。処理された生体サンプル中のタンパク質の量を、未処理生体サンプル中のタンパク質の量と比較する。両サンプルのタンパク質量の差が、処理サンプル中の試験化合物に対する毒性反応を示唆する。

【0228】

当分野で周知の方法でマイクロアレイを準備して使用し、分析する。(例えば、Brennan, T. M. 他 (1995) 米国特許第5,474,796号; Schena, M. 他 (1996) Proc. Natl. Acad. Sci. 93:10614-10619; Baldeschweiler 他 (1995) PCT出願番号WO95/251116; Shalon, D. 他 (1995) PCT出願番号WO95/35505; Heller, R. A. 他 (1997) Proc. Natl. Acad. Sci. 94:2150-2155; 及び Heller, M. J. 他 (1997) 米国特許第5,605,662号を参照)。様々なタイプのマイクロアレイが周知であり、詳細については、DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, Londonに記載されている。また、この文献を引用することを以て本明細書の一部とする。

10

20

30

40

【0229】

本発明の別の実施例ではまた、NHRECをコードする核酸配列を用いて、天然のゲノム配列をマッピングするのに有用なハイブリダイゼーションプローブを作製することが可能である。コーディング配列または非コーディング配列の何れかを用いることができるが、或る例では、コーディング配列より非コード配列が好ましい。例えば、多重遺伝子ファミリーのメンバー間にコーディング配列が保存されていることにより、染色体マッピング時に望ましくない交差ハイブリダイゼーションが生じる可能性がある。この配列は、特定の染色体、染色体の特定領域または人工の染色体、例えば、ヒト人工染色体(HAC)、酵母人工染色体(YAC)、細菌人工染色体(BAC)、細菌P1産物、或いは単一染色体cDNAライブラリに対してマッピングされる(Harrington, J. J. 他 (1997) Nat Genet. 15:345-355、Price, C. M. (1993) Blood Rev. 7:127-134、Trask, B. J. (1991) Trends Genet. 7:149-154等を参照)。一度マッピングすると、本発明の核酸配列を用いて、例えば病状の遺伝と特定の染色体領域やまたは制限断片長多型(RFLP)の遺伝とが相関するような遺伝子連鎖地図を作成可能である(Lander, E. S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357を参照)。

【0230】

in situ 蛍光ハイブリダイゼーション(FISH)は、他の物理的及び遺伝子地図データと相関し得る(例えば、Heinz-Ulrich, 他による(1995) in Meyers, 前出, pp. 965-968を参照)。遺伝子地図データの例は、種々の科学誌あるいはOnline Mendelian Inheritance in Man(OMIM)のワールドワイドウェブのサイトで見付けることができる。物理的な染色体地図上のNHRECをコードする遺伝子の位置と特定の疾患との相関性、或いは特定の疾患に対する素因が、このような疾患と関連するDNA領域の決定に役立つため、更なる位置を決定するクローニングが行われる。

【0231】

染色体標本の in situ ハイブリダイゼーション、及び確定した染色体マーカーを用い

50

た結合分析などの物理的マッピング技術を用いて、遺伝子地図を拡張することもできる。マウスなどの別の哺乳動物の染色体上に遺伝子を配置させることにより、たとえ正確なヒト染色体の位置が分かっているなくても、関連するマーカーが明らかになる場合が多い。この情報は、位置クローニング或いは別の遺伝子発見技術を用いて遺伝的疾患の研究をしている研究者にとって価値がある。疾患や症候群に關与する1つ或いは複数の遺伝子の位置が、例えば血管拡張性失調症の11q22-23などの特定の遺伝子領域に遺伝子結合によって大まかに決定されると、その領域に対するどの配列マッピングも、さらなる調査のための関連する遺伝子或いは調節遺伝子を表す（例えば、Gatti, R. A. 他による(1988) Nature 336: 577-580を参照）。また、目的の本発明のヌクレオチド配列を用いて、正常者、保有者、即ち感染者の間の、転位置、反転などによる染色体位置の違いを検出することもある。 10

【0232】

本発明の別の実施例では、NHREC、その触媒作用断片或いは免疫原断片またはそのオリゴペプチドを、種々の任意の薬剤スクリーニング技術における化合物のライブラリのスクリーニングに用いることができる。このようなスクリーニングに用いる断片は、溶液に遊離、固体支持物に固定、細胞の表面上に保持、或いは細胞内に存在する。NHRECと検査する薬剤との結合による複合体の形成を測定してもよい。

【0233】

薬剤スクリーニングに用いる別の方法は、目的のタンパク質に対して、好適な結合親和性を有する化合物のスクリーニング処理能力を高めるために用いられる（例えば、Geyse 20en, 他による(1984) PCT出願番号 WO84/03564を参照）。この方法では、相当な数の異なる小さな試験用化合物が、プラスチックピン或いは他の基板の上に合成される。試験用化合物は、NHREC、或いはその断片と反応してから洗浄される。次に、結合されたNHRECが、当分野で周知の方法で検出される。精製されたNHRECはまた、前記した薬剤をスクリーニングする技術に用いられるプレート上で直接被覆することもできる。別法では、非中和抗体を用いて、ペプチドを捕らえ、固体支持物に固定することもできる。

【0234】

別の実施例では、NHRECと結合可能な中和抗体がNHRECと結合するため試験用化合物と特に競合する、競合的薬剤スクリーニングアッセイを用いることができる。この方法では、抗体が、NHRECと1つ以上の抗原決定因子を共有するどのペプチドの存在も 30検出する。

【0235】

別の実施例では、発展途上の分子生物学技術にNHRECをコードするヌクレオチド配列を用いて、限定はされないが、現在知られているトリプレット暗号及び特異的な塩基対相互作用などのヌクレオチド配列の特性に依存する新しい技術を提供することができる。

【0236】

当分野の技術者であれば、更なる説明がなくても前述の説明だけで最大限に本発明を利用できるであろう。したがって、以下に記載する特定の好適な実施例は、例示目的であって本発明を限定するものではない。 40

【0237】

前述した及び以下に記載する全ての特許出願、特許、刊行物、特に米国特許出願第60/178, 578号、同第60/185, 079号、および同第60/189, 999号に言及することをもって本明細書の一部とする。

【0238】

(実施例)

1 cDNAライブラリの作製

インサイトcDNAはLIFESeq GOLD データベース (Incyte Genomics, Palo Alto CA) に含まれているcDNAライブラリに由来し、表 4の列5に示されている。組織の一部をホモジナイズしてグアニジニウムイソチオシアネ 50

ート溶液に溶解する一方、この組織の別の一部をホモジナイズしてフェノールに溶解するか、或いはTRIzol (Life Technologies)、グアニジニウムイソチオシアネート及びフェノールの単相溶液などの好適な変性剤の混合液に溶解した。この溶解物を塩化セシウムにおいて遠心分離によって、或いはクロロホルムで抽出した。イソプロパノール或いは酢酸ナトリウムのどちらかとエタノール、或いは別の方法でこの溶解物からRNAを沈殿させた。

【0239】

RNAの純度を高めるためにRNAのフェノールによる抽出及び沈殿を必要な回数繰り返した。場合によっては、DNA分解酵素でRNAを処理する。殆どのライブラリでは、オリゴd(T)連結常磁性粒子(Promega)またはOLIGOTEXラテックス粒子(QIAGEN, Valencia CA)、OLIGOTEX mRNA精製キット(QIAGEN)を用いてポリ(A+) RNAを単離した。別法では、POLY(A) PURE mRNA精製キット(Ambion, Austin TX)などの別のRNA単離キットを用いて組織溶解物から直接単離した。

10

【0240】

ある場合には、Stratagene社にRNAを提供し、Stratagene社が対応するcDNAライブラリを作製した。そうでない場合は、UNIZAPベクターシステム(Stratagene)またはSUPERSCRIPT プラスミドシステム(Life Technologies)を用いて当分野で周知の推奨方法または類似の方法でcDNAを合成してcDNAライブラリを作製した。(例えば、Ausubel, 1997, 前出, ユニット5.1-6.6を参照)。逆転写は、オリゴd(T)またはランダムプライマーを用いて開始した。合成オリゴヌクレオチドアダプターを二本鎖cDNAに結合させてから、好適な1つの制限酵素或いは複数の制限酵素でcDNAを消化した。殆どのライブラリでは、SEPHACRYL S 1000または SEPHAROSE CL2B、SEPHAROSE CL4Bカラムクロマトグラフィー(Amersham Pharmacia Biotech)、アガロースゲル電気泳動法によってcDNAの大きさ(300~1000bp)を選択した。PBLUESCRIPTプラスミド(Stratagene)またはpSPORT1プラスミド(Life Technologies)、pcDNA2.1プラスミド(Invitrogen Carlsbad CA)、PBK-CMVプラスミド(Stratagene)、pINCYプラスミド(Incyte Pharmaceuticals, Palo Alto CA)、またはそれらの誘導体などの好適なプラスミドのポリリンカーの適合性制限酵素部位にcDNAを結合させた。この組換えプラスミドを、Stratagene社のXL1-Blue, XL1-BlueMRF、SOLR、またはLife Technologies社のDH5αまたはDH10B、ELECTROMAX DH10Bを含むコンピテント大腸菌細胞に導入し組み込んだ。

20

30

【0241】

2 cDNAクローンの単離

上記実施例1に記載したように得たプラスミドを、UNIZAPベクターシステム(Stratagene)或いは細胞溶解を利用したin vivo切除によって宿主細胞から回収した。MagicまたはWIZARD Minipreps DNA精製システム(Promega)、及びAGTC Miniprep精製キット(Edge Biosystems, Gaithersburg MD)、QIAGEN社のQIAWELL 8 Plasmid、QIAWELL 8 Plus Plasmid、QIAWELL 8 Ultra Plasmid 精製システム、REAL Prep 96プラスミドキットの内の少なくとも1つを用いてプラスミドを精製した。沈殿させた後、0.1mlの蒸留水に再懸濁して、凍結乾燥して或いは凍結乾燥しないで4℃で保管した。

40

【0242】

別法では、ハイスループットの直接結合PCR法によって宿主細胞溶解物からプラスミドDNAを増幅した。(Rao, V. B. (1994) Anal. Biochem.

50

216:1-14)。宿主細胞の溶解及び熱サイクリング過程を単一反応混合液で行った。サンプルを処理してから384-ウェルプレートに移して保管し、増幅したプラスミドDNAの濃度をPICO GREEN色素 (Molecular Probes, Eugene OR) 及びFluoroskan II 蛍光スキャナ (Lab systems Oy, Helsinki, Finland) を用いて蛍光定量的に測定した。

【0243】

3 シークエンシング及び分析

実施例2に記載したようにプラスミドから回収したインサイトcDNAを、以下に示すようにシークエンシングした。cDNAのシークエンシング反応は標準的な方法で行うか、またはHYDRAマイクロディスペンサー (Robbins Scientific) 或いはMICROLAB 2200 (Hamilton) 液体移送装置と共にABI CATALYST 800 (PE Biosystems) サーマルサイクラー或いはPTC-200 thermal cycler (MJ Research) などのハイスループット装置を用いて行った。cDNAのシークエンシング反応は、Amersham Pharmacia Biotech社の試薬、またはABI PRISM BIGDYE Terminator cycle sequencing ready reactionキット (PE Biosystems) などのABIシークエンシングキットに含まれる試薬を用いて行った。cDNAシークエンシングの反応物の電気泳動的による分離及び標識したポリヌクレオチドの検出は、MEGABACE 1000 DNAシークエンシングシステム (Molecular Dynamics)、標準ABIプロトコル及び塩基対呼び出しソフトウェアを用いるABI PRISM 373または377シークエンシングシステム (PE Biosystems)、または当分野で周知のその他の配列解析システムを用いて行った。cDNA配列内の読み枠は、標準的な方法 (Ausubel, 1997, 前出, unit 7.7) を用いて決定した。cDNA配列の幾つかを選択して、実施例8に記載した方法で配列を伸長した。

【0244】

インサイトcDNAに由来する本ポリヌクレオチド配列の確認は、BLAST、動的プログラミング、およびジヌクレオチドの分布による解析 (dinucleotide nearest neighbor analysis) に基づいたプログラム及びアルゴリズムを用いて、ベクター、リンカー、およびポリA配列を取り除き、更にあいまいな塩基対をマスクすることで行った。次に、インサイトcDNA配列およびそれらの翻訳を、公共のデータベースであるGenBankの霊長類、げっ歯類、哺乳類、脊椎動物、および真核生物のデータベース、およびBLOCKS、PRINTS、DOMO、PRODOM、およびPFAMなどの隠れマルコフモデル (HMM) を基にしたタンパク質ファミリーのデータベースから選択した配列に対して問合せた (HMMは、遺伝子ファミリーのコンセンサス主構造を分析する確率的手法である。例えば、Eddy, S. R. (1996) Curr. Opin. Struct. Biol. 6: 361-365を参照)。このような問合せは、BLAST、FASTA、BLIMPS、およびHMMRに基づいたプログラムを用いて行った。インサイトcDNA配列を組み立てて、完全長ポリヌクレオチド配列を作製した。或いは、GenBank cDNAs、GenBank EST、ステッチ配列 (stitched sequence)、ストレッチ配列 (stretched sequences)、またはGenscan-推定コード配列 (実施例4および5を参照) を用いて、インサイトcDNA群を完全長の配列に伸長した。配列の組み立ては、Phred、Phrap、およびConsedに基づいたプログラムを用いて行い、GeneMark、BLAST、およびFASTAに基づいたプログラムを用いて、オープンリーディングフレームを決定するべくcDNA群をスクリーニングした。完全長のポリヌクレオチド配列を翻訳して対応する完全長ポリペプチド配列を得た。別法では、本発明のポリヌクレオチドは、完全長翻訳ポリヌクレオチドの任意のメチオニン残基から始まり得る。次に、完全長ポリペプチド配列をGenBankタンパク質データベース (genpept)、SwissProt、BLOCKS、PRINTS、DOMO、PR

ODOM、Prosite、およびPFAMなどの隠れマルコフモデル（HMM）に基づいたタンパク質ファミリーデータベースに対して問合せて分析した。これらの完全長ポリヌクレオチド配列はまた、MACDNASIS PRO ソフトウェア（Hitachi Software Engineering, South San Francisco CA）およびLASERGENEソフトウェア（DNASTAR）を用いて分析した。ポリヌクレオチド配列およびポリペプチド配列のアラインメントを、アラインメントした配列間のパーセント同一性も計算するMEGALIGNマルチシークエンスアラインメントプログラム（DNASTAR）に組み込まれたCLUSTALアルゴリズムによって指定されたデフォルトパラメータを用いて作成した。

【0245】

表7は、インサイトcDNAおよび完全長配列の組み立て、および組み立てた配列の分析に利用したツール、プログラム、およびアルゴリズム、並びにそれらの説明、引用文献、閾値パラメータを簡単に示す。表7の列1は用いたツール、プログラム、およびアルゴリズム、列2はそれらの簡単な説明、列3は引用することで本明細書の一部とした引用文献、列4の記載されている部分は2つの配列の一致の程度を評価するために用いたスコア、確率値、およびその他のパラメータを示す（スコアが高くなれば高くなるほど即ち確率値が低ければ低いほど、配列間の相同性が高くなる）。

【0246】

完全長のポリヌクレオチド配列およびポリペプチド配列の組み立て及び分析に用いる上記のプログラムは、SEQ ID NO: 4-6のポリヌクレオチド配列断片の同定にも利用できる。ハイブリダイゼーション及び増幅技術に有用である約20～約4000ヌクレオチドの断片を表4の列4に示した。

【0247】

4 ゲノムDNAからのコード配列の同定および編集

推定上の核ホルモン受容体は、公共のゲノム配列データベース（例えば、gbpriやgbhtg）においてGenscan遺伝子同定プログラムを実行して初めに同定された。Genscanは、様々な生物に由来するゲノムDNA配列を分析するための汎用遺伝子同定プログラムである（Burge, C. および S. Karlin (1997) J. Mol. Biol. 268 : 78-94, Burge, C. および S. Karlin (1998) Curr. Opin. Struct. Biol. 8 : 346-354を参照）。このプログラムは推定エキソンを連結して、メチオニンから停止コードンまで伸長した組み立てcDNA配列を構築する。Genscanにより得られる配列は、FASTAデータベースのポリヌクレオチド配列およびポリペプチド配列になる。Genscanによって一回で解析できる配列の最大長さは30kbに設定されている。これらのGenscan推定cDNA配列の内、どの配列が核ホルモン受容体をコードするかを決定するために、コードされたポリペプチドをPFAMモデルにおいて核ホルモン受容体について問合せて分析した。潜在的なヒト核ホルモン受容体が、核ホルモン受容体としてアノテーションが付けられたインサイトcDNA配列に対する相同性を基に同定された。次に、これらの選択されたGenscan推定配列を、BLAST解析を用いてgeneptおよびgbpri公共データベースの配列と比較した。必要に応じて、Genscan推定cDNA配列を、geneptにおいてBLASTで最もヒットした配列と比較して、Genscan推定配列における余分なエキソンや省いてしまったエキソンなどのエラーを修正し、編集した。BLAST解析を用いてGenscan推定cDNA配列を含むインサイトcDNAまたは公共のcDNAを見つけ出すことにより、転写の証拠が得られる。インサイトcDNAがGenscan推定cDNA配列を含む場合、この情報を用いてGenscan推定配列を修正或いは確認できる。完全長ポリヌクレオチド配列は、実施例3に説明した組み立て方法でGenscan推定コード配列とインサイトcDNAおよび／または公共のcDNA配列を組み立てて作製した。別法では、完全長ポリヌクレオチド配列は、その全てが編集した或いは未編集のGenscan推定コード配列から作製した。

10

20

30

40

50

【0248】

5 cDNA配列データを用いたゲノム配列データの組み立てステッチ配列 (Stitched Sequence)

部分的なcDNA配列を、実施例4に記載したGenscan遺伝子同定プログラムによって推定されたエキソンで伸長した。実施例3に記載されたように組み立てられた部分的なcDNAをゲノムDNAにマッピングし、関連するcDNAおよび1或いは複数のゲノム配列に由来する関連する推定Genscanエキソンを含む複数のクラスターに入れた。各クラスターを、グラフ理論および動的計画法に基づいたアルゴリズムを用いて、cDNAおよびゲノム情報を統合して分析し、後に確認される潜在的なスプライスバリエーションを生成し、編集或いは伸長して完全長の配列を作製した。或るクラスターの2つ以上の配列に或る区間の全長が存在する配列区間を同定し、推移 (transitivity) により同定した区間を同等と考える。例えば、或る区間がcDNAおよび2つのゲノム配列のそれぞれに存在する場合、これら3つ全ての区間を同等と考える。この方法によって、関連しないが連続するゲノム配列をcDNA配列によって繋ぎ1つにする。このようにして同定された区間を、親配列 (parent sequence) に沿って現われるようにステッチアルゴリズムで縫い合わせ、可能な最も長い配列および変異配列を作製する。或るタイプ (cDNAとcDNA、またはゲノム配列とゲノム配列) の親配列に沿って連結される区間と区間との繋ぎ合わせは、親配列のタイプが異なる (cDNAとゲノム配列) 連結より好ましい。得られたステッチ配列を翻訳し、BLAST解析でgenpeptおよびgbpr公共データベースにおける配列と比較した。Genscanによって推定された不適当なエキソンを、genpeptにおいてBLASTで最もヒットした配列と比較して修正する。このような配列を更なるcDNA配列で伸長し、必要に応じてゲノムDNAで検査した。

10

20

【0249】

ストレッチ配列 (Stretched Sequence)

部分的なDNA配列をBLAST解析に基づいたアルゴリズムで完全長に伸長した。まず、実施例3に記載したように組み立てた部分的なcDNAを、BLASTプログラムを用いてGenBankの霊長類、げっ歯類、哺乳類、脊椎動物、および真核生物のデータベースなどの公共のデータベースに対して問い合わせた。次に、GenBankの相同性の最も高いタンパク質を、実施例4に記載したインサイトcDNA或いはGenscanエキソン推定配列の何れかと比較した。得られた複数の高スコアのセグメント対 (HSP) を用いてキメラタンパク質を作製し、GenBankの相同タンパク質上に翻訳した配列をマッピングした。元のGenBankの相同タンパク質に対して、キメラタンパク質に挿入や欠失が起こり得る。公共のヒトゲノムデータベースから相同ゲノム配列を探し出すために、GenBankの相同タンパク質およびキメラタンパク質の両方をプローブとして用いた。このようにして、部分的なDNA配列を相同ゲノム配列の付加によりストレッチすなわち伸長した。完全な遺伝子を含んでいるか得られたストレッチ配列を検査した。

30

【0250】

6 NHRECをコードするポリヌクレオチドの染色体マッピング

SEQ ID NO: 4-6を組み立てるために用いた配列を、BLAST及びSmith-Watermanアルゴリズムを用いて、インサイトLIFESEQデータベース及び公共のドメインデータベースの配列と比較した。SEQ ID NO: 4-6と一致するこれらのデータベースの配列を、Phrap (表7) などの構築アルゴリズムを使用して、連続及び重複した配列のクラスターに組み入れた。Stanford Human Genome Center (SHGC)、Whitehead Institute for Genome Research (WIGR) 及びGenethonなどの公共の情報源から入手できる放射線ハイブリッド (radiation hybrid) 及び遺伝子マッピングのデータを用いて、クラスター化した配列がすでにマッピングされているかを調べる。クラスターにマッピングされた配列が含まれている場合は、そのクラスターの全ての配列 (特定のSEQ ID NOを含む) をそのマッピング位置に割り当てた。

40

50

【0251】

遺伝子地図の位置は、範囲、区間、またはヒト染色体によって表される。センチモルガンで示したマッピング位置の範囲は、染色体の短腕（p）の末端から測定した（センチモルガン（cM）は、同一染色体上の遺伝子間の乗換え率に基づいた距離を表す単位である。平均すると、1 cMはヒトの染色体の1メガベースに概ね等しいが、組換え率の高い部分と低い部分があるため、大きく変化し得る）。距離cMは、配列がそれぞれのクラスターに含まれている放射線ハイブリッドマーカーの境界を検出できるGeneThonによってマッピングされた遺伝子マーカーに基づいている。NCBI「GeneMap99」（<http://www.ncbi.nlm.nih.gov/genemap>）などの公衆が入手可能なヒト遺伝子マップおよびその他の情報源を用いて、上記した区間が既に同定されている疾患遺伝子マップ内若しくは近傍に位置するかを決定できる。

【0252】

7 ポリヌクレオチド発現の分析

ノーザン分析は、遺伝子の転写物の存在を検出するために用いられる実験用技術であり、特定の細胞種或いは組織からのRNAが結合されている膜への標識されたヌクレオチド配列のハイブリダイゼーションを伴う（例えば、Sambrook, 前出, 7章; 及び Ausubel, F. M. 他、前出, 4章及び16章を参照）。

【0253】

BLASTに用いる類似のコンピュータ技術を用いて、GenBank或いはLIFESEQ (Incyte Pharmaceuticals) のようなcDNAデータベース内の同一或いは関連する分子を検索する。この分析は多くの膜系ハイブリダイゼーションより非常に速度が速い。さらにコンピュータ検索の感度を変更して、任意の特定の一致が、厳密な一致或いは相同的一致の何れかとして分類されるかを確定することができる。検索の基準は、

【0254】

【数1】

(BLAST スコア×配列一致率)

5×(長さ(配列1), 長さ(配列2))の最小値

として定義される積スコアである。積スコアは、0～100の標準化された値であり、以下のように求める。BLASTスコアにヌクレオチド配列の一致率を乗じ、その積を2つの配列の短い方の長さの5倍で除する。高スコアのセグメントの対(HSP)において一致する各塩基に+5のスコアを割り当て、各不適性塩基対に-4を割り当てることにより、BLASTスコアを計算する。2つの配列は、2以上のHSPを共有し得る（ギャップにより離隔される）。2以上のHSPがある場合には、最高BLASTスコアの塩基対を用いて積スコアを計算する。積スコアは、BLASTアラインメントの断片的重複と質とのバランスを表す。例えば積スコア100は、比較した2つの配列の短い方の長さ全体にわたって100%一致する場合にのみ得られる。積スコア70は、100%の同一性で重畳が70%であるか、或いは88%の同一性で重畳が100%であるかの何れかの場合である。積スコア50は、100%の同一性で重畳が50%であるか、或いは79%の同一性で重畳が100%であるかの何れかの場合である。

【0255】

或いは、NHRECをコードするポリヌクレオチド配列は、由来する組織に対して分析する。例えば、ある完全長の配列は、少なくとも部分的にインサイトcDNA配列をオーバーラップさせて組み立てられる（実施例3を参照）。各cDNA配列は、ヒト組織から作製されたcDNAライブラリに由来する。各ヒト組織は、心血管系、結合組織、消化系、胚構造、内分泌系、外分泌腺、女性生殖器、男性生殖器、生殖細胞、血液および免疫系、肺、筋骨格系、神経系、脾臓、呼吸器系、感覚器官、皮膚、顎口腔系、分類不能/混合、または尿管などの1つの生物/組織のカテゴリーに分類される。各カテゴリーにおけるラ

ライブラリ数をカウントし、その合計数を全カテゴリーのライブラリ数で除す。同様に、各ヒト組織は、癌、細胞系、発生、炎症、神経、外傷、心血管、プール (p o o l e d) などの1つの疾患／症状のカテゴリーに分類され、各カテゴリーにおけるライブラリ数をカウントし、その合計数を全カテゴリーのライブラリ数で除す。得られるパーセンテージは、NHRECをコードするcDNAの疾患特異的な発現を反映する。cDNA配列およびcDNAライブラリ／組織の情報は、L I F E S E Q G O L D データベース (I n c y t e G e n o m i c s , P a l o A l t o C A) から得ることができる。

【0256】

8 NHRECをコードするポリヌクレオチドの伸長

完全長のポリヌクレオチド配列は、完全長分子の好適な断片から設計したオリゴヌクレオチドプライマーを用いてその完全長分子の好適な断片を伸長して作製した。一方のプライマーは既知の断片の5'の伸長を開始するために合成し、他方のプライマーは既知の断片の3'の伸長を開始するために合成した。開始プライマーは、O L I G O 4 . 0 6 ソフトウェア (N a t i o n a l B i o s c i e n c e s) 或いは他の適切なプログラムを用いて、約22個から約30個のヌクレオチドの長さで約50%以上のGC含量を有し、かつ約68～72℃の温度で標的配列にアニールするように設計した。ヘアピン構造及びプライマー-プライマー二量体が生じないようにヌクレオチドを伸長した。

【0257】

選択されたヒトcDNAライブラリを用いてこの配列を伸長した。2段階以上の伸長が必要な場合、若しくは望ましい場合は、追加或いはネスト化プライマーの組を設計する。

【0258】

当分野で既知の方法を利用したPCR法で高い忠実度で増幅した。PCRはP T C - 2 0 0 t h e r m a l c y c l e r (M J R e s e a r c h , I n c .) 用いて96ウェルブロックプレートで行った。反応混合液は、鋳型DNA及び200 nmolの各プライマー、 Mg^{2+} と $(NH_4)_2SO_4$ とβ-メルカプトエタノールを含むバッファー、Taq DNAポリメラーゼ (A m e r s h a m P h a r m a c i a B i o t e c h) 、E L O N G A S E 酵素 (L i f e T e c h n o l o g i e s) 、P f u DNAポリメラーゼ (S t r a t a g e n e) を含む。プライマーの組、P C I A と P C I B に対して以下のパラメータで増幅を行った。

ステップ1	94℃で3分間	10
ステップ2	94℃で15秒	
ステップ3	60℃で1分間	
ステップ4	68℃で2分間	
ステップ5	ステップ2、3、及び4を20回繰り返す	
ステップ6	68℃で5分間	
ステップ7	4℃で保管	

別法では、プライマーの組、T7とSK+に対して以下のパラメータで増幅を行った。

ステップ1	94℃で3分間	
ステップ2	94℃で15秒	
ステップ3	57℃で1分間	40
ステップ4	68℃で2分間	
ステップ5	ステップ2、3、及び4を20回繰り返す	
ステップ6	68℃で5分間	
ステップ7	4℃で保管。	

【0259】

各ウェルのDNA濃度は、1X TE及び0.5 μlの希釈していないPCR産物に溶解した100 μlのP I C O G R E E N 定量試薬 (0.25% (v/v) P I C O G R E E N ; M o l e c u l a r P r o b e s , E u g e n e O R) を不透明な蛍光光度計プレート (C o m i n g C o s t a r , A c t o n M A) の各ウェルに分配してDNAが試薬と結合できるようにして測定する。このプレートをF l u o r o s k a n I I

(L a b s y s t e m s O y, H e l s i n k i, F i n l a n d) でスキャンして、サンプルの蛍光を計測してDNAの濃度を定量化する。反応混合物の5～10 μ lのアリコットを1%のアガロースミニゲル上での電気泳動によって解析し、何れの反応物が配列を伸長することに成功したかを決定する。

【0260】

伸長したヌクレオチドを脱塩及び濃縮してから384ウェルプレートに移し、C v i J I コレラウィルスエンドヌクレアーゼ (M o l e c u l a r B i o l o g y R e s e a r c h, M a d i s o n W I) で消化し、pUC 18ベクター (A m e r s h a m P h a r m a c i a B i o t e c h) に再連結する前に音波処理またはせん断を行った。シ
ョットガンシーケンシングのために、消化したヌクレオチドを低濃度 (0.6～0.8 10
%) のアガロースゲル上に分離して断片を切断し、寒天をA g a r A C E (P r o m e g a) で消化した。T4リガーゼ (N e w E n g l a n d B i o l a b s, B e v e r l y M A) を用いて伸長したクローンをpUC 18ベクター (A m e r s h a m P h a r m a c i a B i o t e c h) に再連結し、P f u DNAポリメラーゼ (S t r a t a g e n e) で制限部位の伸び出しを処理してコンピテント大腸菌細胞に形質移入した。形質移入した細胞を選択して抗生物質を含む培地に移し、それぞれのコロニーを切りとってLB/2Xカルベニシリン培養液の384ウェルプレートに37℃で一晩培養した。

【0261】

細胞を溶解して、T a q DNAポリメラーゼ (A m e r s h a m P h a r m a c i a B i o t e c h) 及びP f u DNAポリメラーゼ (S t r a t a g e n e) を用いて以 20
下の手順でDNAをPCR増幅した。

- ステップ1 94℃で3分間
- ステップ2 94℃で15秒
- ステップ3 60℃で1分間
- ステップ4 72℃で2分間
- ステップ5 ステップ2、3、及び4を29回繰り返す
- ステップ6 72℃で5分間
- ステップ7 4℃で保管。

上記したようにP I C O G R E E N試薬 (M o l e c u l a r P r o b e s) でDNAを定量化した。DNA回収率の悪いサンプルは、上記した条件で再び増幅した。サンプル 30
を20%のジメチルサルホサイド (d i m e t h y s u l p h o x i d e) (1:2, v/v) で希釈し、D Y E N A M I C D I R E C Tキット (A m e r s h a m P h a r m a c i a B i o t e c h) またはA B I P R I S M B I G D Y E T e r m i n a t o r c y c l e s e q u e n c i n g r e a d y r e a c t i o nキット (A p p l i e d B i o s y s t e m s) を用いてシーケンシングした。

【0262】

同様に上述の手順で、完全長のポリヌクレオチド配列を検査したり、或いは完全長のポリヌクレオチド配列を利用して、この伸長のために設計したオリゴヌクレオチドと好適なゲノムライブラリを用いて5'調節配列を得た。

【0263】

9 個々のハイブリダイゼーションプローブの標識化及び使用法

S E Q I D N O: 4-6から導き出されたハイブリダイゼーションプローブを用いて、cDNA、mRNA、またはゲノムDNAをスクリーニングする。約20塩基対からなるオリゴヌクレオチドの標識について特に記すが、より大きなcDNAフラグメントの場合でも基本的に同じ手順を用いる。オリゴヌクレオチドを、O L I G O 4. 06ソフトウェア (N a t i o n a l B i o s c i e n c e) のような最新式のソフトウェアを用いてデザインし、50 pmolの各オリゴマーと、250 μ Ciの [γ -³²P] アデノシン三リン酸 (A m e r s h a m, C h i c a g o, I L) 及びT4ポリヌクレオチドキナーゼ (D u P o n t N E N, B o s t o n M A) とを組み合わせて用いることにより標識する。標識されたオリゴヌクレオチドを、S E P H A D E X G-25超精細排除デキ 50

ストランビードカラム (Amersham Pharmacia Biotech) を用いて実質的に精製する。毎分 10^7 カウントの標識されたプローブを含むアリコットを、次のエンドヌクレアーゼ、Ase I、Bgl II、Eco RI、Pst I、Xba I 或いは Pvu II (DuPont NEN) の 1 つを用いて切断したヒトゲノム DNA の典型的な膜ベースのハイブリダイゼーション解析において用いる。

【0264】

各切断物からの DNA を、0.7% アガロースゲル上で分画して、ナイロン製メンブラン (Nyttran Plus, Schleicher & Schuell, Durham NH) に転写する。ハイブリダイゼーションは 40°C で 16 時間かけて行う。非特異的シグナルを取り除くため、例えば、最大 0.1x クエン酸ナトリウム食塩水及び 0.5% ドデシル硫酸ナトリウムの条件の下、プロットを順次室温にて洗浄する。ハイブリダイゼーションパターンをオートラジオグラフィー或いは別のイメージ化手段で視覚化して比較する。

10

【0265】

10 マイクロアレイ

マイクロアレイ上のアレイエレメントの連結または合成は、フォトリソグラフィ、ピエゾプリント (インクジェットプリンター、前出の Baldeschweiler 等を参照)、機械的マイクロスポッティング技術及びこれらから派生したものを用いて達成することが可能である。上記各技術において基板は、均一な非多孔性の固体とするべきである (Schena (1999), 前出)。推奨する基板には、シリコン、シリカ、スライドガラス、ガラスチップ及びシリコンウエハがある。別法では、ドットプロット法またはスロットプロット法に類似のアレイを利用して、熱や紫外線、または化学的或いは機械的な結合手段で基板の表面にエレメントを配置して結合させることができる。通常のアレイは利用可能な方法や機械を用いて作製でき、任意の適正な数のエレメントを含めることができる (Schena, M. 他 (1995) Science 270: 467-470, Shalton, D. 他 (1996) Genome Res. 6: 639-645, Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16: 27-31. を参照)。

20

【0266】

完全長 cDNA、発現遺伝子配列断片 (EST)、或いはそれらの断片やオリゴマーが、マイクロアレイのエレメントとなり得る。ハイブリダイゼーションに好適な断片やオリゴマーを、LASERGENE ソフトウェア (DNASTAR) などの当分野で周知のソフトウェアを用いて選択することが可能である。このアレイエレメントを、生体サンプル中のポリヌクレオチドとハイブリダイズさせる。生体サンプル中のポリヌクレオチドは、検出を容易にするために蛍光標識またはその他の分子タグに結合する。ハイブリダイゼーションの後、生体サンプルからハイブリダイズしなかったヌクレオチドを除去し、蛍光スキャナを用いて各アレイエレメントにおけるハイブリダイゼーションを検出する。別法では、レーザー脱離及び質量スペクトロメトリーを用いてもハイブリダイゼーションを検出し得る。マイクロアレイ上のエレメントにハイブリダイズする各ポリヌクレオチドの相補性の程度及び相対的存在量は、算定することができる。一実施例におけるマイクロアレイの調整及び使用について、以下に詳述する。

30

40

【0267】

組織または細胞サンプルの調製

グアニジウムチオシアネート法を用いて組織サンプルから全 RNA を単離し、オリゴ (dT) セルロース法を用いてポリ (A)⁺ RNA を精製する。各ポリ (A)⁺ RNA サンプルは、MMLV 逆転写酵素、0.05 pg/ μl のオリゴ (dT) プライマー (21 mer)、1x 第 1 鎖緩衝液、0.03 単位/ μl の RNアーゼインヒビター、500 μM dATP、500 μM dGTP、500 μM dTTP、40 μM dCTP、40 μM dCTP-Cy3 (BDS) または dCTP-Cy5 (Amersham Pharmacia Biotech) を用いて逆転写する。この逆転写反応は、GEMBRIGHT

50

キット (Incycyte) を用いて、200 ng のポリ (A) ⁺ RNA を含む 25 ml 容量で行う。特異的なコントロールポリ (A) ⁺ RNA は、in vitro 転写により非コーディング酵母ゲノム DNA から合成する。370℃で2時間インキュベートした後、各反応サンプル (一方は Cy3 標識、他方は Cy5 標識) は、2.5 ml の 0.5 M 水酸化ナトリウムで処理し、850℃で20分間インキュベートし、反応を停止させて RNA を変性する。サンプルは、2つの連続する CHROMA SPIN 30 ゲル濾過スピンカラム (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) を用いて精製する。結合後、2つの反応サンプルを、1 ml のグリコーゲン (1 mg/ml)、60 ml の酢酸ナトリウム及び 300 ml の 100% エタノールを用いてエタノール沈殿させる。サンプルは次に、Speed VAC (Savant Instruments Inc., Holbrook NY) を用いて乾燥して仕上げ、14 µl 5×SSC/0.2% SDS 中で再懸濁する。

【0268】

マイクロアレイの準備

本発明の配列を用いて、アレイエレメントを作製する。各アレイエレメントは、クローン化 cDNA 挿入断片を含むベクターを含有する細菌性細胞から増幅する。PCR 増幅は、cDNA 挿入断片に隣接するベクター配列に相補的なプライマーを用いる。30 サイクルの PCR によって、1~2 ng の初期量から 5 µg を超える最終量までアレイエレメントを増幅する。増幅されたアレイエレメントは、SEPHACRYL-400 (Amersham Pharmacia Biotech) を用いて精製する。

【0269】

精製したアレイエレメントを、ポリマーコートされたスライドガラス上に固定する。顕微鏡スライドガラス (Corning) は、処理中及び処理後に大量の蒸留水での洗浄と、0.1% の SDS 及びアセトン中で超音波による洗浄を行う。スライドガラスは、4% フッ化水素酸 (VWR Scientific Products Corporation (VWR), West Chester PA) 中でエッチングし、蒸留水中で広範囲にわたって洗浄し、95% エタノール中の 0.05% アミノプロピルシラン (Sigma) でコーティングする。コーティングしたスライドガラスは、110℃の天火で硬化させる。

【0270】

米国特許第 5,807,522 号に記載されている方法を用いて、コーティングしたガラス基板にアレイエレメントを付加する。この特許に引用することを以て本明細書の一部とする。平均濃度が 100 ng/µl のアレイエレメント DNA 1 µl を高速機械装置により開放型キャピラリープリンティングエレメント (open capillary printing element) に充填する。次にこの装置が、スライド毎に約 5 nl のアレイエレメントサンプルを分注する。

【0271】

マイクロアレイには、STRATALINKER UV クロスリンカー (Stratagene) を用いて UV 架橋する。マイクロアレイは、室温において 0.2% SDS で 1 回洗浄し、蒸留水で 3 回洗浄する。非特異的な結合部位は、リン酸緩衝生理食塩水 (PBS) (Tropix, Inc., Bedford MA) における 0.2% カゼイン中で 60℃で30分間マイクロアレイをインキュベートし、その後上述したように 0.2% SDS 及び蒸留水で洗浄することによってブロックする。

【0272】

ハイブリダイゼーション

ハイブリダイゼーション反応液は、5×SSC、0.2% SDS ハイブリダイゼーション緩衝液に Cy3 及び Cy5 標識した cDNA 合成産物を各 0.2 µg 含む 9 µl のサンプル混合体を含めたものである。サンプル混合液を、65℃で5分間加熱し、マイクロアレイ表面上に一定量分注してから 1.8 cm² のカバーガラスで覆う。このアレイを、顕微鏡スライドより僅かに大きいキャビティを有する防水チェンバーに移す。チェンバーの

角に140 μ lの5×SSCを加えて、チャンバー内を湿度100%に保持する。このアレイを含むチャンバーを、60℃で約6.5時間インキュベートする。アレイは、第1洗浄緩衝液中(1×SSC, 0.1%SDS)において45℃で10分間、第2洗浄緩衝液中(0.1×SSC)において45℃で10分間それぞれ3回洗浄し、その後乾燥させる。

【0273】

検出

レポーター標識されたハイブリダイゼーション複合体は、Cy3を励起するための488nm、及びCy3を励起するための632nmのスペクトル線を生成し得るInnova 70混合ガス10Wレーザー(Coherent, Inc., Santa Clara CA)を備えた顕微鏡で検出する。20倍の顕微鏡対物レンズ(Nikon, Inc., Melville NY)を用いて、アレイ上に励起レーザー光を集中させる。このアレイを含むスライドを顕微鏡のコンピュータ制御X-Yステージに置き、対物レンズを通してラスタスキャンする。本実施例で用いた1.8cm×1.8cmのアレイは、20 μ mの解像度でスキャンする。

10

【0274】

2つの異なるスキャンにおいて、混合ガスマルチラインレーザーは2つの蛍光体を連続的に励起する。放射された光は、波長に基づいて2つの蛍光体に対応する2つの光電子増倍管検出器(PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ)に分割される。アレイと光電子増倍管との間に配設された好適なフィルタを用いて信号をフィルタリングする。用いる蛍光体の最大発光は、Cy3では565nm、Cy5では650nmである。装置は両方の蛍光体からのスペクトルを同時に記録できるが、レーザー源に好適なフィルタを用いて、蛍光体1つにつき1回スキャンし、各アレイを通常2回スキャンする。

20

【0275】

スキャンの感度は通常、既知濃度のサンプル混合体に添加されるcDNAコントロール種により生成されるシグナル強度を用いて校正する。アレイ上の特定の位置には相補的DNA配列を含め、その位置におけるシグナルの強度がハイブリダイズする種の重量比1:100, 000に相関するようにする。異なる試料(例えば検査細胞及びコントロール細胞を代表する)からの2つのサンプルを、各々異なる蛍光体で標識し、他と異なって発現する遺伝子を同定するために単一のアレイにハイブリダイズさせる場合には、校正は2つの蛍光体を有する校正するcDNAのサンプルを標識して、ハイブリダイゼーション混合液に各々等量を加えて行う。

30

【0276】

光電子増倍管の出力は、IBMコンパチブルPCコンピュータにインストールされた12ビットRTI-835Hアナログ-デジタル(AID)変換ボード(Analog Devices, Inc., Norwood MA)を用いてデジタル化される。デジタル化されたデータは、リニア20色変換を用いてシグナル強度が青色(低シグナル)から赤色(高シグナル)までの擬似カラー範囲にマッピングされるイメージとして表示される。データはまた、定量的に分析される。2つの異なる蛍光体を同時に励起して測定する場合には、各蛍光体の発光スペクトルを用いて、先ずデータは蛍光体間の光学的漏話(重複発光スペクトルに起因する)に対して補正される。

40

【0277】

グリッドを蛍光シグナルイメージ上に重畳して、各スポットからのシグナルがグリッドの各エレメントに中央に位置するようにする。各エレメント内の蛍光シグナルを統合し、シグナルの平均強度に対応する数値を得る。シグナル分析に用いるソフトウェアは、GEMTOOLS遺伝子発現分析プログラム(Incyte)である。

【0278】

1.1 相補的ポリヌクレオチド

NHRECをコードする配列或いはその任意の一部に対して相補的な配列は、天然のNH

50

R E C の発現を低下させるため即ち阻害するために用いられる。約 15 ～ 約 30 個の塩基対を含むオリゴヌクレオチドの使用について記すが、より小さな或いはより大きな配列の断片の場合でも本質的に同じ方法を用いることができる。O l i g o 4 . 0 6 ソフトウェア (N a t i o n a l B i o s c i e n c e s) 及び N H R E C のコーディング配列を用いて、適切なオリゴヌクレオチドを設計する。転写を阻害するためには、最も独特な 5' 配列から相補的なオリゴヌクレオチドを設計し、これを用いてプロモーターがコーディング配列に結合するのを阻害する。翻訳を阻害するためには、相補的なオリゴヌクレオチドを設計して、リボソームが N H R E C をコードする転写物に結合するのを阻害する。

【0279】

12 N H R E C の発現

N H R E C の発現及び精製は、細菌若しくはウイルスを基にした発現系を用いて行うことができる。細菌で N H R E C が発現するために、抗生物質耐性及び c D N A の転写レベルを高める誘導性のプロモーターを含む好適なベクターに c D N A をサブクロニングする。このようなプロモーターには、l a c オペレーター調節エレメントに関連する T 5 または T 7 バクテリオファージプロモーター及び t r p - l a c (t a c) ハイブリッドプロモーターが含まれるが、これらに限定されるものではない。組換えベクターを、B L 2 1 (D E 3) などの好適な細菌宿主に形質転換する。抗生物質耐性をもつ細菌が、イソプロピル β -D チオガラクトピラノシド (I P T G) で誘発されると N H R E C を発現する。真核細胞での N H R E C の発現は、昆虫細胞株または哺乳動物細胞株に一般にバキュロウイルスとして知られている A u t o g r a p h i c a c a l i f o r n i c a 核多面性ウイルス (A c M N P V) を感染させて行う。バキュロウイルスの非必須ポリヘドリン遺伝子を、相同組換え或いは転移プラスミドの媒介を伴う細菌の媒介による遺伝子転移のどちらかによって、N H R E C をコードする c D N A と置換する。ウイルスの感染力は維持され、強いポリヘドリンプロモーターによって高いレベルの c D N A の転写が行われる。組換えバキュロウイルスは、多くの場合は S p o d o p t e r a f r u g i p e r d a (S f 9) 昆虫細胞に感染に用いられるが、ヒト肝細胞の感染にも用いられることもある。後者の感染の場合は、バキュロウイルスの更なる遺伝的変更が必要になる。(例えば、E n g e l h a r d , E . K . 他 (1994) P r o c . N a t l . A c a d . S c i . U S A 91:3224-3227; S a n d i g , V . 他 (1996) H u m . G e n e T h e r . 7:1937-1945. を参照)。

【0280】

殆どの発現系では、N H R E C が、例えばグルタチオン S トランスフェラーゼ (G S T) 、または F L A G や 6 - H i s などのペプチドエピトープ標識で合成された融合タンパク質となるため、未精製の細胞溶解物からの組換え融合タンパク質の親和性ベースの精製が素早く 1 回で行うことができる。S c h i s t o s o m a j a p o n i c u m からの 26 キロダルトンの酵素 G S T によって、タンパク質の活性及び抗原性を維持した状態で固定されたグルタチオンで融合タンパク質の精製が可能となる (A m e r s h a m P h a r m a c i a B i o t e c h) 。精製の後、G S T 部分を特定の操作部位で N H R E C からタンパク質分解的に切断できる。アミノ酸 8 個のペプチドである F L A G で、市販のモノクローナル及びポリクローナル抗 F L A G 抗体 (E a s t m a n K o d a k) を用いた免疫親和性の精製が可能となる。6 個の連続するヒスチジン残基のストレッチである 6 - H i s によって、金属キレート樹脂 (Q I A G E N) で精製が可能となる。タンパク質の発現及び精製の方法は、A u s u b e l (1995, 前出, c h 10, 16) に記載されている。これらの方法で精製した N H R E C を直接用いて以下の実施例 16 および 17 のアッセイを行うことができる。

【0281】

13 機能のアッセイ

N H R E C の機能は、哺乳動物細胞培養系において生理学的に高められたレベルでの N H R E C をコードする配列の発現によって評価する。c D N A を、c D N A を高いレベルで発現する強いプロモーターを含む哺乳動物発現ベクターにサブクロニングする。このよ

10

20

30

40

50

うなベクターには、pCMV SPORTTM (Life Technologies.) 及び pCR 3.1 (Invitrogen, Carlsbad, CA) が含まれ、どちらもサイトメガロウイルスプロモーターを含んでいる。5～10 µg の組換えベクターを、例えば内皮由来か造血由来のヒト細胞株にリポソーム製剤或いは電気穿孔法によって一時的に形質移入する。更に、標識タンパク質をコードする配列を含む 1～2 µg のプラスミドを同時に形質移入する。標識タンパク質の発現により、形質移入された細胞と形質移入されていない細胞とを区別できる。また、標識タンパク質の発現によって、cDNA の組換えベクターからの発現を正確に予想できる。このような標識タンパク質には、緑色蛍光タンパク質 (GFP; Clontech)、及び CD64 または CD64-GFP 融合タンパク質が含まれる。レーザー光学に基づいた技術を利用した自動流動細胞計測法 (FCM) を用いて、GFP または CD64-GFP を発現する形質移入された細胞を同定し、その細胞のアポトーシス状態や他の細胞特性を評価する。また、FCM で、先行した或いは同時の細胞死の現象を診断する蛍光分子の取り込みを検出して計量する。これらの現象には、プロピジウムヨウ化物での DNA の染色によって計測される核 DNA 内容物の変化と、プロモデオキシウリジンの取り込み量の低下によって計測される DNA 合成の下方調節と、特異的な抗体との反応性によって計測される細胞表面及び細胞内のタンパク質の発現の変化と、蛍光複合アネキシン V タンパク質の細胞表面への結合によって計測される原形質膜組成の変化とが含まれる。流動細胞計測法は、Ormerod, M. G. による (1994) Flow Cytometry Oxford, New York, NY. に記載されている。

10

20

【0282】

遺伝子発現における NHREC の影響は、NHREC をコードする配列と CD64 または CD64-GFP のどちらかが形質移入された高度に精製された細胞集団を用いて評価することができる。CD64 または CD64-GFP は形質転換された細胞表面で発現し、ヒト免疫グロブリン G (IgG) の保存された領域と結合する。形質転換された細胞と形質転換されない細胞とは、ヒト IgG か CD64 に対する抗体のどちらかで被覆された磁気ビードを用いて分離することができる (DYNALE, Lake Success, NY)。mRNA は、当分野で周知の方法で細胞から精製することができる。NHREC 及び目的の他の遺伝子をコードする mRNA の発現は、ノーザン分析やマイクロアレイ技術で分析することができる。

30

【0283】

1.4 NHREC に特異的な抗体の作製

ポリアクリルアミドゲル電気泳動法 (PAGE; 例えば, Harrington, M. G. (1990) Methods Enzymol. 1816-3088-495 を参照) または他の精製技術で実質的に精製された NHREC を用いて、標準的なプロトコルでウサギを免疫化して抗体を作り出す。

【0284】

別法では、NHREC アミノ酸配列を LASERGENE ソフトウェア (DNASTAR) を用いて解析して免疫原性の高い領域を決定し、対応するオリゴペプチドを合成してこれを用いて当業者に周知の方法で抗体を生産する。C 末端付近の、或いは隣接する親水性領域内のエピトープなどの適切なエピトープの選択については、当分野で周知である (例えば、前出の Ausubel, 1995, 11 章を参照)。

40

【0285】

通常、約 15 残基の長さのオリゴペプチドを、Applied Biosystems の ABI 431A ペプチドシンセサイザー (PE Biosystems) を用いて fmoc 法のケミストリにより合成し、N-マレイミドベンゾイル-N-ヒドロキシスクシンイミドエステル (MBS) を用いた反応により KLH (Sigma-Aldrich, St. Louis MO) に結合させて、免疫原性を高める (例えば、前出の Ausubel, 1995 を参照)。フロイントの完全アジュバントにおいてオリゴペプチド-KLH 複合体を用いてウサギを免疫化する。得られた抗血清の抗ペプチド活性及び抗 NHRE

50

C活性を検査するには、ペプチドまたはNHRECを基板に結合し、1%BSAを用いてブロッキング処理し、ウサギ抗血清と反応させて洗浄し、さらに放射性ヨウ素標識されたヤギ抗ウサギIgGと反応させる。

【0286】

15 特異的抗体を用いる天然NHRECの精製

天然NHREC或いは組換えNHRECを、NHRECに特異的な抗体を用いるイムノアフィニティークロマトグラフィーにより実質的に精製する。イムノアフィニティークラムは、CNBr-活性化SEPHAROSE (Amersham Pharmacia Biotech) のような活性化クロマトグラフィー用レジンと抗NHREC抗体とを共有結合させることにより形成する。結合の後、そのレジンを製造者の使用説明書に従ってブロッ
10

【0287】

NHRECを含む培養液をイムノアフィニティークラムに通し、NHRECを優先的に吸着できる条件で（例えば、界面活性剤の存在下において高イオン強度のバッファーで）そのカラムを洗浄する。そのカラムを、抗体とNHRECとの結合を切るような条件で（例えば、pH2~3のバッファー、或いは高濃度の尿素またはチオシアン酸塩イオンのようなカオトロピックイオンで）溶出させ、NHRECを回収する。

【0288】

16 NHRECと相互作用する分子の同定

NHRECまたは生物学的に活性なその断片を、¹²⁵I Bolton-Hunter 試薬（例えば、Bolton A. E. 及びW. M. Hunter (1973) Biochem. J. 133:529を参照）で標識する。マルチウェルプレートに予め配列しておいた候補の分子を、標識したNHRECと共にインキュベートし、洗浄して、標識したNHREC複合体を有する全てのウェルをアッセイする。様々なNHREC濃度で得られたデータを用いて、候補分子と結合したNHRECの数量及び親和性、会合について
20

【0289】

別法では、NHRECと相互作用する分子を、Fields, S. 及びO. Song (1989, Nature 340:245-246) に記載の酵母2-ハイブリッドシステム (yeast two-hybrid system) やMATCHMAKERシステム (Clontech) などの2-ハイブリッドシステムに基づいた市販のキットを用いて分析する。
30

【0290】

NHRECはまた、ハイスループット型の酵母2-ハイブリッドシステムを使用するPATHCALLINGプロセス (CuraGen Corp., New Haven CT) に用いて、遺伝子の2つの大きなライブラリによってコードされるタンパク質間の全ての相互作用を決定することができる (Nandabalan, K. 他 (2000) 米国特許第6,057,101号)。

【0291】

17 NHREC活性の実証

NHRECの受容体活性のアッセイは、リガンド/受容体仲介性の細胞増殖の調節についてのプロトタイプのアッセイに基づいている。このアッセイは、NHRECを発現するシスマウス3T3細胞において新規に合成されたDNAの量を測定する。NHRECをコードするcDNAを含む好適な哺乳動物発現ベクターを、当分野で周知のトランスフェクション法を用いて静止状態の3T3培養細胞に加える。形質転換細胞を、³Hチミジン及び様々な量のNHRECリガンドの存在化でインキュベートする。酸性沈殿DNA内へ取り込まれた³Hチミジンを、トリチウムラジオアイソトープカウンタを用いて適当な時間間隔で測定する。組み込まれた量が新規に合成されたDNAの量に正比例する。少なくとも100倍のNHRECリガンド濃度範囲に対する用量反応曲線が直線であることは受容体活性を示唆するものである。1ml当たりの活性単位1ユニットは、50%
40
50

の反応レベルを示すNHRECの濃度と定義し、100%の応答レベルは $[^3\text{H}]$ チミジンの酸溶解性DNA内への最大取り込み量を示すものである（例えばMcKay, I. 及び I. Leigh, eds. (1993) Growth Factors: A Practical Approach, Oxford University Press, New York NY, p. 73参照）。

【0292】

別法では、NHRECの受容体活性は、COS7、HeLa、又はCHO等の哺乳動物細胞系のNHRECをコードする真核生物発現ベクターでの形質転換による細胞増殖の増大によって測定する。真核生物発現ベクターは市販されており、これらのベクターを細胞に導入する技術は当分野で周知である。形質転換した後、細胞系のNHRECの発現が許容される条件下でこの形質転換細胞を48～72時間インキュベートする。次に、位相差顕微鏡で形質転換細胞とコントロール細胞の分裂指数を比較する。分裂指数の上昇は、NHRECの受容体活性を示す。

10

【0293】

別法では、NHRECの受容体活性は、 ^{125}I 標識したNHRECの存在下で候補リガンド分子を用いるリガンド結合アッセイで測定する。NHRECを ^{125}I Bolton-Hunter試薬で標識する（Bolton and Hunter、前出）。候補リガンド分子をマルチウェルプレートの各ウェルに注入し、標識したNHRECと共にインキュベートし、次に洗浄し、標識したNHREC複合体が存在するウェルをアッセイする。様ような濃度のNHRECを用いて得たデータを利用して、リガンド分子と結合したNHRECの数、親和性、及び関連を測定する。測定した結合レベルがNHREC活性のレベルに比例する。

20

【0294】

NHRECの転写調節活性を、そのレポーター遺伝子の転写を刺激する能力によって測定する（Liu, H. Y. 他 (1997) EMBO J. 16 : 5289-5298）。このアッセイは、十分に特徴付けられたレポーター遺伝子作成物であるLexA_p-LacZを用いる。このLexA_p-LacZは、大腸菌LacZ酵素をコードする配列に結合されたLexA DNA転写調節エレメント（LexA_p）からなる。融合遺伝子の作成及び発現、融合遺伝子の細胞への導入、及びLacZ酵素活性の測定方法は当分野で周知である。NHRECをコードする配列を、LexA転写因子に由来するDNA結合ドメインとNHRECとからなる融合タンパク質、LexA-NHRECの合成を誘導するプラスミドにクローニングする。得られたLexA-NHREC融合タンパク質をコードするプラスミドを、LexA_p-LacZレポーター遺伝子を含むプラスミドと共に酵母細胞内に導入する。コントロール細胞に対するLexA-NHREC形質転換細胞に関連するLacZ酵素活性の量がNHRECによって刺激された転写の量に比例する。

30

【0295】

当業者は、本発明の範囲及び精神から逸脱することなく本発明の記載した方法及びシステムの種々の改変を行うことができるであろう。特定の好適な実施例に基づいて本発明を説明したが、本発明の範囲が、そのような特定の実施例に不当に制限されるべきではないことを理解されたい。実際に、分子生物学或いは関連する分野の専門家には明らかな、本明細書に記載の本発明の実施例の様々な改変は、特許請求の範囲に含まれる。

40

【0296】

（表の簡単な説明）

表1は、本発明の完全長ポリヌクレオチド配列及びポリペプチド配列に対する系統的な名称を示す。

【0297】

表2は、本発明のポリペプチドに最も近いGenBankの相同体のGenBankの識別番号およびアノテーションを示す。ポリペプチドとそのGenBankの相同体との間の一致を表す確率値スコアも示す。

50

【0298】

表3は、推定上のモチーフおよびドメインを含むポリペプチド配列の構造的な特徴、並びにポリペプチドの分析に用いた方法、アルゴリズム、および検索可能なデータベースを示す。

【0299】

表4は、ポリヌクレオチド配列の組み立てに用いたcDNA断片およびゲノムDNA断片のリスト、並びにポリヌクレオチド配列の選択された断片のリストを示す。

【0300】

表5は、本発明のポリヌクレオチドの代表的なcDNAライブラリを示す。

【0301】

表6は、表5に示すcDNAライブラリの作製に用いた組織およびベクターを示す付録である。

【0302】

表7は、本発明のポリヌクレオチド及びポリペプチドの分析に用いたツール、プログラム、及びアルゴリズム、並びにその説明、引用文献、閾値パラメータを示す。

【表1】

表1

インサイト プロジェクトID	ポリベブチド SEQ ID NO:	インサイト ポリベブチドID	ポリヌクレオチド SEQ ID NO:	インサイト ポリヌクレオチドID
3462174	1	3462174CD1	4	3462174CB1
1322075	2	1322075CD1	5	1322075CB1
7472269	3	7472269CD1	6	7472269CB1

10

20

30

40

表2

ポリペプチド SEQ ID NO:	インサイトID ポリペプチドID	GenBank ID NO:	確率スコア	GenBank 相同体
2	1322075CD1	g2088668	8.40E-287	Achlva ambisexualis アンヘリクサー・ルックス・ロイド受容体に類似 Wilson, R. 5 (1994) Nature 358:32-38
3	7472269CD1	g868032	3.80E-236	farnesoid X活性化受容体 (ドブネズミ)

10

20

30

40

表3

SEQ ID NO.	インサイト ポリペプチド ID	アミノ酸 残基数	潜在的リン酸化部位	潜在的 グリコシル化部位	シグネチャ(signature)配列、モチーフ、 及びドメイン	分析方法及び データベース
1	3462174CD1	623	S48 S222 T309 S353 S377 S408 S170 S215 T295 T309 T331 S415 T552		ロインジッパー: L513-E540	MOTIFS
					プログステロン受容体 PD033811: A52-G139 (P-value=0.0077)	BLAST-PRODOM
					プログステロン受容体 DM07164 P07812 1-408: E33-G283 (P-value=0.0041)	BLAST-DMC
					Achlya ambisexualisアンセリジオールステロイド受容体 に類似したATP結合タンパク質: PD24759:L189-D375, PD014232:E523-R655, PD007775:C255-D488, PD041582:Y107-I188 ナトリウム:アラニン共輸送体: BL00873D:L95-Y107 ATP/GTP結合部位 (P-loop): A289-S296 RBDドメイン: R966-9967 ドーパミン受容体シグネチャモチーフ: PR00242E:A766-S780	BLIMES-BLOCKS MOTIFS MOTIFS BLIMES-PRINTS
2	1322075CD1	1025	T223 S231 S241 T321 T442 T459 S464 S541 T608 T666 S679 T736 T751 T753 S809 T932 T110 S156 S193 S272 T275 T419 S424 T451 S456 S468 T686 T716 T929 S987 S990 S1005 Y335 Y708 Y963	N82 N108 N446 N491 N982	ピタミンド受容体シグネチャ PR00350: C127-T143, C144-C163, A388-E407, H434-M457 ステロイドホルモン受容体 PR00398: C127-C144, K148-K164, C169-C187, H299-H318 受容体タンパク質X転写調節DNA結合 核Znフィンガー-レチノイド相互作用: PD020886:N5-E125 核ホルモン受容体DNA結合領域 DM00047 I49018 128-413:S117-E407 ステロイドフィンガー:C127-R152 Znフィンガー、C4型 (2ドメイン): E125-G200 核ホルモン受容体リガンド結合ドメイン hormone_rec: E294-N449 核ホルモン受容体 BL00031: C127-A159, Y161-M192	BLAST-PRODOM BLAST-PRODOM BLIMES-BLOCKS MOTIFS HMMER-PFAM HMMER-PFAM BLIMES-BLOCKS
					核ホルモン受容体DNA結合領域シグネチャ ステロイドフィンガー: P107-M173 C4型ステロイド受容体 PR00047: C127-T143, T143-N158, R176-L184, L184-M192	PROFILESAN BLIMES-PRINTS
3	7472269CD1	477	T15 S20 T223 S242 S284 S328 S376 T383 T109 S209 T239 T274 T307 S458 Y97			BLIMES-PRINTS BLIMES-PRINTS BLAST-PRODOM BLAST-DMC MOTIFS HMMER-PFAM HMMER-PFAM BLIMES-BLOCKS

10

20

30

40

表4

ポリヌクレオチド SEQ ID NO.	インサイト ポリヌクレオチド ID	配列長	選択された断片	配列断片	5' 位置	3' 位置
4	3462174CB1	2274	2139-2274, 1389-1446	6812329H1 (ADRETUR01)	1808	2274
				92368743	1	353
				6496889H1 (COLNNOT41)	644	913
				3462174H1 (293TF2T01)	1756	597
				6196653H1 (PITUNON01)	1116	1598
				94096054.v113.gb.2.nt	247	2158
				1322075F6 (BLADNOT04)	2044	2828
				1476875F6 (CORPNOT02)	2713	3507
				1337114X16 (COLNNOT13)	1701	2598
				1337114X23H1 (COLNNOT13)	1165	1814
5	1322075CB1	4154	1-1754, 2376-2592, 4134-4154	3591082H1 (293TF5T01)	1	303
				222775R1 (PANCNOT01)	2999	3837
				1337114X11 (COLNNOT13)	7	702
				311188H1 (LUNGNOT02)	3701	4154
				1337114X12 (COLNNOT13)	320	1014
				267166R1 (HNT2NOT01)	656	1596
				9565350	1705	2220
				60146067D1	1	284
				71218724V1	1429	2033
				70831194V1	566	1262
6	7472269CB1	2220	1061-1143, 2152-2220, 387-414	71218077V1	970	1429
				70831860V1	1411	1926
				70831870V1	1497	2154
				71220470V1	207	817
				6631050H1 (LIVRNOT21)	336	824

10

20

30

40

表5

ポリヌクレオチド SEQ ID NO:	インサイト プロジェクトID	代表的ライブラリ
4	3462174CB1	ADREURO1
5	1322075CB1	COLNNO113
6	7472269CB1	STOMNO101

10

20

30

40

表6

ライブラリ	ベクター	ライブラリの説明
ADRETUR01	PCDNA2.1	このランダムプライムライブラリは、52歳白人男性の腎尿管切除および腎病変の局部破壊の際に採取した副腎腫瘍組織の左上極から単離したRNAを用いて作製した。病理学的には、副腎実質のほぼ全体および隣接した脂肪組織に及ぶ腫瘍を形成したグレード3の副腎皮質腺腫であった。転移性腫瘍結節が肺門領域に見られた。腎静脈にも腫瘍が及び、腎静脈の切除した縁にも腫瘍の発生が見られた。副腎皮質腺腫および血栓の断片が下大静脈に見られた。既往症には、過度の体重の減少があった。家族歴には、皮膚癌、1型糖尿病、および神経性の抑うつ症が含まれていた。
COLNNOT13	pINCY	このライブラリは中程度慢性潰瘍性大腸炎の28歳の白人男性の上行結腸結腸より単離したRNAを用いて作製された。
STOMN0T01	PELUESCRIPT	このライブラリは、心肺停止により死亡した55歳の白人男性の胃組織から単離したRNAを用いて作製した。

10

20

30

40

表7-1

プログラム名	説明	引用文献	パラメーター閾値
ABI FACTURA	核酸配列においてベクター配列を除去して不特定の塩基をマスクするプログラム。	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	Fast Data Finder は、アミノ酸または核酸配列の比較及び注釈付け (annotation) に有用である。	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	不一致<90%
ABI AutoAssembler BLAST	核酸配列を構築するプログラム。 Basic Local Alignment Search Tool は、アミノ酸及び核酸配列の配列類似性検索に有用であり、blastp 及び blastn、blastx、tblastn、tblastx の 5 つのファンクションがある。	Applied Biosystems, Foster City, CA. Altschul, S.F. 他 (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. 他 (1997) Nucleic Acids Res. 25: 3389-3402.	ESTs: 確率値=1.0E-8 以下 完全長配列: 確率値=1.0E-10 以下
FASTA	Pearson 及び Lipman アルゴリズムは、問い合わせの配列と同種の配列群との類似性を検索する。FASTA は、fasta 及び tfasta、fastx、tfastx、sssearch の少なくとも 5 つのファンクションを含む。	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E 値=1.06E-6 構築された ESTs: fasta 同一性=95%以上、一致長さ=200 塩基以上、fastx E 値=1.0E-8 以下 完全長配列: fastx スコア=100 以上
BLIMPS	BLOCKS IMPROVED Searcher は、BLOCKS 及び PRINTS、DOMO、PRODOM、PFAM データベースにおける配列に対して、ある配列の一致性を調べ、遺伝子ファミリー及び配列相同性、構造的フィンガープリント領域を探索する。	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.C.; Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; Artwood, I.K. 他 (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	確率値=1.0E-3 以下
HMMER	PFAM などのタンパク質ファミリーコンセンサス配列の隠れマルコフモデル(HMM)に基づいたデータベースに対して問い合わせ配列を検索するアルゴリズム。	Krogh, A. 他 (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. 他 (1998) Nucleic Acids Res. 26:320-322. Durbin, R. 他 (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	PFAM ヒット: 確率値=1.0E-3 以下 シグナルペプチドヒット: スコア=0 以上

【表 8】

表7-2

プログラム名	説明	引用文献	パラメーター閾値
ProfileScan	Prosite で定義された配列パターンと一致するタンパク質配列における構造及び配列のモチーフを検索するアルゴリズム。	Gribskov, M. 他 (1988) CABIOS 4:61-66; Gribskov, 他 (1989) Methods Enzymol. 183:146-159; Bairoch, A. 他 (1997) Nucleic Acids Res. 25:217-221	標準化された質のスコア≧特定の Prositeモチーフに対する GCG 指定 "HIGH" 値 通常、スコア=1.4-2.1
Phred	高い感度及び確率で自動配列決定機のトレースを調べる塩基読出しアルゴリズムである。	Ewing, B. 他 (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	Smith-Waterman アルゴリズムの効率的なインプリメンテーションに基づく SWAT や CrossMatch を含む Phils Revised Assembly プログラムであって、配列相同性の検索及び DNA 配列の構築に有用である。	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; Green, P., University of Washington, Seattle, WA.	スコア=120 以上 一致長さ=56 以上
Consed	Phrap で構築したものの表示及び編集をするためのグラフィックツールである。	Gordon, D. 他 (1998) Genome Res. 8:195-202.	
SPScan	タンパク質配列をスキヤンして分泌シグナルペプチドの存在を調べる重み付けマトリクス解析プログラムである。	Nielson, H. 他 (1997) Protein Engineering 10: 1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	スコア=3.5 以上
TMAP	重み付けマトリクスを用いたタンパク質配列上の膜貫通セグメントの明確化および向きを決定するためのプログラム	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	隠れマルコフモデル (HMM) を用いたタンパク質配列上の膜貫通セグメントの明確化および向きを決定するためのプログラム。	Sonnhammer, E.L. 他 (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow 他 eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
Motifs	Prosite で定義された配列と一致したパターンについてアミノ酸配列を検索するプログラムである。	Bairoch 他, (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

10

20

30

40

【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
2 August 2001 (02.08.2001)

PCT

(10) International Publication Number
WO 01/55392 A2(51) International Patent Classification: **C12N 15/12**,
15/10, C07K 14/705, 16/28, 14/72, C12Q 1/68, A61K
38/17, G01N 33/53, 33/50

(21) International Application Number: PCT/US01/02741

(22) International Filing Date: 26 January 2001 (26.01.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/178,578 28 January 2000 (28.01.2000) US
60/185,079 25 February 2000 (25.02.2000) US
60/189,999 17 March 2000 (17.03.2000) US(63) Related by continuation (CON) or continuation-in-part
(CIP) to earlier applications:US 60/178,578 (CON)
Filed on 28 January 2000 (28.01.2000)
US 60/185,079 (CON)
Filed on 25 February 2000 (25.02.2000)
US 60/189,999 (CON)
Filed on 17 March 2000 (17.03.2000)(71) Applicant (for all designated States except US): INCYTE
GENOMICS, INC. (US/US); 3160 Porter Drive, Palo
Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **MURFORD, Neil**
(GB/US); 105 Wilwood Circle, Durham, CT 06422 (US);
BAUGIN, Mariah, R. (US/US); 14244 Santiago Road,
San Leandro, CA 94577 (US); **TANG, Y. Tom** (CN/US);4230 Ranwick Court, San Jose, CA 95118 (US); **BAND-
MAN, Olga** (US/US); 366 Anna Avenue, Mountain View,
CA 94043 (US); **TRIBOULEY, Catherine, M.** (FR/US);
1121 Tennessee Street, #5, San Francisco, CA 94107 (US);
DAS, Debopriya (IN/US); 1179 Bonita Avenue, Apt. 3,
Mountain View, CA 94040 (US); **FOLICKY, Jennifer,
L.** (US/US); 4864 National Avenue, San Jose, CA 95124
(US); **LU, Terence, P.** (CA/US); 1452 Beach Park Boulle-
vard, Apt. 115, Foster City, CA 94040-1941 (US).(74) Agents: **HAMLET-COX, Diana** et al.; Incyte Genomics,
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,
DE, DK, DM, DZ, EP, ES, FI, GB, GR, GP, 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, TR,
TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.(84) Designated States (regional): ARIPO patent (GH, 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, IT,
LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— with an international search report and to be republished
upon receipt of that reportFor two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

WO 01/55392 A2

(54) Title: NUCLEAR HORMONE RECEPTORS

(57) Abstract: The invention provides human nuclear hormone receptors (NHRs) and polynucleotides which identify and encode NHRs. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating or preventing disorders associated with aberrant expression of NHRs.

WO 01/55392

PCT/US01/02741

NUCLEAR HORMONE RECEPTORS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of nuclear hormone receptors and to the use of these sequences in the diagnosis, treatment, and prevention of metabolic, endocrine, reproductive, developmental, immunological, and cell proliferative disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of nuclear hormone receptors.

BACKGROUND OF THE INVENTION

The nuclear hormone receptors, also known as the nuclear receptors or the intracellular receptors, constitute a protein superfamily whose members are both receptors and transcriptional regulators. Nuclear hormone receptors rely on both their receptor function and their transcriptional regulatory function to affect a broad array of biological processes, including development, homeostasis, cell proliferation, and cell differentiation. (Reviewed in Mangelsdorf, D.J. et al. (1995) *Cell* 83:835-840; Wen, D.X. and D.P. McDonnell (1995) *Curr. Opin. Biotechnol.* 6:582-589; Perlmann, T. and R.M. Evans (1997) *Cell* 90:391-397; Tenhagen, S. and A. Baniahmad (1997) *Int. J. Biochem. Cell Biol.* 29:1325-1341; Moras, D. and H. Gronemeyer (1998) *Curr. Opin. Cell Biol.* 10:384-391; Willy, P.J. and D.J. Mangelsdorf (1998) in: Hormones and Signaling (ed: B.W. O'Malley) vol. 1, Academic Press, San Diego CA, pp. 307-358; Weatherman, R.V. et al. (1999) *Annu. Rev. Biochem.* 68:559-581.)

Nuclear hormone receptors as receptors

Generally, the term receptor describes a protein that specifically recognizes other molecules. As receptors, nuclear hormone receptors specifically recognize and bind to their cognate ligands. Although nuclear hormone receptors are located intracellularly, many receptors are extracellular cell surface proteins which bind extracellular ligands. Such extracellular receptors produce cellular responses affecting growth, differentiation, endocytosis, and the immune response. Other receptors facilitate the selective transport of proteins out of the endoplasmic reticulum and localize enzymes to particular regions of the cell. Transcriptional regulation by nuclear hormone receptors, propagation of cellular signals by extracellular receptors, and transport and localization of proteins by other receptors, all rely upon specific interactions between the receptors and a variety of cellular components. In many cases, the identity of the cognate ligand to which a receptor binds is unknown. Such receptors are termed orphan receptors. This term also applies to those nuclear hormone receptors which carry out their transcriptional regulatory functions without binding any ligands.

35

WO 01/55392

PCT/US01/02741

Nuclear hormone receptors as transcriptional regulators

Multicellular organisms are comprised of diverse cell types that differ dramatically both in structure and function. The identity of a cell is determined by its characteristic pattern of gene expression, and different cell types express overlapping but distinctive sets of genes throughout development. Spatial and temporal regulation of gene expression is critical for the control of cell proliferation, cell differentiation, apoptosis, and other processes that contribute to organismal development. As transcriptional regulators, nuclear hormone receptors play key roles in controlling these fundamental biological processes. Other transcriptional regulators affect gene expression in response to extracellular signals that mediate cell-cell communication and that coordinate the activities of different cell types.

In general, transcriptional regulators such as nuclear hormone receptors initiate, activate, repress, or terminate gene transcription by binding to the promoter, enhancer, and upstream regulatory regions of a gene in a sequence-specific manner. However, some transcriptional regulators bind regulatory elements within or downstream of a gene's coding region. Transcriptional regulatory proteins may bind to a specific region of DNA singly, or in a complex with other accessory factors. (Reviewed in Lewin, B. (1990) in: Genes IV, Oxford University Press, New York NY, and Cell Press, Cambridge MA, pp. 554-570.)

Mechanism of nuclear hormone receptor function

In the unliganded state, a nuclear hormone receptor exists in association with a multiprotein complex of chaperones, including heat shock proteins such as hsp90 and immunophilins such as hsp56. These chaperones maintain the ligand-free receptor in an inactive state which is amenable to binding of free ligand, and prevent the ligand-free receptor from translocating to the nucleus. Upon activation by its cognate ligand, the receptor may form a homodimer or heterodimer which translocates to the nucleus, binds to specific DNA sequences, and exerts its transcriptional regulatory function. In order to effectively carry out its regulatory roles, an activated nuclear hormone receptor dissociates from a histone deacetylase-containing corepressor complex and associates with a histone acetyltransferase-containing coactivator complex (Xu, L. et al. (1999) *Curr. Opin. Genet. Dev.* 9:140-147). The association of the activated receptor with coactivator proteins results in remodeling of chromatin so that it adopts an open transcriptionally active state, providing access to the transcriptional regulatory elements of the activated nuclear receptor (Lemon, B.D. and L.P. Freedman (1999) *Curr. Opin. Genet. Dev.* 9:499-504).

Structure of nuclear hormone receptors

Nuclear hormone receptors function as signal transducers by converting hormonal signals into transcriptional responses. In general, nuclear hormone receptors consist of a variable amino-terminal domain, a highly conserved DNA-binding domain, and a conserved C-terminal ligand-

WO 01/55392

PCT/US01/02741

binding domain. In the steroid-binding nuclear hormone receptors, the amino-terminal domain harbors a trans-activation element termed AF-1. Some nuclear hormone receptors also contain a trans-activation element in the ligand-binding domain termed AF-2. The DNA-binding and ligand-binding domains of nuclear hormone receptors may contain dimerization elements, and the DNA-binding domain may contain a nuclear localization signal (Weatherman, R.V. et al. (1999) *Annu. Rev. Biochem.* 68:559-581).

The DNA-binding domain of nuclear hormone receptors is composed of two zinc finger motifs which mediate recognition of specific DNA sequences. A zinc finger motif contains periodically spaced cysteine and histidine residues which coordinate Zn^{2+} . Examples of this sequence pattern include the C2H2-type, C4-type, and C3HC4-type ("RING" finger) zinc fingers, and the PHD domain (Lewin, *supra*; Aasland, R. et al. (1995) *Trends Biochem. Sci.* 20:56-59). A zinc finger motif contains an α helix and an antiparallel β sheet whose proximity and conformation are maintained by the zinc ion. Contact with DNA is made by the arginine preceding the α helix and by the second, third, and sixth residues of the α helix. Zinc finger motifs may be repeated in a tandem array within a protein such that the α helix of each zinc finger in the protein makes contact with the major groove of the DNA double helix. This repeated contact between the protein and the DNA produces a strong and specific DNA-protein interaction. The strength and specificity of the interaction can be regulated by the number of zinc finger motifs within the protein. Although zinc fingers were originally identified in DNA-binding proteins as regions that interact directly with DNA, they have since been found in proteins that do not bind to DNA. (See, e.g., Lodish, H. et al. (1995) *Molecular Cell Biology*, Scientific American Books, New York NY, pp. 447-451.)

The ligand-binding domain of nuclear hormone receptors is responsible for binding to ligands, coactivator proteins, and corepressor proteins. This domain is composed of three layers of α helices, with the central layer consisting of two helices containing many hydrophobic side chains (Moras, D. and H. Gronemeyer (1998) *Curr. Opin. Cell Biol.* 10:384-391). These two central α helices thus create a hydrophobic pocket which is the site of ligand binding. A ligand bound in this hydrophobic ligand-binding site is completely buried inside the receptor protein and is not exposed to solvent. This suggests that large conformational changes in the ligand-binding domain would accompany binding of a ligand. One of the α helices of the ligand-binding domain provides many of the inter-subunit contacts in dimers of nuclear receptors. This α helix contacts the ligand when it is bound in the ligand-binding pocket, suggesting that ligand binding can affect formation of receptor dimers (Weatherman, R.V. et al. (1999) *Annu. Rev. Biochem.* 68:559-581).

Classes of nuclear hormone receptors

Nuclear hormone receptors can be grouped into three broad classes: the steroid receptors, the RXR-heterodimeric receptors, and the orphan nuclear hormone receptors. The steroid receptors bind

WO 01/55392

PCT/US01/02741

to steroid hormones, and this class includes the androgen receptor, mineralocorticoid receptor, estragen receptor, glucocorticoid receptor, and progesterone receptor. The RXR-heterodimeric receptors bind to nonsteroid ligands, and this class includes the thyroid hormone receptor, retinoic acid receptor, vitamin D receptor, ecdysone receptor, and peroxisome proliferator activated receptor.

- 5 The orphan nuclear hormone receptors include steroidogenic factor 1, nerve growth factor-induced receptor, and X-linked orphan receptor DAX-1.

The steroid hormone receptors are activated upon binding to specific steroid hormones. The conformational change induced by ligand binding leads to dissociation of the receptor from heat shock proteins and formation of receptor homodimers which recognize specific palindromic DNA sequences called hormone response elements (HREs). Upon binding to an HRE, a steroid hormone receptor homodimer can regulate the transcription of target genes.

For example, the progesterone receptor (PR) is a steroid hormone receptor which is activated by progesterone, a 4-pregnene-3,20-dione derived from cholesterol which is a critical oscillating component of the female reproductive cycle. These oscillations correlate with anatomical and morphological changes including menstruation and pregnancy. The activities of progesterone are mediated through PR. In the cytoplasm, PR associates with several other proteins and factors known as the PR heterocomplex. This heterocomplex includes heat shock proteins and immunophilins such as hsp70, hsp90, hsp27, p59 (hsp56), p48, and p23 (Johnson, J.L. et al. (1994) Mol. Cell. Biol. 14:1956-1963). Upon binding progesterone, activated PR translocates to the nucleus, binds to canonical DNA transcriptional elements, and regulates progesterone-regulated genes implicated in differentiation and the cell cycle (Moutsatsou, P and C.E. Sokeris (1997) Ann. N.Y. Acad. Sci. 816:99-115). The PR antagonist RU 486, which can be used to terminate a pregnancy, is an example of a commercial therapeutic targeted toward a steroid hormone receptor.

The RXR-heterodimeric nuclear receptors are distinguished from the steroid hormone receptors in that members of the former group bind to their target DNA sequences upon formation of heterodimers with retinoid X receptors (RXRs) (Mangelsdorf, D.J. and R.M. Evans (1995) Cell 83:841-850). Three different isoforms of RXR have been identified (Minucci, S. and K. Ozato (1996) Curr. Opin. Genet. Dev. 6:567-574). The retinoic acid receptors (RARs) are examples of RXR-heterodimeric nuclear receptors. Retinoic acid (RA) is a biologically active metabolite of vitamin A (retinol), a fat-soluble vitamin found mainly in fish liver oils, liver, egg yolk, butter, and cream. While 9-cis-RA binds to RARs and RXRs, all-trans-RA binds only to RXRs. RAR/RXR heterodimers bind with high affinity to specific DNA sequences known as retinoic acid response elements (RAREs), thus acting as regulators of RA-dependent transcription.

Peroxisome proliferator activated receptors (PPARs) are therapeutically important RXR-heterodimeric nuclear receptors which are induced by fatty acids and eicosanoids. There are three

WO 01/55392

PCT/US01/02741

known isotypes of PPAR, each with specific expression patterns, and these PPARs are involved in the regulation of genes involved in systemic homeostasis of glucose and lipids (Kiewer, S.A. and T.M. Willson (1998) *Curr. Opin. Genet. Dev.* 8:576-581; Michalik, L. and W. Wahli (1999) *Curr. Opin. Biotechnol.* 10:564-570). As such, PPARs are therapeutic targets for disorders such as diabetes, 5 dyslipidemia, and obesity (Smith, S.A. (1996) *Pharmacol. Rev. Commun.* 8:57-64; Willson, T.M. and W. Wahli (1997) *Curr. Opin. Chem. Biol.* 1:235-241; Barroso, I. et al. (1999) *Nature* 402:880-883).

The orphan nuclear receptors either have no known activating ligand, or can exert their transcriptional regulatory activities without benefit of ligand binding. For example, in *Caenorhabditis elegans*, the X-chromosome encoded nuclear hormone receptor homolog *SEX-1* regulates 10 transcription of the sex determination gene *xol-1* (Carmi, I. et al. (1998) *Nature* 396:168-173). Rather than relying on ligand binding, *SEX-1* acts as a transcriptional regulator in a dose-dependent manner, in effect controlling sexual differentiation through an X-chromosome-counting mechanism.

Some nuclear hormone receptors lack the conventional DNA-binding domain typically associated with the nuclear hormone receptor family. DAX-1 is one such nuclear hormone receptor 15 lacking the conventional DNA-binding domain, and mutations in DAX-1 have been shown to cause X-linked adrenal hypoplasia congenita (Zanaria, E.F. et al. (1994) *Nature* 372:635-641). DAX-1 is an orphan nuclear receptor which interacts directly with steroidogenic factor 1 (SF-1) (Ito, M. et al. (1997) *Mol. Cell. Biol.* 17:1476-1483), and DAX-1 is capable of modulating the action of SF-1 in sex-specific gene expression (Nachtigal, M.W. et al. (1998) *Cell* 445-454). SF-1 is an orphan nuclear 20 receptor which acts as a transcription factor for several steroidogenic enzyme genes in the adrenal gland and gonads (Lala, D.S. et al. (1992) *Mol. Endocrinol.* 6:1249-1258; Lynch, J.P. et al. (1993) *Mol. Endocrinol.* 7:776-786; Clemens, J.W. et al. (1994) *Endocrinology* 134:1499-1508), and can also regulate several genes expressed in pituitary gonadotrope cells (Barnhart, K.M. and P.L. Mellon (1994) *Mol. Endocrinol.* 8:878-885; Ingraham, H.A. et al. (1994) *Genes Dev.* 8:2302-2312; 25 Halvorson, L.M. et al. (1996) *J. Biol. Chem.* 271:6645-6650; Keri, R.A. and J.H. Nilson (1996) *J. Biol. Chem.* 271:10782-10785).

SF-1 also acts as a potent transactivator of small heterodimer partner (SHP; short heterodimer partner) (Lee, Y.K. et al. (1999) *J. Biol. Chem.* 274:20869-20873). SHP is another example of a nuclear hormone receptor lacking the conventional DNA-binding domain (Seol, W. et al. (1996) 30 *Science* 272:1336-1339; Lee, H.-K. et al. (1998) *J. Biol. Chem.* 273:14398-14402). SHP interacts with many members of the nuclear hormone receptor family, including retinoid receptors, estrogen receptor, thyroid hormone receptor, and the orphan receptor CAR. SHP acts as an inhibitor of estrogen receptor-mediated transcriptional activation by competing with coactivators for binding to 35 transactivation by the orphan receptor hepatocyte nuclear factor 4, and by retinoid X receptor (Lee,

WO 01/55392

PCT/US01/02741

Y.K. et al. (2000) Mol. Cell. Biol. 20:187-195).

Consequences of defective transcription regulation

Many neoplastic disorders in humans can be attributed to inappropriate gene expression.

- 5 Malignant cell growth may result from either excessive expression of tumor promoting genes or insufficient expression of tumor suppressor genes (Cleary, M.L. (1992) Cancer Surv. 15:89-104). Chromosomal translocations may also produce chimeric loci which fuse the coding sequence of one gene with the regulatory regions of a second unrelated gene. Such an arrangement likely results in inappropriate gene transcription, potentially contributing to malignancy.

- 10 In addition, the immune system responds to infection or trauma by activating a cascade of events that coordinate the progressive selection, amplification, and mobilization of cellular defense mechanisms. A complex and balanced program of gene activation and repression is involved in this process. However, hyperactivity of the immune system as a result of improper or insufficient regulation of gene expression may result in considerable tissue or organ damage. This damage is well documented in immunological responses associated with arthritis, allergens, heart attack, stroke, and infections. (See, e.g., Isselbacher et al. (1996) Harrison's Principles of Internal Medicine, 13/e, McGraw Hill, Inc. and Teron Data Systems Software.)

- Furthermore, the growth of multicellular organisms is based upon the induction and coordination of cell differentiation at the appropriate stages of development. Central to this process is differential gene expression, which confers the distinct identities of cells and tissues throughout the body. Failure to regulate gene expression during development could result in developmental disorders.

- The discovery of new nuclear hormone receptors and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of metabolic, endocrine, reproductive, developmental, immunological, and cell proliferative disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of nuclear hormone receptors.

SUMMARY OF THE INVENTION

- 30 The invention features purified polypeptides, nuclear hormone receptors, referred to collectively as "NHREC" and individually as "NHREC-1," "NHREC-2," and "NHREC-3." In one aspect, the invention provides an isolated polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active

WO 01/55392

PCT/US01/02741

fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-3.

5 The invention further provides an isolated polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of an amino acid sequence selected from the group
10 consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-3. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:4-6.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter
15 sequence operably linked to a polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ
20 ID NO:1-3, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide comprising an amino acid
25 sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of an amino acid sequence selected from the group
30 consisting of SEQ ID NO:1-3. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a
35 polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino

WO 01/55392

PCI/US01/02741

acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-3.

The invention further provides an isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

WO 01/55392

PCT/US01/02741

The invention further provides a composition comprising an effective amount of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional NHRHC, comprising administering to a patient in need of such treatment the composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional NHRHC, comprising administering to a patient in need of such treatment the composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with

WO 01/55392

PCT/US01/02741

overexpression of functional NHREC, comprising administering to a patient in need of such treatment the composition.

The invention further provides a method of screening for a compound that specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:4-6, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of i) a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-

WO 01/55392

PCT/US01/02741

6, ii) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, iii) a polynucleotide sequence complementary to i), iv) a polynucleotide sequence complementary to ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence selected from the group consisting of i) a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, ii) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, iii) a polynucleotide sequence complementary to i), iv) a polynucleotide sequence complementary to ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the present invention.

Table 2 shows the GenBank identification number and annotation of the nearest GenBank homolog for polypeptides of the invention. The probability score for the match between each polypeptide and its GenBank homolog is also shown.

Table 3 shows structural features of polypeptide sequences of the invention, including predicted motifs and domains, along with the methods, algorithms, and searchable databases used for analysis of the polypeptides.

Table 4 lists the cDNA and genomic DNA fragments which were used to assemble polynucleotide sequences of the invention, along with selected fragments of the polynucleotide sequences.

Table 5 shows the representative cDNA library for polynucleotides of the invention.

Table 6 provides an appendix which describes the tissues and vectors used for construction of the cDNA libraries shown in Table 5.

Table 7 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.

WO 01/55392

PCT/US01/02741

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"NHREC" refers to the amino acid sequences of substantially purified NHREC obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of NHREC. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of NHREC either by directly interacting with NHREC or by acting on components of the biological pathway in which NHREC participates.

An "allelic variant" is an alternative form of the gene encoding NHREC. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

WO 01/55392

PCT/US01/02741

"Altered" nucleic acid sequences encoding NHREC include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as NHREC or a polypeptide with at least one functional characteristic of NHREC. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding NHREC, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding NHREC. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent NHREC. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of NHREC is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of NHREC. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of NHREC either by directly interacting with NHREC or by acting on components of the biological pathway in which NHREC participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, Fab', and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind NHREC polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired.

WO 01/55392

PCT/US01/02741

Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic NHREC, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding NHREC or fragments of NHREC may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl

WO 01/55392

PCT/US01/02741

sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (Applied Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

	Original Residue	Conservative Substitution
15	Ala	Gly, Ser
	Arg	His, Lys
	Asn	Asp, Gln, His
	Asp	Asn, Glu
	Cys	Ala, Ser
20	Gln	Asn, Glu, His
	Glu	Asp, Gln, His
	Gly	Ala
	His	Asn, Arg, Gln, Glu
	Ile	Leu, Val
25	Leu	Ile, Val
	Lys	Arg, Gln, Glu
	Met	Leu, Ile
	Phe	His, Met, Leu, Trp, Tyr
	Ser	Cys, Thr
30	Thr	Ser, Val
	Trp	Phe, Tyr
	Tyr	His, Phe, Trp
	Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide can include, for example, replacement of hydrogen by an

WO 01/58392

PCT/US01/02741

alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

5 A "detectable label" refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

A "fragment" is a unique portion of NHREC or the polynucleotide encoding NHREC which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:4-6 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:4-6, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:4-6 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:4-6 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:4-6 and the region of SEQ ID NO:4-6 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-3 is encoded by a fragment of SEQ ID NO:4-6. A fragment of SEQ ID NO:1-3 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-3. For example, a fragment of SEQ ID NO:1-3 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-3. The precise length of a fragment of SEQ ID NO:1-3 and the region of SEQ ID NO:1-3 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A "full length" polynucleotide sequence is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full length" polynucleotide sequence encodes a "full length" polypeptide sequence.

WO 01/55392

PCT/US01/02743

"Homology" refers to sequence similarity or, interchangeably, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62
Reward for match: 1
Penalty for mismatch: -2
Open Gap: 5 and Extension Gap: 2 penalties
Gap x drop-off: 50
Expect: 10

WO 01/55392

PCT/US01/02741

*Word Size: 11**Filter: on*

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12c sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

*Matrix: BLOSUM62**Open Gap: 11 and Extension Gap: 1 penalties**Gap x drop-off: 50**Expect: 10**Word Size: 3*

WO 01/55392

PCT/US01/02741

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence. for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size and which contain all of the elements required for chromosome replication, segregation and maintenance.

The term "humanized antibody" refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. et al. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present

WO 01/55392

PCT/US01/02741

invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C₀t or R₀t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of NHREC which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of NHREC which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, or other chemical compounds on a substrate.

The terms "element" and "array element" refer to a polynucleotide, polypeptide, or other chemical compound having a unique and defined position on a microarray.

The term "modulate" refers to a change in the activity of NHREC. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of NHREC.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide,

WO 01/55392

PCT/US01/02741

polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an NHREC may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of NHREC.

"Probe" refers to nucleic acid sequences encoding NHREC, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes.

"Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel, F.M. et al. (1987) Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis, M. et al. (1990) PCR

WO 01/55392

PCT/US01/02741

Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

5 Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the
10 PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which
15 sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments,
20 thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to
25 identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the
30 artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to
35 transform a cell.

WO 01/55392

PCT/US01/02741

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be used to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

"Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing NHREC, nucleic acids encoding NHREC, or fragments thereof may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells,

WO 01/55392

PCT/US01/02741

trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

A "transcript image" refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

"Transformation" describes a process by which exogenous DNA is introduced into a recipient
 5 cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term
 10 "transformed cells" includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic
 15 acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The
 20 transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989).
 25 supra.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at
 30 least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternative splicing of exons during mRNA processing. The corresponding
 35 polypeptide may possess additional functional domains or lack domains that are present in the

WO 01/55392

PCT/US01/02741

reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human nuclear hormone receptors (NHRs), the polynucleotides encoding NHRs, and the use of these compositions for the diagnosis, treatment, or prevention of metabolic, endocrine, reproductive, developmental, immunological, and cell proliferative disorders.

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the invention. Each polynucleotide and its corresponding polypeptide are correlated to a single Incyte project identification number (Incyte Project ID). Each polypeptide sequence is denoted by both a polypeptide sequence identification number (Polypeptide SEQ ID NO.) and an Incyte polypeptide sequence number (Incyte Polypeptide ID) as shown. Each polynucleotide sequence is denoted by both a polynucleotide sequence identification number (Polynucleotide SEQ ID NO.) and an Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) as shown.

Table 2 shows sequences with homology to the polypeptides of the invention as identified by BLAST analysis against the GenBank protein (genpept) database. Columns 1 and 2 show the polypeptide sequence identification number (Polypeptide SEQ ID NO.) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for polypeptides of the invention. Column 3 shows the GenBank identification number (Genbank ID NO.) of the nearest GenBank homolog. Column 4 shows the probability score for the match between each polypeptide and its GenBank homolog. Column 5 shows the annotation of the GenBank homolog along with relevant citations where applicable, all of which are expressly incorporated by reference herein.

Table 3 shows various structural features of the polypeptides of the invention. Columns 1 and

WO 01/55392

PCT/US01/02741

2 show the polypeptide sequence identification number (SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the
 5 MOTIFS program of the GCG sequence analysis software package (Genetics Computer Group, Madison WI). Column 6 shows amino acid residues comprising signature sequences, domains, and motifs. Column 7 shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these
 10 properties establish that the claimed polypeptides are nuclear hormone receptors. For example, SEQ ID NO:3 is 91% identical to *Rattus norvegicus* farnesoid X activated receptor (GenBank ID g868032) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 3.8e-236, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:3 also contains a nuclear hormone receptor ligand-
 15 binding domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILES SCAN analyses provide further corroborative evidence that SEQ ID NO:3 is a steroid hormone receptor. SEQ ID NO:1 and SEQ ID NO:2 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-3 are
 20 described in Table 7.

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Columns 1 and 2 list the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and the corresponding Incyte polynucleotide
 25 consensus sequence number (Incyte Polynucleotide ID) for each polynucleotide of the invention. Column 3 shows the length of each polynucleotide sequence in basepairs. Column 4 lists fragments of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:4-6 or that distinguish between SEQ ID NO:4-6 and related polynucleotide sequences. Column 5 shows identification numbers corresponding to cDNA
 30 sequences, coding sequences (exons) predicted from genomic DNA, and/or sequence assemblages comprised of both cDNA and genomic DNA. These sequences were used to assemble the full length polynucleotide sequences of the invention. Columns 6 and 7 of Table 4 show the nucleotide start (5') and stop (3') positions of the cDNA and genomic sequences in column 5 relative to their respective full length sequences.

35 The identification numbers in Column 5 of Table 4 may refer specifically, for example, to

WO 01/55392

PCT/US01/02741

Incyte cDNAs along with their corresponding cDNA libraries. For example, 6812329H1 is the identification number of an Incyte cDNA sequence, and ADRETUR01 is the cDNA library from which it is derived. Incyte cDNAs for which cDNA libraries are not indicated were derived from pooled cDNA libraries (e.g., 71218724V1). Alternatively, the identification numbers in column 5 may refer to GenBank cDNAs or ESTs (e.g., g2368743) which contributed to the assembly of the full length polynucleotide sequences. Alternatively, the identification numbers in column 5 may refer to coding regions predicted by Genscan analysis of genomic DNA. For example, g4096054.v113.gs_2.nt is the identification number of a Genscan-predicted coding sequence, with g4096054 being the GenBank identification number of the sequence to which Genscan was applied. The Genscan-predicted coding sequences may have been edited prior to assembly. (See Example IV.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. (See Example V.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon-stretching" algorithm. (See Example V.) In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in column 5 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

Table 5 shows the representative cDNA libraries for those full length polynucleotide sequences which were assembled using Incyte cDNA sequences. The representative cDNA library is the Incyte cDNA library which is most frequently represented by the Incyte cDNA sequences which were used to assemble and confirm the above polynucleotide sequences. The tissues and vectors which were used to construct the cDNA libraries shown in Table 5 are described in Table 6.

The invention also encompasses NHREC variants. A preferred NHREC variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the NHREC amino acid sequence, and which contains at least one functional or structural characteristic of NHREC.

The invention also encompasses polynucleotides which encode NHREC. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:4-6, which encodes NHREC. The polynucleotide sequences of SEQ ID NO:4-6, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding NHREC. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide

WO 01/55392

PCT/US01/02741

sequence encoding NHREC. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:4-6 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:4-6. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of NHREC.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding NHREC, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring NHREC, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode NHREC and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring NHREC under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding NHREC or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding NHREC and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode NHREC and NHREC derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding NHREC or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:4-6 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

WO 01/55392

PCT/US01/02741

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Applied Biosystems), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) *Short Protocols in Molecular Biology*, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) *Molecular Biology and Biotechnology*, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding NHREC may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) *PCR Methods Appl.* 2:318-322.)

Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) *Nucleic Acids Res.* 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) *PCR Methods Appl.* 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) *Nucleic Acids Res.* 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

WO 01/55392

PCT/US01/02741

When screening for full length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence
5 into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the
10 emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Applied Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode NHREC may be cloned in recombinant DNA molecules that direct expression of NHREC, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express NHREC.
15

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter NHREC-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction
20 sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or
30 improve the biological properties of NHREC, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired
35 properties. These preferred variants may then be pooled and further subjected to recursive rounds of

WO 01/55392

PCT/US01/02741

DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of

5 homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding NHREC may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) *Nucleic Acids Symp. Ser.* 7:215-223; and Horn, T. et al. (1980) *Nucleic Acids Symp. Ser.* 7:225-232.)

10 Alternatively, NHREC itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g., Creighton, T. (1984) *Proteins, Structures and Molecular Properties*, W.H. Freeman, New York NY, pp.55-60; and Roberge, J.Y. et al. (1995) *Science* 269:202-204.) Automated synthesis may

15 be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of NHREC, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) *Methods Enzymol.* 182:392-421.)

20 The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, *supra*, pp. 28-53.)

In order to express a biologically active NHREC, the nucleotide sequences encoding NHREC or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which

25 contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding NHREC. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of

30 sequences encoding NHREC. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding NHREC and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-

35 frame ATG initiation codon should be provided by the vector. Exogenous translational elements and

WO 01/55392

PCT/US01/02741

initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression
 5 vectors containing sequences encoding NHREC and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. (See, e.g., Sambrook, J. et al. (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York NY, ch. 9, 13, and
 10 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding NHREC. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors
 15 (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, *supra*; Ausubel, *supra*; Van Hecke, G. and S.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509; Engelhard, E.K. et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:3224-3227; Sandig, V. et al. (1996) *Hum. Gene Ther.* 7:1937-1945; Takamatsu, N. (1987) *EMBO J.* 6:307-311; *The McGraw Hill Yearbook of Science and Technology* (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659; and Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-353.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population.
 20 (See, e.g., Di Nicola, M. et al. (1998) *Cancer Gen. Ther.* 5(6):350-356; Yu, M. et al. (1993) *Proc. Natl. Acad. Sci. USA* 90(13):6340-6344; Butler, R.M. et al. (1985) *Nature* 317(6040):813-815; McGregor, D.P. et al. (1994) *Mol. Immunol.* 31(3):219-226; and Verma, I.M. and N. Sonnia (1997) *Nature* 389:239-242.) The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending
 30 upon the use intended for polynucleotide sequences encoding NHREC. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding NHREC can be achieved using a multifunctional *E. coli* vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSFORT1 plasmid (Life Technologies). Ligation of sequences encoding NHREC into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of
 35 transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for

WO 01/55392

PCT/US01/02741

in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Hecke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of NHREC are needed, e.g. for the production of antibodies, vectors which direct high level expression of NHREC may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of NHREC. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast *Saccharomyces cerevisiae* or *Pichia pastoris*. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Austubel, 1995, supra; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of NHREC. Transcription of sequences encoding NHREC may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding NHREC may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses NHREC in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression

WO 01/55392

PCT/US01/02741

of NHREC in cell lines is preferred. For example, sequences encoding NHREC can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk⁻* and *apv⁻* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, L. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorosulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β glucuronidase and its substrate β -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding NHREC is inserted within a marker gene sequence, transformed cells containing sequences encoding NHREC can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding NHREC under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding NHREC and that express NHREC may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

WO 01/55392

PCT/US01/02741

Immunological methods for detecting and measuring the expression of NHREC using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing 5 monoclonal antibodies reactive to two non-interfering epitopes on NHREC is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J. E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana 10 Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding NHREC include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. 15 Alternatively, the sequences encoding NHREC, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega 20 (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding NHREC may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein 25 produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode NHREC may be designed to contain signal sequences which direct secretion of NHREC through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the 30 inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for 35 post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the

WO 01/55392

PCT/US01/02741

American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding NHREC may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric NHREC protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of NHREC activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the NHREC encoding sequence and the heterologous protein sequence, so that NHREC may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, *supra*, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled NHREC may be achieved *in vitro* using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ³⁵S-methionine.

NHREC of the present invention or fragments thereof may be used to screen for compounds that specifically bind to NHREC. At least one and up to a plurality of test compounds may be screened for specific binding to NHREC. Examples of test compounds include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

In one embodiment, the compound thus identified is closely related to the natural ligand of NHREC, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, e.g., Coligan, J.E. et al. (1991) *Current Protocols in Immunology* 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which NHREC binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case, the compound can be rationally designed using known techniques. In one embodiment, screening for

WO 01/55392

PCT/US01/02741

these compounds involves producing appropriate cells which express NHREC, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing NHREC or cell membrane fractions which contain NHREC are then contacted with a test compound and binding, stimulation, or inhibition of activity of either NHREC or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with NHREC, either in solution or affixed to a solid support, and detecting the binding of NHREC to the compound.

Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

NHREC of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of NHREC. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for NHREC activity, wherein NHREC is combined with at least one test compound, and the activity of NHREC in the presence of a test compound is compared with the activity of NHREC in the absence of the test compound. A change in the activity of NHREC in the presence of the test compound is indicative of a compound that modulates the activity of NHREC. Alternatively, a test compound is combined with an *in vitro* or cell-free system comprising NHREC under conditions suitable for NHREC activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of NHREC may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

In another embodiment, polynucleotides encoding NHREC or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent Number 5,175,383 and U.S. Patent Number 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (*neo*; Capecchi, M.R. (1989) *Science* 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) *Clin. Invest.* 97:1999-2002; Wagner, K.U. et al. (1997) *Nucleic*

WO 01/55392

PCT/US01/02741

Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding NHREC may also be manipulated in vitro in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding NHREC can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding NHREC is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress NHREC, e.g., by secreting NHREC in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

20 THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of NHREC and nuclear hormone receptors. In addition, the expression of NHREC is closely associated with digestive tract disorders and with gastrointestinal tissue. Therefore, NHREC appears to play a role in metabolic, endocrine, reproductive, developmental, immunological, and cell proliferative disorders. In the treatment of disorders associated with increased NHREC expression or activity, it is desirable to decrease the expression or activity of NHREC. In the treatment of disorders associated with decreased NHREC expression or activity, it is desirable to increase the expression or activity of NHREC.

Therefore, in one embodiment, NHREC or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of NHREC. Examples of such disorders include, but are not limited to, a metabolic disorder such as Addison's disease, cerebrotendinous xanthomatosis, congenital adrenal hyperplasia, coumarin resistance, cystic fibrosis, diabetes, fatty hepatocirrhosis, fructose-1,6-diphosphatase deficiency, galactosemia, goiter, glucagonoma, glycogen storage diseases, hereditary fructose intolerance, hyperadrenism, hypoadrenism, hyperparathyroidism, hypoparathyroidism, hypercholesterolemia,

WO 01/55392

PCT/US01/02741

hyperthyroidism, hypoglycemia, hypothyroidism, hyperlipidemia, hyperlipemia, lipid myopathies, lipodystrophies, lysosomal storage diseases, manganosis, neuraminidase deficiency, obesity, pentosuria, phenylketonuria, and pseudovitamin D-deficiency rickets; an endocrine disorder such as a disorder of the hypothalamus and pituitary resulting from lesions such as primary brain tumors, an

5 adenoma, infarction associated with pregnancy, hypophysectomy, aneurysms, vascular malformations, thrombosis, infections, immunological disorders, and complications due to head trauma, disorders associated with hypopituitarism including hypogonadism, Sheehan syndrome, diabetes insipidus, Kallman's disease, Hand-Schuller-Christian disease, Letterer-Siwe disease, sarcoidosis, empty sella syndrome, and dwarfism, disorders associated with hyperpituitarism

10 including acromegaly, gigantism, and syndrome of inappropriate antidiuretic hormone (ADH) secretion (SIADH) often caused by benign adenoma, disorders associated with hypothyroidism including goiter, myxedema, acute thyroiditis associated with bacterial infection, subacute thyroiditis associated with viral infection, autoimmune thyroiditis (Hashimoto's disease), and cretinism, disorders associated with hyperthyroidism including thyrotoxicosis and its various forms, Grave's

15 disease, pretibial myxedema, toxic multinodular goiter, thyroid carcinoma, and Plummer's disease, disorders associated with hyperparathyroidism including Conn disease (chronic hypercalcemia), pancreatic disorders such as Type I or Type II diabetes mellitus and associated complications, disorders associated with the adrenals such as hyperplasia, carcinoma, or adenoma of the adrenal cortex, hypertension associated with alkalosis, amyloidosis, hypokalemia, Cushing's disease, Liddle's

20 syndrome, and Arnold-Healy-Gordon syndrome, pheochromocytoma tumors, and Addison's disease, disorders associated with gonadal steroid hormones such as: in women, abnormal prolactin production, infertility, endometriosis, perturbations of the menstrual cycle, polycystic ovarian disease, hyperprolactinemia, isolated gonadotropin deficiency, amenorrhea, galactorrhea, hermaphroditism, hirsutism and virilization, breast cancer, and, in post-menopausal women, osteoporosis, and, in men,

25 Leydig cell deficiency, male climacteric phase, and germinal cell aplasia, hypergonadal disorders associated with Leydig cell tumors, androgen resistance associated with absence of androgen receptors, syndrome of 5 α -reductase, and gynecomastia; a reproductive disorder such as a disorder of prolactin production, infertility, including tubal disease, ovulatory defects, and endometriosis, a

30 ovarian hyperstimulation syndrome, an endometrial or ovarian tumor, a uterine fibroid, autoimmune disorders, an ectopic pregnancy, and teratogenesis, cancer of the breast, fibrocystic breast disease, and galactorrhea, a disruption of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia; a developmental disorder such as renal tubular

35 acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular

WO 01/55392

PCT/US01/02741

dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucopolysaccharidosis, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly, craniocleidocranial dysplasia, congenital glaucoma, cataract, and sensorineural hearing loss; an immunological disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxicity, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hyper eosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and a cancer including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus.

In another embodiment, a vector capable of expressing NHREC or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of NHREC including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified NHREC in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of NHREC including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of NHREC may be administered to a subject to treat or prevent a disorder associated with decreased expression or

WO 01/55392

PCT/US01/02741

activity of NHREC including, but not limited to, those listed above.

In a further embodiment, an antagonist of NHREC may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of NHREC. Examples of such disorders include, but are not limited to, those metabolic, endocrine, reproductive, developmental, immunological, and cell proliferative disorders described above. In one aspect, an antibody which specifically binds NHREC may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express NHREC.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding NHREC may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of NHREC including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of NHREC may be produced using methods which are generally known in the art. In particular, purified NHREC may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind NHREC. Antibodies to NHREC may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with NHREC or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolucithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and *Corynebacterium parvum* are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to NHREC have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches

WO 01/55392

PCT/US01/02741

of NHREC amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to NHREC may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) *Nature* 256:495-497; Kozbor, D. et al. (1985) *J. Immunol. Methods* 81:31-42; Cote, R.J. et al. (1983) *Proc. Natl. Acad. Sci. USA* 80:2026-2030; and Cole, S.P. et al. (1984) *Mol. Cell Biol.* 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) *Proc. Natl. Acad. Sci. USA* 81:6851-6853; Neuberger, M.S. et al. (1984) *Nature* 312:604-608; and Takeda, S. et al. (1985) *Nature* 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce NHREC-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotype composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) *Proc. Natl. Acad. Sci. USA* 88:10134-10137.)

Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:3833-3837; Winter, G. et al. (1991) *Nature* 349:293-299.)

Antibody fragments which contain specific binding sites for NHREC may also be generated. For example, such fragments include, but are not limited to, F(ab')₂ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')₂ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) *Science* 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between NHREC and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering NHREC epitopes is generally used, but a competitive binding assay may also be employed (Pound, *supra*).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay

WO 01/55392

PCT/US01/02741

techniques may be used to assess the affinity of antibodies for NHREC. Affinity is expressed as an association constant, K_a , which is defined as the molar concentration of NHREC-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their
 5 affinities for multiple NHREC epitopes, represents the average affinity, or avidity, of the antibodies for NHREC. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular NHREC epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the NHREC-antibody complex must withstand rigorous manipulations. Low-affinity antibody
 10 preparations with K_a ranging from about 10^6 to 10^9 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of NHREC, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

15 The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of NHREC-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity,
 20 and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding NHREC, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules
 25 (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding NHREC. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding NHREC. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

30 In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) *J. Allergy Clin. Immunol.* 102(3):469-475; and Scanlon, K.J. et al. (1995)
 35 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral

WO 01/55392

PCT/US01/02741

vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) *Blood* 76:271; Ausubel, *supra*; Uckert, W. and W. Walther (1994) *Pharmacol. Ther.* 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) *Br. Med. Bull.* 51(1):217-225; Boado, R.J. et al. (1998) *J. Pharm. Sci.* 87(11):1308-1315; and Morris, M.C. et al. (1997) *Nucleic Acids Res.* 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding NHREC may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) *Science* 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) *Science* 270:475-480; Bordignon, C. et al. (1995) *Science* 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) *Cell* 75:207-216; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:643-666; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:667-703), thalassemias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and N. Somia (1997) *Nature* 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) *Nature* 335:395-396; Foeschla, E. et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as *Candida albicans* and *Paracoccidioides brasiliensis*; and protozoan parasites such as *Plasmodium falciparum* and *Trypanosoma cruzi*). In the case where a genetic deficiency in NHREC expression or regulation causes disease, the expression of NHREC from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in NHREC are treated by constructing mammalian expression vectors encoding NHREC and introducing these vectors by mechanical means into NHREC-deficient cells. Mechanical transfer technologies for use with cells *in vivo* or *ex vitro* include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) *Annu. Rev. Biochem.* 62:191-217; Ivics, Z. (1997) *Cell* 91:501-510; Boulay, J.-L. and H. Récipon (1998) *Curr. Opin. Biotechnol.* 9:445-450).

Expression vectors that may be effective for the expression of NHREC include, but are not limited to, the pCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA),

WO 01/55392

PCT/US01/02741

PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). NHREC may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β -actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) *Proc. Natl. Acad. Sci. USA* 89:5547-5551); Gossen, M. et al. (1995) *Science* 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) *Curr. Opin. Biotechnol.* 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and Blau, H.M. *supra*)), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding NHREC from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) *Virology* 52:456-467), or by electroporation (Neumann, E. et al. (1982) *EMBO J.* 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to NHREC expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding NHREC under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) *Proc. Natl. Acad. Sci. USA* 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) *J. Virol.* 61:1647-1650; Bender, M.A. et al. (1987) *J. Virol.* 61:1639-1646; Adam, M.A. and A.D. Miller (1988) *J. Virol.* 62:3802-3806; Duil, T. et al. (1998) *J. Virol.* 72:8463-8471; Zufferey, R. et al. (1998) *J. Virol.* 72:9873-9880). U.S. Patent Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4⁺ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in

WO 01/55392

PCT/US01/02741

the art of gene therapy and have been well documented (Ranga, U. et al. (1997) *J. Virol.* 71:7020-7029; Bauer, G. et al. (1997) *Blood* 89:2259-2267; Bonyhadi, M.L. (1997) *J. Virol.* 71:4707-4716; Ranga, U. et al. (1998) *Proc. Natl. Acad. Sci. USA* 95:1201-1206; Su, L. (1997) *Blood* 89:2283-2290).

5 In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding NHRFC to cells which have one or more genetic abnormalities with respect to the expression of NHRFC. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) *Transplantation* 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent Number 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinuzzi, P.A. et al. (1999) *Annu. Rev. Nutr.* 19:511-544 and Verma, I.M. and N. Sonnia (1997) *Nature* 18:389:239-242, both incorporated by reference herein.

15 In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding NHRFC to target cells which have one or more genetic abnormalities with respect to the expression of NHRFC. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing NHRFC to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) *Exp. Eye Res.* 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) *J. Virol.* 73:519-532 and Xu, H. et al. (1994) *Dev. Biol.* 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding NHRFC to target cells. The biology of the prototypic alphavirus,

WO 01/55392

PCT/US01/02741

Serfliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) *Curr. Opin. Biotechnol.* 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for NHREC into the alphavirus genome in place of the capsid-coding region results in the production of a large number of NHREC-coding RNAs and the synthesis of high levels of NHREC in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) *Virology* 228:74-83). The wide host range of alphaviruses will allow the introduction of NHREC into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, *Molecular and Immunologic Approaches*, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding NHREC.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUU, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for

WO 01/55392

PCT/US01/02741

secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared
 5 by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding NHREC. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA
 10 constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase
 15 linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, quosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a
 20 compound which is effective in altering expression of a polynucleotide encoding NHREC. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-
 25 macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased NHREC expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding NHREC may be therapeutically useful, and in the treatment of disorders
 30 associated with decreased NHREC expression or activity, a compound which specifically promotes expression of the polynucleotide encoding NHREC may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in
 altering expression of a specific polynucleotide. A test compound may be obtained by any method
 commonly known in the art, including chemical modification of a compound known to be effective in
 35 altering polynucleotide expression; selection from an existing, commercially-available or proprietary

WO 01/55392

PCT/US01/02741

library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding NHRHC is exposed to at least one test compound thus obtained. The sample
 5 may comprise, for example, an intact or permeabilized cell, or an *in vitro* cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding NHRHC are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding NHRHC. The amount of hybridization may be quantified, thus
 10 forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a *Schizosaccharomyces pombe* gene expression
 15 system (Adkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide
 20 sequence (Bruce, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruce, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use *in vivo*, *in vitro*, and *ex vivo*. For *ex vivo* therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient.
 25 Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and
 30 monkeys.

An additional embodiment of the invention relates to the administration of a composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of
 35 Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may

WO 01/55392

PCT/US01/02741

consist of NHREC, antibodies to NHREC, and mimetics, agonists, antagonists, or inhibitors of NHREC.

The compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

Compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

Compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising NHREC or fragments thereof. For example, liposomal preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, NHREC or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example NHREC or fragments thereof, antibodies of NHREC, and agonists, antagonists or inhibitors of NHREC, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED₅₀ (the dose therapeutically effective in 50% of the population) or LD₅₀ (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the

WO 01/55392

PCT/US01/02741

therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity.

5 The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

15 Normal dosage amounts may vary from about 0.1 μ g to 100,000 μ g, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

20 DIAGNOSTICS

In another embodiment, antibodies which specifically bind NHREC may be used for the diagnosis of disorders characterized by expression of NHREC, or in assays to monitor patients being treated with NHREC or agonists, antagonists, or inhibitors of NHREC. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics.

25 Diagnostic assays for NHREC include methods which utilize the antibody and a label to detect NHREC in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

30 A variety of protocols for measuring NHREC, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of NHREC expression. Normal or standard values for NHREC expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to NHREC under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of NHREC

WO 01/55392

PCT/US01/02741

expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding NHREC may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of NHREC may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of NHREC, and to monitor regulation of NHREC levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding NHREC or closely related molecules may be used to identify nucleic acid sequences which encode NHREC. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding NHREC, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the NHREC encoding sequences. The hybridization probes of the subject invention may be DNA or RNA, and may be derived from the sequence of SEQ ID NO. 4-6 or from genomic sequences including promoters, enhancers, and introns of the NHREC gene.

Means for producing specific hybridization probes for DNAs encoding NHREC include the cloning of polynucleotide sequences encoding NHREC or NHREC derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ³²P or ³⁵S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding NHREC may be used for the diagnosis of disorders associated with expression of NHREC. Examples of such disorders include, but are not limited to, a metabolic disorder such as Addison's disease, cerebrotendinous xanthomatosis, congenital adrenal hyperplasia, coumarin resistance, cystic fibrosis, diabetes, fatty hepatocirrhosis, fructose-1,6-diphosphatase deficiency, galactosemia, goiter, glucagonoma, glycogen storage diseases, hereditary fructose intolerance, hyperadrenalism, hypoadrenalism, hyperparathyroidism, hypoparathyroidism, hypercholesterolemia, hyperthyroidism, hypoglycemia, hypothyroidism, hyperlipidemia, hyperlipemia, lipid myopathies, lipodystrophies, lysosomal storage diseases,

WO 01/55392

PCT/US01/02741

mannosidosis, neuraminidase deficiency, obesity, pentosuria, phenylketonuria, and pseudovitamin D-deficiency rickets; an endocrine disorder such as a disorder of the hypothalamus and pituitary resulting from lesions such as primary brain tumors, an adenoma, infarction associated with pregnancy, hypophysectomy, aneurysms, vascular malformations, thrombosis, infections,

5 immunological disorders, and complications due to head trauma, disorders associated with hypopituitarism including hypogonadism, Sheehan syndrome, diabetes insipidus, Kallman's disease, Hand-Schüller-Christian disease, Letterer-Siwe disease, sarcoidosis, empty sella syndrome, and dwarfism, disorders associated with hyperpituitarism including acromegaly, gigantism, and syndrome of inappropriate antidiuretic hormone (ADH) secretion (SIADH) often caused by benign adenoma.

10 disorders associated with hypothyroidism including goiter, myxedema, acute thyroiditis associated with bacterial infection, subacute thyroiditis associated with viral infection, autoimmune thyroiditis (Hashimoto's disease), and cretinism, disorders associated with hyperthyroidism including thyrotoxicosis and its various forms, Grave's disease, pretibial myxedema, toxic multinodular goiter, thyroid carcinoma, and Plummer's disease, disorders associated with hyperparathyroidism including

15 Conn disease (chronic hypercalcemia), pancreatic disorders such as Type I or Type II diabetes mellitus and associated complications, disorders associated with the adrenals such as hyperplasia, carcinoma, or adenoma of the adrenal cortex, hypertension associated with alkalosis, amyloidosis, hypokalemia, Cushing's disease, Liddle's syndrome, and Arnold-Healy-Gordon syndrome, pheochromocytoma tumors, and Addison's disease, disorders associated with gonadal steroid hormones such as: in

20 women, abnormal prolactin production, infertility, endometriosis, perturbations of the menstrual cycle, polycystic ovarian disease, hyperprolactinemia, isolated gonadotropin deficiency, amenorrhea, galactorrhea, hermaphroditism, hirsutism and virilization, breast cancer, and, in post-menopausal women, osteoporosis, and, in men, Leydig cell deficiency, male climacteric phase, and germinal cell aplasia, hypergonadal disorders associated with Leydig cell tumors, androgen resistance associated

25 with absence of androgen receptors, syndrome of 5 α -reductase, and gynecomastia; a reproductive disorder such as a disorder of prolactin production, infertility, including tubal disease, ovulatory defects, and endometriosis, a disruption of the estrous cycle, a disruption of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, an endometrial or ovarian tumor, a uterine fibroid, autoimmune disorders, an ectopic pregnancy, and teratogenesis, cancer of the breast,

30 fibrocystic breast disease, and galactorrhea, a disruption of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia; a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome,

35

PCT/US01/02741

In a particular aspect, the nucleotide sequences encoding NHREC may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding NHREC may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding NHREC in the

35 a control sample then the presence of altered levels of nucleotide sequences encoding NHREC in the

WO 01/55392

PCT/US01/02741

sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of NHREC, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding NHREC, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding NHREC may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced *in vitro*. Oligomers will preferably contain a fragment of a polynucleotide encoding NHREC, or a fragment of a polynucleotide complementary to the polynucleotide encoding NHREC, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding NHREC may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP,

WO 01/55392

PCT/US01/02741

oligonucleotide primers derived from the polynucleotide sequences encoding NHRHC are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In TSCCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed *in silico* SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of NHRHC include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C., et al. (1993) *J. Immunol. Methods* 159:235-244; Duplax, C. et al. (1993) *Anal. Biochem.* 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, NHRHC, fragments of NHRHC, or antibodies specific for NHRHC may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described

WO 01/55392

PCT/US01/02741

above.

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent Number 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression *in vivo*, as in the case of a tissue or biopsy sample, or *in vitro*, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with *in vitro* model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471, expressly incorporated by reference herein). If a test compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at <http://www.niehs.nih.gov/oc/news/toxchip.htm>.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

WO 01/55392

PCT/US01/02741

In one embodiment, the toxicity of a test compound is assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, *supra*). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for NHREC to quantify the levels of NHREC expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) Anal. Biochem. 270:103-111; Mendoza, L.G. et al. (1999) Biotechniques 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at

WO 01/55392

PCT/US01/02741

each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Neillhamer (1997) *Electrophoresis* 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) *Proc. Natl. Acad. Sci. USA* 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding NHREC may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among

WO 01/55392

PCT/US01/02741

members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355; Price, C.M. (1993) *Blood Rev.* 7:127-134; and Trask, B.J. (1991) *Trends Genet.* 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, for example, Lander, E.S. and D. Botstein (1986) *Proc. Natl. Acad. Sci. USA* 83:7353-7357.)

Fluorescent *in situ* hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, *supra*, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding NHREC on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) *Nature* 336:577-580.) The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, NHREC, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between NHREC and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) *PCT*

WO 01/55392

PCT/US01/02741

application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with NHREC, or fragments thereof, and washed. Bound NHREC is then detected by methods well known in the art. Purified NHREC can also be coated directly onto plates for use in the aforementioned drug screening techniques.

5 Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding NHREC specifically compete with a test compound for binding NHREC. In this manner, antibodies can be used to detect the presence of any peptide which shares
10 one or more antigenic determinants with NHREC.

In additional embodiments, the nucleotide sequences which encode NHREC may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

15 Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications and publications, mentioned above and below, in particular U.S. Ser. No. 60/178,578, U.S. Ser. No. 60/185,079, and U.S. Ser. No. 60/189,999, are expressly incorporated by reference herein.
20

EXAMPLES

I. Construction of cDNA Libraries

25 Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA) and shown in Table 4, column 5. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or
30 extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A)⁺ RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN,
35 Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was

WO 01/55392

PCT/US01/02741

isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPTM plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, *supra*, units 5.1-6.6.) Reverse transcription was initiated using oligo(dT) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), pCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), or pINCY (Incyte Genomics, Palo Alto CA), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by *in vivo* excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal

WO 01/55392

PCT/US01/02741

cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

- 5 Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 10 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

- The polynucleotide sequences derived from Incyte cDNAs were validated by removing vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The 15 Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and hidden Markov model (HMM)-based protein family databases such as PFAM. (HMM is a probabilistic approach which analyzes consensus primary structures of gene families. See, for example, Eddy, S.R. (1996) *Curr. Opin. Struct. Biol.* 6:361-365.) 20 The queries were performed using programs based on BLAST, FASTA, BLIMPS, and HMMER. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or Genscan-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and 25 Consed, and cDNA assemblages were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length polypeptide sequences. Alternatively, a polypeptide of the invention may begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the 30 GenBank protein databases (genpept), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and hidden Markov model (HMM)-based protein family databases such as PFAM. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the 35 CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program

WO 01/55392

PCT/US01/02741

(DNASTAR), which also calculates the percent identity between aligned sequences.

Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score or the lower the probability value, the greater the identity between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:4-6. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 4.

IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative nuclear hormone receptors were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpr and ghgt). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (See Burge, C. and S. Karlin (1997) J. Mol. Biol. 268:78-94, and Burge, C. and S. Karlin (1998) Curr. Opin. Struct. Biol. 8:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode nuclear hormone receptors, the encoded polypeptides were analyzed by querying against PFAM models for nuclear hormone receptors. Potential nuclear hormone receptors were also identified by homology to Incyte cDNA sequences that had been annotated as nuclear hormone receptors. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpr public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or

WO 01/55392

PCT/US01/02741

unedited Genscan-predicted coding sequences.

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

5 "Stitched" Sequences

Partial cDNA sequences were extended with exons predicted by the Genscan gene identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpri public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

25 "Stretched" Sequences

Partial DNA sequences were extended to full length with an algorithm based on BLAST analysis. First, partial cDNAs assembled as described in Example III were queried against public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases using the BLAST program. The nearest GenBank protein homolog was then compared by BLAST analysis to either bccyte cDNA sequences or GenScan exon predicted sequences described in Example IV. A chimeric protein was generated by using the resultant high-scoring segment pairs (HSPs) to map the translated sequences onto the GenBank protein homolog. Insertions or deletions may occur in the chimeric protein with respect to the original GenBank protein homolog. The GenBank protein homolog, the chimeric protein, or both were used as probes to search for homologous genomic sequences from the public human genome databases. Partial DNA sequences

WO 01/55392

PCT/US01/02741

were therefore "stretched" or extended by the addition of homologous genomic sequences. The resultant stretched sequences were examined to determine whether it contained a complete gene.

VI. Chromosomal Mapping of NHREC Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:4-6 were compared with sequences
 5 from the Incyte LIFtSEQ database and public domain databases using BLAST and other
 implementations of the Smith-Waterman algorithm. Sequences from these databases that matched
 SEQ ID NO:4-6 were assembled into clusters of contiguous and overlapping sequences using
 assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available
 from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for
 10 Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences
 had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment
 of all sequences of that cluster, including its particular SEQ ID NO., to that map location.

Map locations are represented by ranges, or intervals, or human chromosomes. The map
 position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-
 15 arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between
 chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in
 humans, although this can vary widely due to hot and cold spots of recombination.) The cM
 distances are based on genetic markers mapped by Généthon which provide boundaries for radiation
 hybrid markers whose sequences were included in each of the clusters. Human genome maps and
 20 other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site
 (<http://www.ncbi.nlm.nih.gov/genemap/>), can be employed to determine if previously identified
 disease genes map within or in proximity to the intervals indicated above.

VII. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a
 25 gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs
 from a particular cell type or tissue have been bound. (See, e.g., Sambrook, *supra*, ch. 7; Ausubel
 (1995) *supra*, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related
 molecules in cDNA databases such as GenBank or LIFtSEQ (Incyte Genomics). This analysis is
 30 much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the
 computer search can be modified to determine whether any particular match is categorized as exact or
 similar. The basis of the search is the product score, which is defined as:

WO 01/55392

PCT/US01/02741

$$\frac{\text{BLAST Score} \times \text{Percent Identity}}{5 \times \text{minimum}\{\text{length}(\text{Seq. 1}), \text{length}(\text{Seq. 2})\}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

Alternatively, polynucleotide sequences encoding NHREC are analyzed with respect to the tissue sources from which they were derived. For example, some full length sequences are assembled, at least in part, with overlapping Incyte cDNA sequences (see Example III). Each cDNA sequence is derived from a cDNA library constructed from a human tissue. Each human tissue is classified into one of the following organ/tissue categories: cardiovascular system; connective tissue; digestive system; embryonic structures; endocrine system; exocrine glands; genitalia, female; genitalia, male; germ cells; hematopoietic and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract. The number of libraries in each category is counted and divided by the total number of libraries across all categories. Similarly, each human tissue is classified into one of the following disease/condition categories: cancer, cell line, developmental, inflammation, neurological, trauma, cardiovascular, pooled, and other, and the number of libraries in each category is counted and divided by the total number of libraries across all categories. The resulting percentages reflect the tissue- and disease-specific expression of cDNA encoding NHREC. cDNA sequences and cDNA library/tissue information are found in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

VIII. Extension of NHREC Encoding Polynucleotides

Full length polynucleotide sequences were also produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer was

WO 01/55392

PCT/US01/02741

synthesized to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+} , $(NH_4)_2SO_4$, and 2-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PC1 A and PC1 B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 µl to 10 µl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviII cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

WO 01/55392

PCT/US01/02741

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

In like manner, full length polynucleotide sequences are verified using the above procedure or are used to obtain 5' regulatory sequences using the above procedure along with oligonucleotides designed for such extension, and an appropriate genomic library.

IX. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:4-6 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μ Ci of [γ - 32 P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

X. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, *supra*), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Schna (1999).

WO 01/55392

PCT/US01/02741

supra). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Schena, M. et al. (1995) *Science* 270:467-470; Shalon, D. et al. (1996) *Genome Res.* 6:639-645; Marshall, A. and J. Hodgson (1998) *Nat. Biotechnol.* 16:27-31.)

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorption and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)⁺ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)⁺ RNA sample is reverse transcribed using MM1.V reverse-transcriptase, 0.05 pg/μl oligo-(dT) primer (2 liner), 1X first strand buffer, 0.03 units/μl RNase inhibitor, 500 μM dATP, 500 μM dGTP, 500 μM dTTP, 40 μM dCTP, 40 μM dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)⁺ RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)⁺ RNAs are synthesized by *in vitro* transcription from non-coding yeast genomic DNA. After incubation at 37°C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to the stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 μl 5X SSC/0.2% SDS.

Microarray Preparation

WO 01/55392

PCT/US01/02741

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 μ g. Amplified array elements are then purified using SEPIACRYL-400 (Amersham Pharmacia Biotech).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in US Patent No. 5,807,522, incorporated herein by reference. 1 μ l of the array element DNA, at an average concentration of 100 ng/ μ l, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60°C followed by washes in 0.2% SDS and distilled water as before.

Hybridization

Hybridization reactions contain 9 μ l of sample mixture consisting of 0.2 μ g each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65°C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm² coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 μ l of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60°C. The arrays are washed for 10 min at 45°C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45°C in a second wash buffer (0.1X SSC), and dried.

Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is

WO 01/55392

PCT/US01/02741

focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

5 In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is
10 typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that
15 location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

20 The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and
25 measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The
30 software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

XI. Complementary Polynucleotides

Sequences complementary to the NHREC-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring NHREC. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same
35 procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are

WO 01/55392

PCT/US01/02741

designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of NHREC. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the NHREC-encoding transcript.

XII. Expression of NHREC

Expression and purification of NHREC is achieved using bacterial or virus-based expression systems. For expression of NHREC in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*lac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express NHREC upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of NHREC in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant *Autographica californica* nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding NHREC by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect *Spodoptera frugiperda* (SF9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, NHREC is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from *Schistosoma japonicum*, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from NHREC at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, *supra*, ch. 10 and 16). Purified NHREC obtained by these methods can be used directly in the assays shown in Examples XVI and XVII, where applicable.

WO 01/55392

PCT/US01/02741

XIII. Functional Assays

NHREC function is assessed by expressing the sequences encoding NHREC at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT (Life Technologies) and PCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μ g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP, Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of NHREC on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding NHREC and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding NHREC and other genes of interest can be analyzed by northern analysis or microarray techniques.

XIV. Production of NHREC Specific Antibodies

NHREC substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the NHREC amino acid sequence is analyzed using LASHRGHNI software

WO 01/55392

PCT/US01/02741

(DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, *supra*, ch. 11.)

- 5 Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using Fmoc chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, *supra*.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for
 10 antipeptide and anti-NHREC activity by, for example, binding the peptide or NHREC to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XV. Purification of Naturally Occurring NHREC Using Specific Antibodies

- Naturally occurring or recombinant NHREC is substantially purified by immunoaffinity
 15 chromatography using antibodies specific for NHREC. An immunoaffinity column is constructed by covalently coupling anti-NHREC antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

- Media containing NHREC are passed over the immunoaffinity column, and the column is
 20 washed under conditions that allow the preferential absorbance of NHREC (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/NHREC binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and NHREC is collected.

XVI. Identification of Molecules Which Interact with NHREC

- NHREC, or biologically active fragments thereof, are labeled with ¹²⁵I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) Candidate molecules
 25 previously arrayed in the wells of a multi-well plate are incubated with the labeled NHREC, washed, and any wells with labeled NHREC complex are assayed. Data obtained using different concentrations of NHREC are used to calculate values for the number, affinity, and association of
 30 NHREC with the candidate molecules.

Alternatively, molecules interacting with NHREC are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989) *Nature* 340:245-246, or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

- NHREC may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT)
 35 which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions

WO 01/55392

PCT/US01/02741

between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

XVII. Demonstration of NHREC Activity

An assay for receptor activity of NHREC is based on a prototypical assay for ligand/receptor-mediated modulation of cell proliferation. This assay measures the amount of newly synthesized DNA in Swiss mouse 3T3 cells expressing NHREC. An appropriate mammalian expression vector containing cDNA encoding NHREC is added to quiescent 3T3 cultured cells using transfection methods well known in the art. The transfected cells are incubated in the presence of [³H]thymidine and varying amounts of NHREC ligand. Incorporation of [³H]thymidine into acid-precipitable DNA is measured over an appropriate time interval using a tritium radioisotope counter, and the amount incorporated is directly proportional to the amount of newly synthesized DNA. A linear dose-response curve over at least a hundred-fold NHREC ligand concentration range is indicative of receptor activity. One unit of activity per milliliter is defined as the concentration of NHREC producing a 50% response level, where 100% represents maximal incorporation of [³H]thymidine into acid-precipitable DNA. (See, e.g., McKay, I. and I. Leigh, eds. (1993) Growth Factors: A Practical Approach, Oxford University Press, New York NY, p. 73.)

Alternatively, receptor activity of NHREC is measured by the increase in cell proliferation resulting from transformation of a mammalian cell line such as COS7, HeLa or CHO with an eukaryotic expression vector encoding NHREC. Eukaryotic expression vectors are commercially available, and the techniques to introduce them into cells are well known to those skilled in the art. The cells are incubated for 48-72 hours after transformation under conditions appropriate for the cell line to allow expression of NHREC. Phase microscopy is then used to compare the mitotic index of transformed versus control cells. An increase in the mitotic index is an indication of receptor activity of NHREC.

Alternatively, receptor activity of NHREC is determined in a ligand-binding assay using candidate ligand molecules in the presence of ¹²⁵I-labeled NHREC. NHREC is labeled with ¹²⁵I Bolton-Hunter reagent (Bolton and Hunter, *supra*). Candidate ligand molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled NHREC, washed, and any wells with labeled NHREC complex are assayed. Data obtained using different concentrations of NHREC are used to calculate values for the number, affinity, and association of NHREC with the ligand molecules. The level of binding measured is proportional to the level of NHREC activity.

Transcriptional regulatory activity of NHREC is measured by its ability to stimulate transcription of a reporter gene (Liu, H.Y. et al. (1997) EMBO J. 16:5289-5298). The assay entails the use of a well characterized reporter gene construct, LexA_{op}-LacZ, that consists of LexA DNA transcriptional control elements (LexA_{op}) fused to sequences encoding the *E. coli* LacZ enzyme. The

WO 01/55392

PCT/US01/02741

methods for constructing and expressing fusion genes, introducing them into cells, and measuring LacZ enzyme activity, are well known to those skilled in the art. Sequences encoding NIREC are cloned into a plasmid that directs the synthesis of a fusion protein, LexA-NIREC, consisting of NIREC and a DNA-binding domain derived from the LexA transcription factor. The resulting
5 plasmid, encoding a LexA-NIREC fusion protein, is introduced into yeast cells along with a plasmid containing the LexA_q-LacZ reporter gene. The amount of LacZ enzyme activity associated with LexA-NIREC transfected cells, relative to control cells, is proportional to the amount of transcription stimulated by the NIREC.

10 Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific
15 embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

WO 01/55392

PCT/US01/02741

Table 1

Incyte SEQ ID NO.	Polypeptide SEQ ID NO.	Incyte Polypeptide ID	Polynucleotide SEQ ID NO.	Incyte Polynucleotide ID
3462174	1	3462174CD1	4	3462174CB1
1322075	2	1322075CD1	5	1322075CB1
7472269	3	7472269CD1	6	7472269CB1

WO 01/55392

PCT/US01/02741

Table 2

Polypeptide Seq. ID NO.	Incyte Polypeptide ID	GenBank ID NO.	Probability score	GenBank Homolog
2	1322075CD1	92088669	8.40E-287	similar to Ach-ya ambisexualis antheridial steroid receptor Nelson, J. et al. (1994) Nature 370:684-687
3	7472769CD1	9386012	3.90E-236	steroid X activated receptor (Rattus norvegicus)

Table 3

[illegible]

Table 4

Polymérisé de dans le no.	Polycy- clic acrylate	Souche Lentille	Sélecté Fragment(s)	Position	Position
2	3462175C81	2274	2139-2274, 1189-1446	1868 644 913 1756 1597 1146 2118 2428 3007 2398 304 393 702	2274 353 913 1597 1598 2118 2428 3007 2398 304 393 702
5	1322075C81	4154	1-1754, 2356-2592, 4154-4154	1754 2356 4154	1754 2356 4154
6	747268C81	2220	106-1143, 2152-2220, 387-414	106 2152 387	1143 2220 414

WO 01/55392

PCT/US01/02741

Table 5

Polynucleotide	Inxyle	Representative Library
SEQ ID NO: 4	Protest IP	
5	346214GB1	AURETUM01
6	332075GB1	COLLUMOT13
7	777269GB1	STOMOT01

WO 01/55392

PCT/US01/02741

Table 6

Library	Vector	Library Description
ADN57401	PCDNA3.1	This random primed library was constructed using RNA isolated from left upper pole, adrenal gland tumor tissue removed from a 52-year-old Caucasian male during nephroureterectomy and local destruction of renal lesion. Pathology indicated grade 3 adrenal cortical carcinoma forming a mass that infiltrated the renal capsule and into adjacent structures. A renal vein thrombus was noted. The tumor was infiltrated by tumor and the neoplastic process was present at the resection margin of the renal vein. Fragments of adrenal cortical carcinoma and thrombus were found in the inferior vena cava. Patient history included abnormal weight loss. Family history included skin cancer, type 1 diabetes, and neurotic depression.
COLN0013	pINCY	Library was constructed using RNA isolated from ascending colon tissue of a 28-year-old Caucasian male with moderate chronic ulcerative colitis.
SPON0001	PSLRESRIP2	Library was constructed using RNA isolated from the stomach tissue of a 55 year old Caucasian male, who died from colon adenocarcinoma.

WO 01/55392

PCT/US01/02741

Table 7

Program	Description	Reference	Parameter Threshold
ABIFACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA	
ABIPARACHUTER	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracore Inc., Pasadena, CA	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, blastn, blastx, fastx, and search.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 181:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489	ESTs: fasta blast value= 1.0E-6 Assembled ESTs: fasta identity= 95% or greater and Match length=200 bases or greater; fasta blast value= 1.0E-8 or less Full Length sequences: fasta score=100 or greater
BLUMPS	A BLUMPS IMPROVED Searcher that matches a sequence against those in BLUMPS, PRINTS, DOMO, PRODOM, and PFAM databases in search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Altschul, S.F. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1534; Samadpour, E.L. et al. (1988) Nucleic Acids Res. 16:320-322; Durbin, R. et al. (1998) Our World View in a Nucleotide, Cambridge Univ. Press, pp. 1-530.	PFAM hits: Probability value= 1.0E-3 or less Signal peptide hits: Score= 0 or greater

Table 7 (cont.)

Program	Description	Reference	Parameter Threshold	WO 01/55392
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gibbskov, M. et al. (1988) CABIOS 4:61-66; Gibbskov, M. et al. (1989) Methods Enzymol. 183: 146-159; Baroch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality scores QCG-specified "HKGP" value for that particular Prosite motif. Generally, score=1.4-2.1.	
Phred	A base calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.		
Phrap	A TruS Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score=120 or greater; Match length= 36 or greater	
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.		
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nelson, H. et al. (1997) Protein Engineering 10:1-12; Clavero, J.M. and S. Andie (1997) CABIOS 12:431-439.	Score=3.5 or greater	
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.		
TMHMM2ER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Samhammer, E.L. et al. (1995) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.		
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI		PCT/US01/02741

WO 01/55392

PCT/US01/02741

What is claimed is:

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:
 - 5 a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-3,
 - b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3,
 - c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and
 - 10 d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-3.
2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-3.
- 15 3. An isolated polynucleotide encoding a polypeptide of claim 1.
4. An isolated polynucleotide encoding a polypeptide of claim 2.
5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID
 - 20 NO:4-6.
6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
- 25 7. A cell transformed with a recombinant polynucleotide of claim 6.
8. A transgenic organism comprising a recombinant polynucleotide of claim 6.
9. A method for producing a polypeptide of claim 1, the method comprising:
 - 30 a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and
 - b) recovering the polypeptide so expressed.
- 35

WO 01/55392

PCT/US01/02741

10. An isolated antibody which specifically binds to a polypeptide of claim 1.
11. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:
- 5 a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6,
b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6,
c) a polynucleotide sequence complementary to a),
d) a polynucleotide sequence complementary to b), and
10 e) an RNA equivalent of a)-d).
12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.
- 15 13. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:
a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization
20 complex is formed between said probe and said target polynucleotide or fragments thereof, and
b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.
14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.
- 25 15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:
a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and
30 b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.
16. A composition comprising an effective amount of a polypeptide of claim 1 and a pharmaceutically acceptable excipient.
- 35

WO 01/55392

PCT/US01/02741

17. A composition of claim 16, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-3.

18. A method for treating a disease or condition associated with decreased expression of functional NHRFC, comprising administering to a patient in need of such treatment the composition of claim 16.

19. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

- 10 a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
 b) detecting agonist activity in the sample.

20. A composition comprising an agonist compound identified by a method of claim 19 and a pharmaceutically acceptable excipient.

21. A method for treating a disease or condition associated with decreased expression of functional NHRFC, comprising administering to a patient in need of such treatment a composition of claim 20.

22. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
 b) detecting antagonist activity in the sample.

23. A composition comprising an antagonist compound identified by a method of claim 22 and a pharmaceutically acceptable excipient.

24. A method for treating a disease or condition associated with overexpression of functional NHRFC, comprising administering to a patient in need of such treatment a composition of claim 23.

25. A method of screening for a compound that specifically binds to the polypeptide of claim 1, said method comprising the steps of:

- a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and
35 b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a

WO 01/55392

PCT/US01/02741

compound that specifically binds to the polypeptide of claim 1.

26. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, said method comprising:

- 5 a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1.
- b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and
- c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

27. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

- 15 a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,
- b) detecting altered expression of the target polynucleotide, and
- 20 c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

28. A method for assessing toxicity of a test compound, said method comprising:

- a) treating a biological sample containing nucleic acids with the test compound;
- 25 b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 11 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 11 or fragment thereof;
- 30 c) quantifying the amount of hybridization complex; and
- d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

WO 01/55392

PCT/US01/02741

<110> INCYTE GENOMICS, INC.
 BURLINGAME, Calif.
 BAUGHN, Mariah R.
 TANG, Y. Tom
 SANDMAN, Olga
 TRIBOLET, Catherine M.
 DAS, Debopriya
 POLICKY, Jennifer L.
 LO, Terence P.

<120> NUCLEAR HORMONE RECEPTORS

<130> PI 0021 PCT

<140> To Be Assigned

<141> Herewith

<160> 6

<170> PERL Program

<210> 1

<211> 623

<212> PPT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3462174CD1

<400> 1

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

WO/01/55392

PCT/US01/02741

```

Met Leu Leu Gly Leu Asp Pro Gly Asp Pro Gly Ala Ala Pro Ala
245 250 255
Ser Val Pro Thr Ala Pro Thr Pro Phe Pro Ala Pro Glu Arg Pro
260 265 270
Asn Val Val Leu Trp Thr Val Pro Leu Gly His Thr Gly Thr Ala
275 280 285
Thr Thr Ala Ala Ala Ala Ser His Pro Thr His Tyr Asp Ala Leu
290 295 300
Ile Leu Val Thr Pro Gly Ala Pro Thr Glu Lys Asp Trp Ala Gln
305 310 315
Val Gln Ala Leu Leu Leu Pro Asp Ala Pro Leu Val Cys Val Arg
320 325 330
Thr Asp Gly Glu Gly Glu Asp Pro Glu Cys Leu Gly Glu Gly Lys
335 340 345
Met Glu Asn Pro Lys Gly Glu Ser Leu Lys Asn Ala Gly Gly Gly
350 355 360
Gly Leu Glu Asn Ala Leu Ser Lys Gly Arg Glu Lys Cys Ser Ala
365 370 375
Gly Ser Gln Lys Ala Gly Ser Gly Glu Gly Pro Gly Lys Ala Gly
380 385 390
Ser Glu Gly Leu Gln Gln Val Val Gly Met Lys Lys Ser Gly Gly
395 400 405
Gly Asp Ser Glu Arg Ala Ala Ala Leu Ser Pro Glu Asp Glu Thr
410 415 420
Trp Glu Val Leu Glu Glu Ala Pro Pro Pro Val Phe Pro Leu Arg
425 430 435
Pro Gly Gly Leu Pro Gly Leu Cys Glu Trp Leu Arg Arg Ala Leu
440 445 450
Pro Pro Ala Gln Ala Gly Ala Leu Leu Leu Ala Leu Pro Pro Ala
455 460 465
Ser Pro Ser Ala Ala Arg Thr Lys Ala Ala Ala Leu Arg Ala Gly
470 475 480
Ala Trp Arg Pro Ala Leu Leu Ala Ser Leu Ala Ala Ala Ala Ala
485 490 495
Pro Leu Pro Gly Leu Gly Trp Ala Cys Asp Val Ala Leu Leu Arg
500 505 510
Gly Gln Leu Ala Glu Trp Arg Arg Gly Leu Gly Leu Glu Pro Thr
515 520 525
Ala Leu Ala Arg Arg Glu Arg Ala Leu Gly Leu Ala Ser Gly Gln
530 535 540
Leu Ala Ala Arg Ala His Phe Pro Gly Pro Val Thr Arg Ala Glu
545 550 555
Val Glu Ala Arg Leu Gly Ala Trp Ala Gly Glu Gly Thr Ala Gly
560 565 570
Gly Ala Ala Leu Gly Ala Leu Ser Phe Leu Trp Pro Ala Gly Gly
575 580 585
Ala Ala Ala Thr Gly Gly Leu Gly Tyr Arg Ala Ala His Gly Val
590 595 600
Leu Leu Gln Ala Leu Asp Glu Met Arg Ala Asp Ala Glu Ala Val
605 610 615
Leu Ala Pro Pro Glu Pro Ala Gln
620

```

```

<210> 2
<211> 1025
<212> PPT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Inocyte ID No: 1322075CD1
<400> 2

```

WO 01/55392

PCT/US01/02741

Met His Arg Lys Lys Val Asp Asn Arg Ile Arg Ile Leu Ile Glu
 1 5 10 15
 Asn Gly Val Ala Glu Arg Gln Arg Ser Leu Phe Val Val Val Gly
 20 25 30
 Asp Arg Gly Lys Asp Gln Val Val Ile Leu His His Met Leu Ser
 35 40 45
 Lys Ala Ala Val Lys Ala Arg Pro Ser Val Leu Trp Cys Tyr Lys
 50 55 60
 Lys Glu Leu Gly Phe Ser Ile Leu Arg Lys Lys Arg Met Arg Gln
 65 70 75
 Leu Glu Lys Lys Ile Lys Asn Gly Thr Leu Asn Ile Lys Gln Asp
 80 85 90
 Asp Pro Phe Glu Leu Phe Ile Ala Ala Thr Asn Ile Arg Tyr Cys
 95 100 105
 Tyr Tyr Asn Glu Thr His Lys Ile Leu Gly Asn Thr Phe Gly Met
 110 115 120
 Cys Val Leu Gln Asp Phe Glu Ala Leu Thr Pro Asn Leu Leu Ala
 125 130 135
 Arg Thr Val Glu Thr Val Glu Gly Gly Glu Leu Val Val Ile Leu
 140 145 150
 Leu Arg Thr Met Asn Ser Leu Lys Gln Leu Tyr Thr Val Thr Met
 155 160 165
 Asp Val His Ser Arg Tyr Arg Thr Glu Ala His Gln Asp Val Val
 170 175 180
 Gly Arg Phe Asn Glu Arg Phe Ile Leu Ser Leu Ala Ser Cys Lys
 185 190 195
 Lys Cys Leu Val Ile Asp Asp Gln Leu Asn Ile Leu Pro Ile Ser
 200 205 210
 Ser His Val Ala Thr Met Glu Ala Leu Pro Pro Gln Thr Pro Asp
 215 220 225
 Glu Ser Leu Gly Pro Ser Asp Leu Glu Leu Arg Glu Leu Lys Glu
 230 235 240
 Ser Leu Gln Asp Thr Gln Pro Val Gly Val Leu Val Asp Cys Cys
 245 250 255
 Lys Thr Leu Asp Gln Ala Lys Ala Val Leu Lys Phe Ile Glu Gly
 260 265 270
 Ile Ser Glu Lys Thr Leu Arg Ser Thr Val Ala Leu Thr Ala Ala
 275 280 285
 Arg Gly Arg Gly Lys Ser Ala Ala Leu Gly Leu Ala Ile Ala Gly
 290 295 300
 Ala Val Ala Phe Gly Tyr Ser Asn Ile Phe Val Thr Ser Pro Ser
 305 310 315
 Pro Asp Asn Leu His Thr Leu Phe Glu Phe Val Phe Lys Gly Phe
 320 325 330
 Asp Ala Leu Gln Tyr Gln Glu His Leu Asp Tyr Glu Ile Ile Gln
 335 340 345
 Ser Leu Asn Pro Glu Phe Asn Lys Ala Val Ile Arg Val Asn Val
 350 355 360
 Phe Arg Glu His Arg Gln Thr Ile Gln Tyr Ile His Pro Ala Asp
 365 370 375
 Ala Val Lys Leu Gly Gln Ala Glu Leu Val Val Ile Asp Glu Ala
 380 385 390
 Ala Ala Ile Pro Leu Pro Leu Val Lys Ser Leu Leu Gly Pro Tyr
 395 400 405
 Leu Val Phe Met Ala Ser Thr Ile Asn Gly Tyr Glu Gly Thr Gly
 410 415 420
 Arg Ser Leu Ser Leu Lys Leu Ile Gln Gln Leu Arg Gln Gln Ser
 425 430 435
 Ala Gln Ser Gln Val Ser Thr Thr Ala Glu Asn Lys Thr Thr Thr
 440 445 450
 Thr Ala Arg Leu Ala Ser Ala Arg Thr Leu His Glu Val Ser Pro
 455 460 465
 Gln Glu Ser Ile Arg Tyr Ala Pro Gly Asp Ala Val Glu Lys Trp

WO 01/55392

PCT/US01/02741

```

470
Leu Asn Asp Leu Leu Cys Leu Asp Cys Leu Asn Ile Thr Arg Ile 480
485
Val Ser Gly Cys Pro Leu Pro Glu Ala Cys Glu Leu Tyr Tyr Val 495
500
Asn Arg Asp Thr Leu Phe Cys Tyr His Lys Ala Ser Glu Val Phe 510
515
Leu Gln Arg Leu Met Ala Leu Tyr Val Ala Ser His Tyr Lys Asn 525
530
Ser Pro Asn Asp Leu Gln Met Leu Ser Asp Ala Pro Ala His His 540
545
Leu Phe Cys Leu Leu Pro Pro Val Pro Pro Thr Gln Asn Ala Leu 555
560
Pro Glu Val Leu Ala Val Ile Gln Val Cys Leu Glu Gly Glu Ile 570
575
Ser Arg Gln Ser Ile Leu Asn Ser Leu Ser Arg Gly Lys Lys Ala 585
590
Ser Gly Asp Leu Ile Pro Trp Thr Val Ser Glu Gln Phe Gln Asp 600
605
Pro Asp Phe Gly Gly Leu Ser Gly Gly Arg Val Val Arg Ile Ala 615
620
Val His Pro Asp Tyr Gln Gly Met Gly Tyr Gly Ser Arg Ala Leu 630
635
Gln Leu Leu Gln Met Tyr Tyr Glu Gly Arg Phe Pro Cys Leu Glu 645
650
Glu Lys Val Leu Glu Thr Pro Gln Glu Ile His Thr Val Ser Ser 655
665
Glu Ala Val Ser Leu Leu Glu Glu Val Ile Thr Pro Arg Lys Asp 675
680
Leu Pro Pro Leu Leu Leu Lys Leu Asn Glu Arg Pro Ala Glu Arg 685
695
Leu Asp Tyr Leu Gly Val Ser Tyr Gly Leu Thr Pro Arg Leu Leu 705
710
Lys Phe Trp Lys Arg Ala Gly Phe Val Pro Val Tyr Leu Arg Gln 720
725
Thr Pro Asn Asp Leu Thr Gly Glu His Ser Cys Ile Met Leu Lys 735
740
Thr Leu Thr Asp Glu Asp Glu Ala Asp Gln Gly Gly Trp Leu Ala 750
755
Ala Phe Trp Lys Asp Phe Arg Arg Arg Phe Leu Ala Leu Leu Ser 765
770
Tyr Gln Phe Ser Thr Phe Ser Pro Ser Leu Ala Leu Asn Ile Ile 775
785
Gln Asn Arg Asn Met Gly Lys Pro Ala Gln Pro Ala Leu Ser Arg 795
800
Glu Glu Leu Glu Ala Leu Phe Leu Pro Tyr Asp Leu Lys Arg Leu 810
815
Glu Met Tyr Ser Arg Asn Met Val Asp Tyr His Leu Ile Met Asp 825
830
Met Ile Pro Ala Ile Ser Leu Ile Tyr Phe Leu Asn Gln Leu Gly 835
845
Asp Leu Ala Leu Ser Ala Ala Gln Ser Ala Leu Leu Leu Gly Ile 855
860
Gly Leu Gln His Lys Ser Val Asp Glu Leu Glu Lys Glu Ile Glu 865
875
Leu Pro Ser Gly Gln Leu Met Gly Leu Phe Asn Arg Ile Ile Arg 885
890
Lys Val Val Lys Leu Phe Asn Glu Val Gln Glu Lys Ala Ile Glu 905
910
Glu Gln Met Val Ala Ala Lys Asp Val Val Met Glu Pro Thr Met 915
920
Lys Thr Leu Ser Asp Asp Leu Asp Glu Ala Ala Lys Glu Phe Gln 935
940
945

```

WO 01/55392

PCT/US01/02741

Glu Lys His Lys Lys Glu Val Gly Lys Leu Lys Ser Met Asp Leu
 950 955 960
 Ser Glu Tyr Ile Ile Arg Gly Asp Asp Glu Glu Trp Asn Glu Val
 965 970 975
 Leu Asn Lys Ala Gly Pro Asn Ala Ser Ile Ile Ser Leu Lys Ser
 980 985 990
 Asp Lys Lys Arg Lys Leu Glu Ala Lys Gln Glu Pro Lys Gln Ser
 995 1000 1005
 Lys Lys Leu Lys Asn Arg Glu Thr Lys Asn Lys Lys Asp Met Lys
 1010 1015 1020
 Leu Lys Arg Lys Lys
 1025

<210> 3
 <211> 477
 <212> PRT
 <213> Homo sapiens
 <220>
 <221> misc_feature
 <223> Incyte ID No: 7472369CD1

<400> 3
 Met Gly Ser Lys Met Asn Leu Ile Glu His Ser His Leu Pro Thr
 1 5 10 15
 Thr Asp Glu Phe Ser Phe Ser Glu Asn Leu Phe Gly Val Leu Thr
 20 25 30
 Glu Glu Val Ala Gly Pro Leu Gly Gln Asn Leu Glu Val Glu Pro
 35 40 45
 Tyr Ser Gln Tyr Ser Asn Val Gln Phe Pro Gln Val Gln Pro Gln
 50 55 60
 Ile Ser Ser Ser Ser Tyr Tyr Ser Asn Leu Gly Phe Tyr Pro Gln
 65 70 75
 Gln Pro Glu Glu Trp Tyr Ser Pro Gly Ile Tyr Glu Leu Arg Arg
 80 85 90
 Met Pro Ala Glu Thr Leu Tyr Gln Gly Glu Thr Glu Val Ala Glu
 95 100 105
 Met Pro Val Thr Lys Lys Pro Arg Met Gly Ala Ser Ala Gly Arg
 110 115 120
 Ile Lys Gly Asp Glu Leu Cys Val Val Cys Gly Asp Arg Ala Ser
 125 130 135
 Gly Tyr His Tyr Asn Ala Leu Thr Cys Glu Gly Cys Lys Gly Phe
 140 145 150
 Phe Arg Arg Ser Ile Thr Lys Asn Ala Val Tyr Lys Cys Lys Asn
 155 160 165
 Gly Gly Asn Cys Val Met Asp Met Tyr Met Arg Arg Lys Cys Gln
 170 175 180
 Glu Cys Arg Leu Arg Lys Cys Lys Glu Met Gly Met Leu Ala Glu
 185 190 195
 Cys Met Tyr Thr Gly Leu Leu Thr Glu Ile Gln Cys Lys Ser Lys
 200 205 210
 Arg Leu Arg Lys Asn Val Lys Gln His Ala Asp Gln Thr Val Asn
 215 220 225
 Glu Asp Ser Glu Gly Arg Asp Leu Arg Gln Val Thr Ser Thr Thr
 230 235 240
 Lys Ser Cys Arg Glu Lys Thr Glu Leu Thr Pro Asp Gln Gln Thr
 245 250 255
 Leu Leu His Phe Ile Met Asp Ser Tyr Asn Lys Gln Arg Met Pro
 260 265 270
 Gln Glu Ile Thr Asn Lys Ile Val Leu Lys Glu Glu Phe Ser Ala
 275 280 285
 Glu Glu Asn Phe Leu Ile Leu Thr Glu Met Ala Thr Asn His Val
 290 295 300

WO 01/55392

PCT/US01/02741

Gln Val Leu Val Glu Phe Thr Lys Lys Leu Pro Gly Phe Gln Thr
 305 310 315
 Leu Asp His Glu Asp Gln Ile Ala Leu Leu Lys Gly Ser Ala Val
 220 325 330
 Glu Ala Met Phe Leu Arg Ser Ala Glu Ile Phe Asn Lys Lys Leu
 335 340 345
 Pro Ser Gly His Ser Asp Leu Leu Glu Glu Arg Ile Arg Asn Ser
 350 355 360
 Gly Ile Ser Asp Glu Tyr Ile Thr Pro Met Phe Ser Phe Tyr Lys
 365 370 375
 Ser Ile Gly Glu Leu Lys Met Thr Gln Glu Glu Tyr Ala Leu Leu
 380 385 390
 Thr Ala Ile Val Ile Leu Ser Pro Asp Arg Gln Tyr Ile Lys Asp
 395 400 405
 Arg Glu Ala Val Glu Lys Leu Gln Glu Pro Leu Leu Asp Val Leu
 410 415 420
 Gln Lys Leu Cys Lys Ile His Gln Pro Glu Asn Pro Gln His Phe
 425 430 435
 Ala Cys Leu Leu Gly Arg Asp Thr Glu Leu Arg Thr Phe Asn His
 440 445 450
 His His Ala Glu Met Leu Met Ser Trp Arg Val Asn Asp His Lys
 455 460 465
 Phe Thr Pro Leu Leu Cys Glu Ile Trp Asp Val Gln
 470 475

<210> 4
 <211> 3274
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incode ID No: 3462174CB1

<400> 4
 atttcagagt ttttcaagca gaasgggaaa cgtttccatc tagagggagg gaggacagac 60
 aggaacgctc ggattctggg ggaatagggt ctgggctctg gattttttag gtcttgagag 120
 gtgacactca ggaacggaag gggacagggt gaaggactga ggcctcggtt cggagtgggg 180
 cggctaggaa aggaaaaggg taggtgctcc ctgacccagg tctctcttan cgttgcttt 240
 ccagccatcg ttccgctgca ggggtgctcg accgctctgt tcttggggcu lccgggcttg 300
 ggggaagtcc cgtgcatcgc agcgtctgtc gacaaagatg tggagagcct caggggcccc 360
 gagggaacgc ggaactccgg ggttcccaag ctanngcrg cggggccagg cctttttctg 420
 gggaagctga gctcccccgc cgcacacccg gggcccttgg cggcgggaag caactgctg 480
 ctactgttgt tcccggagcc cgaagggaac ggggaacgt tggctcagc cctgggagag 540
 gcagggcttg cggccttggc cggaggagac ctgctactgg ctgtgcgga cctccgctc 600
 ggggtatcgc agactgcgc ccaggccctg gctcagacag cagctctact caacagctcg 660
 ggggtaggag ctggcgacat gtttttgcct cggcggaact ggggcagcag cggcggcttc 720
 gaggagctag agcgcctccg ggcgggctcg cggcgccagg caggagcgcct cgggagcttc 780
 ctgccaacgg cccaggatgg ctctggagct tgggtgtag caggagctaga ggtatgcat 840
 gaggcctttg agacagcggt ccttgaggct ggcctgctgt ggttcgcctc aggcctggag 900
 cgccttgcca ggcacagcgt agacctggcc ctggttagca agcctgaagt gggccttgrg 960
 gtcgacatgc tgcctggatt ggaatctggc gacctaggcg ctgcgcctgc ttcggctccc 1020
 acagaccccc ctccctccc agcccccagg ccccccattg tgcctgcttg gacgtgcat 1080
 ctgggcccac cggggaactgc ccccccctgg gctgcgcctc ctacacccac gactacgac 1140
 ggcctatcct tgcctacccc tggggctccc actggaagag actgggcccc ggtccagggc 1200
 ctgctgctac cagatggccc tctgtctctg ggggacag agggacaggg cggagatccg 1260
 gactgtctcg gaggaggcaa catgaataat ccaaggggcg agagcttaa gaacccagct 1320
 ggaaggggga tggagaaatgc actcagtaag ggaagggaga aatgtatgct tggatccgag 1380
 aaagtgagca cccgggaagc cctcgggaag gctggcagcg agggcttaca gcatgtctgc 1440
 ggcctgagaa aatccggctg gggcgactca gaggggcccg ctgccttaag cccggaggac 1500
 ggaagctggg aggtgtctga ggaaggcccg cgcctagctg tcccctaac gctcggggga 1560
 cccccgggc tatgcgaatg gctgcggcga ggcctcccc cagccagggc aggggcctcg 1620
 ctgctggcgt tgcacacagc atctcccgag gctgcccga ccaaggctcg ggcgttcgca 1680

WO 01/55392

PCT/US01/02741

```

gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 1740
gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 1800
gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 1860
gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 1920
gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 1980
gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 2040
gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 2100
gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 2160
gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 2220
gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 2274

```

```

<210> 5
<211> 4154
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> Incyte ID No: 1322075CB1

```

```

<226>
<221> unsure
<222> 4021, 4121
<223> a, t, c, g, or other

```

```

<400> 5
gtggcgaggac caggcttact acgtgacccg gacacacagg Atacgtctagg ggcagtcagg 60
tggtcctctct ctttcggagt tgttcctgtgc tccacacgtgc ttcctctctc ccacttgcctg 120
ggatccccccg ggtcggggcg gcatgaataa ttttcaaca tgcctcgaa aaaggctggat 160
aacggaatcc ggattctcat tgagaatgga gtatctgagc ggcgaagatc tctcttctgt 240
gtagctgggg atcgaggaaa agatcaggctg gtatctcttc atccatgtct atccaaagca 300
gtctgtgagg ctggcctctc agtgcctgctg gctctcaggg aagggcctggg gcttagcatt 360
ctccggaaa aaagaaatgcg acagctgcag agaaatcaa agaaatgga acatgaacaa 420
aacggggagc accccttctg actctctata gacgaccaa acattctgca ctgctactac 480
aacggagccc acaagatccc gggcaatacc ttctgcatct gtgtctctga ggaatctgaa 540
gccctaaact caaaacttgt ggcacagact gtgaaacacg tggaaagctg tgggctagt 600
gtctctctcc tacggacct gaactcactc aagcaattgt ccacagtgac tatggatgt 660
cattccaggt acagaactga ggcctcatcg gatgtgtgag gaacatttaa tgaaaggttt 720
attctgtctc tggcctcttg taagaagtgt ctctctcttg atcaccagct caacatctct 780
cccctctctc cccacgtctg caccatggag gccctctctc cccagacttc ggtatgaggt 840
cttctctctc ctgctctgga gctgaggagg ttgaaggaga gcttgcagga caccagagct 900
cttctctctc tgtgtgactg ctgtgaagct ctgaacagg ccgaagctgt cttgaatctt 960
atcgaggaga ctctggaaga gacctgagg agtactgttg cactcagag tctctgagga 1020
gggggaaaaa ctgcagagct gggattgaga atctctgggg gctgagcatt tgggacatct 1080
aatatctttg ttacctctcc aagccttgat aactctcat ctctgttga atttgattt 1140
aaaggatttg atgctctgca atacagggaa catctggatt atgaaattat ccagctctca 1200
aaccttgat ttacaaagc agtgacacaga gtaactgtat ttctggaaca caggcagact 1260
attcagata tacatctctg agtgcctgtg aggtgtggcc aggtctgaat agtctgatt 1320
gatgaagctg ccgcatctcc cctcctcttg gtgaagagcc taactgagcc ctactctgt 1380
ttctctgcat ccaacataa tggctatgag tgcacttgcc ggtcactctc cctcaagta 1440
attcagcagc tctgtcaaca gacgcccag agccaggtca gacacctgc tgaagataag 1500
accacagaga cagcagagat ggcctcagcg cggacactgc atcaggttly ucccagagag 1560
tcaacacagat acgcccctgg ggaagcagtg gaaagctgag tgaatgactt gctgtgctg 1620
gattgactca acatcactcg gatagctcga ggtctctctt tgcctgagc tctgtgactg 1680
tactagata atagagatct cctcttctgt taacacagag cctctgagat ttctctctca 1740
gggtctctgg cctctcagc ggcctctctc taacaaagat ctctctgag ttctctctca 1800
ctctctctgg cactctgca ccatctctct tgcctctctg ctctctctg ctctctctca 1860
aacgctctct cagaagctgt tgcctctctc caggtctgag ttgaagcaga gattctctca 1920
cagctcactc tgaacagctt gctctcagag aagaaggtct caggggagctt gattctctca 1980
acagtgctcag aacagttcaa agatccagag ttctgggtct tctctggtg aggggtctct 2040
cgcattctgt ttacccagga ttatcaaggg atgggtctat gacgctctg cctgagagct 2100
ctgcagagct acctgaagc caggtttctc tctctggagg aagaggtct tgaagacaa 2160
caggaaattc acacgttaag cagcagagct gtcagctctc tggaaagagct cactactctc 2220

```

WO 01/55392

PCT/US01/02741

```

cggaaggacc  tgcctccctt  actccctcaa  ttgaatgaga  ggcctgcgga  accgctggat  2280
taccctgggt  ttctctatgg  ctggaccccc  aggtctccca  agtctctgaa  accgctggat  2340
ttgtttccct  ttatctctag  ccgaaccccc  aatgacctga  ccggaagaca  ctctgtcatc  2400
atgctgaaga  cgtccactga  cggagatgag  gctgacctga  gaggtctggt  tccagccttc  2460
cggaaggatt  tccgaacggc  gtctctagcc  ttgtctctct  accagttcag  taccctctct  2520
ccctccctgg  ctccgaacat  cttccgaacc  aggaacatgg  ggaagccagg  cctgctctgc  2580
ctgagccggg  agggctctga  agcacccttc  ctccctctatg  accggaaggg  gctggggatg  2640
tattcaagg  atatggtgga  ctatcccttc  ctccctctatg  agatccctgc  cctctctctc  2700
atctattctc  tgaacccagg  gggggacccc  cctctctctg  cggctcagtc  ggtctctctc  2760
ttggggatgg  gctgtcagca  caagctcttg  gacccagctg  caagctcttg  tgaacccccc  2820
tcggggcagg  tgaatgggac  tttnaacccg  atccctctga  aagttctgaa  cctctctaat  2880
gaagttccag  aaaaagccat  tgaagagcag  atggtctgag  cgaagatgt  ggtctctgag  2940
cccccgatga  agacccctag  tgaaccccta  gatgaagcag  caaagatct  tccagggaaa  3000
cacaagaagg  aagtgggaaa  gctgaagagc  atggacctct  ctgaatctat  aatccgtggg  3060
cagcatgag  agtggatga  agtttllgaa  aagctctggc  cgaacccctc  gatcctcag  3120
ctgaagag  acaagaagag  gaantttag  gcccaacag  aaccccaac  gaggcagag  3180
ttgaagag  gaggagag  gaacaaaaa  gatcctgaa  tgaacccccc  gaaacccccc  3240
aggaagag  gggcatctgc  gcttctctat  gggagagat  tctcactaac  tgaacccccc  3300
ctggctggag  tgttaaaagc  aacgagag  cccggcagac  ctggagagct  gggcctctct  3360
cggctctctg  gctgtctctg  ctggaagctc  accctgctca  aaggaagag  cactcccaaa  3420
cggctctctg  tgaagctctg  cggctctctg  ctggaagctc  aaggaagag  cactcccaaa  3480
ctctctctgc  cagctctctg  gctctctctg  cctctctctg  ctctctctg  ctctctctg  3540
ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  3600
ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  3660
ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  3720
ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  3780
ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  3840
ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  3900
ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  3960
ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  4020
ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  4080
ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  4140
ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  ctctctctgc  4200

```

<210> 6

<211> 2220

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> Intron 13 No: 7472269051

<400> 6

```

tccctccctc  attgtctctc  cgaattatcc  taattcgaaa  ttggattctg  agcattctga  50
gcaaaatcgc  tggggtctgg  agaggagagc  tcaagtccag  atccctccag  ggccttcaaa  120
gtccatctct  gacccaaaac  aatccaaagg  ggtagaagac  atctatagag  ggtggaaga  180
agaaagagag  actcagaagc  atagctcaaa  gtgaacacgc  ctctctctag  ttctctggat  240
ttctctctga  cattctctca  agatgaagac  ctgaacacgc  ttggattctg  ttctctggat  300
accataaaga  aagtgcattt  caattgaaac  attctgagtg  gatcaaaat  gattctctat  360
gaacattctc  atttaccctc  cagagatgaa  ttctctctct  ctgaaaatct  attctggtgt  420
taacacgagc  aagtggcagg  tctctctgga  cagaacctgg  aagtggagac  atattctgaa  480
taacacgagc  ttcaattctc  ccaagttcaa  ccacagatct  cctctctatc  ctattattcc  540
aaactgggtt  ttatccctcc  gcaagctgaa  gagggtgact  cctctctgaa  atattgaaac  600
aggtctctgc  caggtgagac  ttctctctcc  ggaagaaact  aggtgagcga  atattgaaac  660
acaaagagag  cctgagatgg  cctgagatga  gggagagaca  aggggagaga  gctgctctgt  720
gttttctgag  acagagcctc  tgaatccaac  taaatctgac  tgaacccccc  gggggtcaaa  780
ggtttctctc  ggaagagcct  taccacaaac  gctctgtaca  agtgaaaaa  cgggagcaca  840
tgtgtgtatg  atattgacac  cgaagagaaa  tgcagagagt  gtcagatcag  gaaatgcaaa  900
gagatgggaa  tgtctgtgca  atgtatgct  acaggtctct  taactgaat  caggtgtaaa  960
tctaaagcag  tgaagaaaaa  tgtgaagcag  cctgagagac  agacccgtaa  tgaagagagt  1020
gaaggtctgt  atttgcgaca  agtgacctgc  acaacaaagt  catgagcaga  gaaactgaa  1080
ctcaccacag  ataacagag  tctctctact  tctattatcg  attcatataa  caaacagagg  1140

```

WO 01/55392

PCT/US01/02741

```

atgcctcagg aaataacaaa laaaattgta ttaaagaag aattcagtc agaagaaac 1200
tttctcattt tgacgggaat ggaacccaat catgtacagg tctttgtgga attcacaana 1260
aagctaacag gatttcagac ttgggaccac gaagaccaga ttgctttgct gaaagggctc 1320
gggttgaa ctatcttctt tggtracgt gagatttcca ataaqaact tccgtctggg 1380
cattctgacc tattggaaaa aaqaattcga aatagtggta cctccgata acataaaca 1440
cctatgtcta gctttataa aagratggg gaactaaaa cgcctcaaga gagtatgt 1500
ctgcttccag caattgttat cctgtctcca gatagacac acataaaga taagaggca 1560
gcaaggaagc tccaggagcc acttcttgat gctgacaaa agttgtgta gattaccag 1620
cctgaanaac ctcaacatt tgcctgtctc ctgggtcgc tgaactaatt acggacattc 1680
aatcatcacc acgcctgagat gctgatgta tggagagtaa acgaccaca gtttaccaca 1740
cttctctgtg aaattctgga cgtgcaatga tggggattac aggggagggg ccttagctct 1800
tttctctctt catattaatc tgaatataa ctttcttta tttcactgc accagtttc 1860
actcaagaaa tcttgatgaa tatttatgtt gtaattacat gtgtacttc cacaactgca 1920
aatactgggc tagctagaac aaatttctct acatttgtt ttaaaaggct ccagggaatc 1980
ctgcatctta attggcaagc cctgtttgct taattaaatt gattgttact tcaattctat 2040
ctgttgactc agggaaatc tcatcttgct cactctacca tattgcatat attttattca 2100
agagttgat tcaatcttgg caactaagca aacataatgg caacaaaaa aaactcgaat 2160
cgatctaaa aaaaagatca tcaattaac agtctaata tatttctta atgattgcaa 2220

```


【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
2 August 2001 (02.08.2001)

PCT

(10) International Publication Number
WO 01/55392 A3(51) International Patent Classification: C12N 15/12,
US10, C07K 14/705, 14/72, 16/28, C12Q 1/68, A61K
38/17, G01N 3/56, 3/53

(21) International Application Number: PCT/US01/02741

(22) International Filing Date: 26 January 2001 (26.01.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/178,578 28 January 2000 (28.01.2000) US
60/185,079 25 February 2000 (25.02.2000) US
60/189,999 17 March 2000 (17.03.2000) US(63) Related by continuation (CON) or continuation-in-part
(CIP) to earlier applications:
US 60/178,578 (CON)
Filed on 28 January 2000 (28.01.2000)
US 60/185,079 (CON)
Filed on 25 February 2000 (25.02.2000)
US 60/189,999 (CON)
Filed on 17 March 2000 (17.03.2000)(71) Applicant (for all designated States except US): INCYTE
GENOMICS, INC. [US/US], 3160 Porter Drive, Palo
Alto, CA 94304 (US).

(72) Inventors: and

(75) Inventors/Applicants (for US only): BURFORD, Neil
[GB/US], 105 Wildwood Circle, Durham, CT 06422 (US);
BAUGHN, Mariah, R. [US/US], 14244 Santiago Road,
San Leandro, CA 94577 (US); TANG, Y. Tan [CN/US];4290 Ramack Court, San Jose, CA 95118 (US); BAND-
MAN, Olga [US/US], 366 Anna Avenue, Mountain View,
CA 94043 (US); TRIBOULEY, Catherine, M. [FR/US];
1121 Tennessee Street, #5, San Francisco, CA 94107 (US);
DAS, Debopriya [IN/US], 1179 Bonita Avenue, Apt. 3,
Mountain View, CA 94040 (US); POLICKY, Jennifer,
L. [US/US], 4881 National Avenue, San Jose, CA 95124
(US); LO, Terence, P. [CA/US], 1452 Birch Park Boul-
vard, Apt. 115, Foster City, CA 94034-1941 (US).(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics,
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,
DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, GR,
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.(84) Designated States (regional): ARIPO patent (GH, 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, 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:
14 March 2002For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

WO 01/55392 A3

(54) Title: NUCLEAR HORMONE RECEPTORS

(57) Abstract: The invention provides human nuclear hormone receptors (NHRs) and polynucleotides which identify and encode NHRs. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating or preventing disorders associated with abnormal expression of NHRs.

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International Application No. PC./US 01/02741
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/12 C12N15/10 C07K14/705 C07K14/72 C07K16/28 C12Q1/68 A61K38/17 G01N33/50 G01N33/53		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12N C07K C12Q A61K G01N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EP0-Internal		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL [Online] Entry/Acc.no. AC006276. 6 January 1999 (1999-01-06) LAMERDIN, J.E. ET AL.: "Homo sapiens chromosome 19, cosmid R28379, complete sequence." XP002174692 * see nt. 13088-14585 *	11,12
X	DATABASE EMBL [Online] Entry/Acc.no. A1125175. 22 September 1998 (1998-09-22) HILLIER, L. ET AL.: "am55h07.x1 Johnston frontal cortex Homo sapiens cDNA clone IMAGE:1539517 3', mRNA sequence." XP002174693 the whole document --- -/-	11,12
☒ Further documents are listed in the continuation of box C. ☒ Patent family members are listed in annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document published on or after the international filing date "L" document which may throw doubts on priority claims or which is cited to establish the publication state of another claim or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date (as not in coming with the specification but cited to understand the purpose or theory underlying the invention) "X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "Z" document member of the same patent family		
Date of the actual completion of the international search 13 August 2001		Date of making of the international search report 30.11.2001
Name and mailing address of the ISA European Patent Office, P.O. Box 50118, 8001 München 2 NL - 2280 HV Rijswijk Tel: +31-701 340-2040, Tx: 31 651 epo n, Fax: +31-701 340-3016		Authorized officer Smalt, R

Form: PCT/ISA/210 (multiple sheets) July 1997

INTERNATIONAL SEARCH REPORT

International Application No.
PG/1, US 01/02741

C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 95 13373 A (ARCH DEV CORP ; LIAO SHUTSUNG (US); SONG CHING (US)) 18 May 1995 (1995-05-18) the whole document ---	
A	WO 99 35245 A (SALK INST FOR BIOLOGICAL STUDI) 15 July 1999 (1999-07-15) the whole document ---	
A	WO 99 48915 A (GLAXO GROUP LTD ; KIEWER STEVEN ANTHONY (US); WILLSON TIMOTHY MARK) 30 September 1999 (1999-09-30) the whole document -----	

INTERNATIONAL SEARCH REPORT	national application No. PCT/US 01/02741
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. ☒ Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely: Although claim 18 is 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. 2. ☒ Claims Nos.: 20,21,23,24 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: see FURTHER INFORMATION sheet PCT/ISA/210 3. ☐ Claims Nos.: 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. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims. 2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee. 3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.: 4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-28 all partly	
Remark on Protest ☐ The additional search fees were accompanied by the applicant's protest. ☐ No protest accompanied the payment of additional search fees.	

International Application No. PCT/US 01/02741

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-28, all partially

Human nuclear receptor polypeptide with amino acid sequence according to seq.ID.1, variants and immunologically or biologically active fragments thereof, polynucleotide encoding it as represented by seq.ID.4, variants and complements thereof and nucleic acids comprising at least 60 consecutive nucleotides thereof, optionally comprising a functionally linked promoter, cell or transgenic organism comprising said nucleic acid, method for producing said polypeptide, isolated antibody directed at said polypeptide, methods for detecting the polypeptide and the polynucleotide, pharmaceutical composition comprising said polypeptide, methods for screening for (anti)agonists, specific binding agents, or modulators of the activity of said polypeptide, method for screening for an agent which alters the expression of said polynucleotide, and use of said method to access the toxicity of a compound.

2. Claims: 1-28, all partially

Human nuclear receptor polypeptide with amino acid sequence according to seq.ID.2, variants and immunologically or biologically active fragments thereof, polynucleotide encoding it as represented by seq.ID.5, variants and complements thereof and nucleic acids comprising at least 60 consecutive nucleotides thereof, optionally comprising a functionally linked promoter, cell or transgenic organism comprising said nucleic acid, method for producing said polypeptide, isolated antibody directed at said polypeptide, methods for detecting the polypeptide and the polynucleotide, pharmaceutical composition comprising said polypeptide, methods for screening for (anti)agonists, specific binding agents, or modulators of the activity of said polypeptide, method for screening for an agent which alters the expression of said polynucleotide, and use of said method to access the toxicity of a compound.

3. Claims: 1-28, all partially

Human nuclear receptor polypeptide with amino acid sequence according to seq.ID.3, variants and immunologically or biologically active fragments thereof, polynucleotide encoding it as represented by seq.ID.6, variants and complements thereof and nucleic acids comprising at least 60 consecutive nucleotides thereof, optionally comprising a functionally linked promoter, cell or transgenic organism comprising said nucleic acid, method for producing said polypeptide, isolated antibody directed at said polypeptide, methods for detecting the polypeptide and the

International Application No. PCT/US 01/02741

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

polynucleotide, pharmaceutical composition comprising said polypeptide, methods for screening for (ant)agonists, specific binding agents, or modulators of the activity of said polypeptide, method for screening for an agent which alters the expression of said polynucleotide, and use of said method to access the toxicity of a compound.

International Application No. PCT/US 01/02741

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box 1.2

Claims Nos.: 20,21,23,24

Present claims 20,21,23 and 24 relate to a compound defined by reference to a desirable characteristic or property, namely being (ant)agonistic to the polypeptide(s) of claim 1.

The application provides no support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for such compounds. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search for said claims is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the product by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search for said claims impossible.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT
 Information on patent family members

 International Application No.
 PCT/US 01/02741

Parent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9513373 A	18-05-1995	AU 1173995 A WO 9513373 A1 US 5639616 A	29-05-1995 18-05-1995 17-06-1997
WO 9935246 A	15-07-1999	AU 2218899 A EP 1051479 A1 WO 9935246 A1	26-07-1999 15-11-2000 15-07-1999
WO 9948915 A	30-09-1999	AU 3211699 A EP 1066320 A1 WO 9948915 A1	18-10-1999 10-01-2001 30-09-1999

Form PCT/ISA/210 (BREVET INTERNETION) May 1997

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 15/00	A 6 1 P 35/00	4 C 0 8 4
A 6 1 P 35/00	A 6 1 P 37/00	4 H 0 4 5
A 6 1 P 37/00	C 0 7 K 14/72	
C 0 7 K 14/72	C 0 7 K 16/28	
C 0 7 K 16/28	C 1 2 N 1/15	
C 1 2 N 1/15	C 1 2 N 1/19	
C 1 2 N 1/19	C 1 2 N 1/21	
C 1 2 N 1/21	C 1 2 P 21/02	C
C 1 2 N 5/10	C 1 2 Q 1/68	A
C 1 2 P 21/02	G 0 1 N 33/15	Z
C 1 2 Q 1/68	G 0 1 N 33/50	Z
G 0 1 N 33/15	G 0 1 N 33/53	D
G 0 1 N 33/50	G 0 1 N 33/53	M
G 0 1 N 33/53	G 0 1 N 33/566	
G 0 1 N 33/566	C 1 2 N 5/00	A
	A 6 1 K 37/02	

(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,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,S,G,SI,SK,SL,TJ,TM,TR,TT,TZ,UA,UG,US,UZ,VN,YU,ZA,ZW

- (72)発明者 ボーグン、マライア・アール
アメリカ合衆国カリフォルニア州 9 4 5 7 7 ・サンレアンドロ・サンティアゴロード 1 4 2 4 4
- (72)発明者 タング、ワイ・トム
アメリカ合衆国カリフォルニア州 9 5 1 1 8 ・サンノゼ・ランウィックコート 4 2 3 0
- (72)発明者 バンドマン、オルガ
アメリカ合衆国カリフォルニア州 9 4 0 4 3 ・マウンテンビュー・アンナアベニュー 3 6 6
- (72)発明者 トリボレー、キャサリーン・エム
アメリカ合衆国カリフォルニア州 9 4 1 0 7 ・サンフランシスコ・# 5 ・テネシーストリート 1 1 2 1
- (72)発明者 ダス、デボプリヤ
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・アパートメント 3 ・ボニータアベニュー 1 1 7 9
- (72)発明者 ポリッキー、ジェニファー・エル
アメリカ合衆国カリフォルニア州 9 5 1 2 4 ・サンノゼ・ナショナルアベニュー 4 8 6 4
- (72)発明者 ロー、テレンス・ピー
アメリカ合衆国カリフォルニア州 9 4 4 0 4 - 1 9 4 1 ・フォスターシティ・アパートメント 1 1 5 ・ビーチパークブルーバード 1 4 5 2

F ターム(参考) 2G045 AA34 AA35 BB01 BB07 BB10 BB50 BB51 CB01 DA13 DA36
FB02 FB05 FB08 FB12 GC15
4B024 AA01 AA11 BA61 BA63 CA04 DA02 DA05 DA06 DA11 EA02
EA04 GA11 HA12
4B063 QA01 QA18 QQ42 QQ53 QR32 QR56 QR62 QR77 QR82 QX02
4B064 AG20 AG26 CA02 CA05 CA10 CA11 CA19 CC24 DA01 DA13
4B065 AA01X AA26X AA57X AA87X AA93Y AB01 AB04 CA24 CA44 CA46

4C084	AA02	AA03	AA07	AA17	BA01	BA08	BA22	CA18	MA13	MA16
	MA43	MA52	MA56	MA57	MA59	MA60	MA63	MA66	NA14	ZA812
	ZB052	ZB262	ZC032	ZC212						
4H045	AA10	AA11	AA20	AA30	BA09	CA40	DA50	DA75	EA20	EA50
	FA72	FA74								

专利名称(译)	核激素受体		
公开(公告)号	JP2004500819A	公开(公告)日	2004-01-15
申请号	JP2001554421	申请日	2001-01-26
[标]申请(专利权)人(译)	洞察Genomics公司		
申请(专利权)人(译)	洞察基因组公司		
[标]发明人	バーフォードニール ボーグンマライアアール タングワイトム バンドマンオルガ トリボレーキャサリーンエム ダスデボプリヤ ポリッキージェニファーエル ローテレンスピー		
发明人	バーフォード、ニール ボーグン、マライア・アール タング、ワイ・トム バンドマン、オルガ トリボレー、キャサリーン・エム ダス、デボプリヤ ポリッキー、ジェニファー・エル ロー、テレンス・ピー		
IPC分类号	G01N33/50 A61K38/00 A61K45/00 A61P3/00 A61P5/00 A61P15/00 A61P35/00 A61P37/00 C07K14/705 C07K14/72 C07K16/28 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/09 C12N15/12 C12P21/02 C12Q1/68 G01N33/15 G01N33/53 G01N33/566		
CPC分类号	A61K38/00 A61P3/00 A61P15/00 C07K14/70567 C07K14/721		
FI分类号	C12N15/00.ZNA.A A61K45/00 A61P3/00 A61P5/00 A61P15/00 A61P35/00 A61P37/00 C07K14/72 C07K16/28 C12N1/15 C12N1/19 C12N1/21 C12P21/02.C C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N33/566 C12N5/00.A A61K37/02		
F-TERM分类号	2G045/AA34 2G045/AA35 2G045/BB01 2G045/BB07 2G045/BB10 2G045/BB50 2G045/BB51 2G045/CB01 2G045/DA13 2G045/DA36 2G045/FB02 2G045/FB05 2G045/FB08 2G045/FB12 2G045/GC15 4B024/AA01 4B024/AA11 4B024/BA61 4B024/BA63 4B024/CA04 4B024/DA02 4B024/DA05 4B024/DA06 4B024/DA11 4B024/EA02 4B024/EA04 4B024/GA11 4B024/HA12 4B063/QA01 4B063/QA18 4B063/QQ42 4B063/QQ53 4B063/QR32 4B063/QR56 4B063/QR62 4B063/QR77 4B063/QR82 4B063/QX02 4B064/AG20 4B064/AG26 4B064/CA02 4B064/CA05 4B064/CA10 4B064/CA11 4B064/CA19 4B064/CC24 4B064/DA01 4B064/DA13 4B065/AA01X 4B065/AA26X 4B065/AA57X 4B065/AA87X 4B065/AA93Y 4B065/AB01 4B065/AB04 4B065/CA24 4B065/CA44 4B065/CA46 4C084/AA02 4C084/AA03 4C084/AA07 4C084/AA17 4C084/BA01 4C084/BA08 4C084/BA22 4C084/CA18 4C084/MA13 4C084/MA16 4C084/MA43 4C084/MA52 4C084/MA56 4C084/MA57 4C084/MA59 4C084/MA60 4C084/MA63 4C084/MA66 4C084/NA14 4C084/ZA812 4C084/ZB052 4C084/ZB262 4C084/ZC032 4C084/ZC212 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30 4H045/BA09 4H045/CA40 4H045/DA50 4H045/DA75 4H045/EA20 4H045/EA50 4H045/FA72 4H045/FA74		
優先権	60/178578 2000-01-28 US 60/185079 2000-02-25 US 60/189999 2000-03-17 US		

摘要(译)

本发明提供了鉴定和编码NHREC的人核激素受体 (NHREC) 和多核苷酸。 本发明还提供表达载体，宿主细胞，抗体，激动剂和拮抗剂。 本发明还提供了用于诊断，治疗或预防与NHREC异常表达有关的疾病的方法。

(P2004-500619A)			
(43) 公表日 平成16年1月15日 (2004.1.15)			
(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	2 G 0 4 5
A 6 1 K 38/00	A 6 1 K 45/00		4 B 0 2 4
A 6 1 K 45/00	A 6 1 P 3/00		4 B 0 6 3
A 6 1 P 3/00	A 6 1 P 5/00		4 B 0 6 4
A 6 1 P 5/00	A 6 1 P 15/00		4 B 0 6 5
審査請求 未請求		予備審査請求 有	(全 186 頁) 最終頁に続く
(21) 出願番号	特願2001-554421 (P2001-554421)	(71) 出願人	301005050
(36) (22) 出願日	平成13年1月26日 (2001.1.26)		インサイト・ゲノミックス・インコーポレ
(35) 翻訳文提出日	平成14年7月29日 (2002.7.29)		イテッド
(36) 国際出願番号	PCT/US2001/002741		アメリカ合衆国カリフォルニア州9430
(37) 国際公開番号	W02001/055392		4・パロアルト・ボータードドライブ 31
(37) 国際公開日	平成13年8月2日 (2001.8.2)		60
(31) 優先権主張番号	60/178,578	(74) 代理人	100089266
(32) 優先日	平成12年1月28日 (2000.1.28)		弁理士 大島 陽一
(33) 優先権主張国	米国 (US)	(72) 発明者	バーフォード、ニール
(31) 優先権主張番号	60/185,079		アメリカ合衆国コネチカット州06422
(32) 優先日	平成12年2月25日 (2000.2.25)		・ダラム・ワイルドウッドサークル 10
(33) 優先権主張国	米国 (US)		5
(31) 優先権主張番号	60/189,999		
(32) 優先日	平成12年3月17日 (2000.3.17)		
(33) 優先権主張国	米国 (US)		