

(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2004-520027

(P2004-520027A)

(43) 公表日 平成16年7月8日(2004.7.8)

(51) Int.Cl.⁷

F I

テーマコード (参考)

C 1 2 N 15/09

C 1 2 N 15/00 Z N A A

C 0 7 K 16/40

C 0 7 K 16/40

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

審査請求 未請求 予備審査請求 有 (全 271 頁) 最終頁に続く

(21) 出願番号 特願2002-552162 (P2002-552162)

(71) 出願人 597011463

(86) (22) 出願日 平成13年12月18日 (2001.12.18)

ノバルティス アクチエンゲゼルシャフト

(85) 翻訳文提出日 平成15年6月19日 (2003.6.19)

スイス国、4 0 5 6 バーゼル、リヒトシ

(86) 国際出願番号 PCT/EP2001/014928

ユトラーセ 3 5

(87) 国際公開番号 W02002/050285

(74) 代理人 100062144

(87) 国際公開日 平成14年6月27日 (2002.6.27)

弁理士 青山 稔

(31) 優先権主張番号 60/256,827

(74) 代理人 100067035

(32) 優先日 平成12年12月20日 (2000.12.20)

弁理士 岩崎 光隆

(33) 優先権主張国 米国 (US)

(74) 代理人 100064610

(31) 優先権主張番号 60/293,089

弁理士 中嶋 正二

(32) 優先日 平成13年5月23日 (2001.5.23)

(74) 代理人 100072730

(33) 優先権主張国 米国 (US)

弁理士 小島 一晃

(31) 優先権主張番号 60/317,984

(32) 優先日 平成13年9月6日 (2001.9.6)

(33) 優先権主張国 米国 (US)

最終頁に続く

(54) 【発明の名称】 ヒストンデアセチラーゼ関連遺伝子およびタンパク質

(57) 【要約】

H D A C 関連遺伝子および遺伝子産物が開示されている。特に、本発明は、既知の H D A C に相同性が高くそして本明細書において H D A C 9 と呼ばれるタンパク質および変異体、そのようなタンパク質をコードする核酸分子、該タンパク質をコードする抗体、ならびに H D A C 活性または遺伝子発現の異常に関連した病状の診断方法に関する。

【特許請求の範囲】

【請求項 1】

配列番号 1、配列番号 5 または配列番号 6 で示されるアミノ酸配列を含む単離ポリペプチド。

【請求項 2】

配列番号 1、配列番号 5 または配列番号 6 で示されるアミノ酸配列からなる単離ポリペプチド。

【請求項 3】

請求項 1 または 2 に記載のポリペプチドをコードする核酸配列を含む単離 DNA。

【請求項 4】

請求項 3 に記載の単離 DNA の少なくとも 1 つのフラグメントを含むベクター分子。

【請求項 5】

転写制御配列を含む請求項 4 に記載のベクター分子。

【請求項 6】

請求項 5 に記載のベクター分子を含む宿主細胞。

【請求項 7】

請求項 3 に記載の単離 DNA であって、

(1) 請求項 2 において定義されたポリペプチドをコードする完全 cDNA 配列である、配列番号 2、7 または 8 で示されるヌクレオチド配列；

(2) 請求項 2 において定義されたポリペプチドをコードする cDNA 配列のオープンリーディングフレームである、配列番号 3 で示されたヌクレオチド配列；

(3) 配列番号 3 で示されたヌクレオチド配列にハイストリンジェント条件下でハイブリダイズできるヌクレオチド配列；および

(4) 請求項 2 において定義されたポリペプチドをコードする内在性ヒトゲノム DNA である、配列番号 4 で示されたヌクレオチド配列、
からなる群から選択されるヌクレオチド配列を含む単離 DNA。

【請求項 8】

請求項 7 に記載の単離 DNA 分子の少なくとも 1 つフラグメントを含むベクター分子。

【請求項 9】

転写制御配列を含む、請求項 8 に記載のベクター分子。

【請求項 10】

請求項 9 に記載のベクター分子を含む宿主細胞。

【請求項 11】

インビトロで増殖可能であり、そして請求項 1 もしくは 2 に記載のポリペプチドを培養増殖で生成する能力のある宿主細胞であって、そこでは該細胞は請求項 2 に記載のポリペプチドをコードする天然型内在性ヒト遺伝子の転写制御配列でない少なくとも 1 つの転写制御配列を含み、そこでは該 1 もしくはそれ以上の転写制御配列は請求項 1 もしくは 2 に記載のポリペプチドをコードする DNA の転写を制御する、宿主細胞。

【請求項 12】

ヒトにおける異常細胞増殖、癌、アテローム性動脈硬化症、炎症性腸疾患、宿主炎症もしくは免疫応答、または乾癬を含む遺伝子発現制御の異常に関連する病状を診断する方法であって；

ヒトからの適切な組織もしくは細胞において請求項 2 で定義されたポリペプチドをコードする天然型内在性ヒト遺伝子から転写されるメッセンジャー RNA の転写の異常を検出すること含み、そこでは転写の異常が該病状の診断となる、方法。

【請求項 13】

天然型内在性ヒト遺伝子が配列番号 4、7 または 8 において示されるヌクレオチド配列を含む、請求項 12 に記載の方法。

【請求項 14】

請求項 12 に記載の方法であって、適切な組織もしくは細胞の試料を、または該組織もし

10

20

30

40

50

くは細胞に由来する単離RNAもしくはDNA分子を、請求項3において定義された単離ヌクレオチド配列とハイストリンジェント条件下でハイブリダイズするところの少なくとも約15～20ヌクレオチド長の単離ヌクレオチド配列と接触させることを含む、方法。

【請求項15】

ヒトにおけるHDAC9発現または活性の異常に関連した病状の診断方法であって：
該病状を患うヒトからの適当な組織または細胞中の、配列番号1、5または6で示されるアミノ酸配列またはそのフラグメントを含むポリペプチドの量を測定することを含み、ここではHDAC9発現または活性の異常に関連した該病状を患わないヒトからの対応する組織中の該ポリペプチドまたはそのフラグメントの量と比較した該ポリペプチドまたはそのフラグメントの異常の存在が病状を患うヒトの診断となる、方法。

10

【請求項16】

請求項15に記載の方法であって、検出工程が、適当な組織または細胞を、配列番号1、5または6で示されるアミノ酸配列またはそのフラグメントを含むポリペプチドに特異的に結合する抗体と接触させること、および該適当な組織または細胞においてポリペプチドとの該抗体の特異的結合を検出することを含み、ここで、ポリペプチドへの特異的結合の検出が配列番号1、5または6で示されるアミノ酸配列またはそのフラグメントを含むポリペプチドの存在を示す、方法。

【請求項17】

配列番号1、5または6で示されるアミノ酸配列を含むポリペプチドまたは該ポリペプチドのフラグメントに特異的に結合する、抗体またはそのフラグメント。

20

【請求項18】

FabまたはF(ab')₂フラグメントである、請求項17に記載の抗体フラグメント。

【請求項19】

ポリクローナル抗体である、請求項17に記載の抗体。

【請求項20】

モノクローナル抗体である、請求項17に記載の抗体。

【請求項21】

請求項1または2において定義されたポリペプチドの製造方法であって：
宿主細胞におけるポリペプチドの発現に十分な条件下で配列番号1、5もしくは6で示されるアミノ酸配列を含むポリペプチドをコードする外来性ポリヌクレオチドを含む発現ベクターが組み込まれた宿主細胞を培養し、それにより発現されるポリペプチドの産生を引き起こす、方法。

30

【請求項22】

請求項21に記載の方法であって、さらに、細胞が生成したポリペプチドを回収することを含む方法。

【請求項23】

外来性ポリヌクレオチドが配列番号1、5もしくは6で示されるアミノ酸配列からなるポリヌクレオチドをコードする、請求項21に記載の方法。

【請求項24】

外来性のポリヌクレオチドが配列番号2、7または8で示されるヌクレオチド配列を含む、請求項21に記載の方法。

40

【請求項25】

外来性ポリヌクレオチドが配列番号3で示されるヌクレオチド配列を含む、請求項21に記載の方法。

【請求項26】

外来性ポリヌクレオチドが配列番号3で示されるヌクレオチド配列からなる、請求項21に記載の方法。

【請求項27】

外来性ポリヌクレオチドが配列番号4で示されるヌクレオチド配列を含む、請求項24に記載の方法。

50

【発明の詳細な説明】

【技術分野】

【0001】

本発明はヒストンデアセチラーゼ遺伝子および遺伝子産物に関する。特に、本発明は、既知の酵母ヒストンデアセチラーゼ1 (hda1) クラス I I ヒストンデアセチラーゼ (H D A C) に高い相同性を有するタンパク質、このようなタンパク質をコードする核酸分子、該タンパク質を認識する抗体、および、例えば、細胞増殖の異常、癌、アテローム性動脈硬化症、炎症性腸疾患、宿主炎症性もしくは免疫性応答、または乾癬を含む H D A C の異常に関連する病状の診断方法に関する。

【背景技術】

10

【0002】

ヒストンのアセチル化は、D N A への転写因子の接近可能性を改変することにより遺伝子発現を調節する主要な制御メカニズムである。ヒストンのアセチル化は、ヌクレオソームの集合の間および D N A 合成の間に起こるリジンの遊離 Σ -アミノ基の可逆的修飾である。ヒストンのアセチル化レベルの変化は、また、転写活性化および不活性化 (silencing) の間にも起こる。ヒストンのアセチル化は、一般に、転写の活性化に関連する。一方、脱アセチル化は転写抑制に関連する。ヒストンのアセチル化のレベルは、競合するヒストンデアセチラーゼおよびデアセチラーゼの間の平衡に起因する (Emiliani, S., Fischle, W., Van Lindt, C., Al-Abed, Y., and Verdin, E., Proc Nat. Acad. Sci., U. S. A., 95, 2795-2800 (1998))。

20

【0003】

H D A C は、転写の制御において重要な役割を果たすことが示されている。H D A C は、転写抑制に関与する複合体の構成要素として機能する。これは H D A C と複合タンパク質複合体との相互作用を仲介するものであり、そしてデアセチラーゼの活性化を要する。H D A C 複合体は、コリプレッサー m S i n 3 A (Kasten, M.M., Dorland, S., Stillman, D.J. Mol. Cell. Biol. 17, 4852-4858 (1997)) および m S i n 3 A -関連タンパク質 (Zhang, Y., Iratni, R., Erdjument-Bromage, H., Tempst, P., Reinberg, D. Cell 89, 357-364 (1997); Zhang, Y., Sun, Z.W., Iratni, R., Erdjument-Bromage, H., Tempst, P., Hampsey, M., Reinberg, D. Mol. Cell. 1, 1021-1031 (1998))、不活性化メディエーター N c o R (Nagy, L., H.- Y. Kao, D. Chakravarti, R. J. Lin, C. A. Hassig, D. E. Ayer, S. L. Schreiber, and R. M. Evans (1997) Cell 89, 373-380 および S M R T (Alland, L. et al., Nature 387:49-55 (1997); Heinzl, T. et al., Nature 387:43-8 (1997)), 転写リプレッサー R b (Brownell, J. E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D. G., Roth, S.Y., and Allis, C. D. (1996) Cell 84, 843-851)、R b 様タンパク質 p 1 0 7 (Ferreira, R., Magnaghi-Jaulin, L., Robin, P., Harel-Bellan, A., Trouche, D. (1998) Proc. Natl. Acad. Sci. USA 95, 10493-10498) および p 1 3 0 (Stiegler, P., De Luca, A. Bagella, L., Giordano, A. (1998) Cancer Res. 38 9, 187-190)、R b 結合タンパク質 (Nicolas, E., Morales, V., Magnaghi-Jaulin, L., Harel-Bellan, A., Richard-Foy, H., Trouche, D. (2000) J. Biol. Chem. 275, 9797-9804, Lai, A., Lee, J.M., Yang, W.M., DeCaprio, J.A., Kaelin, W.G. Jr., Seto, E., Branton, P.E. (1999) Mol. Cell. Biol. 19, 6632-6641)、M a d / M a x (Laherty, C., W.- M. Yang, J.- M. Sun, J. R. Davie, E. Seto, and R. N. Eisenman. (1997) Cell 89, 349-456)、核ホルモン受容体 (Nagy, L., H.- Y. Kao, D. Chakravarti, R. J. Lin, C. A. Hassig, D. E. Ayer, S. L. Schreiber, and R. M. Evans. (1997) Cell 89, 373-380)、ヌクレオソームリモデリング因子 (Xue, Y., Wong, J., Moreno, G.T., Young, M. K., Cote, J., Wang, W. (1998) Mol. Cell. 2, 851-861)、メチルー結合タンパク質 (Fuk s, F., Burgers, W.A., Brehm, A., Hughes-Davies, L., Kouzarides, T. (2000) Nat. Genet. 24, 88-91, Nan, X., Ng, H.H., Johnson, C.A., Laherty C.D., Turner, B.M., Eisenman, R.N., Bird, A. (1998) Nature 393, 386-389, Ghosh, A.K., Steele, R., Ray, R.B. (1999) Biochem. Biophys. Res. Commun. 260, 405-409, Ng, H. H., Zhang, Y.,

30

40

50

Hendrich, B., Johnson, C.A., Turner, B.M., Erdjument-Bromage, H., Tempst, P., Reinberg, D., Bird, A. (1999) *Nat. Genet.* 23, 58-61)、およびDNA修復機構タンパク質(Yarden, R.I., Brody, L.C. (1999) *Proc. Natl. Acad. Sci. U. S. A.* 96, 4983-4988, Cai, R.L., Yan-Neale, Y., Cueto, M.A., Xu, H., Cohen, D. (2000) *J. Biol. Chem.* 275, 27909-27916)を含み得る。

【0004】

さらに、HDAC1はYY1 (Yang, W.-M., Inouye, C., Zeng, Y., Bearss, D., and Seto, E. (1996) *Proc. Natl. Acad. Sci.* 93, 122845-12850)およびSp1 (Doetzlhofer, A., Rotheneder, H., Lagger, G., Koranda, M., Kurtev, V., Brosch, G., Wintersberger, E., Seiser, C. (1999) *Mol. Cell. Biol.* 19, 5504-5511)に直接結合することが見出されており、そしてHDAC4および5はMEF2 (Grozinger, C. M., and Schreiber, S. L. (2000) *Proc. Natl. Acad. Sci.* 97, 7835-7840)に結合する。さらに、HDACは複合体中で一体となって見出されている(Eilers, A.L., Billin, A.N., Liu, J., Ayer, D.E. (1999) *J Biol Chem* 274, 32750-32756, Grozinger, C. M., and Schreiber, S. L. (2000) *Proc. Natl. Acad. Sci.* 97, 7835-7840)。

10

【0005】

2つの明確なクラスの酵母ヒストンデアセチラーゼがサイズおよび配列に基づいて同定された。酵母クラスI HDACはRpd3、Hos1p、およびHos2pを含む。クラスIIは酵母HDAPを含む。さらに、これら2つのクラスのメンバーは、異なる複合体を形成することが見出されている。ヒトHDACは、それらの酵母配列への類似性に基づいて分類される。クラスIヒトHDACはHDAC1~3および8を含む。クラスII HDACはHDAC4~7を含む。クラスI HDACのデアセチラーゼコアは、最初から390までのアミノ酸に存在する。N末端に第2の触媒ドメインを含むHDAC4を除いて、クラスII HDAC触媒ドメインはこれらのペプチドのC末端に位置する(Grozinger, C. M., Hassig, C. A., and Schreiber, S. L. (1999) *Proc. Natl. Acad. Sci. U. S. A.* 96, 4868-4873)。

20

【0006】

クロマチンのアセチル化の機能を研究するために使用されてきた重要なアプローチは、ヒストンデアセチラーゼの特異的インヒビターの使用である。いくつかのクラスの化合物は、HDACを阻害することが同定されている。ヒストンデアセチラーゼインヒビターは、G1/SおよびG2/M細胞周期アレスト、分化(Itazaki, H., K. Nagashima, K. Sugita, H. Yoshida, Y. Kawamura, Y. Yasuda, K. Matsumoto, K. Ishii, N. Uotani, H. Nakai, A. Terui, S. Yoshimatsu, Y. Ikenishi and Y. Nakagawa. (1990) *J. Antibiot.* 12, 1524-1532, Hoshikawa, Y., Kijima, M., Yoshida, M., and Beppu, T. (1991) *Agric. Biol. Chem.* 55, 1491-1497, Hoshikawa, Y., Kwon, H.-J., Yoshida, M., Horinouchi, S., and Beppu, T. (1994) *Exp. Cell Res.* 214, 189-197, Sugita, K., Koizumi, K., and Yoshida, H. (1992) *Cancer Res.* 52, 168-172, Yoshida, M., Y. Hoshikawa, K. Koseki, K. Mori and T. Beppu. (1990) *J. of Antibiot.* 43, 1101-106, Yoshida, M., Nomura, S., and Beppu, T. (1987) *Cancer Res.* 47, 3688-3691)、ならびに変形および正常細胞のアポトーシス(Medina, V., Edmonds, B., Young, G. P., James, R., Appleton, S., Zalewski, P. D. (1997) *Cancer Res.* 57, 3697-3707)および変形の反転(Kwon, H. J., Owa, T., Hassig, C. A., Shimada, J., and Schreiber, S. (1998) *Proc. Natl. Acad. Sci. U. S. A.* 95, 3356-3361, Kim, M.-S., Son, M.-W., Park, Y. I., and Moon, A. (2000) *Cancer Lett.* 157, 23-30)の誘導を含む抗増殖効果を有することが見出されている。

30

40

【0007】

これらの効果は、非リガンド化レチノイン酸受容体PML-RAR α およびPLZF-RAR α の融合体との複合体中のHDACの存在と共に、腫瘍発生能におけるHDACに対する役割を示す(Grignani, F., De Matteis, S., Nervi, C., Tomassoni, L., Gelmetti, V., Cioce, M., Fanelli, M., Ruthardt, M., Ferrara, F. F., Zamir, I., Seiser, C.

50

, Grignani, F., Lazar, M. A., Minucci, S., Pelicci, P. G. (1998) *Nature* 391, 815-818, He, L. Z., Guidez, F., Tribioli, C., Peruzzi, D., Ruthardt, M., Zelent, A., Pandolfi, P. P. (1998) *Nat. Genet.*, 18, 126-35, Lin, R.J., Nagy, L., Inoue, S., Shao, W., Miller, W. H. Jr and Evans, R. M. (1998) *Nature* 391, 811-814)。さらに、ヒストンデアセチラーゼインヒビター、フェニルブチレートおよびトリコスタチンAは、前骨髄球性白血病の処置において有望であり、そしていくつかの他のH D A Cインヒビターは研究中であり、臨床での使用に近づいている(Byrd, J.C., Shinn, C., Ravi, R., Willis, C.R., Waselenko, J.K., Flinn, I.W., Dawson, N.A., Grever, M.R. (1999) *Blood* 94, 1401-1408, Kim, Y.B., Lee, K.H., Sugita, K., Yoshida, M., Horinouchi, S. (1999) *Oncogene* 18, 2461-2470, Cohen, L.A., Amin, S., Marks, P.A., Rifkind, R. A., Desai, D., Richon, V.M. (1999) *Anticancer Res.* 19, 4999-5005)。

10

【0008】

さらに、H D A Cインヒビターであるブチレートは、炎症誘発性サイトカインT N F- α 、T N F- β 、I L-6、およびI L 1- β の発現の減少させることが見出された。これらの効果は、N F k Bの活性化の阻害(Segain JP, Raingeard de la Bletiere D, Bourre ille, A., Leray V., Gervois, N., Rosales, C., Ferrier, L., Bonnet, C., Blottiere, H.M., Galmiche, J.P. (2000) Butyrate inhibits inflammatory responses through N FkappaB inhibition: implications for Crohn's disease. *Gut* 47, 397-403)およびヒストンデアセチラーゼを阻害する能力(Inan M.S., Rasoulpour, R.J., Yin, L., Hubbard, A.K., Rosenberg, D.W., Giardina, C. (2000) The luminal short-chain fatty acid butyrate modulates NF-kappaB activity in a human colonic epithelial cell line. *Gastroenterology* 118, 724-34)に起因するものと考えられている。

20

【0009】

H D A Cインヒビターであるトラポキシンの発見により、トラポキシンの様分子が結合したアフィニティマトリックスカラムを用いて、最初のヒトヒストンデアセチラーゼであるH D A C 1を単離することが可能となった(Taunton, J., Collins, J. L., and Schreiber, S. (1996) *J. Am. Chem. Soc.* 118, 10412-10422)。続いて、7つの他のヒトH D A C酵素の異性体が報告された(Taunton, J., Hassig, C. A. and Schreiber, S.L. (1996). *Science* 272, 408-411, Yang, W. m., Inouye, C., Zeng, Y., Bearss, D., and Seto, D. (1996) *Proc. Natl. Acad. Sci. U.S.A.* 93, 12845-12850, Yang, W. M., Yao, y. L., Sun, J. M., Davie, J. R., and Seto, E. (1997). *J. Biol Chem.* 272, 28001-28007, Emiliani, S., Fischle, W., Van Lint, C., Al-Abed, Y., and Verdin, E. (1998). *Proc. Natl. Acad. Sci. U.S.A.* 95, 2795-27800)。これら8つのH D A Cは、配列相同性に基づいてクラスI (酵母遺伝子R p d 3に類似したH D A C 1~3および8)とクラスII H D A C (酵母遺伝子h d a 1に類似した4~7 (Grozing, C. M., Hassig, C.A., and Schrieber, S. L. (1999). *Proc. Natl. Acad. Sci. U.S.A.* 96, 4983-4988.)とに分けられる。今回、我々は可能性のある新規なH D A C (以下、H D A C 9という。)の単離および特徴化を報告する。これはh d a 1 クラスII H D A Cに対して配列類似性を示す。H D A C 9は、H D A C クラスIおよびクラスIIを橋渡しする特徴を有する。

30

【0010】

40

発明の要約

本発明は、ヒストンデアセチラーゼ、特に新規ヒストンデアセチラーゼH D A C 9に関する。

第1の観点において、本発明は配列番号1、配列番号5または配列番号6で示されるアミノ酸配列を含む単離ポリペプチドを提供する。さらに、本発明は配列番号1、配列番号5または配列番号6で示されるアミノ酸配列からなる単離ポリペプチドを提供する。配列番号1、配列番号5または配列番号6で示されるアミノ酸配列はH D A Cのファミリーの既知のメンバーとかなりの程度の相同性を示す。簡便のため、配列番号1、配列番号5または配列番号6で示されるアミノ酸配列からなるポリペプチドを、ヒストンデアセチラーゼ9またはH D A C 9と称する。このようなポリペプチド、またはそのフラグメントは、種

50

々の正常組織中において発現し、例えば、H D A C 9 は約 3 k b の転写物として正常な精巢、胃、脾臓、小腸、胎盤、肝臓、腎臓、結腸、肺、心臓および脳に存在する。H D A C 9 は筋肉においては検出されないが、このレーンは、また、G A P D H にもハイブリダイズしなかった(図 7)。

【0011】

配列番号 1、配列番号 5 または配列番号 6 で示されるアミノ酸配列を有する単離ポリペプチドのフラグメントは、約 5 ～ 1 4 8 アミノ酸、好ましくは約 1 0 ～ 約 1 4 3 アミノ酸、さらに好ましくは約 2 0 ～ 約 1 0 0 アミノ酸、および最も好ましくは約 2 0 ～ 約 5 0 アミノ酸を含むポリペプチドを含む。このようなフラグメントは、また、本発明の一部を形成する。好ましくは、フラグメントは触媒ドメインを包含し、これはアミノ酸番号 1 ～ 3 9 0 の間に存在すると予想されている。本発明のこの観点にしたがって、ヒト由来の新規ポリペプチド、ならびに生物学的、診断的または治療的に有用なそのフラグメント、変異体および誘導体、フラグメントの変異体および誘導体、および前記のアナログを提供する。 10

【0012】

第 2 の観点において、本発明は上記のポリペプチドをコードするヌクレオチド配列を含む単離 D N A を提供する。特に、本発明は、

(1) 配列番号 2、配列番号 7 または配列番号 8 で示されるヌクレオチド配列を含む単離 D N A ；

(2) 配列番号 3 で示されるヌクレオチド配列を含む単離 D N A ；

(3) 配列番号 3 で示されるヌクレオチド配列にハイストリンジェント条件下でハイブリダイズすることができる単離 D N A ；および 20

(4) 配列番号 4 で示されるヌクレオチド配列を含む単離 D N A ；

を提供する。また、少なくとも約 1 5 塩基、好ましくは少なくとも約 2 0 塩基を含む核酸配列、さらに好ましくは配列番号 2、配列番号 7 または配列番号 8 あるいは配列番号 3 の約 3 0 連続塩基を含む核酸配列も提供される。

【0013】

また、配列番号 2、配列番号 7 または配列番号 8 あるいは配列番号 3 で示される核酸配列の核酸に実質的に類似した核酸も、本発明の範囲内である。好ましい実施態様において、単離 D N A は、本発明の D N A の少なくとも 1 つのフラグメントを含む、特に配列番号 2、配列番号 7 または配列番号 8 あるいは配列番号 3 で示されるヌクレオチド配列からなる 30 D N A を含む、ベクター分子の形態をとる。

【0014】

本発明の第 3 の観点は、ヒトにおける細胞増殖の異常、癌、アテローム性動脈硬化症、炎症性腸疾患、または乾癬に関連した病状を含む（ただし、これらに限定されるわけではない。）遺伝子発現制御の異常に関連した病状の診断方法を含み、該方法はヒトからの適当な組織または細胞中の配列番号 1、配列番号 5 または配列番号 6 で示されるアミノ酸配列からなる新規ポリペプチドをコードする天然型内在性ヒト遺伝子から転写されたメッセンジャー R N A の転写の異常を検出することを含み、ここで、該転写の異常はこのような病状のヒトの苦痛の診断となる。特に、配列番号 1、配列番号 5 または配列番号 6 で示されるアミノ酸配列からなる新規ポリペプチドをコードする該天然型内在性ヒト遺伝子は、配 40 列番号 4 で示されるゲノム性ヌクレオチド配列を含む。本発明の 1 つの実施態様において、該診断方法は、適当な組織または細胞の試料を、あるいは組織または細胞由来の単離 R N A または D N A 分子を、配列番号 1、5 または 6 で示されるアミノ酸配列を有する新規ポリペプチドをコードする単離ヌクレオチド配列とハイストリンジェント条件下でハイブリダイズする少なくとも約 1 5 ～ 2 0 ヌクレオチド長の単離ヌクレオチド配列と、接触させることを含む。

【0015】

本発明のアッセイの観点の別の実施態様は、ヒトにおける H D A C 9 活性の異常に関連した病状の診断方法であって、そのような病状を患うヒトからの一定の組織または細胞におけるデアセチラーゼ活性のレベルを測定することを含み、ここで、H D A C 活性の異常に 50

関連した病状に罹患していないヒトのそれぞれの組織または細胞におけるそのレベルと比較したデアセチラーゼ活性のレベルの異常が該病状を患うヒトの診断となる、方法を提供する。

【0016】

本発明のこの観点の1つの実施態様にしたがって、新規H D A C 9をコードする遺伝子の転写を制御できるアンチセンスポリヌクレオチドが提供され；別の実施態様において、新規H D A C 9をコードする遺伝子の転写を制御できる二本鎖R N Aが提供される。

【0017】

本発明の別の観点により、前記のポリペプチド、ポリペプチドフラグメント、変異体および誘導体、変異体および誘導体のフラグメント、ならびにこれらのアナログの製造方法が提供される。本発明のこの観点の好ましい実施態様において、宿主細胞においてポリペプチドの発現に十分な条件下で、該ポリペプチドをコードする外来性ヌクレオチド配列を含む発現ベクターが組み込まれた宿主細胞を培養すること、これによってポリペプチドの発現を引き起こすこと、および所望により、発現したポリペプチドを回収することを含む、前記H D A C 9の製造方法が提供される。本発明のこの観点の好ましい実施態様において、配列番号1、5または6で示されるアミノ酸配列を含むまたはからなるポリペプチドの製造方法であって、宿主細胞においてポリペプチドの発現に十分な条件下で、配列番号1、5または6で示されるアミノ酸配列を含むまたはからなるポリペプチドをコードする外来性ポリヌクレオチドを含む発現ベクターが組み込まれた宿主細胞を培養すること、これにより発現ポリペプチドの生成を引き起こすこと、そして所望により発現したポリペプチドを回収することを含む方法が提供される。

10

20

【0018】

好ましくは、このような方法のいずれかにおいて、外来性ポリヌクレオチドは、配列番号2、7または8で示されるヌクレオチド配列、配列番号3で示されるヌクレオチド配列、あるいは配列番号4で示されるヌクレオチド配列を含むあるいはからなる。本発明の別の観点にしたがって、とりわけ、研究、生物学的、臨床的および治療目的のための、生成物、組成物、製造方法、ならびに前記のポリペプチドおよびポリヌクレオチドの使用が提供される。

【0019】

本発明のこの観点の一定のさらなる好適な実施態様において、配列番号1、5または6で示されるアミノ酸配列を含むポリペプチド、すなわちすべてのH D A C 9変異体に特異的に結合する抗体またはそのフラグメントが提供される。この点に関して一定の特に好ましい実施態様において、該抗体は、ヒトH D A C 9 ポリペプチドまたはヒトH D A C 9 ポリペプチドの部分に極めて選択的である。

30

【0020】

さらなる観点において、配列番号1、5または6で示されるアミノ酸配列のフラグメントまたは部分に結合する抗体またはそのフラグメントが提供される。

【0021】

別の観点において、病状が対象中のH D A C 9 ポリペプチドの存在下で、H D A C 9 の遺伝子発現の異常、増加または減少、あるいはH D A C 9 ポリペプチドの活性の増加または減少に関連した、配列番号1、5または6で示されるアミノ酸配列、あるいはそのフラグメントまたは部分を有するポリペプチドに結合した抗体の有効量を投与することによる、該対象における該病状の処置方法が提供される。また、対象中のH D A C 9 の存在下で、H D A C 9 遺伝子の発現の異常または増加または減少、あるいはH D A C 9 ポリペプチドの活性の増加または減少に関連した疾患または病状の診断方法であって、例えば、前に記載したH 4 ヒストンアッセイ(Inokoshi, J., Katagiri, M., Arima, S., Tanaka, H., Hayashi, M., Kim, Y.-B., Furumai, R., Yoshida, M., Horinouchi, S., Omura, S. (1999) Biochem. Biophys. Res. Com. 256, 372-376.)を含む、通常の方法の使用を含む方法も提供される。

40

【0022】

50

さらに別の観点において、本発明は、インビトロで増殖可能な宿主細胞、好ましくは、配列番号1、5もしくは6で示されるアミノ酸配列もしくはそのフラグメントを含むポリペプチドを培養増殖で生成することができる脊椎動物細胞、特に哺乳動物細胞もしくは細菌細胞を提供し、そこでは細胞は転写制御性DNA配列を含有し、そこでは転写制御配列は、配列番号1、5もしくは6にしたがうアミノ酸配列もしくはそのフラグメントを有するポリペプチドをコードするRNAの転写を制御する。これは、プラスミドにおけるHDAC9の増殖、および内在性HDAC9プロモーターまたは他の任意の転写制御配列を用いる、ヒトまたは昆虫細胞または細菌におけるDNA、RNAまたはタンパク質の生成を含むが、これらに限定されない。

【0023】

10

本発明のさらに別の観点において、配列番号1、5もしくは6で示されるアミノ酸配列を含むポリペプチドまたは配列番号1、5もしくは6で示されるアミノ酸配列もしくはそのフラグメントを含むポリペプチドをコードするポリヌクレオチドの正常を超えた発現を、患者から得られる体組織試料中に検出するために必要な成分を含むアッセイ方法およびキットが提供され、そのようなキットは例えば、配列番号1、5もしくは6で示されるアミノ酸配列を含むポリペプチドに、もしくはそのフラグメントに、結合する抗体を、または本発明のポリヌクレオチドとハイブリダイズするオリゴヌクレオチドプローブを、含む。好ましい実施態様では、そのようなキットは、また、キットの構成成分が使用されるべき手順の詳細を記した説明書を含む。

【0024】

20

別の観点において、本発明は、HDAC活性または遺伝子発現の異常に関連した疾患を処置するための医薬の製造における、配列番号1、5もしくは6で示されるアミノ酸配列を含むポリペプチドもしくはそのフラグメント、そのようなポリペプチドをコードするポリヌクレオチドもしくはそのフラグメント、または配列番号1、5もしくは6で示されるアミノ酸配列を含む該ポリペプチドもしくはそのフラグメントに結合する抗体の使用を目的とする。

【0025】

別の観点は、HDAC活性または遺伝子発現の異常に関連した疾患の処置のために、適当な医薬担体、賦形剤もしくは希釈剤と一緒に、配列番号1、5もしくは6で示されるアミノ酸配列を含むかまたはからなるポリペプチドもしくはそのフラグメント、そのようなポリペプチドをコードするポリヌクレオチドもしくはそのフラグメント、またはそのようなポリペプチドもしくはそのフラグメントに結合する抗体を含む医薬組成物を目的とする。

30

【0026】

別の観点において、本発明は、配列番号1、5もしくは6で示されるアミノ酸配列を含むポリペプチドに結合できる、および／または配列番号1、5もしくは6で示されるアミノ酸配列を含むポリペプチドの活性を調節できる分子、あるいは配列番号1、5もしくは6で示される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の転写または翻訳を調節する核酸配列に結合できる分子の同定方法を目的とする。このような方法は、例えば、米国特許5,541,070；5,567,317；5,593,853；5,670,326；5,679,582；5,856,083；5,858,657；5,866,341；5,876,946；5,989,814；6,010,861；6,020,141；6,030,779；および6,043,024に記載され、これらすべての文献のすべての内容を引用により本明細書に包含させる。このような方法で同定された分子は、また、本発明の範囲内である。

40

【0027】

関連する観点において、HDACに生物学的に関連する複合体中の、結合タンパク質を同定するための新規HDAC9の使用を目的とする。これまで、HDAC9に結合するタンパク質は知られていない。しかしながら、これらは、HDAC9が他のHDACと相互作用することが前から知られているタンパク質と結合するかどうかを測定することにより特徴付けられ得る(序論を参照。)。例えば、HDAC9複合体の構成成分は、共免疫沈降を

50

含む、通常の方法を用いて測定され得る（実施例 9 を参照。）。

【0028】

さらに別の観点において、本発明は、処置を必要とする対象の 1 もしくはそれ以上の組織中に本発明の核酸を導入する方法であって、その方法の結果、該核酸によりコードされた 1 もしくはそれ以上のタンパク質が発現し、そして／または組織内で細胞により分泌される、方法を目的とする。

【0029】

本発明の他の目的、特徴、利点、および観点は、以下の記載から当業者に明らかであろう。しかしながら、以下の記載および具体的実施例は、本発明の好ましい実施態様を示したものであるが、例示のためだけに与えられたことを理解すべきである。開示される発明の精神および範囲内での多様な変更および修飾は、下記の説明を読むことからおよび本開示の他の部分を読むことから当業者には容易に明らかになるであろう。

10

【0030】

図面の簡単な説明

図 1 は、GENFAM (プロプライエタリ・ソフトウェア) を用いて同定され、そして完全 HDAC9 cDNA 配列に関してデータベースを検索するために使用される 1156 bp オープンリーディングフレームを示す。それぞれの ORF (配列番号 3) はヌクレオチド位置番号 1 で開始し、そしてヌクレオチド位置番号 1156 で終了する。

【0031】

図 2 A および 2 B は、それぞれ、HDAC9 の完全長 cDNA 配列 (配列番号 2) およびそのアミノ酸配列 (配列番号 1) を示す。完全長 cDNA 配列は、ヌクレオチド位置番号 1 で開始し、そしてヌクレオチド位置番号 2022 で終了する。

20

【0032】

図 3 は、クローン 198929 / HDAC9 の配列とアラインした、インシリコ (in silico) のゲノム性 DNA 配列 (AL022328) (配列番号 4) を示す。該アラインメントは、プロプライエタリ・ソフトウェア (Novartis Pharmaceuticals, Summit, NJ) を用いて製造された。

【0033】

図 4 は、HDAC9 予想ペプチドおよび *S. pombe* Hda1 ペプチドのアラインメントを示す。クエリー (query) は HDAC9 ペプチドであり、そして対象 (subject) は *S. pombe* Hda1 ペプチドである。アラインメントは、Clustalw アルゴリズム (Higgins, D.G., Thompson, J.D., Gibson, T.J. (1996) 複数の配列アラインメントのための CLUSTAL の使用。Methods Enzymol 266, 383-402) を用いて製造された。

30

【0034】

図 5 は、HDAC1 および HDAC9 v1 のアラインメントならびに推定される触媒ドメインのアミノ酸および Rb 結合ドメインの位置を示す。触媒ドメインのアミノ酸は実線で囲まれ、そして推定される Rb ドメインのアミノ酸は点線の囲み中に含まれる。アラインメントは、Clustalw アルゴリズム (Higgins, D.G., Thompson, J.D., Gibson, T.J. (1996) 複数の配列アラインメントのための CLUSTAL の使用。Methods Enzymol 266, 383-402) を用いて製造された。

40

【0035】

図 6 は、HDAC 1 ~ 9 v1 のアラインメントを示す。該アラインメントは、Clustalw アルゴリズム (Higgins, D.G., Thompson, J.D., Gibson, T.J. (1996) 複数の配列アラインメントのための CLUSTAL の使用。Methods Enzymol 266, 383-402) を用いて製造された。

【0036】

図 7 は、HDAC9 のノーザン分析を示す。(A) 正常ヒト組織における HDAC9 の分布のノーザンプロット分析。GAPDH を、RNA ローディングのための対照と同一のプロットとハイブリダイズした。(B) 腫瘍および正常組織の比較における HDAC9 のノーザンプロット分析。GAPDH を、RNA ローディングのための対照と同一のプロットとハイブリダイズさせた。

50

【0037】

図8は、18SリボソームRNAと比較した正常ヒト組織および細胞株中のHDAC9の分布のリアルタイムPCR分析を示す。ヒト肺カルシノーマ細胞株であるA549からのRNAを、内部標準として使用した。

【0038】

図9は、クラスII HDAC(HDAC4、5、6、7)とのHDAC9 v1のアラインメントを示す。該アラインメントは、Clustalwアルゴリズム(Higgins, D.G., Thompson, J.D., Gibson, T.J. (1996) 複数の配列アラインメントのためのCLUSTALの使用。Methods Enzymol 266, 383-402)を用いて製造された。触媒ドメインのアミノ酸が実線で囲まれている。

10

【0039】

図10は、クラスI HDAC(HDAC1、2、3、8)とのHDAC9 v1のアラインメントを示す。該アラインメントは、Clustalwアルゴリズム(Higgins, D.G., Thompson, J.D., Gibson, T.J. (1996) 複数の配列アラインメントのためのCLUSTALの使用。Methods Enzymol 266, 383-402)を用いて製造された。触媒ドメインのアミノ酸が実線で囲まれている。

【0040】

図11は、3つのHDAC9配列変異体(HDAC9 v1、HDAC9 v2、およびHDAC9 v3)を示す。HDAC9 v1およびHDAC9 v2は、ヒトESTデータベースを検索することにより見出され、そしてHDAC9 v3はCelera配列データベース中の予想される転写物として見出された。(A)は、該3つのHDAC9変異体ペプチド配列のアラインメントを示す。(B)は、クラスIおよびクラスII HDACペプチド配列の略図を示す。触媒ドメインは黒塗りのボックスであり、そして推定されるLXCXEモチーフは白抜きのボックスである。(C)は、HDAC9 v1およびHDAC9 v2のゲノム構造の略図である。エキソンを黒塗りのボックスで示し、そしてイントロンを黒塗りのボックスの間の線で示す。ボックスおよび線の長さは、エキソンおよびイントロンの長さを示す。

20

【0041】

図12は、HDAC9が酵素的に活性なヒストンデアセチラーゼであることを示す。(A) HDAC9の触媒活性は、HDAC3およびHDAC4の活性に匹敵する。(B)は、HDAC1が本アッセイにおけるヒストン基質の脱アセチル化に関して、HDAC3、HDAC4、およびHDAC9よりも効果的であったことを示す。

30

【0042】

図13は、HDAC9が核タンパク質であることおよびHDAC9-フラッグがインビトロで翻訳されることを示す。

図14は、HDAC9 v3およびHDAC9 v2のDNAおよびペプチド配列を示す。

【0043】

発明の詳細な説明

本明細書で引用した全ての特許出願、特許および参考文献は、その全体を本明細書に包含させる。

40

【0044】

本発明の実施において、分子生物学、微生物学および組換えDNAにおける多くの通常の技術が使用される。これらの技術は周知であり、そして、例えば、Current Protocols in Molecular Biology, Volumes I, II, and III, 1997 (F. M. Ausubel ed.); Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.; DNA Cloning: A Practical Approach, Volumes I and II, 1985 (D. N. Glover ed.); Oligonucleotide Synthesis, 1984 (M. L. Gait ed.); Nucleic Acid Hybridization, 1985, (Hames and Higgins); Transcription and Translation, 1984 (Hames and Higgins eds.); Animal Cell Culture, 1986 (R. I. Freshney ed.); Immobilized Cells and Enzymes, 1986 (IRL Press); Perbal

50

, 1984, A Practical Guide to Molecular Cloning; the series, Methods in Enzymology (Academic Press, Inc.); Gene Transfer Vectors for Mammalian Cells, 1987 (J. H. Miller and M. P. Calos eds., Cold Spring Harbor Laboratory); および Methods in Enzymology Vol. 154 and Vol. 155 (Wu and Grossman, and Wu, eds., respectively)において説明されている。

【0045】

本開示を通して使用される下記の略語を以下に列挙する：

ヒストンデアセチラーゼ(HDAC)、ヒストンデアセチラーゼ様タンパク質(HDLP)。

【0046】

最も広い意味において、ヌクレオチド配列に関して本明細書において使用される場合の“実質的に類似”なる語は、参照ヌクレオチド配列に対応するヌクレオチド配列を意味し、ここで、対応する配列は、参照ヌクレオチド配列によりコードされるポリペプチドと実質的に同一の構造および機能を有するポリペプチド（例えば、該ポリペプチドの機能に影響しない程度のアミノ酸における変化のみが起こる）をコードする。望ましくは、実質的に類似ヌクレオチド配列は、参照ヌクレオチド配列によりコードされるポリペプチドをコードする。実質的に類似ヌクレオチド配列と参照ヌクレオチド配列との間の同一性の百分率は、望ましくは、少なくとも80%、さらに望ましくは少なくとも85%、好ましくは少なくとも90%、さらに好ましくは少なくとも95%、いっそうさらに好ましくは少なくとも99%である。配列の比較を、Clustalw(例えば、Higgins, D.G. et al. Methods Enzymol. 266:383-402 (1996)参照。)を用いて実施する。Clustalwアラインメントは、デフォルトのパラメーターを用いて実施された。

【0047】

参照ヌクレオチド配列に“実質的に類似する”ヌクレオチド配列は、50℃にて2×SSC、0.1%SDS中で洗浄しながら50℃にて7%硫酸ドデシルナトリウム(SDS)、0.5M NaPO₄、1mM EDTAにおいて、さらに望ましくは、50℃にて1×SSC、0.1%SDS中で洗浄しながら50℃にて7%硫酸ドデシルナトリウム(SDS)、0.5M NaPO₄、1mM EDTAにおいて、さらにいっそう望ましくは、50℃にて0.5×SSC、0.1%SDS中で洗浄しながら50℃にて7%硫酸ドデシルナトリウム(SDS)、0.5M NaPO₄、1mM EDTAにおいて、好ましくは、50℃にて0.1×SSC、0.1%SDS中で洗浄しながら50℃にて7%硫酸ドデシルナトリウム(SDS)、0.5M NaPO₄、1mM EDTAにおいて、さらに好ましくは、65℃にて0.1×SSC、0.1%SDS中で洗浄しながら50℃にて7%硫酸ドデシルナトリウム(SDS)、0.5M NaPO₄、1mM EDTAにおいて、参照ヌクレオチドとハイブリダイズし、さらに、機能的に等価な遺伝子産物をコードしている。

【0048】

“mRNAの上昇した転写”は、HDAC9活性の異常に罹患していない対象よりも、このような疾患または病状を患う個体の適当な組織または細胞中に存在する本発明の新規ポリペプチドをコードする天然型内在性ヒト遺伝子から転写されたメッセンジャーRNAの量が多いことを意味する；特に、このような病状に罹患していないヒトにおける対応する組織において見出される量の、少なくとも約2倍、好ましくは少なくとも約5倍、さらに好ましくは少なくとも約10倍、最も好ましくは少なくとも約100倍の量のmRNA。このような上昇したレベルのmRNAは、結果的に、健康な個体と比較した細胞増殖の異常に関連した病状を患う個体におけるこのようなmRNAから翻訳されたタンパク質のレベルの上昇をもたらす得る。また、“mRNAの上昇した転写”は、発現が単独または他の分子と組み合わせたHDAC9により調節される遺伝子から転写されたメッセンジャーRNAの量がより多いことを意味し得ることも理解される。

【0049】

本明細書において使用される“宿主細胞”とは、例えば、エレクトロポレーション、燐酸カルシウム沈降、マイクロインジェクション、形質転換、ウイルス感染などの任意の方法により細胞中に導入された異種のDNAを含む原核もしくは真核細胞を意味する。

10

20

30

40

50

【0050】

本明細書において使用される“異種の”とは、“異なる天然由来の”を意味するか、もしくは非天然の状態を表す。例えば、もし宿主細胞が他の生物から、特に他の生物種から、得られるDNAもしくは遺伝子で形質転換されるならば、その遺伝子はその宿主細胞に関して、そしてまたその遺伝子を保有する宿主細胞の子孫に関して、異種である。同様に、“異種の”とは、同じ天然の元の細胞タイプから得られかつその中に挿入されるが、非天然の状態、例えば異なるコピー数、もしくは異なる調節エレメントによる制御下で存在する、ヌクレオチド配列を意味する。

【0051】

“ベクター”分子は、その中に異種核酸を挿入することができて、次いで適切な宿主細胞中に導入され得る、核酸分子である。ベクターは好ましくは、1つもしくはそれ以上の複製起点、および組換えDNAが挿入され得る1つもしくはそれ以上の部位を有する。ベクターは、しばしば、ベクターを有する細胞をそうでないものから選別できる簡便な手段を有しており、例えばそれらは薬剤耐性遺伝子をコードする。一般的なベクターとしては、プラスミド、ウイルスゲノム、および(主として酵母および細菌中の)“人工染色体”が挙げられる。 10

【0052】

“プラスミド”は、先行する小文字のpおよび／もしくはそれに続く大文字ならびに／または数字により、当業者によく知られた標準的命名規約にしたがって本明細書において示される。本明細書において開示される出発材料のプラスミドは市販で入手可能であるか、無制限に公開で入手可能であるか、または入手可能なプラスミドから周知の公表された手順の日常的な適用により構築され得るかのいずれかである。本発明にしたがって使用され得る多くのプラスミドならびに他のクローニングベクターおよび発現ベクターは周知であり、かつ当業者には容易に入手可能である。さらに、当業者は本発明における使用に適した任意の数の他のプラスミドを容易に構築し得る。本発明におけるこのようなプラスミドならびに他のベクターの性質、構築および使用は、本開示より当業者には容易に明らかであろう。 20

【0053】

“単離(された)”なる用語は、材料がその本来の環境(例えば、それが天然に由来するならば、天然の環境)から除去されることを意味する。例えば、生きている動物中に存在している天然由来のポリヌクレオチドもしくはポリペプチドは単離されていないが、天然の系中に共存する物質の一部もしくは全てから分離された該ポリヌクレオチドもしくはポリペプチドは、たとえその後で天然の系中に再導入されたとしても、単離されている。このようなポリヌクレオチドはベクターの一部であってもよく、そして／またはこのようなポリヌクレオチドもしくはポリペプチドは組成物の一部であってもよく、そしてなおこのようなベクターもしくは組成物がその天然環境の一部でない点において、単離されている。 30

【0054】

本明細書において使用される“転写制御配列”なる用語は、それらが機能し得るように結合された、タンパク質をコードする核酸配列の転写を誘導し、抑制し、もしくはさもなければ制御する、開始配列、エンハンサー配列およびプロモーター配列のような、DNA配列を意味する。 40

【0055】

本明細書において使用される“ヒト転写制御配列”は、本発明の新規HDAC9ポリペプチドをコードするヒト遺伝子と関連して、それがそれぞれのヒト染色体中に見出されるように、通常見出されるそれらの転写制御配列のいずれかである。該用語は、また、単独または他の分子と組み合わされたHDAC9により調節されるヒト遺伝子発現に関連して通常見出される転写制御配列を意味し得ることも理解される。

【0056】

本明細書において使用される“非ヒト転写制御配列”は、ヒトゲノムに見出されない任意 50

の転写制御配列である。

“ポリペプチド”なる用語は、本明細書において、“ポリペプチド類”および“タンパク質(類)”なる用語と互換的に用いられる。

【0057】

本明細書において使用されるように、本発明のポリペプチドの“化学的誘導体”は、通常その分子の一部ではないさらなる化学的部分を含有する本発明のポリペプチドである。そのような部分は分子の溶解性、吸収性、生物学的半減期などを改善し得る。これに代えて、これらの部分は分子の毒性を軽減し、分子の任意の望ましくない副作用などを消失もしくは減弱させ得る。このような効果を仲介する能力のある部分は、例えば、Remington's Pharmaceutical Sciences, 16th ed., Mack Publishing Co., Eatson, Pa. (1980)に開示されている。 10

【0058】

本明細書において使用されるように、“H D A C 9”は、天然、合成、半合成または組換えのどれかの任意の供給源からの任意の種、特にウシ、ヒツジ、ブタ、ネズミ、ウマおよび好ましくはヒトを含む哺乳動物から得られた実質的に精製されたH D A C 9のアミノ酸配列を意味する。

【0059】

本明細書において使用されるように、“H D A C 9 活性”を含む“H D A C 活性”は、³H-標識H 4 ヒストンペプチドを含むヒストンタンパク質を脱アセチル化するH D A C ポリペプチドの能力を意味する。このような活性は、常法にしたがって、例えば Inokoshi, J., Katagiri, M., Arima, S., Tanaka, H., Hayashi, M., Kim, Y.-B., Furumai, R., Yoshida, M., Horinouchi, S., and Omura, S. (1999) Biochem. Biophys. Res. Com. 25 6, 372-376 に記載されたように測定され得る。生物学的に“活性な”タンパク質は、天然の分子の構造的、制御的または生化学的機能を有するタンパク質を意味する。 20

【0060】

本明細書において使用される“アゴニスト”なる語は、H D A C 9 と結合した場合に、H D A C 9 において変化を引き起こしてH D A C 9 の活性を調節する分子を意味する。アゴニストは、H D A C 9 と結合するタンパク質、核酸、炭水化物、または他の任意の分子を含み得る。

【0061】

本明細書において使用される“アンタゴニスト”または“インヒビター”なる語は、H D A C 9 と結合した場合に、H D A C 9 の生物学的活性をブロックまたは調節する分子を意味する。アンタゴニストおよびインヒビターには、H D A C 9 に結合する、天然または合成の、タンパク質、核酸、炭水化物、または他の任意の分子が含まれ得る。 30

【0062】

H D A C 9 は、Celera Human Genome Database、Incyte LIFESEQ(登録商標)データベースおよび公共のHigh Throughput Genomicデータベースにおけるヒストンデアセチラーゼに関連する新規ヒト配列を検索するための、G E N F A Mと呼ばれるプロプライエタリー・コンピューター・ソフトウェアを用いて同定された。1 1 5 6 b p のオープンリーディングフレーム(O R F)を同定および使用して、全組織および後根神経節c D N Aライブラリーからの配列決定されたクロンのデータベースを検索した。各ライブラリーから2つずつ、計4つのクロンが、該O R F(M 6、K 1 0、P 3、F 2 3)を含むことが見出された。これらのクロンのうち、全組織ライブラリーからのM 6 は、配列解析およびインビトロ翻訳の結果として最も完全なc D N Aであることが決定された。 40

【0063】

B L A S T(Altshul S.F. et al Nucleic Acid Res 25:3389-402 (1997))を、c D N A クロンのM 6 を用いてG e n B a n k データベースを検索するために使用した。ゲノム配列A L 0 2 2 3 2 8 は、M 6 c D N A と同一の配列であるエキソンを含むことが見出された。ゲノム配列A L 0 2 2 3 2 8 との、H D A C 9 (2 0 2 2 ~ 8)のアンチセンス配列のC l u s t a l w アラインメントを図3に示す。H D A C 9 と予想されるc D N A の最初 50

の7塩基は、おそらくそれらが次のイントロンへと続きそしてこの配列がおそらくソフトウェアにとって短すぎてアラインメントを測定することができないために、アラインメントしない。cDNAクローンM6の配列は、自動DNA配列決定法(ACGT, Inc., Northbrook, IL)により確認された。ゲノム配列AL022328から予想されたcDNA配列に基づいて、44塩基がM6のN末端から消失していた。続いて、この配列をPCRにより加えた。

【0064】

HDAC9の完全長cDNAから、673アミノ酸のタンパク質が予想される。HDAC9のcDNA配列は2022塩基対の長さである。他の既知のHDACとのHDAC9の類似性の割合を測定するために、Clustalw複数配列アラインメントを、HDAC1~9の完全ペプチド配列を用いて行った。HDAC9は、ヒトHDAC6に対するペプチド配列類似性が37%と最も高い。クラスII HDACとのHDAC9のClustalwアラインメントを、図9に示す。HDAC9は、また、S. pombeからの酵母クラスII配列hdalに対して40%の類似性であった。ヒトHDAC9およびS. pombeのClustalwアラインメントを図4に示す。HDAC9はクラスI HDACに対して比較的類似性が低い($\leq 18\%$)。クラスI HDACに対するHDAC9のClustalwアラインメントを図10に示す。

10

【0065】

HDAC9は、他の既知のすべてのHDACの推定触媒ドメインとのHDAC9のアラインメントに基づいて、約317aa(約6から323)を包含する触媒ドメインを有する。HDAC9の触媒ドメインを同定するために、Clustalwアラインメントを、クラスI HDAC(1~3および8)またはクラスII HDAC(4~7)からのHDAC9完全ペプチドおよび触媒ドメイン配列を用いて個別に行う。従来から、1アミノ酸変異による不活性化ならびにHDAC様タンパク質(HDLP)、Zn²⁺およびHDACインヒビターの複合体により形成される三次元構造に基づいて、13アミノ酸がデアセチラーゼ活性に寄与することが示されていた(Finnin, M. S., Donigian, J. R., Cohen, A., Richon, V. M., Rifkind, R. a., Marks, P. A., Breslow, R., and Pavletich, N. P. (1999) Structures of a histone deacetylase homologue bound to TSA and SAHA inhibitors. Nature 401, 188-193)。

20

【0066】

これらの13アミノ酸は、Pro 22、His 131、His 132、Gly 140、Phe 141、Asp 166、Asp 168、His 170、Asp 173、Phe 198、Asp 258、Leu 265、およびTyr 297を含む。これらの13アミノ酸のうちの12は、HDAC9において保存されている。保存されていないアミノ酸は、Leu 265である。この疎水性残基はTS結合ポケットの一部を形成し、そしてHDAC9において272位のアミノ酸がGluに置換されている。Leu 265は、HDAC8においてはMetに置換され、そしてHDAC6ドメイン1においてはLysに置換されている。これにより、該残基は高度に保存されているわけではなく、他のHDACと同一である必要がないことが示唆される。HDLP、HDAC1およびHDAC2と異なる第2の残基であるAsp 173は、HDAC9の177位においてGlnで置換されており、この相違はHDAC6の触媒ドメイン1においても存在する。さらに、Asp 173は、HDAC4、5、6(ドメイン2)、および7において、Asnで置換されている。この証拠により、これらのAsp 173の置換はHDAC活性に影響しないことが示唆される。

30

40

【0067】

従来から、アミノ酸配列モチーフはHDAC1および2の網膜芽腫タンパク質(Rb)に対する結合に重要であることが見出されている。HDAC1および2とRbの複合体が、E2F応答プロモーターの抑制を誘導する(Brehm, A., Miska, E. A., McCance, D. J., Reid, J. L., Bannister, A. J., and Kouzarides, T. (1998) Nature 391, 597-601)。Rb結合モチーフは、配列モデルLXCXE(式中、“X”は任意のアミノ酸である。)と

50

適合する。該 L X C X E ドメインは、R b の増殖抑制機能に関して必須でないことが見出されているが、H D A C 結合には必要である (Chen, T.-T. and Wang, J. Y. J. (2000) M ol. Cell Biol. 20, 5571-5580)。以前に H D A C 1 において決定された R b 結合ドメインは、アミノ酸 4 1 4 からアミノ酸 4 1 9 に位置し、そして配列 I A C E E である。これまで、他の H D A C が R b に対して結合できるかどうか決定されていなかった。しかしながら、H D A C 9 は推定 R b 結合モチーフ L S C I L を含み、これは H D A C 1 I A C E E とアラインし、そしてアミノ酸 5 6 0 ~ 5 6 4 の間に位置する。R b との H D A C 9 の共免疫沈降は、H D A C 9 におけるこのモチーフの機能を確認するために使用され得る 1 つのストラテジーである。

【0068】

H D A C ファミリーのメンバーのように、H D A C 9 はタンパク質と生物学的に等価な複合体を形成し、そして他の H D A C に関して記載された機能を示し得る。例えば、それは、複合タンパク質複合体との H D A C の相互作用を介する転写抑制に関係する複合体の構成成分として転写制御に関係し得、そしてこれはデアセチラーゼ活性を必要とする。したがって、H D A C 9 の活性または発現の増加は、細胞増殖の異常、癌、アテローム性動脈硬化症、炎症性腸疾患、宿主炎症性または免疫性応答、あるいは乾癬を含む(しかし、これらに限定されるわけではない)数多くの病理学的状態に関連し得る。

【0069】

したがって、D N A / アミノ酸配列および H D A C 9 の予想された構造は、H D A C 活性の異常に関連する病状を改善するのに有用な薬剤(例えば、アンタゴニストまたはインヒビター)を設計するのに有用である。これらには、例えば、それに対して指向された小分子またはタンパク質(例えば、抗体)の使用を介する抗増殖性または抗炎症性薬剤、あるいは H D A C 転写リプレッサー複合体中の関連タンパク質を含まれ得る。さらに、H D A C 9 配列に由来するタンパク質は、また、宿主細胞増殖性または炎症性応答を緩和する治療薬としても使用され得る。

【0070】

新規ポリペプチドの発現パターンを測定するために、種々のヒト組織からの m R N A パネルをノーザン分析にかける。データにより、H D A C 9 がヒト組織中で発現しており、脳、結腸、心臓、腎臓、肝臓、胎盤、小腸、脾臓、胃および精巣において検出可能であることが示唆される。したがって、H D A C 9 は転写された遺伝子を表す。

【0071】

したがって、1 つの観点において、本発明は新規ヒストンデアセチラーゼ (H D A C) に関する。上で概説したように、H D A C 9 は h d a 1 クラス I I H D A C 中の他の H D A C タンパク質に対して極めて類似しているので、H D A C 9 は明らかに H D A C ファミリーのメンバーである。それは、また、H D A C ファミリーと多くの類似性を共有する。

【0072】

本発明は、配列番号 1 で示されるアミノ酸配列を含む単離ポリペプチドに関連する。例えば、このようなポリペプチドは、新規 H D A C 9 のアミノ酸配列を含む融合タンパク質であり得る。別の観点において、本発明は、特に新規 H D A C 9 である、配列番号 1 で示されるアミノ酸配列からなる単離ポリペプチドに関する。

【0073】

本発明は、核酸またはヌクレオチド分子、好ましくは D N A 分子、特に新規 H D A C 9 をコードするものを含む。好ましくは、本発明の単離核酸分子、好ましくは D N A 分子は、配列番号 1、配列番号 5 または配列番号 6 で示されるアミノ酸配列を含むポリペプチドをコードする。同様に、配列番号 1、配列番号 5 または配列番号 6 で示されるアミノ酸配列からなるポリペプチドをコードする単離核酸分子、好ましくは D N A 分子が同様に好ましい。このような核酸またはヌクレオチド、特にこのような D N A 分子は、好ましくは、(1)それぞれ配列番号 1、5 および 6 で示されるアミノ酸配列からなるポリペプチドをコードする完全 c D N A 配列である、配列番号 2、7 または 8 で示されるヌクレオチド配列；

10

20

30

40

50

(2) 配列番号 2 で示される cDNA 配列のオープンリーディングフレームに対応する配列番号 3 で示されるヌクレオチド配列；

(3) 配列番号 3 で示されるヌクレオチド配列とハイストリンジェント条件下でハイブリダイズできるヌクレオチド配列；および

(4) 配列番号 1 で示されるアミノ酸配列からなるポリペプチドをコードする内在性ヒトゲノム DNA に対応する配列番号 4 で示されるヌクレオチド配列；

からなる群から選択されるヌクレオチド配列を含む。このようなハイブリダイズの条件は、上記のように、ハイストリンジェントであるか、または比較的ハイストリンジェントでなくてもよい。核酸がデオキシオリゴヌクレオチド(“オリゴ”)である場合において、ハイストリンジェント条件とは、例えば、37℃(14-塩基オリゴのために)、48℃(17-塩基オリゴのために)、55℃(20-塩基オリゴのために)、および60℃(23-塩基オリゴのために)にて、6×SSC/0.05%ピロリン酸ナトリウム中で洗浄することを意味し得る。さまざまな組成物の核酸のためのこのようなストリンジェント条件の適当な範囲は、Krause and Aaronson (1991), Methods in Enzymology, 200:546-556、ならびに前出の Maniatis et al. において記載されている。

【0074】

これらの核酸分子は、例えば、標的遺伝子の制御において有用な標的遺伝子アンチセンス分子として、および/または標的遺伝子核酸配列の増幅反応におけるアンチセンスプライマーとして、機能し得る。さらに、このような配列は、標的遺伝子の制御に有用な、リボザイムおよび/またはトリプルヘリックス配列の部分として使用され得る。さらになお、このような分子は診断方法の構成要素として使用され得、その結果、HDAC9発現または活性の異常を伴う疾患、例えば、細胞増殖の異常、癌、アテローム性動脈硬化症、炎症性腸疾患、宿主炎症性または免疫性応答、あるいは乾癬を引き起こす対立遺伝子の存在が検出され得る。

【0075】

本発明は、また、

(a) 任意の前記のヌクレオチド配列および/またはそれらに相補的なもの(すなわち、アンチセンス)の少なくとも1つのフラグメントを含むベクター；

(b) コードしている配列の発現を指示する制御エレメントと機能的に結合した任意の前記のコード配列を含む、ベクター分子、好ましくは転写制御配列を含むベクター分子、特に発現ベクター；および

(c) 本明細書において記載したベクター分子または宿主細胞においてコード配列の発現を指示する制御エレメントと機能的に結合した任意の前記のヌクレオチド配列の少なくとも1つのフラグメントを含む遺伝子操作された宿主細胞

を含む。本明細書において使用されるように、制御エレメントには、誘導可能なおよび誘導不可能なプロモーター、エンハンサー、オペレーターおよび発現を駆動および制御する当業者に既知の他のエレメントが含まれるが、これらに限定されない。好ましくは、宿主細胞は脊椎動物の宿主細胞、好ましくは哺乳動物の宿主細胞、例えばヒト細胞またはネズミ細胞、例えばCHOまたはBHK細胞であり得る。同じく好ましくは、宿主細胞は細菌性宿主細胞、特にE.coli細胞であり得る。

【0076】

特に好ましいのは、インビトロで増殖させることができ、そして培養増殖によりHDAC9ポリペプチド、特に配列番号1で示されるアミノ酸配列を含むかからなるポリペプチドを生成する能力のあるところの、とくに上述のタイプの宿主細胞であり、そこでは該細胞は、該ポリペプチドをコードする天然型内在性ヒト遺伝子の転写制御配列ではない1もしくはそれ以上の転写制御配列により制御された構成体中にHDAC9コード配列のあるフラグメントまたは完全配列を含み、そこでは該1もしくはそれ以上の転写制御配列が該ポリペプチドをコードするDNAの転写を制御する。可能性のある転写制御配列には、細菌性またはウイルス性配列が含まれるが、これらに限定されない。

【0077】

10

20

30

40

50

本発明は、該遺伝子の完全配列ならびに本明細書において開示された任意の核酸配列のフラグメントを含む。新規H D A C 9ポリペプチドをコードする核酸配列のフラグメントは、H D A C 9遺伝子に対する高い配列類似性または類似の生物学的活性を有する他の遺伝子を単離するためのc D N Aライブラリーのためのハイブリダイズ用プローブとして使用され得る。このタイプのプローブは、好ましくは少なくとも約30塩基を有し、そして、例えば、約30～約50塩基、約50～約100塩基、約100～約200塩基、または200塩基以上を含み得る。該プローブは、また、完全長転写物ならびにゲノム性クローンまたは制御領域およびプロモーター領域、エキソンおよびイントロンを含む完全H D A C 9遺伝子を含むクローンに対応するc D N Aクローンを同定するために使用され得る。スクリーニングの1つの例は、既知のD N A配列を用いることによりH D A C 9遺伝子のコード領域を単離して、オリゴヌクレオチドプローブを合成することを含む。本発明の遺伝子と相補的な配列を有する標識オリゴヌクレオチドは、ヒトc D N A、ゲノムD N Aまたはm R N Aのライブラリーをスクリーニングして、プローブがハイブリダイズするライブラリーのメンバーを決定するために使用され得る。

10

20

30

40

50

【0078】

上記の遺伝子配列に加えて、例えば他の生物種に存在するであろうような配列のホモログを、当分野において周知の分子生物学的技法により、必要以上の実験(undue experimentation)を行わずに、同定しかつ容易に単離し得る。さらに、そのような遺伝子産物の1またはそれ以上のドメインに対して相同性を有するタンパク質をコードする遺伝子が、ゲノム内の他の遺伝子座に存在してもよい。これらの遺伝子は、また、同様の技法で同定され得る。例えば、新規H D A C 9ポリペプチドをコードする本発明の単離ヌクレオチド配列を標識し、興味のある生物から得られたm R N Aから構築されたc D N Aライブラリーをスクリーニングするのに使用してもよい。c D N Aライブラリーが、標識配列が得られた生物のタイプと異なる生物から得られる場合には、ハイブリダイゼーションはより低いストリンジェント条件下で行われるであろう。これに代えて、標識フラグメントを用いて、興味のある生物から得られるゲノムライブラリーを、ここでも適切にストリンジェント条件を用いてスクリーニングし得る。そのような低いストリンジェント条件は当業者に周知であろうし、そしてライブラリーおよび標識配列が得られる特定の生物に依存して予想の通りに変動するであろう。そのような条件の指針として例えば上で引用されたSambrook et al.を参照されたい。

【0079】

さらに、興味のある遺伝子内のアミノ酸配列に基づいてデザインした二つの縮重したオリゴヌクレオチドプライマープールを用いてP C Rを実施することにより、以前に知られていなかった、差次的発現遺伝子型(differentially expressed gene-type)の配列を単離し得る。反応のテンプレートは、差次的発現対立遺伝子を発現すると知られているか推測されるヒトのもしくは非ヒトの細胞株または組織から調製したm R N Aの逆転写により得られるc D N Aであってもよい。P C R産物をサブクローン化して配列決定し、増幅された配列が差次的発現遺伝子様核酸配列の配列を表わすことを確認し得る。次いで、P C Rフラグメントを用いて、種々の方法により完全c D N Aクローンを単離し得る。例えば、増幅されたフラグメントを標識し、バクテリオファージのc D N Aライブラリーをスクリーニングに用い得る。これに代えて、標識フラグメントを用いてゲノムライブラリーをスクリーニングし得る。

【0080】

P C R技法は、また、完全長c D N A配列を単離するのに利用され得る。例えば、適切な細胞もしくは組織の供給源から、標準的な手順にしたがって、R N Aを単離し得る。第1鎖の合成のプライミングのために増幅フラグメントの5'最末端に特異的なオリゴヌクレオチド・プライマーを用いて、R N Aについて逆転写反応を実行し得る。次いで、得られたR N A/D N Aハイブリッドに、標準的な末端トランスフェラーゼ反応を用いてグアニンで“テール”をつけ、該ハイブリッドをR NアーゼHで消化し、次いで第2鎖の合成をポリCプライマーでプライムし得る。かくして、増幅フラグメント上流のc D N A配列を容

易に単離し得る。使用し得るクローン化ストラテジーの総説については、例えば、上述の Sambrook et al., 1989を参照されたい。

【0081】

同定された遺伝子が正常の、即ち野生型の、遺伝子である場合には、この遺伝子を使用して、遺伝子の変異型対立遺伝子を単離し得る。そのような単離は、遺伝的基盤を有すると知られているか疑われているプロセスまたは障害において好ましい。変異型対立遺伝子は、細胞増殖の異常、癌、アテローム性動脈硬化症、炎症性腸疾患、宿主炎症性もしくは免疫性応答、または乾癬のような病状を含む（しかしこれらに限定されない）H D A C活性の異常に関連する疾患症状に寄与する遺伝子型を有すると知られているか疑われている個人から単離され得る。次いで、変異型対立遺伝子および変異型対立遺伝子産物を下記の診断アッセイシステムにおいて利用し得る。 10

【0082】

変異遺伝子のc D N Aは、例えば、当業者に周知である技術のP C Rを用いて単離され得る。この場合には、第1のc D N A鎖を、変異型対立遺伝子を保有すると考えられる個人において発現していることが知られているか疑われている組織から単離されたm R N Aにオリゴd Tオリゴヌクレオチドをハイブリダイズさせることにより、そして新しい鎖を逆転写酵素で伸張することにより、合成し得る。次いで、c D N Aの第2鎖は、正常遺伝子の5'末端に特異的にハイブリダイズするオリゴヌクレオチドを用いて合成される。これら2つのプライマーを用いて、産物を次いでP C Rで増幅し、適当なベクター中にクローン化し、そして当業者に周知である方法によりD N A配列分析に供する。変異型遺伝子のD N A配列を正常遺伝子のそれと比較することにより、変異型遺伝子産物の機能の消失もしくは変化の原因となる変異を確認することができる。 20

【0083】

これに代えて、ゲノムもしくはc D N Aライブラリーを、変異型対立遺伝子を保有すると知られているか疑われている個人において、目的の遺伝子を発現していることが知られているか疑われている組織からのD N AまたはR N Aをそれぞれ用いて構築し、そしてスクリーニングすることができる。次いで、正常な遺伝子もしくはその任意の適当なフラグメントを標識し、プローブとして用いて、ライブラリー中の対応する変異型対立遺伝子を同定し得る。次いで、この遺伝子を含有するクローンを当分野において日常的に実施される方法により精製し、上述のように配列分析に供し得る。 30

【0084】

これに加えて、変異型対立遺伝子を保有すると疑われているか知られている個人において、目的の遺伝子を発現していることが知られているか疑われている組織から単離されたD N Aまたは同組織から合成されたc D N Aを利用して、発現ライブラリーを構築することができる。この方法で、推定的変異組織から作成される遺伝子産物を発現し、以下に説明するように、正常な遺伝子産物に対する抗体と一緒に標準的な抗体スクリーニング技法を用いてスクリーニングし得る(スクリーニング技法については、例えば、Harlow, E. and Lane, eds., 1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor Press, Cold Spring Harborを参照のこと)。変異により、変化した機能を有する発現遺伝子産物がもたらされる場合には(例えば、ミスセンス変異の結果として)、抗体のポリクローナルセットは変異遺伝子産物と交叉反応すると思われる。そのような標識抗体との反応により検出されるライブラリークローンを精製し、上述のように配列分析に供することができる。 40

【0085】

本発明は、配列番号2、3、4、7または8のいずれかで示されるヌクレオチド配列によりコードされたタンパク質、特に、配列番号1、5もしくは6で示されるアミノ酸配列であるかもしくはこれらを含むポリペプチドまたはそのフラグメントを含む。

【0086】

さらに、本発明は機能的に同等な遺伝子産物を表すタンパク質を含む。そのような同等な差次的発現遺伝子産物は、上述の差次的発現遺伝子によりコードされるアミノ酸配列内に、アミノ酸残基の欠失、付加もしくは置換を含んでもよいが、それはサイレントな変化を 50

もたらし、かくして機能的に同等な差次的発現遺伝子産物を生成する。アミノ酸の置換は、関与する残基の極性、電荷、溶解性、疎水性、親水性および／もしくは両親媒性における類似性に基づいてなされてもよい。

【0087】

例えば、非極性(疎水性)アミノ酸としては、アラニン、ロイシン、イソロイシン、バリン、プロリン、フェニルアラニン、トリプトファンおよびメチオニンが挙げられ；極性中性アミノ酸としては、グリシン、セリン、スレオニン、システイン、チロシン、アスパラギンおよびグルタミンが挙げられ；陽電荷をもつ(塩基性)アミノ酸としては、アルギニン、リジン、およびヒスチジンが挙げられ；負電荷をもつ(酸性)アミノ酸としては、アスパラギン酸およびグルタミン酸が挙げられる。本明細書において使用される“機能的に同等な”とは、上述の差次的発現遺伝子配列によりコードされた内在性の差次的発現遺伝子産物と実質的に類似したインビボおよびインビトロ活性を発揮する能力のあるタンパク質もしくはポリペプチドを意味し得る。

10

【0088】

“機能的に同等な”は、また、内在性の差次的発現遺伝子産物の対応する部分がするであろう方式と実質的に類似した様式で他の細胞内もしくは細胞外分子と相互作用する能力のあるタンパク質もしくはポリペプチドを意味し得る。例えば、“機能的に同等な”ペプチドは、イムノアッセイにおいて、内在性タンパク質の対応するペプチド(すなわち、そのアミノ酸配列を修飾して“機能的に同等な”ペプチドを達成したペプチド)への、もしくは内在性タンパク質自身への、抗体の結合を減少させることができるであろうが、そこでは該抗体は内在性タンパク質の対応するペプチドに対するものである。等モル濃度の機能的に同等なペプチドは、対応するペプチドの前記結合を少なくとも約5%、好ましくは約5%~10%、さらに好ましくは約10%~25%、なおさらに好ましくは約25%~50%、そして最も好ましくは約40%~50%だけ減少させるであろう。

20

【0089】

本発明のポリペプチドは、当分野において周知である技法を用いる組換えDNA技術により製造され得る。したがって、本発明のポリペプチドを製造する方法が提供され、該方法は、配列番号1、5もしくは6で示されるアミノ酸配列を含むポリペプチドをコードする外来性ポリヌクレオチドを含有する発現ベクターがその中に組み込まれた宿主細胞を、その宿主細胞中でポリペプチドを発現させるに十分な条件下で培養することを含み、それによって発現ポリペプチドの生成を引き起こす。所望により、該方法はさらに該細胞により生成されたポリペプチドを回収することを含む。そのような方法の好ましい実施態様において、該外来性ポリヌクレオチドは、配列番号1、5もしくは6で示されるアミノ酸配列からなるポリペプチドをコードする。好ましくは、該外来性ポリヌクレオチドは、配列番号2、配列番号3、配列番号4、配列番号7もしくは配列番号8のいずれかで示されるヌクレオチド配列を含む。配列番号3で示されるヌクレオチド配列、すなわちオープンリーディングフレームを使用する場合には、その配列は、ベクターに挿入されるとき、1つもしくはそれ以上の適切な翻訳停止コドンが、好ましくはcDNA配列中のヌクレオチド2021で始まる天然型内在性停止コドンTGAが、後ろに続いてよい。

30

【0090】

かくして、それぞれのヌクレオチド配列をコードする核酸を発現させることにより、本発明のポリペプチドおよびペプチドを調製する方法が、本明細書において記載される。当業者に周知である方法を用いて、タンパク質コード配列および適当な転写／翻訳制御シグナルを含有する発現ベクターを構築することができる。これらの方法としては、例えば、インビトロ組換えDNA技法、合成技法およびインビボ組換え／遺伝子組換えが挙げられる。例えば、上記Sambrook et al., 1989および上記Ausubel et al., 1989に記載されている技法を参照されたい。これに代えて、タンパク質配列の差次的発現遺伝子をコードする能力のあるRNAを、例えば合成機(synthesizers)を用いて化学的に合成してもよい。例えば、"Oligonucleotide Synthesis", 1984, Gait, M. J. ed., IRL Press, Oxford(出典明示によりその全体を本明細書の一部とする)に記載されている技法を参照されたい。

40

50

【0091】

種々の宿主-発現ベクターシステムを利用して、本発明のH D A C 9 遺伝子コード配列を発現させ得る。そのような宿主-発現システムは、目的のコード配列を生成し次いで精製し得るビヒクルを表すが、また、適切なヌクレオチドコード配列を形質転換もしくは形質移入したときに、インサイツ(in situ)で本発明のH D A C 9 遺伝子タンパク質を発現し得る細胞を表す。これらとしては、差次的発現遺伝子タンパク質コード配列を含有する、組換えバクテリオファージDNA、プラスミドDNAもしくはコスミドDNA発現ベクターで形質転換した細菌(例えば、E.coli、B.subtilis)のような微生物；差次的発現遺伝子タンパク質コード配列を含有する、組換え酵母発現ベクターで形質転換した酵母(例えば、Saccharomyces、Pichia)；差次的発現遺伝子タンパク質コード配列を含有する、組換えウイルス発現ベクター(例えば、バキュロウイルス)で感染もしくは形質移入させた昆虫細胞システム；組換え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、C a M V；タバコモザイクウイルス、T M V)で感染させた、もしくは、タンパク質コード配列を含有する、プラスミド(例えば、T i プラスミド)を含む組換えベクターで形質転換させた、植物細胞システム；または哺乳動物細胞のゲノムから得られるプロモーター(例えばメタロチオネインプロモーター)もしくは哺乳動物ウイルスから得られるプロモーター(例えばアデノウイルス後期プロモーター、ワクシニアウイルス7.5 KプロモーターもしくはC M Vプロモーター)を含有する組換え発現構築体を収容する哺乳動物細胞システム(例えばC O S、C H O、B H K、2 9 3、3 T 3)が挙げられるが、これらに限定されない。

【0092】

また、その細胞に本来存在するH D A C 9 コード遺伝子から細胞により本発明のH D A C 9 の発現を実施することができる。そのような発現の方法は、例えば、米国特許5,641,670；5,733,761；5,968,502；および5,994,127(これらは全て出典明示により、これらの全体を本明細書の一部とする。)に詳述されている。米国特許5,641,670；5,733,761；5,968,502；および5,994,127のいずれかの方法によりH D A C 9 を発現するように誘導されている細胞を、生きている動物中で望ましい組織中に植え込んで、該組織内のH D A C 9 の局所的濃度を増加させることができる。

【0093】

細菌のシステムにおいて、発現されるタンパク質の使用目的に応じて、多数の発現ベクターを有利に選択し得る。例えば、抗体の作成のためにもしくはペプチドライブラリーをスクリーニングするために、大量のそのようなタンパク質を生成したい場合には、例えば、容易に精製できる融合タンパク質産物の高レベルの発現を指令するベクターが望ましいであろう。この点において、ヘキサヒスチジンタグ、例えばp C D N A 3.1 / H i s (Invitrogen, Carlsbad, CA)を含むEpiTag vectosを含む融合タンパク質が使用され得る。他のベクターとして、E.coli発現ベクターp U R 2 7 8 (Ruther et al., 1983, EMB J. 2: 1791)[該タンパク質コード配列は、融合タンパク質が生成され得るように、l a c Z コード領域とフレーム中のベクター中で個別に結合され得る]；p I Nベクター(Inouye & Inouye, 1985, Nucleic Acids Res. 13: 3101-3109; Van Heeke & Schuster, 1989, J. Biol . Chem. 264: 5503-5509)などが挙げられるが、これらに限定されない。

【0094】

p G E Xベクターは、また、外来ポリペプチドをグルタチオン S-トランスフェラーゼ(G S T)と共に融合タンパク質として発現させるために使用され得る。一般に、そのような融合タンパク質は可溶性であり、そして溶解した細胞からグルタチオン-アガロースビーズへの吸着とこれに続く遊離グルタチオン存在下での溶出により容易に精製され得る。p G E Xベクターは、クローン化された目的遺伝子タンパク質をG S T部分から放出できるように、トロニンもしくは第X a 因子プロテアーゼ切断部位を含有するようデザインされている。F l a g タグ、例えば3 X F l a g (Sigma, St. Louis, MO)またはm y c タグ、例えばp C D N A 3.1 / m y c - H i s (Invitrogen, Carlsbad, CA)を含む融合

タンパク質が使用され得る。これらの融合物は、抗体が入手できないタンパク質の共免疫沈降およびウエスタン検出を可能にする。

【0095】

プロモーター領域は、候補プロモーターフラグメント；即ちプロモーターを含有し得るフラグメント、を導入するための1つ以上の制限部位の下流で、クロラムフェニコールアセチルトランスフェラーゼ(“CAT”)のようなプロモーター領域を欠くレポーター転写ユニット、またはルシフェラーゼ転写ユニットを含有するベクターを用いて任意の望ましい遺伝子から選択され得る。例えば、cat遺伝子の上流の制限部位におけるプロモーター含有フラグメントのベクター中への導入は、標準的なCATアッセイにより検出することができるCAT活性の生成を惹き起こす。この目的に適したベクターは周知であり、そして容易に入手可能である。二つのそのようなベクターはpKK232-8およびpCM7である。かくして、本発明のポリヌクレオチドの発現用のプロモーターには、周知かつ容易に入手可能なプロモーターのみならず、レポーター遺伝子を用いて前述の技法により容易に得られ得るプロモーターも含まれる。

10

【0096】

本発明にしたがうポリヌクレオチドおよびポリペプチドの発現に適した公知の細菌プロモーターの中には、大腸菌のlacIおよびlacZプロモーター、T3およびT7プロモーター、T5tacプロモーター、ラムダPR、PLプロモーターならびにtrpプロモーターがある。この点に関して適当な公知の真核細胞のプロモーターの中には、CMVの直接前初期プロモーター、HSVチミジンキナーゼプロモーター、初期および後期SV40プロモーター、Rous肉腫ウイルス(“RSV”)のそれらのようなレトロウイルスLTRのプロモーター、ならびにマウスメタロチオネイン-Iプロモーターのようなメタロチオネインプロモーターがある。例えば、プラスミド構築体は、β-ガラクトシダーゼ、クロラムフェニコールアセチルトランスフェラーゼ、緑色蛍光タンパク質またはルシフェラーゼのコード領域をコードするレポーター転写ユニットと融合したHDAC9転写制御配列を含有し得る。

20

【0097】

この構築体は、HDAC9の転写を調節する小分子をスクリーニングするために使用され得る。このような小分子は、潜在的な治療薬である。さらに、HDAC9レポーター遺伝子は、蛍光リポーターおよびイメージング技術、例えば蛍光顕微鏡またはBiophotonicインビボイメージング、リアルタイムで視覚的および定量的測定を提供する技術(Xenogen, Palo Alto, CA)を用いて哺乳動物細胞または異種移植片におけるHDAC9治療の効果を試験するために使用され得る。正常、罹患した、または薬物処置した組織または細胞中のこれらのレポーターにおける変化は、HDAC9発現または活性における変化のインジケーターであり得る。

30

【0098】

昆虫のシステムにおいて、Autographa californicaの核多角体病ウイルス(AcNPV)は、外来遺伝子発現ベクターとして使用され得る幾つかの昆虫システムの1つである。このウイルスは、Spodoptera frugiperda細胞中で増殖する。コード配列をウイルスの非必須領域(例えば多角体遺伝子)の中に個別にクローン化し、AcNPVプロモーター(例えば多角体プロモーター)の制御下に置き得る。コード配列の成功的挿入は、ポリヘドリン遺伝子の失活および閉じ込められていない組換えウイルス(すなわち、ポリヘドリン遺伝子によってコードされるタンパク質性の外被を欠いているウイルス)の生成をもたらす。次いで、これらの組換えウイルスを用いて、挿入遺伝子を発現するSpodoptera frugiperda細胞に感染させる(例えば、Smith et al., 1983, J. Virol. 46: 584; Smith, 米国特許第4,215,051号を参照のこと)。

40

【0099】

哺乳動物宿主細胞において、多数のウイルスに基づく発現システムを利用し得る。アデノウイルスを発現ベクターとして用いる場合には、興味のあるコード配列をアデノウイルスの転写／翻訳制御複合体、例えば後期プロモーターおよび三連リーダー配列に結合させ得

50

る。次いで、このキメラ遺伝子をインビトロもしくはインビボ組換えによりアデノウイルスのゲノム中に挿入し得る。ウイルスゲノムの非必須領域(例えば、E 1もしくはE 3領域)への挿入は、感染宿主中で生存可能かつ望ましいタンパク質を発現する能力のある、組換えウイルスをもたらすであろう(例えば、Logan & Shenk, 1984, Proc. Natl. Acad. Sci. USA 81: 3655-3659を参照のこと)。また、挿入された遺伝子コード配列の効率的翻訳のために特定の開始シグナルが必要とされ得る。

【0100】

これらのシグナルとしては、A T G開始コドンおよび隣接の配列が挙げられる。それ自身の開始コドンおよび隣接の配列を含有する遺伝子全体を適切な発現ベクター中に挿入する場合には、さらに別の翻訳制御シグナルは必要でないであろう。しかしながら、配列をコードする遺伝子の一部のみが挿入される場合には、おそらくA T G開始コドンを含む外来性翻訳制御シグナルが提供されなければならない。さらに、挿入物全体の翻訳を確保するためには、開始コドンは望ましいコード配列のリーディングフレームと合っていなければならない。これらの外来性翻訳制御シグナルおよび開始コドンは、天然のおよび合成の両方の、種々の起源のものでよい。

10

【0101】

発現の効率は、適切な転写エンハンサーエレメント、転写ターミネーター等を含めることにより増大され得る(Bittner et al., 1987, Methods in Enzymol. 153: 516-544を参照のこと)。他の一般的なシステムは、S V 4 0、レトロウイルスもしくはアデノ随伴ウイルスに基づくものである。宿主細胞内での発現用の適切なベクターおよびプロモーターの選択は周知の手順であり、そして発現ベクターの構築、宿主内へのベクターの導入、および宿主自身内での発現に不可欠な技法は当分野において日常的技術である。一般的に、組換え体発現ベクターは、複製起点、下流の構造配列の転写を指令するための高度に発現される遺伝子から得られるプロモーター、およびベクターへの曝露後にベクター含有細胞の単離を可能にする選択マーカを含むであろう。

20

【0102】

加えて、挿入配列の発現を調整するか、もしくは遺伝子産物を特定の望ましい方法で修飾およびプロセッシングする、宿主細胞株を選択し得る。タンパク質産物のそのような修飾(例えばグリコシル化)およびプロセッシング(例えば切断)は、タンパク質の機能にとって重要であり得る。種々の宿主細胞は、タンパク質の翻訳後のプロセッシングおよび修飾のための特性的かつ特異的なメカニズムを有する。適切な細胞株もしくは宿主系を選んで、発現された外来性タンパク質の正確な修飾およびプロセッシングを確実にすることができる。この目的のために、一次転写物の適正なプロセッシング、遺伝子産物のグリコシル化および磷酸化のための細胞の仕組み(cellular machinery)を有する真核生物の宿主細胞を用い得る。そのような哺乳動物宿主細胞としては、C H O、V E R O、B H K、H e L a、C O S、M D C K、2 9 3、3 T 3、W I 3 8などが挙げられるが、これらに限定されない。

30

【0103】

組換えタンパク質の長期間、高収率な生成のためには、安定的な発現が好ましい。例えば、差次的発現遺伝子タンパク質を安定的に発現する細胞株を工学操作し得る。ウイルス性の複製起点を含有する発現ベクターを用いるよりもむしろ、適切な発現制御エレメント(例えば、プロモーター、エンハンサー、配列、転写ターミネーター、ポリアデニル化部位、等)で制御されるD N Aおよび選択マーカにより、宿主細胞を形質転換することができる。外来性D N Aの導入に続いて、工学操作された細胞は強化培地で1～2日間増殖させてもよく、次いで選択培地に切り替えられる。組換えプラスミド中の選択マーカが選択抵抗性を与え、そして細胞がプラスミドをその染色体中に安定的に組み込み、かつ増殖巣を形成するために増殖することを可能にし、これを今度は細胞株中にクローン化し伸張させることができる。この方法は、差次的発現遺伝子タンパク質を発現する細胞株を工学操作するのに有利に用いられる。そのような工学操作された細胞株は、発現されたタンパク質の内在的活性に影響を及ぼす化合物のスクリーニングおよび評価に特に有用であろう。

40

【0104】

それぞれ $t k^-$ 、 $h g p r t^-$ もしくは $a p r t^-$ 細胞内で使用され得る単純ヘルペスウイルスのチミジンキナーゼ (Wigler, et al., 1977, Cell 11:223)、ヒポキサンチン-グアニンホスホリボシルトランスフェラーゼ (Szybalska & Szybalski, 1962, Proc. Natl. Acad. Sci. USA 48: 2026) および アデニンホスホリボシルトランスフェラーゼ (Lowy, et al., 1980, Cell 22: 817) の遺伝子を含むが、これらに限定されない、多数の選択システムが使用され得る。また、メトトレキセートに対する抵抗性を付与する $d h f r$ (Wigler, et al., 1980, Natl. Acad. Sci. USA 77:3567; O'Hare, et al., 1981, Proc. Natl. Acad. Sci. USA 78: 1527); ミコフェノール酸耐性を付与する $g p t$ (Mulligan & Berg, 1981, Proc. Natl. Acad. Sci. USA 78: 2072); アミノグリコシド G-418 に対する耐性を付与する $n e o$ (Colberre-Garapin, et al., 1981, J. Mol. Biol. 150:1); およびヒグロマイシン抵抗性を付与する $h y g r o$ (Santerre, et al., 1984, Gene 30: 147) の遺伝子に対する選択の基盤として、代謝拮抗剤抵抗性を用いることができる。

10

【0105】

これに代わる融合タンパク質システムは、ヒト細胞株で発現される非変性融合タンパク質の迅速な精製を可能にする (Janknecht, et al., 1991, Proc. Natl. Acad. Sci. USA 88: 8972-8976)。このシステムでは、目的の遺伝子は、該遺伝子のオープンリーディングフレームが6個のヒスチジン残基からなるアミノ末端タグに翻訳的に融合されるように、ワクシニアの組換えプラスミドの中にサブクローン化される。組換えワクシニアウイルスを感染させた細胞からの抽出物を $N i^{2+}$ ニトリロ酢酸-アガロースカラムに装てんし、そしてヒスチジンタグをつけたタンパク質をイミダゾール含有緩衝液で選択的に溶出する。

20

【0106】

以下に説明するようなアッセイシステムにおける成分として用いるときには、本発明のタンパク質を直接的もしくは間接的に標識して、タンパク質と試験物質との間で形成された複合体の検出を容易にし得る。^{1 2 5} I のような放射性同位元素; 基質に曝露すると検出可能な比色シグナルもしくは光を発生する酵素標識システム; および蛍光標識を含むが、これらに限定されない、種々の適当な標識システムのいずれかを使用し得る。

【0107】

組換え DNA 技術を用いて、このようなアッセイシステム用に本発明のタンパク質を生成する場合には、標識、固相化、検出および/もしくは単離を容易にすることができる、融合タンパク質を工学操作するのが有利であり得る。

30

間接標識は、本発明のポリペプチドに特異的に結合する標識抗体のような、タンパク質の使用を伴う。そのような抗体としては、ポリクローナル、モノクローナル、キメラ、単鎖、F a b フラグメント、および F a b 発現ライブラリーにより生成されるフラグメントが挙げられるが、これらに限定されない。

【0108】

別の実施態様において、H D A C 9 タンパク質をコードする配列を含む核酸もしくはその機能的誘導体を投与して、遺伝子治療により、正常な生物学的機能、例えば正常な転写制御を促進する。遺伝子治療とは、対象への核酸の投与により実施される治療をいう。本発明のこの実施態様において、核酸は、正常な転写制御を促進することにより治療効果を仲介する、そのコードされたタンパク質を生成する。

40

【0109】

当分野において使用可能な遺伝子治療のいずれかを、本発明にしたがって使用することができる。例示的な方法は以下に説明されている。

【0110】

好ましい観点において、治療薬は、適当な宿主中で H D A C 9 タンパク質またはそのフラグメントもしくはキメラタンパク質を発現する、発現ベクターの一部であるところの H D A C 9 核酸を含む。特に、そのような核酸は H D A C 9 コード領域に機能し得るように結合したプロモーターを有しており、該プロモーターは誘導性もしくは構成的、そして所望により組織特異的である。別の特定の実施態様において、その中で H D A C 9 コード領域

50

配列および任意の他の望ましい配列がゲノム中の望ましい部位で相同的組換えを促進する領域で隣接している核酸分子が用いられ、かくしてH D A C 9 核酸の染色体内発現が提供される(Koller and Smithies, 1989, Proc. Natl. Acad. Sci. USA 86:8932-8935; Zijlstra et al., 1989, Nature 342:435-438)。

【0111】

患者への核酸の送達は、直接的（この場合には、患者は核酸もしくは核酸を保有するベクターに直接曝露される。）または間接的（この場合には、細胞は先ずインビトロで核酸により形質転換され、次いで患者に移植される。）のいずれかであり得る。これら二つのアプローチは、それぞれインビボもしくはエキスビボ(ex vivo)遺伝子治療として知られている。

10

【0112】

特異的な実施態様において、核酸はインビボで直接投与され、そこでは発現してコードされた産物を生成する。これは当分野において公知の多数の方法のいずれかにより達成されることができ、それは、例えば、それを適切な核酸発現ベクターの一部として構築し、そして、例えば、欠失もしくは弱毒レトロウイルスまたは他のウイルスベクター（例えば、米国特許第4,980,286号および上述の他のものを参照のこと）を用いる感染により、もしくは裸のDNAの直接的注入により、もしくは微粒子衝撃（例えば、遺伝子銃; Biolistic, Dupont)の使用により、または脂質もしくは細胞表面受容体もしくは形質移入剤、リボゾーム内への封入、微粒子もしくはマイクロカプセルによるコーティングにより、それが細胞内となるように、投与することによるか、または核に入ることが知られているペプチドに結合した形でそれを投与することによるか、受容体—伸介エンドサイトーシスを受けるリガンドに結合した形でそれを投与することによる（例えば、米国特許5,166,320; 5,728,399; 5,874,297; および6,030,954（これらのすべてを出典明示によりそれらの全体を本明細書の一部とする）を参照のこと）（これは受容体の特異的に発現する細胞型を標的にするために用いることができる）、等である。

20

【0113】

別の実施態様において、核酸—リガンド複合体を形成することができて、その中で、リガンドはエンドゾームを破壊する融合誘導性ウイルスペプチドを含み、核酸がリソソームによる分解を避けることを可能にする。さらに別の実施態様において、核酸は、特異的な受容体を標的にすることにより、細胞特異的な取り込みと発現をインビボで標的にされることができ（例えば、PCT公開WO 92/06180; WO 92/22635; WO 92/20316; WO 93/14188; およびWO 93/20221を参照のこと）。これに代えて、核酸を細胞内に導入し、相同的組換えにより宿主細胞のDNA内に組み込んで発現させることができる（例えば、米国特許5,413,923; 5,416,260; および5,574,205; ならびにZijlstra et al., 1989, Nature 342:435-438を参照のこと）。

30

【0114】

特定の実施態様において、H D A C 9 核酸を含有するウイルスベクターが用いられる。例えば、レトロウイルスベクターを用いることができる（例えば、米国特許5,219,740; 5,604,090; および5,834,182を参照のこと）。これらのレトロウイルスベクターは、ウイルスゲノムをパッケージして宿主細胞のDNAに組み込むために必要でない、レトロウイルス配列を取り除くように修飾されている。遺伝子治療に用いられるH D A C 9 核酸はベクター内にクローン化され、これは患者への遺伝子の送達を促進する。

40

【0115】

アデノウイルスは遺伝子治療に用いることのできる他のウイルスベクターである。アデノウイルスは気道上皮に遺伝子を送達するためのとりわけ魅力的なビヒクルである。アデノウイルスは天然では気道上皮に感染し、そこで軽度の疾患を引き起こす。アデノウイルスに基づく送達システムの他の目標は、肝臓、中枢神経系、内皮細胞および筋肉である。アデノウイルスは非分裂細胞に感染する能力があるという利点を有している。アデノウイル

50

スに基づく遺伝子治療を実施する方法は、例えば、米国特許 5,824,544 ; 5,868,040 ; 5,871,722 ; 5,880,102 ; 5,882,877 ; 5,885,808 ; 5,932,210 ; 5,981,225 ; 5,994,106 ; 5,994,132 ; 5,994,134 ; 6,001,557 ; および 6,033,884 (これらの全てを、出典明示によりそれらの全体を本明細書の一部とする)に記載されている。

【0116】

アデノ随伴ウイルス(AAV)は、また、遺伝子治療への使用が提案されている。AAVの生成し、利用する方法は、例えば、米国特許 5,173,414 ; 5,252,479 ; 5,552,311 ; 5,658,785 ; 5,763,416 ; 5,773,289 ; 5,843,742 ; 5,869,040 ; 5,942,496 ; および 5,948,675 (これらの全てを、出典明示によりそれらの全体を本明細書の一部とする)に記載されている。 10

【0117】

遺伝子治療へのもう1つのアプローチは、エレクトロポレーション、リポフェクション、燐酸カルシウム仲介形質移入もしくはウイルス感染のような方法により、遺伝子を組織培養の細胞に移入することを伴う。通常、移入の方法は選択マーカーの細胞への移入を含む。次いで、該細胞を、移入された遺伝子を取り込みそして発現しているそれらの細胞を単離するための選択下に置く。次いで、それらの細胞を患者に送達する。

【0118】

この実施態様では、核酸を先ず細胞に導入した後に、得られた組換え細胞をインビボで投与する。そのような導入は当分野において公知である任意の方法によって行なうことができ、その方法としては、形質移入、エレクトロポレーション、マイクロインジェクション、核酸配列を含有するウイルスもしくはバクテリオファージベクターによる感染、細胞融合、染色体仲介性遺伝子導入、マイクロセル仲介性遺伝子導入、スフェロプラスト融合、等が挙げられるが、これらに限定されない。外来性遺伝子を細胞に導入するための多数の技法は、当分野において公知であり、そしてレシピエント細胞の必要な発生上および生理的機能が損なわれない限り、本発明にしたがって用いられ得る。その技法は、核酸が細胞により発現可能であり、そして好ましくは、細胞の子孫によって遺伝性でありかつ発現可能であるように、核酸の細胞への安定な移入を提供すべきである。 20

【0119】

ここで得られる組換え細胞を、当分野において公知である種々の方法により患者に送達することができる。好ましい実施態様では、上皮細胞を例えば皮下に注射する。もう1つの実施態様では、組換え皮膚細胞を皮膚移植片として患者に適用し得る。好ましくは、組換え血液細胞(例えば、造血幹細胞もしくは前駆細胞)が静脈内投与される。使用を想定する細胞の量は、所望の効果、患者の状態などに依存し、当業者により決定されることができる。 30

【0120】

遺伝子治療の目的で核酸を導入できる細胞は、任意の望ましい、入手可能な細胞型をも包含し、そして上皮細胞、内皮細胞、ケラチン生成細胞、線維芽細胞、筋肉細胞、肝細胞；Tリンパ球、Bリンパ球、単球、マクロファージ、好中球、好酸球、巨核球、顆粒球のような血液細胞；種々の幹細胞もしくは前駆細胞、特に、例えば骨髄、臍帯血、末梢血、胎児肝などから得られるような造血幹細胞および前駆細胞、が挙げられるが、これらに限定されない。 40

好ましい実施態様では、遺伝子治療に用いられる細胞は患者にとって自己由来のものである。

【0121】

組換え細胞が遺伝子治療に用いられる実施態様では、HDAC9核酸はそれが細胞およびその子孫により発現可能であるように細胞中に導入され、次いで、組換え細胞が治療効果のためにインビボで投与される。特定の態様では、幹細胞もしくは前駆細胞が用いられる。単離されそしてインビトロで維持されることができる任意の幹細胞および／もしくは前駆細胞は、本発明のこの実施態様にしたがって潜在的に使用されることができる。そのよう 50

な幹細胞としては、造血幹細胞(HSC)、皮膚および腸管の内層のような上皮組織の幹細胞、胚性心筋細胞、肝臓幹細胞(例えばWO 94/08598参照)、ならびに神経幹細胞(Stemple and Anderson, 1992, Cell 71:973-985)が挙げられるが、これらに限定されない。

【0122】

上皮幹細胞(ESC)もしくはケラチン生成細胞は、皮膚および腸管の内層のような組織から公知の手順(Rheiwald, 1980, Meth. Cell Bio. 21A:229)により得られ得る。皮膚のような重層上皮組織では、再生が基底層に最も近い層である基底層内の幹細胞の分裂によって起こる。腸管の内層内の幹細胞は、この組織の迅速な再生速度を提供する。患者もしくはドナーの皮膚もしくは腸管の内層から得られるESCもしくはケラチン生成細胞を、組織培養で増殖させることができる(Pittelkow and Scott, 1986, Mayo Clinic Proc. 61:771)。もしESCがドナーにより提供されるならば、宿主対移植片反応性を抑制する方法(例えば、放射線照射、中等度の免疫抑制を促進する薬剤もしくは抗体の投与)をまた用いることができる。

【0123】

造血幹細胞(HSC)に関しては、インビトロでHSCの単離、増殖および維持を提供する任意の技法を本発明のこの実施態様において用いることができる。これを達成し得る技法としては、(a)将来の宿主もしくはドナーから単離された骨髓細胞からのHSC培養液の単離および確立、または(b)同種異系もしくは異種であり得る、先に確立された長期HSC培養の使用、が挙げられる。非自家性HSCは、好ましくは将来の宿主/患者の移植免疫反応を抑制する方法と一緒に用いられる。本発明の特別な実施態様では、ヒト骨髓細胞を後部腸骨稜(posterior iliac crest)から針穿刺吸引により得ることができる(例えば、Kodo et al., 1984, J. Clin. Invest. 73:1377-1384を参照のこと)。本発明の好ましい実施態様では、HSCは高度に富化した、もしくは実質的に純粋な形で作成することができる。この富化は長期培養の前、期間中、もしくは後で達成されることができ、そして、当分野において公知である任意の技法によって行われることができる。骨髓細胞の長期培養は、例えば、修飾Dexterの細胞培養技法(Dexter et al., 1977, J. Cell Physiol. 91:335)もしくはWitlock-Witteの培養技法(Witlock and Witte, 1982, Proc. Natl. Acad. Sci. USA 79:3608-3612)を用いて確立され、そして維持され得る。

【0124】

特定の実施態様において、遺伝子治療の目的で導入される核酸は、核酸の発現が適切な転写インデューサーの存在もしくは不存在を制御することにより制御可能であるように、コード領域に機能し得るように結合された誘導性プロモーターを含む。

【0125】

本発明のさらなる実施態様は、配列番号1、5もしくは6で示されるアミノ酸配列を含むポリペプチド、または該ポリペプチドのフラグメントに特異的に結合する精製抗体またはそのフラグメントに関するものである。好ましい実施態様はそのような抗体のフラグメントに関するものであり、該フラグメントはFabもしくはFab'₂である。特に、抗体はポリクローナル抗体もしくはモノクローナル抗体であり得る。

【0126】

1つもしくはそれ以上の差次的発現遺伝子エピトープを特異的に認識できる抗体を作成する方法が本明細書において記載される。そのような抗体としては、ポリクローナル抗体、モノクローナル抗体(mAb)、ヒト化もしくはキメラ抗体、単鎖抗体、Fabフラグメント、Fab'₂フラグメント、Fab発現ライブラリーで生成したフラグメント、抗イディオタイプ(anti-Id)抗体、および上記のいずれかのエピトープ結合フラグメントを挙げることができるが、これらに限定されない。そのような抗体は、例えば、フィンガープリント、標的、生物学的試料中の遺伝子の検出に、もしくはこれに代えて異常な標的遺伝子活性を阻害する方法として使用され得る。かくして、そのような抗体を心臓血管疾患の処置方法の一部として利用し得、そして/または患者のHDAC9ポリペプチドの異常なレベルをもしくは異常な型のHDAC9ポリペプチドの存在を試験し得るところの

10

20

30

40

50

診断技法の一部として使用し得る。

【0127】

H D A C 9 ポリペプチドに対する抗体の生成のために、H D A C 9 ポリペプチドもしくはその一部の注射により、種々の宿主動物を免疫し得る。そのような宿主動物としては、2、3の例を挙げると、ウサギ、マウスおよびラットが挙げられ得るが、これらに限定されない。免疫応答を増強するために、宿主動物種に応じて種々のアジュバントを用いることができるが、これらとしては、フロインドの(完全および不完全)アジュバント、水酸化アルミニウムのような鉱物ゲル、リゾレシチンのような界面活性剤、プルロニックポリオール、ポリアニオン、ペプチド、オイルエマルジョン、スカシガイのヘモシアニン、ジニトロフェノール、ならびに B C G (bacille Calmette-Guerin) および *Corynebacterium parvum* のような潜在的に有用なヒトアジュバントが挙げられるが、これらに限定されない。

【0128】

ポリクローナル抗体は、標的遺伝子産物もしくはその抗原的に機能を有する誘導体のような抗原により免疫した動物の血清から得られる、抗体分子の不均一な集団である。ポリクローナル抗体の生成には、上述のような宿主動物を、また上述のようにアジュバントを補充した H D A C 9 ポリペプチドもしくはその一部による注射により免疫し得る。

【0129】

特定抗原に対する抗体の均一な集団であるモノクローナル抗体は、持続性細胞株の培養により抗体分子の生成を提供する、任意の技法によって得られ得る。これらとしては、Kohler および Milstein のハイブリドーマ技法(1975, *Nature* 256:495-497; および米国特許第 4,376,110 号)、ヒト B 細胞ハイブリドーマ技法(Kosbor et al., 1983, *Immunology Today* 4:72; Cole et al., 1983, *Proc. Natl. Acad. Sci. USA* 80:2026-2030)、および E B V-ハイブリドーマ技法(Cole et al., 1985, *Monoclonal Antibodies And Cancer Therapy*, Alan R. Liss, Inc., pp.77-96)が挙げられるが、これらに限定されない。そのような抗体は、I g G、I g M、I g E、I g A、I g D およびその任意のサブクラスを含む、任意のイムノグロブリンクラスのものでもよい。本発明の m A b を生成するハイブリドーマはインビトロもしくはインビボで培養され得る。インビトロでの高力価の m A b の生成に関して、これが現在のところ好ましい生成方法とされている。

【0130】

加えて、適切な抗原特異性を持つマウス抗体分子からの遺伝子を適切な生物活性のヒト抗体分子からの遺伝子と共にスプライシングすることにより、“キメラ抗体”を生成するために開発された技法(Morrison et al., 1984, *Proc. Natl. Acad. Sci.*, 81:6851-6855; Neuberger et al., 1984, *Nature*, 312:604-608; Takeda et al., 1985, *Nature*, 314:452-454)を用いることができる。キメラ抗体は、マウス m A b から得られる可変領域もしくは超可変領域およびヒトイムノグロブリンの定常領域を有するそれらのような、種々の部分が種々の動物種から得られる分子である。

【0131】

これに代えて、単鎖抗体の生成のために記述された技法(米国特許第 4,946,778 号; Bird, 1988, *Science* 242:423-426; Huston et al., 1988, *Proc. Natl. Acad. Sci. U S A* 85:5879-5883; and Ward et al., 1989, *Nature* 334:544-546)を適応して、差次的に発現する遺伝子-単鎖抗体を生成することができる。単鎖抗体は、アミノ酸架橋を介して F v 領域の重鎖および軽鎖フラグメントを結合させることにより形成されて、単鎖ポリペプチドをもたらす。

【0132】

最も好ましくは、“ヒト化抗体”の生成に有用な技法を適応して、本明細書において開示されたポリペプチド、フラグメント、誘導体および機能的同等物に対する抗体を生成させることができる。そのような技法は、米国特許第 5,932,448 号; 第 5,693,762 号; 第 5,693,761 号; 第 5,585,089 号; 第 5,530,101 号; 第 5,910,771 号; 第 5,569,825 号; 第 5,625,126 号; 第 5,633,425 号; 第 5,789,650 号; 第 5,545,580 号; 第 5,661,016 号; および第 5,

770,429号(これらは全体的な出典明示によりその全体を本明細書の一部とする)に開示されている。

【0133】

特定のエピトープを認識する抗体フラグメントは公知の技法により作成され得る。例えば、そのようなフラグメントとしては、抗体分子のペプシン消化により生成されることのできるF(ab')₂フラグメント、およびF(ab')₂フラグメントのジスルフィド架橋を還元することにより作成されるFabフラグメントが挙げられるが、これらに限定されない。これに代えて、Fab発現ライブラリーを構築して(Huse et al., 1989, Science 246: 1275-1281)、所望の特異性を有するモノクローナルFabフラグメントの迅速かつ容易な同定を可能にし得る。

10

【0134】

好ましくは、本発明の抗体は、ヒトにおけるHDAC9発現または活性の異常、例えば細胞増殖の異常、癌、アテローム動脈硬化症、炎症性腸疾患、宿主炎症性もしくは免疫性応答、または乾癬に関連する病状を診断する方法において使用され得、該方法は、HDAC9活性の異常に関連した病状を患うヒトからの適切な組織または細胞中で、配列番号1、5または6で示されるアミノ酸配列もしくはそのフラグメントを含むポリペプチドの量を測定することを含み、そこでは、HDAC9活性の異常に関連した病状を患っていないヒトからのそれぞれの組織中の該ポリペプチドもしくはそのフラグメントの量に比べて上昇した量の該ポリペプチドもしくはそのフラグメントの存在は、そのような病状に該ヒトが罹患していることの診断となる。

20

【0135】

そのような方法は本発明のさらなる実施態様を形成する。好ましくは、該検出ステップは該適切な組織もしくは細胞を、配列番号1、5または6で示されるアミノ酸配列もしくはそのフラグメントを含むポリペプチドに特異的に結合するところの抗体と接触させること、および該抗体と該適切な組織もしくは細胞中のポリペプチドとの特異的結合を検出することを含み、そこではポリペプチドへの特異的結合の検出は配列番号1、5または6で示されるアミノ酸配列またはそのフラグメントを含むポリペプチドの存在を示す。

【0136】

検出の容易のためにはサンドイッチアッセイが特に好ましく、それには多数の変法が存在するが、その全てが本発明に包括されるものとする。

30

【0137】

例えば、典型的なフォワードアッセイでは、非標識抗体を固体の基盤に固定し、結合した分子と被検試料を接触させる。適当な期間のインキュベーションの後、抗体-抗原の二元複合体の形成を起こさせるのに十分な期間。この時点で、検出可能なシグナルを誘発する能力のあるレセプター分子で標識された第2抗体を次いで添加してインキュベートし、抗体-抗原-標識抗体の三元複合体の形成に十分な時間を与える。いかなる未反応物質も洗浄して除き、そして抗原の存在をシグナルの観察で測定するか、もしくは既知量の抗原を含有する対照試料と比較して定量し得る。

【0138】

フォワードアッセイの変法としては、結合した抗体に試料および抗体を同時に添加する同時アッセイ、もしくは標識抗体と被検試料とを先ず一緒にしてインキュベートし、表面に結合した非標識抗体に添加する逆アッセイが挙げられる。これらの技法は当業者に周知であり、そして小さな変法の可能性は容易に明らかであろう。本明細書において使用されるように、“サンドイッチアッセイ”は、基本的な二サイト技法に基づく全ての変法を包含することとする。本発明のイムノアッセイについて、唯一の制限因子は、標識抗体がHDAC9ポリペプチドもしくはそのフラグメントに特異的である抗体でなければならないということである。

40

【0139】

このタイプのアッセイにおいて最も汎用されるレセプター分子は、酵素、蛍光物質もしくは放射性核種を含有する分子のいずれかである。酵素イムノアッセイの場合には、酵素を

50

、通常グルタルアルデヒドもしくは過ヨウ素酸(periodate)を用いて第2抗体に結合させる。容易に認識されるように、しかしながら、多様な異なる連結反応技法が存在し、これらは当業者に周知である。汎用される酵素としては、なかんずく西洋わさびペルオキシダーゼ、グルコースオキシダーゼ、ペーターガラクトシダーゼおよびアルカリホスファターゼが挙げられる。特定の酵素と共に用いられる基質は、対応する酵素による加水分解によって検出可能な色調変化を生成するように一般に選ばれる。

【0140】

例えば、p-ニトロフェニルホスフェートはアルカリホスファターゼ複合体と共に使用するのに適しており；ペルオキシダーゼ複合体には1,2-フェニレンジアミンもしくはトリジンが汎用される。また、蛍光基質を使用することも可能であり、これは上述の発色性基質の代わりに蛍光産物を与える。次いで、適切な基質を含有する溶液を三元複合体に添加する。第2抗体と結合した酵素と基質が反応して定性的な視覚的シグナルを与え、これを通常分光測定法的にさらに定量して、血清試料中に存在するH D A C 9の量の判定を得ることができる。

10

【0141】

これに代えて、フルオレセインもしくはローダミンのような蛍光性化合物を、抗体の結合能を変えることなく抗体に化学的に結合させ得る。特定波長の光による照射で活性化された場合、蛍光色素で標識された抗体は光エネルギーを吸収して分子中に励起状態が誘起され、次いで特徴的なより長波長の光が放射される。発光は、光学顕微鏡により目視で検出可能な特徴的な色調として現れる。免疫蛍光およびE I A技法は両方とも当分野において非常に良く確立されていて、本発明の方法にとって特に好ましい。しかしながら、放射性同位元素、化学発光もしくは生物発光分子のようなレポーター分子も、また、使用され得る。要求される使用に合わせるための手順の変更の仕方は、当業者に容易に明らかであろう。

20

【0142】

本発明はまた、診断用試薬としての本発明のポリヌクレオチドの使用に関する。特に、本発明はヒトにおけるH D A C 9発現または活性の異常、例えば、細胞増殖の異常、癌、アテローム性動脈硬化症、炎症性腸疾患、宿主炎症性もしくは免疫性応答、または乾癬に関連した病状を診断する方法に関し、これは、配列番号1、5または6で示されるアミノ酸配列からなるポリペプチドをコードする天然型内在性ヒト遺伝子から転写されるメッセンジャーRNAの上昇した転写を、ヒトからの適切な組織もしくは細胞内で検出することを含み、そこでは上昇した転写は、該ヒトがH D A C 9発現または活性の異常に関連した病状に罹患していることの診断となる。

30

【0143】

特に、該天然型内在性ヒト遺伝子は配列番号4、7または8で示されるヌクレオチド配列を含む。好ましい実施態様では、そのような方法は、該適切な組織もしくは細胞の試料を、またはその組織もしくは細胞から得られる単離RNAもしくはDNA分子を、配列番号1、5または6で示されるアミノ酸配列からなるポリペプチドをコードする単離ヌクレオチド配列とハイストリンジェント条件下でハイブリダイズする、少なくとも約20ヌクレオチド長の単離ヌクレオチド配列と接触させることを含む。

40

【0144】

機能障害と関連する、配列番号4、7または8のポリヌクレオチドを特徴とする遺伝子の変異型の検出は、この遺伝子の過少発現、過大発現、または変化した位置的もしくは時間的発現によりもたらされるところの疾患、もしくは疾患に罹りやすさの診断に追加し、または、該診断を規定することができる、診断の手段を提供する。この遺伝子の変異を保有する個人は種々の技法によりDNAレベルで検出され得る。

【0145】

診断用の核酸、特にmRNAは、血液、尿、唾液、組織生検もしくは剖検材料からのような対象の細胞から入手し得る。ゲノムDNAは検出のために直接使用され得るか、または、分析前にPCRもしくはその他の増幅技法を用いることにより酵素的に増幅され得る。

50

R N Aもしくはc D N Aはまた、同様に用いられ得る。欠失および挿入は、正常な遺伝子型と比較して増幅産物のサイズの変化により検出され得る。本発明のH D A C 9ポリペプチドをコードする標識ヌクレオチド配列に増幅D N Aをハイブリダイズさせることにより、点突然変異を同定することができる。

【0146】

完全にマッチした配列は、R Nアーゼ消化もしくは融解温度の違いによりミスマッチの二重鎖から区別されることができる。D N A配列の差はまた、変性剤の有無で、ゲル中でのD N Aフラグメントの電気泳動の移動度の変化により、もしくは直接的なD N A配列決定により検出され得る(例えば、Myers et al., Science (1985) 230:1242)。特定の位置での配列変化は、また、R NアーゼおよびS 1保護のようなヌクレアーゼ保護アッセイ、もしくは化学的切断法(Cotton et al., Proc Natl Acad Sci USA (1985) 85:4397-4401を参照のこと)によって明らかにされ得る。もう1つの実施態様では、本発明のH D A C 9ポリペプチドをコードするヌクレオチド配列もしくはそのようなヌクレオチド配列のフラグメントを含むオリゴヌクレオチドプローブのアレイを構築して、例えば遺伝的変異の効率よいスクリーニングを実施することができる。アレイ技法の方法は周知で汎用性を有し、そして、遺伝子発現、遺伝連鎖および遺伝的変異性を含む分子遺伝学における様々な問題に対処するために使用され得る(例えば：M. Chee et al., Science, Vol 274, pp 610-613 (1996)を参照のこと)。

10

【0147】

診断用アッセイは、記載の方法によるH D A C 9遺伝子中の変異の検出を通して、疾患の罹りやすさを診断もしくは決定する方法を提供する。加えて、そのような疾患は、対象から得られる試料からポリペプチドもしくはm R N Aの異常に減少したもしくは増加したレベルを測定することを含む方法により診断され得る。減少したもしくは増加した発現は、例えば核酸増幅、例を挙げればP C R、R T-P C R、R Nアーゼ保護、ノーザンブロットティングおよび他のハイブリダイゼーション法のようなポリヌクレオチド定量用に当分野において周知である任意の方法を用いてR N Aのレベルで測定されることができる。宿主から得られる試料中における、本発明のポリペプチドのような、タンパク質のレベルを測定するために使用できるアッセイ技法は当業者に周知である。そのようなアッセイ法としては、ラジオイムノアッセイ、拮抗的結合アッセイ、ウェスタンブロット分析およびE L I S Aアッセイが挙げられる。

20

30

【0148】

かくして、別の態様において、本発明は：

- (a)本発明のポリヌクレオチド、好ましくは配列番号2、3、4、7もしくは8のヌクレオチド配列、またはそのフラグメント；
- (b)(a)のヌクレオチド配列と相補的なヌクレオチド配列；
- (c)本発明のポリペプチド、好ましくは配列番号1、5もしくは6のポリペプチド、またはそのフラグメント；あるいは
- (d)本発明のポリペプチドに対する、好ましくは配列番号1、5もしくは6のポリペプチドに対する、抗体を含む、診断用キットに関する。

40

【0149】

そのような任意のキットにおいて、(a)、(b)、(c)もしくは(d)は実質的な成分を含み得ることは認識されるであろう。そのようなキットは、疾患、特にH D A C 9発現または活性の異常に関連した疾患または病状、例えば細胞増殖の異常、癌、アテローム性動脈硬化症、炎症性腸疾患、宿主炎症性または免疫性応答、または乾癬の診断、あるいは該疾患への罹りやすさの診断に有用であろう。

【0150】

本発明のヌクレオチド配列はまた、染色体上の位置決定にも有用である。該配列を個別のヒト染色体上の特定の部位に特異的に標的化し、これとハイブリダイズさせることができる。本発明にしたがって関連配列を染色体にマッピングすることは、これらの配列を疾患

50

関連遺伝子と相関させる重要な最初のステップである。一旦配列が正確な染色体の部位にマッピングされれば、染色体上の配列の物理的な位置を遺伝子マップデータと相関させることができる。そのようなデータは、例えば、V. McKusick, Johns Hopkins University Welch Medical Libraryを通してオンラインで入手可能)に見出される。同じ染色体領域にマップされている遺伝子と疾患との関係は、次いで、連鎖分析(物理的に隣接する遺伝子の共遺伝)により同定される。

【0151】

また、罹患した個体と罹患していない個体の間のcDNAもしくはゲノム配列の差を決定することが可能である。もし変異が罹患している個体の全てもしくは一部で観察されるが正常人のいずれにも観察されないならば、該変異は疾患の原因因子と思われる。

10

【0152】

本発明のさらなる実施態様は、上で議論したいずれかの治療効果のための、薬学的に許容される担体、賦形剤もしくは希釈剤と一緒に、医薬組成物の投与に関する。そのような医薬組成物は、HDAC9、そのポリペプチドに対する抗体、HDAC9機能の模倣物、アゴニスト、アンタゴニストまたはインヒビターからなってもよい。組成物は単独で、もしくは安定剤のような少なくとも1つの他の剤と組み合わせて投与され得、それらは生理食塩水、緩衝生理食塩水、デキストロースおよび水を含む(ただし、これらに限定されない)任意の無菌で生体適合性の医薬担体中で投与され得る。組成物は、単独、もしくは他の薬剤、薬物もしくはホルモンと組み合わせて、患者に投与され得る。

【0153】

さらに、上記の治療用タンパク質、アンタゴニスト、抗体、アゴニスト、アンチセンス配列またはベクターは、他の適当な治療剤と組み合わせて投与され得る。併用治療において使用するための適当な剤の選択は、通常の医薬的原則にしたがって、当業者によりなされ得る。組合せ治療剤は、共同的に作用して、上記のさまざまな障害の処置または予防に効果を及ぼし得る。このアプローチを用いて、それぞれ、より少ない投与量の剤で治療効果を達成することができ、したがって、不都合な副作用の可能性を減少させることができる。HDAC9のアンタゴニストおよびアゴニストは、当分野において一般的に知られた方法を用いて製造され得る。

20

【0154】

本発明に包含される医薬組成物は、経口、静脈内、筋肉内、関節内、動脈内、髄内、くも膜下腔内、脳室内、経皮、皮下、腹腔内、鼻腔内、経腸、局所、舌下もしくは直腸の手段を含む(ただし、これらに限定されない)任意の数の経路により投与され得る。

30

【0155】

活性成分に加えて、これらの医薬組成物は活性化合物を薬学的に使用され得る製剤に加工するのを容易にする賦形剤および補助剤を含む適当な薬学的に許容される担体を含有してもよい。処方および投与の技法に関するさらなる詳細は、Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, Pa)の最新版に見出され得る。

【0156】

経口投与用の医薬組成物を、当分野において周知である薬学的に許容される担体を用いて、経口投与に適する剤型に製剤化することができる。そのような担体は、医薬組成物を患者の摂取用に、錠剤、丸剤、糖剤、カプセル剤、液剤、ゲル剤、シロップ剤、スラリー剤、懸濁剤等への製剤化を可能にする。

40

【0157】

経口用医薬製剤は、活性化合物を固体賦形剤と組み合わせ、所望により得られた混合物を粉碎し、そして適当な補助剤を添加した後、顆粒の混合物を加工し、所望により錠剤もしくは糖剤の核を得ることを通して得られ得る。適当な賦形剤は、炭水化物またはタンパク質のような増量剤、例えば、乳糖、蔗糖、マンニトールもしくはソルビトールを含む糖類; コーン、コムギ、コメ、ジャガイモもしくは他の植物から得られる澱粉; メチルセルロース、ヒドロキシプロピルメチルセルロースもしくはカルボキシメチルセルロースナトリウムのようなセルロース; アラビアゴムおよびトラガcantゴムを含むゴム; ならびにゼ

50

ラチンおよびコラーゲンのようなタンパク質である。所望により、架橋ポリビニルピロリドン、寒天、アルギン酸、もしくはその塩、例えばアルギン酸ナトリウムのような崩壊剤もしくは可溶化剤を添加してもよい。

【0158】

糖剤の核は、濃縮糖液のような適切なコーティング剤と一緒に用いてもよく、これはまた、アラビアゴム、タルク、ポリビニルピロリドン、カルボポルゲル(carbopol gel)、ポリエチレングリコール、ならびに／または二酸化チタン、ラッカー溶液、および適当な有機溶剤もしくは溶剤混合物を含有してもよい。製品の識別のために、もしくは活性化化合物の量すなわち投与量を特色づけるために染料もしくは色素を錠剤もしくは糖剤に添加してもよい。

10

【0159】

経口で利用できる医薬製剤としては、ゼラチンで作成した押込嵌めカプセル(push-fit capsule)、ならびにゼラチンおよびグリセリンもしくはソルビトールのようなコーティング剤で作成した密封軟カプセルが挙げられる。押込嵌めカプセルは、乳糖もしくは澱粉のような増量剤もしくは結合剤、タルクもしくはステアリン酸マグネシウムのような滑沢剤および、所望により安定剤、と混合された活性成分を含有することができる。軟カプセルにおいて、活性化化合物を、安定剤が入ったまたは入っていない、脂肪油、液体または液状ポリエチレングリコールのような適当な液体中に溶解または懸濁させてもよい。

【0160】

非経腸投与に適した医薬製剤は、水溶液中で、好ましくはハंकス溶液、リンゲル溶液もしくは生理緩衝食塩水のような、生理的に適合する緩衝液中で、製剤化され得る。水溶性注射懸濁剤は、カルボキシメチルセルロースナトリウム、ソルビトールもしくはデキストランのような、懸濁液の粘性を増加させる物質を含有してもよい。加えて、活性化化合物の懸濁液は適切な油性の注射懸濁液として調製してもよい。適当な親油性溶剤もしくはビヒクルとしては、胡麻油のような脂肪油、またはオレイン酸エチルもしくはトリグリセリドのような合成脂肪酸エステル、またはリポゾームが挙げられる。非脂質性、ポリカチオン性アミノポリマーをまた、送達用のために使用し得る。所望により、懸濁液はまた、適当な安定剤、または高濃度溶液の調製を可能にするために、化合物の溶解度を増加させる薬剤、を含有してもよい。

20

【0161】

局所もしくは鼻腔内投与には、浸透させるべき特定の障壁に対して適当な浸透剤が製剤化において用いられる。そのような浸透剤は当分野において一般的に公知である。本発明の医薬組成物は、当分野において公知の様式で、例えば、従来の混合、溶解、造粒、糖衣がけ、湿式粉碎、乳化、カプセル充てん、封入もしくは凍結乾燥により製造され得る。

30

【0162】

医薬組成物は塩として供給されてもよく、そして塩を、塩酸、硫酸、酢酸、乳酸、酒石酸、リンゴ酸、コハク酸等を含む(ただし、これらに限定されない)多くの酸により形成することができる。塩は、対応する遊離塩基型であるよりも水性もしくは他のプロトン性溶媒中に溶けやすい傾向がある。他の場合には、好ましい製剤は以下のもの:すなわち1～50 mMヒスチジン、0.1%～2%蔗糖、および2～7%マンニトール、のいずれかもしくは全てをpH 4.5～5.5の範囲で含有し得る凍結乾燥粉末でもよく、該製剤を使用前に緩衝液と混合する。

40

【0163】

医薬組成物を調製した後、それらを適切な容器の中に入れ、適用される病状の処置のための表示をする。H D A C 9の投与については、そのような表示は投与の用量、頻度および方法を含むであろう。

本発明における使用に適した医薬組成物としては、意図する目的を達成するための有効量で活性成分が含有されている、組成物が挙げられる。有効用量の決定は、十分に、当業者の能力の範囲内である。

50

【0164】

任意の化合物について、治療上の有効用量を、例えば、腫瘍細胞の細胞培養アッセイで、または通常マウス、ウサギ、イヌもしくはブタの動物モデルで、のどちらかで初めに推定することができる。動物モデルは、また、投与の適切な濃度範囲および経路を決定するために使用され得る。次いで、そのような情報を用いて、ヒトにおける投与のための有用な用量および経路を決定することができる。

【0165】

治療的有效用量とは、症状もしくは病状を軽減する、活性成分、例えばH D A C 9もしくはそのフラグメント、H D A C 9に対する抗体、H D A C 9のアゴニスト、アンタゴニストもしくはインヒビター、の量をいう。治療的有效性および毒性は、細胞培養もしくは実験動物での標準的な薬学的手順、例えばE D 5 0 (集団の5 0 %で治療効果を示す用量)およびL D 5 0 (集団の5 0 %で致死的な用量)、により決定され得る。毒性作用と治療効果との間の用量比が治療係数(therapeutic index)であり、これはL D 5 0 / E D 5 0 の比で表すことができる。大きな治療係数を示す医薬組成物が好ましい。細胞培養アッセイおよび動物試験で得られたデータは、ヒトでの使用のための一定の範囲の投与量の製剤化において用いられる。そのような組成物に含有される投与量は、好ましくは、毒性が殆どまたは全く無しに、E D 5 0 を含む血液濃度の範囲内である。投与量は、用いられる剤型、患者の感受性および投与経路に依存してこの範囲内で変動する。

10

【0166】

正確な投与量は、処置を要する対象に関連した諸要因に照らし合わせて、医師により決定されるであろう。投与量と投与法は、十分なレベルの活性成分を供給するように、あるいは望ましい効果を維持するように調整される。考慮に入れられる要因としては、疾病状態の重症度、対象の全身健康状態、対象の年齢、体重および性別、食事、投与の時刻および頻度、併用薬、反応感受性、ならびに治療に対する認容性／応答が挙げられる。持続性医薬組成物は特定の製剤の半減期およびクリアランス速度に応じて3～4日毎、週1回、もしくは隔週毎に投与されてもよい。

20

【0167】

通常の投与量は、投与経路に応じて0.1～100,000マイクログラム、約1gの総用量まで変動してもよい。特定の投与量および送達方法に関する指針は文献中に提供されており、当分野における医師は一般に入手できる。当業者は、タンパク質もしくはその阻害剤に対するのとは異なる処方ヌクレオチドに対して使用するであろう。同様に、ポリヌクレオチドもしくはポリペプチドの送達は、特定の細胞、病状、部位等に特異的であろう。タンパク質の経口投与に適した医薬製剤は、例えば、米国特許5,008,114; 5,505,962; 5,641,515; 5,681,811; 5,700,486; 5,766,633; 5,792,451; 5,853,748; 5,972,387; 5,976,569; および6,051,561に記載されている。

30

【0168】

以下の実施例は本発明を例示するものであって、いかなる点からもその範囲を限定するものではない。

【0169】

実施例

実施例1: バイオインフォマティクスを用いる新規H D A C 関連ヒトD N A 配列の同定
最初に非重複アミノ酸データベース、続いてCelera Human Genome Database (CHGD)、公開High Throughput Genomic (HTG) データベースおよびIncyte LIFESEQ (商標) データベースのようなあまり大きくないデータベースを検索することによって、既知のH D A C ファミリーのメンバー中から最大の進化的関連を見つけるためのストラテジーに基づく遺伝子ファミリーの新しいメンバーの同定のためのコンピューターソフトウェアを用いてH D A C 9を同定した。Smith-Waterman モデル(Pearson W. R. Comparison of methods for searching protein sequence databases. Protein Sci (1995) 4, 1145-60) および Hidden Markov モデル(すべての位置でアミノ酸の多様性に由来する推定モデル (Eddy S. R.

40

50

Hidden Markov models. Curr Opin Struct Biol (1996) 6, 361-5)を実施した。1156 bpのオープンリーディングフレーム(ORF)を同定および使用して、全組織および後根神経節cDNAライブラリーから配列決定されたクローンのデータベースを検索した。

【0170】

実施例2：全組織および後根神経節cDNAライブラリーの構築

全組織および後根神経節cDNAライブラリーを、ポリA+RNAから構築する。総RNAを31のヒト組織または後根神経節のプールしたサンプルから抽出し、そして製造者(Life Technologies, Rockville, MD)の使用説明書にしたがってTRIzol試薬を用いて単離した。mRNAを、製造者(Promega, Madison, WI)の使用説明書にしたがってPolytract mRNA Isolation System IIIを用いて単離する。総RNAを、ビオチン化オリゴ(dt)プローブとハイブリダイズする。該オリゴ(dt)-mRNAハイブリッドを、ストレプトアビジン・マグネスフェアー・パーティクル(streptavidin magnetosphere particles)上で捕捉し、そしてRnaseの入っていないH₂O中で溶出する。3μlのビオチン化オリゴ(dt)プローブ(50 pmol/μl)および13μlの20×SSCを、Rnaseの入っていない水中で65℃に加熱した60~150μgのRNAに加える。完全に冷却されるまで、この混合物を室温でインキュベートする。ストレプトアビジン・マグネスフェアー・パーティクル(ビーズ)を再懸濁し、そして0.5×SSC中で3回洗浄し、次いで0.5×SSC中で再懸濁する。前のステップからのRNA-オリゴ(dt)ハイブリッドを、これらのビーズに加える。

【0171】

ビーズからポリ-A RNAを分離するために、該ビーズを、Rnaseの入っていない水中で再懸濁し、そして磁氣的に捕捉し、そして次いでビーズからの溶出液をエタノール沈澱する。第1および第2ストランドのcDNA合成を、Life Technologiesからの修正手順(D'Alessio, J. M., Gruber, C.E., Cain, C. R., and Noon, M. C. (1990) Focus 12, 47)を用いて行う。第1ストランド合成を、1×第1ストランド用緩衝液(Life Technologies)/6 mM DTT/600 nM dNTPs/2単位の抗-Rnase中で60℃に加熱した1~5μgのRNAをインキュベートすることにより行う。この混合液を40℃にて2分間インキュベートし、次いで、Superscript II 逆転写酵素(RT)および1μlのDisplay Thermo RT ターミネーター・ミックスを加え、そして混合液を40℃で1時間、続いて60℃で10分間インキュベートする。第2ストランド合成を、DEPC-H₂O/66 nM/1μl E. coli DNA リガーゼ/4μl E. coli DNA ポリメラーゼI/1μl E. coli Rnase H中の、1×第2ストランド用緩衝液(Life Technologies)中で行う。この混合液を10℃で10分間、そして次いで16℃で2時間インキュベートする。

【0172】

この混合液に、2μlのT4 DNAポリメラーゼを加え、そして16℃で5分間インキュベーションを継続する。反応を10μlの0.5 M EDTAで停止させ、フェノール/クロロホルム/イソアミルアルコールで抽出し、そしてエタノール沈澱を行う。Sal IおよびNot Iアダプターを、通常の方法を用いて指向性クローニングのためのライゲーションによりcDNAの5'末端に加える。次いで、該cDNAをサイズ分別カラムにかけ、製造者(Life Technologies, Rockville, MD)の使用説明書にしたがって、>500 bpの長さのcDNAを回収する。cDNAを、通常の方法を用いて、Sal I/Not I消化Gateway適合性pCMV-Sport 6ベクター(Life Technologies, Rockville, MD)にライゲートする。コンピテントDH10B細胞(Life Technologies, Rockville, MD)を、通常の方法を用いて、得られたライブラリーで形質転換する。ライブラリーの半固体増幅を、製造者(Life Technologies, Rockville, MD)の使用説明書にしたがって行う。

【0173】

実施例3：配列番号1、5もしくは6からなる新規HDAC9をコードする完全長cDNAの調製：

10

20

30

40

50

1 1 5 6 塩基対の O R F を使用して、B L A S T を用いる全組織および後根神経節 c D N A ライブラリーからの配列決定されたデータベースを検索した。それぞれのライブラリーから 2 つ、計 4 つのクローンが O R F (M 6、K 1 0、P 3、F 2 3) を含有することが見出された。これらのクローンのうち、全組織のライブラリーからの M 6 は最も完全であるが N 末端からの約 4 4 塩基対を欠失していることが判明した。完全な c D N A に対して予想されるのよりも少し小さいタンパク質が、インビトロでの翻訳により観察された。

【0 1 7 4】

タンパク質が不完全 c D N A のインビトロでの翻訳により観察されるという結果により、H D A C 9 内の別の翻訳開始部位の可能性が示唆される。特に、p C M V S p o r t 6 中の H D A C 9 の配列決定を、自動 A B I S e q u e n c e r (ACGT, Northbrook, IL) を用いて行った。P C R を、ABI Prism BigDye(商標)Terminator Cycle Sequencing Ready Reaction Kit マニュアル中に挙げられた条件を用いて行った、そして下記の通りである：変性、9 6 °C で 3 0 秒間、アニーリング、5 0 °C で 1 5 秒間、伸張、6 0 °C で 4 分間、計 2 5 サイクル。それぞれのシーケンシングにより、2 0 0 ~ 6 0 0 b p の配列が提供される。

10

【0 1 7 5】

第 1 のシーケンシングのための P C R プライマーは、5'-A T T T A G G T G A C A C T A T A G-3'(S p 6、センス鎖)および 5'-T A A T A C G A C T C A C T A T A G G G-3'(T 7、アンチセンス鎖)であった。S p 6 プライマーを用いるシーケンシングの結果は次の通りである。太字の配列は、p C M V S p o r t 6 ベクター配列である。

20

【表 1】

CTggtACCGGTCCGGAATTCCCGGGATATCGTCGACCCACGCGTCCG/GGCTGCT
CCCGGCCGAAGCCCCGAGTGCAGATCGAGCGTCCTGAGCGCCTGACCGCAGCCCT
GGATCGCCTGCGGCAGCGCGGCCTGGAACAGAGGTGTCTGCGGTTGTCAGCCCGCG
AGGCCTCGGAAGAGGAGCTGGGCCTGGTGCACAGCCCAGAGTATGTATCCCTGGTC
AGGGAGACCCAGGTCCTAGGCAAGGAGGAGCTGCAGGCGCTGTCCGGACAGTTCGA
CGCCATCTACTTCCACCCGAGTACCTTTCACTGCGCGCGGCTGGCCGCAGGGGCTGG
ACTGCAGCTGGTGGACGCTGTGCTCACTGGAGCTGTGCA:AAATGGGCTTGCCCTGG
TGAGGCCTCCCGGGCACCATGGCCAGAGGGCGGCTGCAACGGGTTCTGCGTGTTCA
ACAACGTGGCCATAGCAGCTGCACATGCCAAGCAGAAACACGGGCTACACAGGATC
CTCGTCGTGGACTGGGGGATGTGCACCATGGCAGGGGGATCCAGTATCTCTTTGAAG
GATGACCCAGCGTCCTTTACTTCTCCTGGCACCGCTATGAGCATTGGGCGCCTTCT
GGCCTTTCTGCGAGAGTCAGATGAgACGCATGGGGGGCGGGGGACAGGGCCTCGGC
TTCACTGTCAACCTGCCCTGACCAAGTTgGGGGAATGGGGAAACGCTGACTTACGTG
GCTGGCCTTCTTGACCTTGCTGGTTCCAeTGGCCTTTTGGAGTTTGACCTGAgCTGG
GTGCTTGGTcTCgGCAGGGATTTGACTcagcCaTtCgGGACCCTGAgGGGGCAAA

30

40

【0 1 7 6】

T 7 プライマーを用いるシーケンシングの結果は、

【表 2】

TCAAGCCACCAGGTGAGGATGGCACTGCAACATCTTCCACTGAGGCTCCAGCTGCCC
TCTCAGGTACATCAGGGCCTGGACGTCCTCTGGGGAGGCCACAGAGGAAGGGCCTA
GGCTAGGAGGTGCCTCTCCATTCAGCACCCGGGCCAGGATCCCTGCTAGCTGGGGTG
TGGAGTTCTCCTCCAGGAGGGCCAGGACTCGGCCCCCTGCCAGCCCCCGAAGCATTG
CAGCCAGGAGTGCAGCGTGGGGGCCCTGCAGGCCATGGCCAGGCCCCAGCGCCACC
AGCACCAGGTGAGGCTGGAAGCCATAGGCCAGGGGCAGCaCCAAGCCCAAGATGCA
GCTCAGGAAACCACCGGTCATCACTGGCAGTGGCGTGGAGACATGGAACATGGA[T
AGGGCAGcCGCCTCCTTGCCCTGATGTTAGCCACAGACTcCTCCCGTCATGGGCGA
AGTCTGGAGGCCGGTCCAgCTGTtaGGCCACGCACAGAgTCTCTGGGCTCCgtGGGACA
gGCCT:TTTtGAAAAGAtATTtAGGGTGGGTTGTGAacaggGCTGGAATGGCTGGTATAAcC
AcTGtTTAcCTGCCATT

10

であった。

【0177】

第2および第3シーケンシング・プライマーは、前の段階のシーケンシングから得られた配列のプライマーとなるように設計されている。第2プライマーは、5'-GTCATCACTGGCGCAGTGGCGTG-3'(HUF7392、アンチセンス鎖)および5'-TGGACCTGCAGCTGGTG-3'(DF-2、センス鎖)である。DF-2プライマーを用いるシーケンシングの結果は、

20

【表3】

CTGGcAAATGGGCTTGCCCTGG

TGAGGCCTCCCGGGCACCATGGCCAGAGGGCGGCTGCCAACGGGTCTGCGTGTC
AACAACTGGCCATAGCAGCTGCACATGCCAAGCAGAAACACGGGCTACACAGGAT
CCTCGTCGTGGACTGGGATGTGCACCATGGCCAGGGGATCCAGTATCTCTTTGAGGA
TGACCCAGCGTCCTTTACTTCTCCTGGCACCGCTATGAGCATGGGCGCTTCTGGCCT
TTCCTGCGAGAGTCAGATGCAGACGCAGTGGGGCGGGGACAGGGCCTCGGCTTCAC
TGTC AACCTGCCCTGGAACCAGGTTGGGATGGGAAACGCTGACTACGTGGCTGCCTT
cCTGCACCTGCTGCTCCCACTGGCCTTTGAGTTTGACCCTGAGCTGGTGCTGGTCTCG
GCAGGATTTGACTCAGCCATCGGGGACCCTGAGGGGCAAATGCAGGCCACGCCAGA
GTGCTTCGCCCACCTCACACAGCTGCTGCAGGTGCTGGCCGGCGGCCGGGTCTGTGC
CGTGCTGGAGGGCGGCTACCACCTGGAGTCACTGGCGGAGTCAGTGTGCATGACAG
TACAGACGCTGCTGGGTGACCCGGcCCCACCCCTGTCAGGGCCAATGGCGCCATGTC
AGAGTGCCCTAgAgTCATTCAgAGTGCCCGTGCTGCCAGGcCCCGCACTGGAAAgAgG
CTTCAGCAGCAAgATGTGACCGcTGTGCCGATGAACCCCA

30

40

であった。

【0178】

HUF7392プライマーに関するシーケンシング結果は、

【表4】

TGtaTAGGGcAGCCGCTCCTTGCC
 CCTGATGTTcAGCCACAGACTCCTCCCGTCATGGGCGAGG
 TCTGGAGGCCGGTCCAGCTGTCCCAGGGCCACGCACAGCAGCCTCTGGGCTCCGTG
 GGACAGGCCTCTCCGAACAGCCACATCCAGGGTGGCTGCTGCAGCAGAGGCTGGAG
 TGGCTGCTATACCACTGTTcACCTGCCCATCCAGCATCCCATCTAAGAGGTACAGGA
 GCTTCCCAAGTGcAGTGAGGGCCTCCTCCCGGGCCAGGGACTCGTGTGGCCTGGCCC
 AGGCTTCTGTCTCCTCCCTCAGGGGCTGACGCTTCTGTTGGATGACGTCAGGGGGCAG
 AACCAATGTGATATCCGGCGTTGTCAAGGGCAACAGCGGTGCGGACAGAGGGTGCG
 GGGCAGAGGCACgGCTGGTCCAgGAGGGAGCTCGGTGCAgATGCAGcTGCCTTACAC
 ACTGgACCCCCAGGCAGCAGAGGTGGAGGCCTCCCCTCTGGGGAGTG

10

であった。

【0179】

第3シーケンシングプライマーは、5'-AACAGCGGTGC GGACAGA-3'
 '(HUF2A、アンチセンス鎖)および5'-CTGGAGTCACTGGCGGAG-
 3'(DF3A、センス鎖)であった。DF3Aプライマーを用いるシーケンシングの結果
 は、

【表5】

20

AgcaCAGA cGCTgCTGGGTGACCCGGCCCCACCCCTG
 TCAGGGCCAATGGCGCCATGTCAGAGTGCCCTAGAGTCCATCCAGAGTGcCCGTGCT
 GCCCAGGCCCGCACTGGAAGAGCCTCCAGCAGCAAGATGTGACCGCTGTGCCGAT
 GAGCCCCAGCAGCCACTCCCCAGAGGGGAGGCCTCCACCTCTGCTGCCTGGGGGTC
 CAGTGTGTAAGGCAGCTGCATCTGCACCGAGCTCCCTCCTGGACCAGCCGTGCCTCT
 GCCCCGCACCCTCTGTCCGCACCGCTGTTGCCCTGACAACGCCGGATATCACATTGG
 TTCTGCCCCCTGACGTCATCCAACAGGAAGCGTCAGCCCTGAGGGAGGAGACAGAA
 GCCTGGGCCAGGCCACACGAGTCCCTGGCCCGGGAGGAGGCCCTcACTGcACTTGGG
 AAGCTCCTGTACCTcTTAgATGGGATGCTGGATGGGCAGGTGAACAgTGGTATA.

30

であった。

【0180】

HUF2Aプライマーを用いるシーケンシングの結果は、

【表6】

TgcaCGGATGGTCCAGGAGGGAGCTCG
 GTGCAAATGCAGCTGCCTTACACACTGGACCCCCAGGCAGCAgAGGTGGAGGCCTC
 CCCTcTGGGGAGTGGCTGCTGGGGCTCATCGGCACAGCGGTCACATCTTGCTGCTGG
 AGGCTCTTCCAGTGCGGGGGCCTGGGCAGCACGGGGCACTCTGGATGGACTCTAGGGC
 ACTCTGACATGGCGCCATTGGCCCTGACAGGGGTGGGGCCGGGTACCCAGCAGCG
 TCTGTACTGTTCATGCACACTGACTCCGCCAGTGACTCCAGGTGGTAGCCGCCCTCCA
 GCACGGCACAgACCCGGCCGCCGGCCAGCACCTGCAGCAGCTGTGTGAGGTGGGCg
 AAGCACTCTGGCGTGGCCTGCATTTGCCCTCAGGGTCCCCGATGGCTTGAGTCAAA
 TCCTGCCGAGACCAGCACCCAGCTCAGGGTCAAACCTCAAAGGCCAGTGGGAGCAGCA
 GGTGCAGGAAGGCAGCCACgTATCAGCGTTTCCCATCCCAACCTGgTTCCAGGGGCA
 GGTTGAACAGTGAAGCCGAGGGCCCCCTTGTCCCCgCCCCACCTTGCGTCTGCATcGA
 CTCTCGCAGGAAAGGCCAAgAAGCgCCCATgCTATTTT

10

であった。センス鎖およびアンチセンス鎖のシーケンシングの組合せから得られたオーバーラップしている配列を再構築して、H D A C 9 の完全な c D N A 配列を得た。図 2 A を参照。

【0181】

20

B L A S T を、M 6 c D N A 配列を含むゲノム配列を同定するためのクエリー(query)として c D N A クローン M 6 を用いる G e n b a n k データベースを検索するために使用する。この検索結果により、M 6 c D N A と同一の配列のエキソンを含むことが見出されているゲノム配列 A L 0 2 2 3 2 8 が同定された。c D N A クローン M 6 の配列を、自動 D N A シーケンシング(ACGT, Inc. Northbrook, IL)により確認した。図 2 A 参照。

【0182】

N 末端配列の残りの 4 4 b p を、ネステッド(nested)センス鎖プライマー 5'-G C G G T C G A C G C C A C C A T G G G G A C C G C G C T T G T G T A C C A T G A G G A C A T G - 3'および 5'-G T G T A C C A T G A G G A C A T G A C G G C C A C C C G G C T G C T C T G G G A C G A C C C C G A G T G C - 3'ならびに 3'プライマー 5'-G A A C C A A T G T G A T A T C C G G C G T T G - 3'を用いる P C R により加えた。該 5'プライマーをクロニングのために k o z a k 配列および S a l 1 部位に加え、そして該 3'プライマー配列は、H D A C 9 中の E c o R V 部位とオーバーラップする。P C R を、9 4 ° C で 3 0 秒間、6 8 ° C で 3 0 秒間、および 7 2 ° C で 1 分間の 1 サイクル、続いて 9 4 ° C で 3 0 秒間および 7 2 ° C で 1 分間の 2 0 サイクルを用いる増幅のためのステップサイクルファイル(step-cycle file)を用いて行った。

30

【0183】

実施例 3 H D A C 9 配列の変異体 H D A C 9 配列の 3 つの変異体、すなわち、H D A C 9 v 1、H D A C 9 v 2、および H D A C 9 v 3 が見出された。H D A C 9 v 1 は、発見されたオリジナルの配列であり、そして上に記載されている。H D A C 9 v 2 はヒト後根神経節 c D N A ライブラリーにおいておよび A L 0 2 2 3 2 8 ゲノム配列において見出された。H D A C 9 v 3 は、C e l e r a ヒトゲノムデータベースにおいて見出された停止コドンを含む予想された転写物である。H D A C 9 v 1 は 2 0 エキソンを有し、そして H D A C 9 v 2 は 2 0 エキソンを有する。H D A C 9 変異体のペプチド配列の比較により、H D A C 9 v 1 および H D A C 9 v 2 はエキソン 1 7 まで同一であるが、このエキソン以降は異なることが示された。H D A C 9 v 2 は、エキソン 1 7 と 1 8 の間に延長されたイントロン、および H D A C 9 v 1 のエキソン 1 9 を含む延長されたエキソン 1 8 を有するが、4 4 6 位のヌクレオチドにおける単一ヌクレオチド挿入の結果としてエキソン 2 0 を欠く。この挿入は該配列をフレームシフトさせ、そして該ペプチドを 1 1 アミノ酸だけ短くする(図 1 1 A)。H D A C 9 v 1 および H D A C 9 v 2 と比較すると、H D A C 9 v

40

50

3は219～240位のアミノ酸の内部欠失を有し、そして486位のアミノ酸で始まるC末端において異なる。HDAC9は、配列変異体が報告された最初のHDAC酵素である。HDAC9v1は、特記しない限り、特徴付けされた該配列変異体である。

【0184】

実施例4：HDAC関連配列モチーフの同定

M6クローンを、HDAC触媒ドメインならびにRbおよびRb-様タンパク質の結合部位を示すモチーフの存在に関して分析した。HDACは、保存されたアミノ酸を有する触媒ドメインの存在により特徴付けられる。今までに同定された大抵のHDACは1つの触媒ドメインを有するが、例外的にHDAC6は2つのドメインを有する。N末端触媒ドメインはクラスI HDACと結合する。一方、C末端触媒ドメインはクラスII HDACと結合する。N末端触媒ドメインは、PFAM予想および他のHDACの触媒ドメインとのアラインメントに基づいて、HDAC9において見出された。

【0185】

一組の保存されたアミノ酸は、HDAC活性のために極めて重要であり、そしてHDAC1における1アミノ酸変異、ならびにHDAC-様タンパク質(HDLP)、 Zn^{2+} およびHDACインヒビターTSAの複合体により形成される三次元構造に基づいて、HDACインヒビターTSAに関する極めて重要な接触を提供することが前に示された(Hassig CA, Tong JK, Fleischer TC, Owa T, Grable PG, Ayer DE, Schreiber SL. (1998) Proc Natl Acad Sci U S A. 95, 3519-3524; Finnin, M. S., Doniglan, J. R., Cohen, A., Richon, V. M., Rifkind, R. a., Marks, P. A., Breslow, R., and Pavletich, N. P. (1999) Structures of a histone deacetylase homologue bound to TSA and SAHA inhibitors. Nature 401, 188-193)。ヒトHDACに対して配列および酵素活性類似性を有する細菌性タンパク質であり、かつ、構造が解明された唯一のクラスI HDAC-様構造である、HDLPが、HDACのテンプレートとして使用された。

【0186】

いくつかの例外を有する多くのこれらの保存されたアミノ酸が、HDAC9において見出された[表4 (Table 4)]。HDACペプチド配列のアラインメントにより、HDLPにおいて結合ポケットの一部を形成する疎水性残基Leu 265はHDAC9の272位のアミノ酸においてGluで置換されていることが示された。同様に、Leu 265は、また、HDAC8においてMetで、およびHDAC6ドメイン1においてLysで、置換されている。さらに、HDLP中のAsp 173はHDAC9の177位においてGlnで置換されており、当該相違はHDAC6触媒ドメイン1においても見出された。該Aspは、HDAC4、HDAC5、HDAC6ドメイン2、およびHDAC7においてAsnに置換されている。これらのタンパク質におけるアミノ酸置換物は全く酵素的意義を有さないので、HDAC1～8は触媒活性を有する。

【0187】

HDAC9は、クラスIおよびクラスII HDACと配列類似性を有する。HDACは、酵母HDAC Rpd3、Hda1およびSir2との配列類似性、ならびに触媒ドメインの位置により分類されている。HDAC9、酵母HDAC Rpd3、Hda1、分裂酵母からのHda1サブファミリーのメンバー、cryptic loci regulator 3 (Clr3)、およびSir2のペプチド配列のアラインメントにより、HDAC9はClr3と最も高い配列類似性を有することが決定された(表1)。しかしながら、当該配列類似性は、HDAC9に分類するほどは高くない。

【0188】

ヒトHDAC 1～9およびSir 1～7ペプチド配列のアラインメントにより、HDAC9がクラスII ヒトHDAC6と極めて類似性が高いことが示された(表2)。クラスIおよびクラスII HDAC触媒ドメインとHDAC9触媒ドメインとのアラインメントにより、HDAC6触媒ドメイン1がHDAC9と最も高い配列類似性を有することが示された(表3)。

【0189】

HDAC中の触媒ドメインの位置を比較するために、HDACペプチドの触媒ドメインのPFAM予測を行った(図11B)。HDAC9触媒ドメインの位置は、クラスI HDACと同様にN末端であり、そしてアミノ酸4～323のアミノ酸配列に及ぶと推定されている。さらに、クラスI HDACの平均的な長さは443アミノ酸である。一方、クラスII HDACの平均的な長さは1069アミノ酸である。673アミノ酸のHDAC9ペプチドが、クラスIおよびクラスII HDACの間の平均的サイズである(図11B)。

【表7】

表1

HDACクラス	HDAC異性体	HDAC9に対する相同性(%)
クラスI	Rpd3	16
クラスII	Hda1	18
	Clr3	23
クラスIII	Sir2	5

10

【0190】

20

【表8】

表2

HDAC クラス	HDAC異性体	HDAC9に対する相同性(%)
クラスI	HDAC1	14
	HDAC2	15
	HDAC3	15
	HDAC8	22
クラスII	HDAC4	21
	HDAC5	19
	HDAC6	37
	HDAC7	20
クラスIII	Sir1	5
	Sir2	7
	Sir3	11
	Sir4	4
	Sir5	8
	Sir6	10
	Sir7	15

30

40

50

【0191】

【表9】

表3

HDAC クラス	HDAC異性体	HDAC 9に対する相同性(%)
クラス I	HDAC 1	20
	HDAC 2	20
	HDAC 3	20
	HDAC 8	19
クラス II	HDAC 4	39
	HDAC 5	38
	HDAC 6-1	55
	HDAC 6-2	53
	HDAC 7	40

10

20

【0192】

網膜芽腫タンパク質(Rb)遺伝子のタンパク質産物は、DNA合成、細胞サイクル、分化およびアポトーシスを制御する転写レギュレーターであり、そして通常の成長において組織特異的な役割を果たす。Rbは転写因子E2Fと複合体を形成し、該相互作用はリン酸化により制御される。Rbにおける変異は、網膜の癌、網膜芽腫の遺伝的な形態をもたらす。変異は、また、数多くの間葉性および上皮性癌において見出されている。サイクリンD1、cdk4、およびp16を含むRbのリン酸化のレギュレーターに影響する変異が、多くの癌において見出されている。したがって、Rbの機能は腫瘍形成において極めて重要な役割を果たすと考えられている(Sellers, W.R., Kaelin, W.G. Jr. (1997) J. Clin. Oncol. 15, 3301-3312, DiCiommo, D., Gallie, B.L., Bremner, R.(2000) Semin. Cancer Biol. 10, 255-269)。Rb-結合モチーフは、アミノ酸配列LXCXE〔式中、“X”は任意のアミノ酸であり得る。〕として前に定義された(Chen, T.-T. and Wang, J. Y. J. (2000) Mol. Cell Biol. 20, 5571-5580)。HDAC1中の該LXCXEドメインは、Rbの成長抑制機能にとって重要ではないが、RbへのHDACの結合に必要であることが見出された。2つの推定的Rb-結合モチーフがHDAC9において見出された(図11A、緑色のボックス)。LLCVAはアミノ酸510および515の間に位置し、そしてLSCILはアミノ酸560および564の間に位置する。これらはともにHDAC9v1およびHDAC9v2中に存在する。

30

【0193】

【表10】

40

表 4

HDAC 異性体	HDAC 異性体の触媒ドメインにおけるアミノ酸													
	Pro	Tyr	His	His	Gly	Phe	Asp	Asp	His	Asp	Phe	Asp	Leu	Tyr
HDLP	22	91	131	132	140	141	166	168	170	173	198	258	265	297
HDAC1	Pro	Glu	His	His	Gly	Phe	Asp	Asp	His	Asp	Phe	Asp	Leu	Tyr
	29	98	139	140	148	149	174	176	178	181	205	264	271	303
HDAC2	Pro	Glu	His	His	Gly	Phe	Asp	Asp	His	Asp	Phe	Asp	Leu	Tyr
	30	99	140	141	149	150	175	177	179	182	206	265	272	304
HDAC3	Pro	Asp	His	His	Gly	Phe	Asp	Asp	His	Asp	Phe	Asp	Leu	Tyr
	23	92	134	135	143	144	167	168	171	174	199	259	266	298
HDAC4	Pro	Trp	His	His	Gly	Phe	Asp	Asp	His	Asn	Phe	Asp	Leu	His
	592	762	802	803	900	901	838	839	843	846	870	934	943	976
HDAC5	Pro	Trp	His	His	Gly	Phe	Asp	Asp	His	Asn	Phe	Asp	Leu	His
	705	793	832	833	841	842	868	869	873	876	900	964	973	1006
HDAC6-1	Pro	Tyr	His	His	Gly	Tyr	Asp	Asp	His	Gln	Phe	Asp	Lys	Tyr
	106	175	215	216	224	225	251	252	255	258	283	946	353	386
HDAC6-2	Pro	Tyr	His	His	Gly	Phe	Asp	Asp	His	Asn	Phe	Asp	Leu	Tyr
	501	570	593	594	602	603	647	648	651	654	679	742	749	782
HDAC7	Pro	Tyr	His	His	Gly	Phe	Asp	Asp	His	Asn	Phe	Asp	Leu	His
	502	589	629	630	638	639	668	669	673	676	700	764	773	806
HDAC8	N/A	Tyr	His	His	Gly	Phe	Asp	Asp	His	Asp	Phe	Asp	Met	Tyr
		100	141	142	150	151	176	177	180	183	208	267	274	306
HDAC9	Pro	Tyr	His	His	Gly	Phe	Asp	Asp	His	Gln	Phe	Asp	Glu	Tyr
	21	94	134	135	143	144	170	173	174	177	205	265	272	305

非保存的アミノ酸 (太字)。アラニンメントなし (N/A)

10

20

30

40

50

【0194】

実施例 5: 正常組織における HDAC9 の mRNA の分布

正常組織における HDAC9 の mRNA の分布を、ノーザン分析を用いて調べる。プローブを、製造者 (Amersham, Piscataway, NJ) の使用説明書にしたがって Redi-Prime ランダムヌクレオチド標識キットを用いて、750bp の EcoRV/NotI HDAC9 フラグメントを ³²P-標識することにより製造する。12 の正常組織からのポリ A+RNA (Origene Technologies, Rockville, MD) を含むノーザンプロット、ならびに適合化した腫瘍と正常 cDNA のアレイ (Clontech, Palo Alto, Ca) を、[³²P] で標識した 750bp の EcoRV/NotI HDAC9 フラグメントでプローブ化し、そしてハイストリンジェント条件下 (68℃) で洗浄する。ハイブリダイズしたプロットを、6

8℃にて $2 \times \text{SSC} / 0.1\% \text{SDS}$ 中で15分間2回洗浄し、続いて68℃にて $0.1 \times \text{SSC} / 0.1\% \text{SDS}$ 中で30分間2回洗浄する。該プロットを、増感スクリーンとともにフィルムに18時間曝露する。これにより、約3.0 KbのHDAC9 mRNAが脳、結腸、心臓、腎臓、肝臓、肺、胎盤、小腸、脾臓、胃および精巣において検出されることが示される。HDAC9メッセージは筋肉において検出されないが、GAPDHもまた検出されなかった。図7参照。

【0195】

BLASTを用いる同様のコンピューター技術(Altshul, S.F. 1993, 1990参照)を用いて、GenBankまたはLIFESEQ(商標)データベースのようなヌクレオチドデータベース中の同一または関連する分子を検索する。該検索の基本は、

10

【数1】

%配列同一性 $\times$ %最大BLASTスコア

100

として定義されるプロダクト・スコア(product score)である。

【0196】

該プロダクト・スコアは、2つの配列間の同一性の度合いおよび適合配列の長さの両方を考慮している。例えば、40のプロダクト・スコアでは、該適合は誤差1～2%以内で正確であり；そして70では、該適合は正確であろう。ホモログ分子は、通常、15～40の間のプロダクト・スコアを示すものを選択することにより同定されるが、もっと低いスコアであっても関連分子を同定し得る。

20

【0197】

ノーザン分析の結果は、HDAC9をコードしている転写物が存在するライブラリーのリストとして報告されている。存在量(abundance)および存在率も報告されている。存在量は、直接的に、特定の転写物がcDNAライブラリー中に現れる回数を反映し、そして存在率は、cDNAライブラリーにおいて試験された配列の総数で除した存在量である。

【0198】

この場合において、LIFESEQ(商標)データベース(Incyte Pharmaceuticals, Inc. Palo Alto, Calif)の電子的ノーザン分析により、表5において見られるHDAC9配列の組織分布が示される。これらの結果は、HDAC9をコードする転写物が存在するcDNAライブラリーのリストとして報告されている。さまざまな組織特異的かつ混合組織源からの20ライブラリー中のHDAC9の存在により、他のHDACファミリーのメンバーのようにHDAC9は広範な組織において発現している遺伝子として見出され得ることが示される。この結果は、12の正常組織からのmRNAに対するHDAC9プローブのノーザンハイブリダイゼーションにより支持される(図7参照)。

30

【0199】

【表11】

表 5. LIFESEQ(商標)データベースから電子的に測定された組織分布

組織の分類	
心血管系	
結合組織	
消化系	
胚性構造(Embryonic Structures)	10
内分泌系	
外分泌腺	
生殖器、雌	
生殖器、雄	
生殖細胞	
血液および免疫細胞	20
肝臓	
筋骨格系	
神経系	
脾臓	
呼吸器系	
感覚器	30
皮膚	
顎口腔系	
未分類／混合	
尿管	

【0200】

実施例 6：ヒト正常組織および細胞株における HDAC9 分布のリアルタイム PCR 測定 40
 リアルタイム PCR。培養細胞株からの総 RNA を、製造者(Qiagen, Valencia CA)のプロ
 トコルにしたがって Rneasy 96 キットを用いて単離した。ヒト組織からの RNA
 A を購入し(Clontech Inc, Palo Alto, Ca)、そして組織源を下記の表 6 に示す。

【表 12】

表6 リアルタイムPCR分析のためのRNAの組織源

組織	ドナーの性別	ドナーの年齢の範囲(年)	プールされたサンプルの数
脳1	M	57	1
脳2	F	16 & 36	2
小脳	M	64	1
脊髄	M/F	17~72	31
胎児脳	M/F	20~23週	8
気管	M/F	17~70	84
肝臓1	M	27	1
肝臓2	M/F	15 & 35	2
胎児肝臓	?	15~24週	?
胃	M/F	23~61	15
膵臓	M/F	17~69	18
結腸	M	35 & 50	2
小腸	M/F	25 & 30	2
腎臓	M/F	24~55	8
脊髄	M/F	18~68	24
脾臓	M	22~60	7
胸腺	M	6~45	9
甲状腺	M/F	10~46	4
副腎	M	32~50	6
唾液腺	M/F	13~78	43
乳腺	F	23~47	8
骨格筋	M/F	23~56	10
精巣	M	28~64	25
前立腺1	M	26~64	23
前立腺2	M	14~60	10
胎盤	F	22~41	15

組織の後の数字は同一の組織のタイプからの別個のサンプルを表す；雄(M)、雌(F)

【0201】

ヒト細胞株、H1299 ヒト肺カルシノーマ、T24 膀胱カルシノーマ、SJRH30 筋性横紋筋肉腫、SJS A-1 骨肉腫、ヒト繊維芽細胞、およびA549 ヒト肺カルシノーマを、American Type Tissue Culture Collectionから入手した。総RNAを、製造者(Qiagen, Valencia, CA)の使用説明書にしたがって、RNAイーザー・キット(easy kit)を用いてヒト細胞株から単離した。RNAを、ABI Prism Sequence Detection System

10

20

30

40

50

m上でR T-P C Rを用いて定量した。H D A C 9の検出のために使用されたプライマーは、フォワードプライマー5'-G G A T C C A G T A T C T C T T T G A G G A T G A C-3'、リバープライマー5'-A G A A G C G C C C A T G C T C A T A-3'、およびT a q m a nプローブ5'-A G C G T C C T T T A C T T C T C C T G G C A C C G-3'であった。

【0202】

該Taqman Reaction System (Eurogentec, Belgium)を、製造者により指示された比率であるが0.25 U/μl 逆転写酵素(MultiScribe ABI, Perkin Elmer, Branchburg NJ)および0.08 U/μl RNaseOUT RNaseインヒビター(Life Technologies, Gaithersburg, MD)を補った、25 μl 反応液中の10 ngの総RNAとともに使用した。該逆反応を、m R N Aの逆転写反応のための48℃での5分間のインキュベーションで開始し、続いて、95℃で10分間インキュベーションして逆転写酵素を不活性化し、そして同時に‘hot-start’熱安定性DNAポリメラーゼを活性化する。この後、95℃で15秒および60℃で60秒に変更した2ステップPCR反応を50サイクルした。

10

【0203】

評価を、A B I 配列検出ソフトウェア(バージョン1.6.3)を用いて行った。R T-P C Rアッセイを、製造者のプロトコルにしたがってM a x i s c r i p tキット(AMBIION Inc., Austin, TX)を用いて、T 7 RNAポリメラーゼ反応でインビトロにて転写されたc R N Aで標準化した。R T-P C Rアッセイを、A 5 4 9肺腫瘍細胞から単離された総R N Aの希釈シリーズで標準化した。R T-P C Rと並行して、各反応におけるR N Aの総量を、製造者の使用説明書にしたがって、基準としてキットに備えられた哺乳類リボソームRNAを用いて、R i b o G r e e nキット(Molecular Probes Inc., address)を用いる蛍光測定アッセイ中で定量した。

20

【0204】

リアルタイムPCRを、また、18SリボソームRNAのレベルと比較した組織および腫瘍細胞株におけるH D A C 9の分布およびレベルを調べるために使用した。ヒトA 5 4 9肺カルシノーマ細胞株からのRNAを、サンプル中の総RNAのレベルのための内部標準として任意に選択した。A 5 4 9細胞内のH D A C 9および18S rRNAのレベルを100%に設定し、そして他の組織および細胞株におけるH D A C 9および18S rRNAのレベルをA 5 4 9 RNA中のこれらの遺伝子のレベルのパーセントとして測定した。

30

【0205】

18SリボソームRNAのレベルは、すべてのRNAサンプルのA 5 4 9内部標準の82%~126%の範囲であり、このことは分析された組織サンプル中に同じくらいの量のRNAが存在することを示唆している。H D A C 9は広範な組織においてリアルタイムPCRによりさまざまなレベルで検出され(図8)、これはノーザンブロット分析の確認となる(図7)。正常組織において、H D A C 9は、胎児脳(89.4%)、小脳(53.8%)、および胸腺(58.9%)において最も高いレベルで検出された。腫瘍細胞株において、H D A C 9は、S J R H 30細胞(85.0%)において最も高いレベルで検出された(図8)。これらの結果は、H D A C 9がいくつかの組織においてRNAレベルに関して差次的に発現していることを示唆している。

40

【0206】

実施例7：H D A C 酵素アッセイ

H D A C 9-フラッグの調製。フラッグエピトープタグ配列を、PCRによりH D A C 9 v 1の3'末端に付加した。PCRプライマーは、5'-A C G C C G G A T A T C A C A T T G G T T C T G C-3'および5'-G C G G A A T T C T T A T T A T T T A T C A T C A T C A T C T T T A T A A T C C C C G T C G A C A G C C A C C A G G T G A G G A T G G C A-3'であった。フラッグ-タグ化H D A C 9 v 1を、第1プライマーのE c o R V部位を用いて再構築し、そしてヒト発現ベクターp C D N A 3.1(-)のX b a IおよびE c o R I部位中にサブクローン化した(Invitrogen, Carlsbad, CA)。

50

【0207】

H D A C 活性アッセイ。H D A C 活性アッセイを前に記載されたように行う (Emiliani, S., Fischle, W., Van Lint, C., Al-Abed, Y., and Verdin, E. (1998) Proc. Natl. Acad. Sci. U.S.A. 95, 2795-2800)。100 mm 皿中の 50 % コンフルエントまで増殖させた 5×10^6 個の 293 細胞を、製造者の使用説明書にしたがって Geneporter トランスフェクションキットを用いて 30 μ g の C 末端フラグタグ化 H D A C 1、H D A C 3、H D A C 4、H D A C 6、H D A C 7、または H D A C 9 とともにトランスフェクトする。

【0208】

細胞培養培地を、トランスフェクションの 5 時間後に交換する。トランスフェクションの 48 時間後、細胞を冷 P B S 中で洗浄し、そして 1 ml の I P 緩衝液 (50 mM T r i s - H C l p H 7.5、120 mM N a C l、0.5 mM E D T A、0.5 % N P - 40) 中にこすり落とし、ロッカー (a rocker) 上で 20 分間インキュベートする。細胞片を、14 K で 20 分間の遠心でペレット化させる。上清を、事前に I P 緩衝液中のプロテイン G ビーズ (Pharmacia Biotech) で 1 時間透明化させる。免疫沈降は、事前に透明化した上清を、 α -F L A G M 2 アガロースアフィニティーゲル (Sigma) で 4 $^{\circ}$ C にて 2 時間または抗-H D A C 2 (Santa Cruz) で 1 時間のどちらかでインキュベートし、続いてプロテイン G ビーズで 4 $^{\circ}$ C にて 1 時間のインキュベーションすることにより行われた。次いで、ビーズを I P 緩衝液中で 5 分間 3 回洗浄し、そして次いで、高塩濃度の I P 緩衝液 (50 mM T r i s - H C l p H 7.5、1000 mM N a C l、0.5 mM E D T A、0.5 % N P - 40) 中で 4 $^{\circ}$ C にて 3 回洗浄する。次いで、I P S を 1 ml の H D - 緩衝液 (10 mM T r i s - H C l p H 8.0、10 mM N a C l、10 % グリセリン) 中で 2 分間 2 回洗浄する。

【0209】

トラボキシン阻害を測定する場合、I p s を H D - 緩衝液中の 0.3、3、30 および 300 nM T P X とともに 20 分間インキュベートする。上清を、30 μ l の H D - 緩衝液または H D - 緩衝液中の T P X 中の 100000 c p m の基質 ($[^3\text{H}]$ -A c (H 4 1 - 24) S G R G K G G K G L G K G G A K R H R K V L R D、インビトロ/BOP 化学を用いて化学的アセチル化されている) とともにインキュベートし、チューブを軽くたたくことによりセファロースを再懸濁させ、そして E p p e n d o r f 5436 T h e r m o m i x e r 中で、最高速度で 37 $^{\circ}$ C にて 2 時間振盪する。170 μ l の H D - 緩衝液および 50 μ l の停止-混合液 (1 M H C l、0.16 M H A c) を加え、15 分間ボルテックスし、次いで 600 μ l の酢酸エチルを加え、そして 45 分間ボルテックスし、次いで 14000 g にて 7 分間遠心分離する。次いで、540 μ l の有機 (上) 相を、通常の方法を用いて 5 ml のシンチレーション液中で計測する。

【0210】

H D A C 9 は触媒活性を有する。H D A C 9 が触媒活性を有するかどうかを調べるため、ならびに H D A C 1、H D A C 3 および H D A C 4 の既知の触媒活性と H D A C 9 の活性を比較するために、基質として免疫沈降 H D A C 9 および ^3H -アセチル化ヒストン H 4 ペプチドを用いるインビトロのヒストンデアセチラーゼアッセイを行った。触媒活性を欠く H D A C - 関連タンパク質、H D R P / M I T R / H D A C C を、ネガティブコントロールとして使用した (Zhou, X., Richon, V.M., Rifkind, R.A., Marks, P.A. (2000) Identification of a transcriptional repressor related to the noncatalytic domain of histone deacetylases 4 and 5. Proc Natl Acad Sci U S A 97, 1056-61)。これらの結果により、H D A C 9 は H D A C 3 および H D A C 4 と同等のレベルでヒストンペプチド基質を脱アセチル化することができるが (図 12 A)、H D A C 1 は本アッセイにおいてさらに効果的である (図 12 B) ことが示された。

【0211】

実施例 8 H D A C 9 発現および細胞局在化

製造者 (Promega, Madison, WI) の使用説明書にしたがって、1 μ g の M 6 クローン、2 μ

1 の ^{3 5} S-メチオニンおよび Sp6 TNT Quick Coupled Transcription/Translation System を用いて、H D A C 9 をインビトロで発現させる。タンパク質を通常の方法にしたがって S D S-P A G E ゲル上で電気泳動し、そして Storm phosphorimager により視覚化する。完全 H D A C 9 配列の分子量は、VectorNTI Suite ソフトウェア (Informax, North Bethesda, MD) を用いてインシリコ (in silico) で 72 k d a と推定される。二重線が 10 % S D S-P A G E ゲル上で観察された。H D A C 1 がインビトロで翻訳された場合にも二重線が観察される。これらの二重線により、第 2 の翻訳開始部位の存在の可能性が示唆される。さらに、これらの結果により、H D A C 9 が発現遺伝子であることが示唆される。図 13 を参照。

【0212】

10

1 × 10⁵ 個の C o s 7 細胞を、チャンバースライド上で培養する。細胞を該スライド上にて、製造者の使用説明書にしたがって血清なしの培地中の Gene p o r t e r 2 を用いて、2 μ g のフラッグエピトープタグ化 H D A C 9 または細胞質発現タンパク質 (Ena-flag) でトランスフェクトする。細胞培養培地を、トランスフェクションの 24 時間後に交換する。トランスフェクションの 48 時間後、細胞を P B S で 3 回洗浄し、5 % ホルムアルデヒド中で 15 分間固定し、P B S 中で 2 回洗浄し、そして 0.5 % Triton-X-100 含有 P B S 中の 10 % ウシ胎児血清 (Sigma) 中で室温にて 30 分間ブロックして、該細胞を浸透可能にする。細胞を P B S 中でさらに 2 回洗浄し、そして次いで、25 m g / m l の抗-フラグ- F I T C コンジュゲートで 1 時間インキュベートする。染色細胞を P B S で洗浄し、そして蛍光顕微鏡を用いて写真を撮る。

20

【0213】

H D A C 9 は核タンパク質である。翻訳された H D A C 9 ペプチド配列は、72 K d a のタンパク質と予想され、そしてこれをインビトロでの翻訳により確認した (図 13 A)。H D A C 9 の細胞局在を調べるために、フラッグエピトープタグ化 H D A C 9、E n a b l e d (E n a) または p C M V 4 フラッグを C o s 7 および 293 細胞中にトランスフェクトするか、または細胞をプラスミドなしで擬トランスフェクトした。フラッグエピトープを、トランスフェクションの 48 時間後に蛍光免疫化学により検出した (図 13 B)。E n a は、軸索誘導中に Roundabout 受容体の軸索反発機能を形質導入する A b l チロシンキナーゼの細胞骨格関連細胞質タンパク質基質である (Gertler FB, Comer AR, Juang JL, Ahern SM, Clark MJ, Liebl EC, Hoffmann FM. (1995) enabled, a dosage-sensitive sup

30

【0214】

実施例 9: H D A C 複合体における関連タンパク質の同定

トランスフェクション。1 × 10⁷ 個の C o s 7 細胞を、0.3 K v / 500 μ F にセッ

40

【0215】

免疫沈降。免疫沈降を記載されたように行う (Grozinger, C. M., Hassig, C. A., and Schreiber, S. L. 1999. Proc. Natl. Acad. Sci. USA 96, 4868-4873)。全細胞抽出物を、完全プロテアーゼインヒビターカクテル (Boehringer-Mannheim) を含む J L B 緩衝液 (50 m M T r i s-H C L、p H 8、150 m M N a C l、10 % グリセリン、0.5 % T

50

r i t o n - X - 1 0 0) 中 に 細 胞 を こ す り 落 と す こ と に よ り 、 ト ラ ン ス フ ェ ク シ ョ ン の 4 8 時 間 後 に 調 製 す る 。 溶 解 を 4 ° C で 1 0 分 間 継 続 し 、 そ し て 次 い で 、 細 胞 片 を 1 4 K で 5 分 間 の 遠 心 分 離 に よ り ペ レ ッ ト 化 す る 。 上 清 を 、 S e p h a r o s e A / G + ア ガ ロ ー ス ビ ーズ (S a n t a C r u z) で 事 前 に 透 明 に す る 。 組 換 え タ ン パ ク 質 を 、 4 ° C に て 2 時 間 の α -F L A G M 2 ア ガ ロ ー ス ア フ ィ ニ テ ィ ー ゲ ル (S i g m a) ま た は 4 ° C に て 1 時 間 の 抗 - H D A C 1 (S a n t a C r u z, S a n t a C r u z, C A) で の イ ン キ ュ ベ ー シ ョ ン に よ り 、 続 い て S e p h a r o s e A / G ビ ーズ と の イ ン キュベーションにより、事前に透明化した上清から免疫沈降する。ウエスタンブロット分析のために、ビーズをM S W B 緩衝液 (5 0 m M T r i s - H C l, p H 8、1 5 0 m M N a C l、1 m M E D T A、0 . 1 % N P - 4 0) で 洗 浄 し 、 そ し て タ ン パ ク 質 を S D S / P A G E に よ り 分 離 す る 。 ウ エ ス タ ン ブ ロ ッ ト を 、 抗 - フ ラ ッ グ M 2 (S i g m a)、H D A C 1 (S a n t a C r u z)、抗 - H D A C 2 (S a n t a C r u z)、抗 - H D A C 6 (S a n t a C r u z)、抗 - R b (P h a r m i n g e n)、ま た は 抗 - m S i n 3 A (T r a n s d u c t i o n L a b s, L e x i n g t o n, K Y) を 用 い て プ ロ ー プ 化 す る 。 10

【0216】

m S i n 3 A 複 合 体 中 の タ ン パ ク 質 と 結 合 し た H D A C 9。 ク ラ ス I の H D A C (た だ し 、 ク ラ ス I I の H D A C で は な い 。) は 、 前 か ら 、 m S i n 3 A 複 合 体 と 結 合 す る こ と が 知 ら れ て い た 。 コ ア H D A C 1 複 合 体 は 、 H D A C 1、H D A C 2、R b A p 4 6、R b A p 4 8 か ら な る 。 こ の コ ア 複 合 体 は 、 R b お よ び E 2 F 複 合 体 を 介 す る 転 写 抑 制 に 関 与 す る m S i n 3 A と 結 合 す る こ と が 見 出 さ れ て い る (L u o R X, P o s t i g o A A, D e a n D C. (1998) R b i n t e r a c t s w i t h h i s t o n e d e a c e t y l a s e t o r e p r e s s t r a n s c r i p t i o n. C e l l. 92, 463-473; M a g n a g h i - J a u l i n L, G r o i s m a n R, N a g u i b n e v a I, R o b i n P, L o r a i n S, L e V i l l a i n J P, T r o a l e n F, T r o u c h e D, H a r e l - B e l l a n A. (1998) R e t i n o b l a s t o m a p r o t e i n r e p r e s s e s t r a n s c r i p t i o n b y r e c r u i t i n g a h i s t o n e d e a c e t y l a s e. N a t u r e. 391, 601-605; B r e h m A, M i s k a E A, M c C a n c e D J, R e i d J L, B a n n i s t e r A J, K o u z a r i d e s T. (1998) R e t i n o b l a s t o m a p r o t e i n r e c r u i t s h i s t o n e d e a c e t y l a s e t o r e p r e s s t r a n s c r i p t i o n. N a t u r e. 391, 597-601)。 20

【0217】

H D A C 9 が こ の 複 合 体 の 一 部 で あ る か ど う か を 調 べ る た め に 、 内 在 性 H D A C 1、H D A C 2、R b、お よ び m S i n 3 タ ン パ ク 質 を 、 フ ラ ッ グ - エ ピ ト ー プ タ グ 化 H D A C 1、H D A C 3、H D A C 4、H D A C 6、H D A C 7 ま た は H D A C 9 で ト ラ ン ス フ ェ ク ト し た 細 胞 か ら 共 免 疫 沈 降 し た 。 ト ラ ン ス フ ェ ク ト さ れ た フ ラ ッ グ エ ピ ト ー プ タ グ 化 H D A C が 細 胞 中 に 検 出 さ れ る こ と を 確 か め る た め に 、 H D A C 発 現 の レ ベ ル を 、 フ ラ ッ グ エ ピ ト ー プ に 対 す る 抗 血 清 と と も に 免 疫 沈 降 お よ び ウ エ ス タ ン ブ ロ ッ ト す る こ と に よ り 検 出 し た 。 S i n 3 複 合 体 の 構 成 成 分 と 結 合 し た H D A C を 測 定 す る た め に 、 S i n 3 複 合 体 中 の 内 在 性 タ ン パ ク 質 を 免 疫 沈 降 し 、 そ し て 結 合 H D A C を ウ エ ス タ ン ブ ロ ッ テ ィ ン グ フ ラ ッ グ エ ピ ト ー プ - 特 異 的 抗 体 に よ り 検 出 す る と 、 H D A C 9 が H D A C 1、H D A C 2、R b お よ び m S i n 3 A に 結 合 し て い る こ と が 見 出 さ れ 、 こ れ に よ り H D A C 9 が m S i n 3 A 複 合 体 の 構 成 成 分 で あ る こ と が 示 唆 さ れ る 。 30

【0218】

S M R T お よ び N C o R と 結 合 し た H D A C 9。 コ リ プ レ ッ サ ー S M R T お よ び N C o R は m S i n 3 コ ア 複 合 体 と 結 合 す る の で 、 H D A C を N C o R お よ び S M R T と 共 免 疫 沈 降 さ せ る 実 験 を 行 っ た (図 1 5)。H D A C 9 は こ れ ら の タ ン パ ク 質 の 両 方 と 共 免 疫 沈 降 し 、 こ れ に よ り H D A C 9 は S M R T お よ び N C o R と 結 合 す る こ と が 示 唆 さ れ る 。 抗 - N C o R で の フ ラ ッ グ 検 出 ブ ロ ッ ト の ウ エ ス タ ン 分 析 に よ り 、 N C o R が 免 疫 沈 降 す る こ と が 示 さ れ る 。 従 来 か ら 報 告 さ れ て い た よ う に 、 S M R T は H D A C 4 お よ び H D A C 6 と 共 免 疫 沈 降 し 、 そ し て H D A C 6 お よ び H D A C 7 は S i n 3 A 複 合 体 と 結 合 し な か っ た 。 40

【0219】

1 4 - 3 - 3 お よ び E r k タ ン パ ク 質 と 結 合 し た H D A C 9。H D A C 4 は 、 従 来 か ら 、 1 4 - 3 - 3 - β 、1 4 - 3 - 3 - ϵ 、C a m K、E r k 1 お よ び E r k 2 タ ン パ ク 質 と 50

結合することが見出されており、これらは細胞質中のH D A C 4を隔離し、そしてリン酸化されたH D A C 4およびH D A C 5が核内に入ってM E F 2活性化転写を抑制することを防ぐ。H D A C 9がこれらのタンパク質と結合するかどうかを調べるために、H D A Cを1 4-3-3およびE r kタンパク質と共免疫沈降させる実験を行った。試験されたすべてのH D A Cは、1 4-3-3およびE r kと結合した。これらの結果により、H D A Cの1 4-3-3およびE r kとの結合は細胞質においてH D A Cの隔離の一般的なメカニズムであり得る。

【0 2 2 0】

H D A C 9の分類。H D A Cは、酵母H D A Cとの配列類似性、配列の長さ、触媒ドメインの位置、細胞の局在性、結合しているタンパク質、およびH D A Cインヒビターに対する感受性により分類される。この研究のデータにより、H D A C 9はクラスIおよびクラスIIのH D A Cの両方の特徴を有することが示唆される。H D A C 9は、クラスII酵母h d a 1サブファミリーメンバーC l r 3およびH D A C 6触媒ドメイン1と配列類似性を有する。さらに、3 K bのH D A C 9転写物は腎臓および精巣においてのみ検出され、これによりクラスII H D A Cのように組織分布が限定されていることが示唆される。H D A C 9はクラスIとクラスII H D A Cの間の長さである。クラスI H D A Cは平均4 4 3 b pの長さであるが、クラスII H D A Cは平均1 0 6 9 b pの長さである。しかしながら、H D A C 9は、クラスII H D A Cにおいて見出されているC末端触媒ドメインとは反対に、N末端触媒ドメインを有することが見出された。

10

【0 2 2 1】

N末端およびC末端触媒ドメインの両方を有するH D A C 6は例外である。さらに、クラスI H D A Cは核タンパク質であるが、クラスII H D A Cは核-細胞質性である。免疫細胞化学により、H D A C 9は、優勢的に核性であり、そして細胞質で発現するE n aタンパク質と比較して、さまざまな垂細胞性区画において検出されることが示された。差次的に発現し得る3 K bのH D A C 9転写物とは対照的に、ノーザン分析により同定される3.5 K bのH D A C 9転写物は、クラスI H D A Cと同様に、正常組織、腫瘍組織および細胞株においていたところで発現していた。さらに、H D A C 9は、前に、H D A C 1、H D A C 2、m S i n 3 AおよびR bを含むクラスI H D A C複合体とのみ結合したタンパク質と共免疫沈降することが見出された。また、H D A C 9は、これまではH D A C 1においてのみ見出されていた推定的C末端L X C X Eモチーフを有する。H D A C 9は、また、N C o RおよびS M R Tと結合することも見出された。この事実により、H D A C 9はクラスIおよびクラスII H D A Cの特徴の橋渡しをする特徴を有することが示唆される。

20

30

【図面の簡単な説明】

【0 2 2 2】

【図1】図1は、G E N F A M(プロプライエタリ・ソフトウェア)を用いて同定され、そして完全H D A C 9 c D N A配列に関してデータベースを検索するために使用される1 1 5 6 b pオープンリーディングフレームを示す。それぞれのO R F(配列番号3)はヌクレオチド位置番号1で開始し、そしてヌクレオチド位置番号1 1 5 6で終了する。

【図2】図2 Aおよび2 Bは、それぞれ、H D A C 9の完全長c D N A配列(配列番号2)およびそのアミノ酸配列(配列番号1)を示す。完全長c D N A配列は、ヌクレオチド位置番号1で開始し、そしてヌクレオチド位置番号2 0 2 2で終了する。

40

【図3】図3は、クローン1 9 8 9 2 9/H D A C 9の配列とアラインした、インシリコ(in silico)のゲノム性D N A配列(A L 0 2 2 3 2 8)(配列番号4)を示す。該アラインメントは、プロプライエタリ・ソフトウェア(Novartis Pharmaceuticals, Summit, NJ)を用いて製造された。

【図4】図4は、H D A C 9予想ペプチドおよびS. pombe Hda1ペプチドのアラインメントを示す。クエリー(query)はH D A C 9ペプチドであり、そして対象(subject)はS. pombe Hda1ペプチドである。アラインメントは、Clustalwアルゴリズム(Higgins, D.G., Thompson, J.D., Gibson, T.J. (1996) 複数の配列アラインメントのためのCLUSTALの使用。

50

Methods Enzymol 266, 383-402)を用いて製造された。

【図5】図5は、HDAC1およびHDAC9v1のアラインメントならびに推定される触媒ドメインのアミノ酸およびRb結合ドメインの位置を示す。触媒ドメインのアミノ酸は実線で囲まれ、そして推定されるRbドメインのアミノ酸は点線の囲み中に含まれる。

アラインメントは、Clustalwアルゴリズム(Higgins, D.G., Thompson, J.D., Gibson, T.J. (1996) 複数の配列アラインメントのためのCLUSTALの使用。Methods Enzymol 266, 383-402)を用いて製造された。

【図6】図6は、HDAC1~9v1のアラインメントを示す。該アラインメントは、Clustalwアルゴリズム(Higgins, D.G., Thompson, J.D., Gibson, T.J. (1996) 複数の配列アラインメントのためのCLUSTALの使用。Methods Enzymol 266, 383-402)を用いて製造された。 10

【図7】図7は、HDAC9のノーザン分析を示す。(A)正常ヒト組織におけるHDAC9の分布のノーザンブロット分析。GAPDHを、RNAローディングのための対照と同一のブロットとハイブリダイズした。(B)腫瘍および正常組織の比較におけるHDAC9のノーザンブロット分析。GAPDHを、RNAローディングのための対照と同一のブロットとハイブリダイズさせた。

【図8】図8は、18SリボソームRNAと比較した正常ヒト組織および細胞株中のHDAC9の分布のリアルタイムPCR分析を示す。ヒト肺カルシノーマ細胞株であるA549からのRNAを、内部標準として使用した。

【図9】図9は、クラスII HDAC(HDAC4、5、6、7)とのHDAC9v1のアラインメントを示す。該アラインメントは、Clustalwアルゴリズム(Higgins, D.G., Thompson, J.D., Gibson, T.J. (1996) 複数の配列アラインメントのためのCLUSTALの使用。Methods Enzymol 266, 383-402)を用いて製造された。触媒ドメインのアミノ酸が実線で囲まれている。 20

【図10】図10は、クラスI HDAC(HDAC1、2、3、8)とのHDAC9v1のアラインメントを示す。該アラインメントは、Clustalwアルゴリズム(Higgins, D.G., Thompson, J.D., Gibson, T.J. (1996) 複数の配列アラインメントのためのCLUSTALの使用。Methods Enzymol 266, 383-402)を用いて製造された。触媒ドメインのアミノ酸が実線で囲まれている。

【図11】図11は、3つのHDAC9配列変異体(HDAC9v1、HDAC9v2、およびHDAC9v3)を示す。HDAC9v1およびHDAC9v2は、ヒトESTデータベースを検索することにより見出され、そしてHDAC9v3はCelera配列データベース中の予想される転写物として見出された。(A)は、該3つのHDAC9変異体ペプチド配列のアラインメントを示す。(B)は、クラスIおよびクラスII HDACペプチド配列の略図を示す。触媒ドメインは黒塗りのボックスであり、そして推定されるLXCXEモチーフは白抜きのボックスである。(C)は、HDAC9v1およびHDAC9v2のゲノム構造の略図である。エキソンを黒塗りのボックスで示し、そしてイントロンを黒塗りのボックスの間の線で示す。ボックスおよび線の長さは、エキソンおよびイントロンの長さを示す。 30

【図12】図12は、HDAC9が酵素的に活性なヒストンデアセチラーゼであることを示す。(A)HDAC9の触媒活性は、HDAC3およびHDAC4の活性に匹敵する。(B)は、HDAC1が本アッセイにおけるヒストン基質の脱アセチル化に関して、HDAC3、HDAC4、およびHDAC9よりも効果的であったことを示す。 40

【図13】図13は、HDAC9が核タンパク質であることおよびHDAC9-フラッグがインビトロで翻訳されることを示す。

【図14】図14は、HDAC9v3およびHDAC9v2のDNAおよびペプチド配列を示す。

【図 5】

Figure 5.

[illegible]

□ 触媒ドメイン残基
 □ 可能性のある RB 結合
 ポケット残基

【図 6. 2】

```

    配列 (7:9) アライン スコア: 19
    配列 (8:9) アライン スコア: 20
Guide tree      file created: [/biointfnv/software/bjbenchsw/tmp/align/1478.dnd]
Start of Multiple Alignment
8グループが存在する。

```

```

Aligning...
グループ1: Sequences: 2      スコア:24259
グループ2: Sequences: 3      スコア138415
グループ3: Sequences: 4      スコア112882
グループ4:                      Delayed
グループ5: Sequences: 2      スコア19847
グループ6: Sequences: 3      スコア7569
グループ7: Sequences: 4      スコア15689
グループ8: Sequences: 8      スコア12641
Sequence19      スコア13452
アライン スコア 36872
CLUSTALW alignment file created /biocininv/software/biocinchenw/tmp/align/1478.out]
CLUSTAL W (1.81) multiple sequence alignment

HDAC5      HNSPNESDGHSGREPSELEILPRTSLMSIPVTVEVFLPRAMPSSMGGGGGGSPSPFVELR
HDAC6      HNSPNESDGHSGREPSELEILPRTSLMSIPVTVEVFLPRAMPSSMGGGGGGSPSPFVELR
HDAC4      HSSGSHPDLSGRDPFVELLNPAKYVNDHSPVTVDVATALPLGVAFVAVPMDLRIDHQSLP
HDAC7      -BELVGGGGPSPFVEVPF-----
HDAC1      -----
HDAC2      -----
HDAC3      -----
HDAC8      -----
HDAC9      -----

HDAC5      GALLVGSVPTLIREQDLOCELLAALQCGGLOGLLFAEFGDQDHILTRGNEVQLQRLKQQ
HDAC6      GALLVGSVPTLIREQDLOCELLAALQCGGLOGLLFAEFGDQDHILTRGNEVQLQRLKQQ
HDAC4      -----VAEPALEIQDLOCELLAALQCGGLOGLLFAEFGDQDHILTRGNEVQLQRLKQQ
HDAC7      -----EPTLLALQFQRLNIDHFLAQLGQ-----QQ
HDAC1      -----EPTLLALQFQRLNIDHFLAQLGQ-----
HDAC2      -----
HDAC3      -----
HDAC8      -----
HDAC9      -----

HDAC5      QDHMLAAQQDHMLAAJQQGLIQGQREQDQDEQLKGLRGLGQLLILNKRKSKESIASYAS
HDAC6      QDHMLAAQQDHMLAAJQQGLIQGQREQDQDEQLKGLRGLGQLLILNKRKSKESIASYAS
HDAC4      QDHMLAAJQQGELLINQR - KLEHNP-----EULEKQHQREQLQGLNKRKSKESIASY
HDAC7      REVETPHILSGMTF-----HPELQVGQDQGLRLILNKRKSKESIASYAS
HDAC1      -----
HDAC2      -----
HDAC3      -----
HDAC8      -----
HDAC9      -----

```

```

NDAC5      TEVKRLRQEFLLSKSEPTPGGLNLSLPQHFACWG--AHMSLDSQSSPQSGPFGTTPBY
NDAC6      TEVKRLRQEFLLSKSEPTPGGLNLSLPQHFACWG--AHMSLDSQSSPQSGPFGTTPBY
NDAC4      TEVKRLQEFVLNKKALAHRLNHCISSEDPATYVGTQTHSSLDSQSSPQSG--VSTBY
NDAC7      SVYKQKLAIEVLKQQAALERTVHPNPGFI-----YTLEP-LETEGATSHLSF
NDAC1      -----
NDAC2      -----

```

【図 6 . 1】

【図 6. 3】

```

HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 RLPLFG-VTHERDDPLAKTASFWLWELKQVAKRRSPLLRRKDGTVISTFKKXA
HDAC6 RLPLFG-VTHERDDPLAKTASFWLWELKQVAKRRSPLLRRKDGTVISTFKKXA
HDAC4 NHPVLG-HYKAKDDPLAKTASFWLWELKQVAKRRSPLLRRKDGTVISTFKKXA
HDAC7 LFPVFLSDPPPHFPLKATVSEFWMLKRYAPK-RELEKRNPLLRK-SPAPSLERKP
HDAC1 -----
HDAC2 -----
HDAC8 -----
HDAC9 -----

HDAC5 VEITGAGCAESVCMGAPGSPSPH-SBSTIANGFTGGSVNIPTDLPCHRALPLDS
HDAC6 VEITGAGCAESVCMGAPGSPSPH-SBSTIANGFTGGSVNIPTDLPCHRALPLDS
HDAC4 LSVT-----DSACSAFGSPSPHNSGSGVEANGIAPVSIPTAISLAHRLVAREG
HDAC7 AETLG-----DESPSSSTPASGCSFNDSEHG-----
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 EFWQFSLTTEPSPFWISLGLQATVTVTHSLTASFKLSTQZAEKQALQSLRGOGTITGX
HDAC6 EFWQFSLTTEPSPFWISLGLQATVTVTHSLTASFKLSTQZAEKQALQSLRGOGTITGX
HDAC4 SAAPLPLTSPSPFWISLGLPATG-----PSAGTAGQUTTERITLFPALQQR-LS-
HDAC7 LFGTGLTFLTSPSPFWISLGLPATG-----PSAGTAGQUTTERITLFPALQQR-LS-
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 FMHTESSIPGCLLVALEGGSPGCHASILQVILLEGQARQSTLIA-----VPLHQGSP
HDAC6 FMHTESSIPGCLLVALEGGSPGCHASILQVILLEGQARQSTLIA-----VPLHQGSP
HDAC4 LFGTGLTFLTSPSPFWISLGLPATG-----PSAGTAGQUTTERITLFPALQQR-LS-
HDAC7 LFGTGLTFLTSPSPFWISLGLPATG-----PSAGTAGQUTTERITLFPALQQR-LS-
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 LVYTCVATSHMTVCKLPRKPLRSTSSPLPQSPALQGLVMOOQOGLFLKQKQ----
HDAC6 LVYTCVATSHMTVCKLPRKPLRSTSSPLPQSPALQGLVMOOQOGLFLKQKQ----
HDAC4 LVYTCVATSHMTVCKLPRKPLRSTSSPLPQSPALQGLVMOOQOGLFLKQKQ----
HDAC7 LVYTCVATSHMTVCKLPRKPLRSTSSPLPQSPALQGLVMOOQOGLFLKQKQ----
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

```

【図 6. 4】

```

HDAC5 QOLQGLKILTKTGLPQPTTTHPIETEELELTCQVLLGEGALTPREGTESESTQEDL
HDAC6 QOLQGLKILTKTGLPQPTTTHPIETEELELTCQVLLGEGALTPREGTESESTQEDL
HDAC4 QOLQGLKILTKTGLPQPTTTHPIETEELELTCQVLLGEGALTPREGTESESTQEDL
HDAC7 QOLQGLKILTKTGLPQPTTTHPIETEELELTCQVLLGEGALTPREGTESESTQEDL
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 FEDEEEDGEEELDCIQADEEGSGAEGPDLEPGAGTKLFSDAOPLQPLQVYQAPL
HDAC6 FEDEEEDGEEELDCIQADEEGSGAEGPDLEPGAGTKLFSDAOPLQPLQVYQAPL
HDAC4 FEDEEEDGEEELDCIQADEEGSGAEGPDLEPGAGTKLFSDAOPLQPLQVYQAPL
HDAC7 FEDEEEDGEEELDCIQADEEGSGAEGPDLEPGAGTKLFSDAOPLQPLQVYQAPL
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 SLATVP-----HQALETQSFPAAPGCHSPFDPQVHML-PTTQVVTITMLXHQCHQCH
HDAC6 SLATVP-----HQALETQSFPAAPGCHSPFDPQVHML-PTTQVVTITMLXHQCHQCH
HDAC4 SLATVP-----HQALETQSFPAAPGCHSPFDPQVHML-PTTQVVTITMLXHQCHQCH
HDAC7 SLATVP-----HQALETQSFPAAPGCHSPFDPQVHML-PTTQVVTITMLXHQCHQCH
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 THVHPEHAGRIQSIWERLQETGLSFCERJRGKAAITLRLQVHSEHTLLYGTFLNRQ
HDAC6 THVHPEHAGRIQSIWERLQETGLSFCERJRGKAAITLRLQVHSEHTLLYGTFLNRQ
HDAC4 THVHPEHAGRIQSIWERLQETGLSFCERJRGKAAITLRLQVHSEHTLLYGTFLNRQ
HDAC7 THVHPEHAGRIQSIWERLQETGLSFCERJRGKAAITLRLQVHSEHTLLYGTFLNRQ
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 XLDKXLLGFIQGMAYVLPCCGIQVDEDTVHNEHSSAVPMVAGCLLEAFKVAAGEL
HDAC6 XLDKXLLGFIQGMAYVLPCCGIQVDEDTVHNEHSSAVPMVAGCLLEAFKVAAGEL
HDAC4 XLDKXLLGFIQGMAYVLPCCGIQVDEDTVHNEHSSAVPMVAGCLLEAFKVAAGEL
HDAC7 XLDKXLLGFIQGMAYVLPCCGIQVDEDTVHNEHSSAVPMVAGCLLEAFKVAAGEL
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 KNGF11RPFQCHAEESTAMGCTFNSVAJ1TAKLLQKQ-----LNGKVLIVDWD1HGHNGT
HDAC6 KNGF11RPFQCHAEESTAMGCTFNSVAJ1TAKLLQKQ-----LNGKVLIVDWD1HGHNGT
HDAC4 KNGF11RPFQCHAEESTAMGCTFNSVAJ1TAKLLQKQ-----LNGKVLIVDWD1HGHNGT
HDAC7 KNGF11RPFQCHAEESTAMGCTFNSVAJ1TAKLLQKQ-----LNGKVLIVDWD1HGHNGT
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

```

【図 6. 5】

```

HDAC7 KNGF11RPFQCHAEESTAMGCTFNSVAJ1TAKLLQKQ-----LNGKVLIVDWD1HGHNGT
HDAC1 AGG-----LHNAKXIASGCTVNDVLAILELKY-----HQRVLIYDIDIHNGDGV
HDAC2 AGG-----LHNAKXIASGCTVNDVLAILELKY-----HQRVLIYDIDIHNGDGV
HDAC3 AGG-----LHNAKXIASGCTVNDVLAILELKY-----HQRVLIYDIDIHNGDGV
HDAC8 AGG-----LHNAKXIASGCTVNDVLAILELKY-----HQRVLIYDIDIHNGDGV
HDAC9 QNGIALVFPFGCHQRAALNGCTVTHVTAALAHAKQKRG-----LHNAKXIASGCTVNDVLAILELKY-----HQRVLIYDIDIHNGDGV

HDAC5 QOATYHDESVLYLSLHRYDNGHFFGS--GAFIEVGGGPGCVTVHNVMTGVDVFIQDV
HDAC6 QOATYHDESVLYLSLHRYDNGHFFGS--GAFIEVGGGPGCVTVHNVMTGVDVFIQDV
HDAC4 QOATYHDESVLYLSLHRYDNGHFFGS--GAFIEVGGGPGCVTVHNVMTGVDVFIQDV
HDAC7 QOATYHDESVLYLSLHRYDNGHFFGS--GAFIEVGGGPGCVTVHNVMTGVDVFIQDV
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 EYLTAFTFVVPV1JHEFSPDVLVSAGDAVEGHLSPGSGVTARCTGHLTRQMLTAG
HDAC6 EYLTAFTFVVPV1JHEFSPDVLVSAGDAVEGHLSPGSGVTARCTGHLTRQMLTAG
HDAC4 EYLTAFTFVVPV1JHEFSPDVLVSAGDAVEGHLSPGSGVTARCTGHLTRQMLTAG
HDAC7 EYLTAFTFVVPV1JHEFSPDVLVSAGDAVEGHLSPGSGVTARCTGHLTRQMLTAG
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 GRVYVALLGGHDLTAJCDASEACVALLSVELD-----PLDEVLQQRPNJNAVATLEKVI
HDAC6 GRVYVALLGGHDLTAJCDASEACVALLSVELD-----PLDEVLQQRPNJNAVATLEKVI
HDAC4 GRVYVALLGGHDLTAJCDASEACVALLSVELD-----PLDEVLQQRPNJNAVATLEKVI
HDAC7 GRVYVALLGGHDLTAJCDASEACVALLSVELD-----PLDEVLQQRPNJNAVATLEKVI
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 EYKHLFPQVIVNQVDTTGTCTVLCGDSLGGC--RLGCTHNTKGBKCVFVYKTHL
HDAC6 EYKHLFPQVIVNQVDTTGTCTVLCGDSLGGC--RLGCTHNTKGBKCVFVYKTHL
HDAC4 EYKHLFPQVIVNQVDTTGTCTVLCGDSLGGC--RLGCTHNTKGBKCVFVYKTHL
HDAC7 EYKHLFPQVIVNQVDTTGTCTVLCGDSLGGC--RLGCTHNTKGBKCVFVYKTHL
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 GRVYVALLGGHDLTAJCDASEACVALLSVELD-----PLDEVLQQRPNJNAVATLEKVI
HDAC6 GRVYVALLGGHDLTAJCDASEACVALLSVELD-----PLDEVLQQRPNJNAVATLEKVI
HDAC4 GRVYVALLGGHDLTAJCDASEACVALLSVELD-----PLDEVLQQRPNJNAVATLEKVI
HDAC7 GRVYVALLGGHDLTAJCDASEACVALLSVELD-----PLDEVLQQRPNJNAVATLEKVI
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

```

【図 6. 6】

```

HDAC8 RTAVALTTPDITLVLFPDVIQGLSALKEETAMAPHSIARLAEALATLGLLTLDDG
HDAC9 RTAVALTTPDITLVLFPDVIQGLSALKEETAMAPHSIARLAEALATLGLLTLDDG

HDAC5 -----
HDAC6 -----
HDAC4 -----
HDAC7 -----
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 -----
HDAC6 -----
HDAC4 -----
HDAC7 -----
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

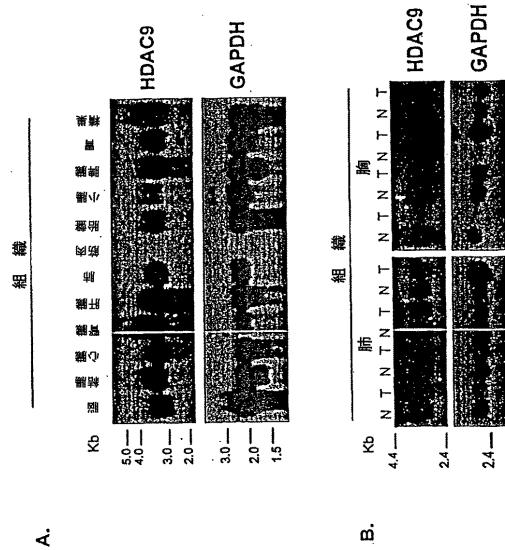
HDAC5 -----
HDAC6 -----
HDAC4 -----
HDAC7 -----
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 -----
HDAC6 -----
HDAC4 -----
HDAC7 -----
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

```

【図 7】

Figure 7



【图 8】

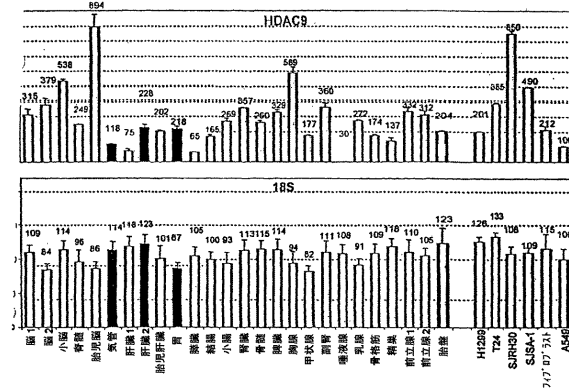


Figure 8

【図 9 . 1】

Figure 9.

配列フォーマットはヒトである。

配列 1:	HDAC4	触媒ドメイン	336 aa
配列 2:	HDAC5	触媒ドメイン	329 aa
配列 3:	HDAC6	触媒ドメイン 1	302 aa
配列 4:	HDAC6	触媒ドメイン 2	181 aa
配列 5:	HDAC7	触媒ドメイン	334 aa
配列 6:	HDAC9	完全ペプチド	673 aa

対によるアラインメントの開始

```

AlignSeq...
Sequences (1172)   アライン スコア: 70
Sequences (1339)   アライン スコア: 42
Sequences (1344)   アライン スコア: 45
Sequences (1185)   アライン スコア: 75
Sequences (1146)   アライン スコア: 37
Sequences (2029)   アライン スコア: 42
Sequences (2344)   アライン スコア: 44
Sequences (2345)   アライン スコア: 72
Sequences (2348)   アライン スコア: 37
Sequences (3174)   アライン スコア: 49
Sequences (3157)   アライン スコア: 40
Sequences (3176)   アライン スコア: 55
Sequences (4285)   アライン スコア: 46
Sequences (4146)   アライン スコア: 41
Sequences (5616)   アライン スコア: 38
Outside score      $file created: [/bioinformatics/biochenwa/tmp/align/5664.dms]

複合アラインメントの開始
5つのグループが存在する。
AlignSeq...
グループ1: Sequences: 2   スコア: 6517
グループ2: Sequences: 3   スコア: 6390
グループ3: Sequences: 4   スコア: 4601
グループ4: Sequences: 2   スコア: 5205
グループ5: Sequences: 6   スコア: 4295
AlignSeq score 15000
CLUSTALW-AlignScore $file created: [/bioinformatics/biochenwa/tmp/align/5664.out]
CLUSTAL W [1.01] multiple sequence alignment

```

【图 9.2】

[illegible]

【図 10. 1】

[illegible]

Figure 10.

配列フォーマットはヒトである。

配列 1:	HDAC1	触媒ドメイン	318	aa
配列 2:	HDAC2	触媒ドメイン	310	aa
配列 3:	HDAC3	触媒ドメイン	310	aa
配列 4:	HDAC8	触媒ドメイン	308	aa
配列 5:	HDAC9	完全ペプチド	673	aa

対によるアラインメントの関

```
aligning...
Sequences: (1:52)   アライン スコア:   92
Sequences: (14:5)  アライン スコア:   65
Sequences: (13:4)  アライン スコア:   42
Sequences: (1:15)  アライン スコア:   20
Sequences: (2:13)  アライン スコア:   64
Sequences: (2:14)  アライン スコア:   33
Sequences: (2:15)  アライン スコア:   29
Sequences: (3:4)   アライン スコア:   32
Sequences: (3:15)  アライン スコア:   21
Sequences: (4:5)   アライン スコア:   18
Guide tree: 1116 created: /bioinformatics/biochenow/mg/align/3.650.und
```

複合アラインメントの開始
4つのグループが存在する。

```

4つのグループが存在する。
Missing...
グループ1: Sequence# 2 スコア: 6624
グループ2: Sequence# 3 スコア: 5626
グループ3: Sequence# 4 スコア: 4870
グループ4: Deleted.

```

```
Sequence(S)      スコア: 2573
アラインメント スコア: 6332
'CLUSTAL' alignment file created 4/15/2019 at 16:56:36
CLUSTAL V: (1.01)  適合配列アラインメント
```

【図 10.3】

[illegible]

HDA C1 触媒ドメイン
 HDA C2 触媒ドメイン
 HDA C3 触媒ドメイン
 HDA C4 触媒ドメイン
 HDA C5 完全ペプチド KDAALSGHVSPLPHVTGCTLSCTLGVLPLATGTFPDLVLVALGPGHGLGPGHAALL

YDACA1 触媒ドメイン
 YDACA2 触媒ドメイン
 YDACA3 触媒ドメイン
 YDACA3 触媒ドメイン
 YDACA8 触媒ドメイン
 YDACA9 完全ペプチド AANDHLAGGVALLLEHSTPQLAGIARLVHGEAPPSLPSSVASPEEDVQALMDLAF

MDAC3	触媒ドメイン	-----
MDAC2	触媒ドメイン	-----
MDAC3	触媒ドメイン	-----
MDAC6	触媒ドメイン	-----
MDAC9	完全ペプチド	LEPQVWGLQCQHPHRYA

【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
27 June 2002 (27.06.2002)

PCT

(10) International Publication Number
WO 02/50285 A2(51) International Patent Classification: C12N 15/55,
9/16, C07K 14/47San Jose, CA 95124 (US); CAI, Richard, Lie [CN/US];
367 Pond Road, Bridgewater, NJ 08807 (US). FIS-
CHER, Denise, Dawn [US/US]; 17 Woodland Road,
Bernardsville, NJ 07924 (US).

(21) International Application Number: PCT/EP01/14928

(22) International Filing Date:
18 December 2001 (18.12.2001)(74) Agent: BECKER, Konrad; Novartis AG, Corporate
Intellectual Property, Patent and Trademark Department,
CH-4002 Basel (CH).

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/256,827 20 December 2000 (20.12.2000) US
60/293,089 23 May 2001 (23.05.2001) US
60/317,984 6 September 2001 (06.09.2001) US(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,
LT, LU, LV, MA, MD, MK, MN, MX, NO, NZ, OM, PH,
PL, PT, RO, RU, SE, SG, SI, SK, TJ, TM, TN, TR, TT, UA,
US, UZ, VN, YU, ZA, ZW.(71) Applicant (for all designated States except AT, US): NO-
VARTIS AG [CH/CH]; Lichstrasse 35, CH-4056 Basel
(CH).(71) Applicant (for AT only): NOVARTIS-ERFINDUNGEN
VERWALTUNGSGESELLSCHAFT M.B.H. [AT/AT];
Brunner Strasse 59, A-1230 Vienna (AT).

Published:

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

(72) Inventors: and

(75) Inventors/Applicants (for US only): COHEN, Dalia
[FR/US]; 236 East Cedar Avenue, Livingston, NJ 07039
(US). BHATIA, Umesh [US/US]; 5212 Union Avenue,For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

WO 02/50285 A2

(54) Title: HISTONE DEACETYLASE-RELATED GENE AND PROTEIN

(57) Abstract: Disclosed is an HDAC related genes and gene products. In particular, the invention relates to a protein and variants
that is highly homologous to known HDACs and referred to herein as HDAC9, nucleic acid molecules that encode such a protein,
antibodies that recognize the protein, and methods for diagnosing conditions related to abnormal HDAC9 activity or gene expression.

WO 02/50285

PCT/EP01/14928

HISTONE DEACETYLASE - RELATED GENE AND PROTEIN**FIELD OF THE INVENTION**

This invention relates to a histone deacetylase gene and gene product. In particular, the invention relates to a protein that is highly homologous to known yeast histone deacetylase 1 (hda1) class II histone deacetylases (HDACs), nucleic acid molecules that encode such a protein, antibodies that recognize the protein, and methods for diagnosing conditions related to abnormal HDAC activity, including, for example, abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, host inflammatory or immune response or psoriasis.

BACKGROUND OF THE INVENTION

Histone acetylation is a major regulatory mechanism that modulates gene expression by altering the accessibility of transcription factors to DNA. Acetylation of histones is a reversible modification of the free Σ -amino group of lysine that occurs during the assembly of nucleosomes and during DNA synthesis. Changes in histone acetylation levels also occur during transcriptional activation and silencing. Acetylation of histones is generally associated with transcriptional activity, whereas deacetylation is associated with transcriptional repression. Histone acetylation levels result from an equilibrium between competing histone acetylases and deacetylases (Emiliani, S., Fischle, W., Van Lindt, C., Al-Abed, Y., and Verdin, E., *Proc Nat. Acad. Sci., U. S. A.*, **95**, 2795-2800 (1998)).

HDACs have been shown to play an important role in the regulation of transcription. HDACs function as components of complexes that are involved in transcriptional repression. This is mediated through interactions of HDACs with multi-protein complexes and requires deacetylase activity. HDAC complexes may contain the co-repressor mSin3A (Kasten, M.M., Dorland, S., Stillman, D.J. *Mol. Cell. Biol.* **17**, 4852-4858 (1997)) and mSin3A-associated proteins (Zhang, Y., Iatni, R., Erdjument-Bromage, H., Tempst, P., Reinberg, D. *Cell* **89**, 357-364 (1997); Zhang, Y., Sun, Z.W., Iatni, R., Erdjument-Bromage, H., Tempst, P., Hampsey, M., Reinberg, D. *Mol. Cell.* **1**, 1021-1031(1998)) silencing mediators NcoR (Nagy, L., H.-Y. Kao,

WO 02/50285

PCT/EP01/14928

- D. Chakravarti, R. J. Lin, C. A. Hassig, D. E. Ayer, S. L. Schreiber, and R. M. Evans (1997) *Cell* **89**, 373-380 and SMRT (Alland, L. et al., *Nature* **387**:49-55 (1997); Heinzel, T. et al., *Nature* **387**:43-8 (1997)), transcriptional repressors Rb (Brownell, J. E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D. G., Roth, S.Y., and Allis, C.D. (1996) *Cell* **84**, 843-851), Rb-like proteins p107 (Ferreira, R., Magnaghi-Jaulin, L., Robin, P., Harel-Bellan, A., Trouche, D. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 10493-10498) and p130 (Stiegler, P., De Luca, A. Bagella, L., Giordano, A. (1998) *Cancer Res.* **389**, 187-190), Rb-associated proteins (Nicolas, E., Morales, V., Magnaghi-Jaulin, L., Harel-Bellan, A., Richard-Foy, H., Trouche, D. (2000) *J. Biol. Chem.* **275**, 9797-9804, Lai, A., Lee, J.M., Yang, W.M., DeCaprio, J.A., Kaelin, W.G. Jr., Seto, E., Branton, P.E. (1999) *Mol. Cell. Biol.* **19**, 6632-6641), Mad/Max (Laherty, C., W.-M. Yang, J.-M. Sun, J. R. Davie, E. Seto, and R. N. Eisenman. (1997) *Cell* **89**, 349-456), nuclear hormone receptors (Nagy, L., H.-Y. Kao, D. Chakravarti, R. J. Lin, C. A. Hassig, D. E. Ayer, S. L. Schreiber, and R. M. Evans. (1997) *Cell* **89**, 373-380), nucleosome remodeling factors (Xue, Y., Wong, J., Moreno, G.T., Young, M.K., Cote, J., Wang, W. (1998) *Mol. Cell.* **2**, 851-861), methyl-binding proteins (Fuks, F., Burgers, W.A., Brehm, A., Hughes-Davies, L., Kouzarides, T. (2000) *Nat. Genet.* **24**, 88-91, Nan, X., Ng, H.H., Johnson, C.A., Laherty C.D., Turner, B.M., Eisenman, R.N., Bird, A. (1998) *Nature* **393**, 386-389, Ghosh, A.K., Steele, R., Ray, R.B. (1999) *Biochem. Biophys. Res. Commun.* **260**, 405-409, Ng, H. H., Zhang, Y., Hendrich, B., Johnson, C.A., Turner, B.M., Erdjument-Bromage, H., Tempst, P., Reinberg, D., Bird, A. (1999) *Nat. Genet.* **23**, 58-61), and DNA repair machinery proteins (Yarden, R.I., Brody, L.C. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 4983-4988, Cai, R.L., Yan-Neale, Y., Cueto, M.A., Xu, H., Cohen, D. (2000) *J. Biol. Chem.* **275**, 27909-27916). Furthermore, HDAC1 has been found to bind directly to YY1 (Yang, W.-M., Inouye, C., Zeng, Y., Bearss, D., and Seto, E. (1996) *Proc. Natl. Acad. Sci.* **93**, 122845-12850) and Sp1 (Doetzlhofer, A., Rotheneder, H., Lagger, G., Koranda, M., Kurtev, V., Brosch, G., Wintersberger, E., Seiser, C. (1999) *Mol. Cell. Biol.* **19**, 5504-5511) and HDACs 4 and 5 bind to MEF2 (Grozinger, C. M., and Schreiber, S. L. (2000) *Proc. Natl. Acad. Sci.* **97**, 7835-7840). In addition, HDACs have been found together in complexes (Eilers, A.L., Billin, A.N., Liu, J., Ayer, D.E. (1999) *J Biol Chem* **274**, 32750-32756, Grozinger, C. M., and Schreiber, S. L. (2000) *Proc. Natl. Acad. Sci.* **97**, 7835-7840).

30

WO 02/50285

PCT/EP01/14928

Two distinct classes of yeast histone deacetylases have been identified based upon size and sequence. Yeast class I HDACs include Rpd3, Hos1p, and Hos2p. Class II contains yeast HDA1p. Furthermore, members of these two classes were found to form different complexes. Human HDACs have been classified based upon their similarity to yeast sequences. Class I

5 human HDACs include HDACs1-3 and 8. Class II HDACs include HDACs 4-7. The deacetylase core of class I HDACs reside in the first ~390 amino acids. Class II HDAC catalytic domains are located in the C-terminal of these peptides, with the exception of HDAC4 that contains a second catalytic domain in the N-terminus (Grozinger, C. M., Hassig, C. A., and Schreiber, S. L. (1999) *Proc. Natl. Acad. Sci. U. S. A.* 96, 4868-4873).

10

An important approach that has been used to study the function of chromatin acetylation is the use of specific inhibitors of histone deacetylase. Several classes of compounds have been identified that inhibit HDAC. Histone deacetylase inhibitors have been found to have anti-proliferative effects, including induction of G1/S and G2/M cell cycle arrest, differentiation

15 (Itazaki, H., K. Nagashima, K. Sugita, H. Yoshida, Y. Kawamura, Y. Yasuda, K. Matsumoto, K. Ishii, N. Uotani, H. Nakai, A. Terui, S. Yoshimatsu, Y. Ikenishi and Y. Nakagawa. (1990) *J. Antibiot.* 12, 1524-1532, Hoshikawa, Y., Kijima, M., Yoshida, M., and Beppu, T. (1991) *Agric. Biol. Chem.* 55, 1491-1497, Hoshikawa, Y., Kwon, H.-J., Yoshida, M., Horinouchi, S., and Beppu, T. (1994) *Exp. Cell Res.* 214, 189-197, Sugita, K., Koizumi, K., and Yoshida, H. (1992) *Cancer Res.* 52, 168-172, Yoshida, M., Y. Hoshikawa, K. Koseki, K. Mori and T. Beppu. (1990) *I. of Antibiot.* 43, 1101-106, Yoshida, M., Nomura, S., and Beppu, T. (1987) *Cancer Res.* 47, 3688-3691, and apoptosis (Medina, V., Edmonds, B., Young, G. P., James, R., Appleton, S., Zalewski, P. D. (1997) *Cancer Res.* 57, 3697-3707) of transformed and normal cells and reversal of transformation (Kwon, H. J., Owa, T., Hassig, C. A., Shimada, J., and Schreiber, S. (1998)

20 *Proc. Natl. Acad. Sci. U. S. A.* 95, 3356-3361, Kim, M.-S., Son, M.-W., Park, Y. I., and Moon, A. (2000) *Cancer Lett.* 157, 23-30). These effects, along with the presence of HDAC in complexes with fusions of unliganded retinoic acid receptors PML-RAR α and PLZF-RAR α indicate a role for HDACs in tumorigenicity (Grignani, F., De Matteis, S., Nervi, C., Tomassoni, L., Gelmetti, V., Cioce, M., Fanelli, M., Ruthardt, M., Ferrara, F. F., Zamir, I., Seiser, C.,

25 Grignani, F., Lazar, M. A., Minucci, S., Pelicci, P. G. (1998) *Nature* 391, 815-818, He, L. Z.,

WO 02/50285

PCT/EP01/14928

- Guidéz, F., Tribioli, C., Peruzzi, D., Ruthardt, M., Zelent, A., Pandolfi, P. P. (1998) *Nat. Genet.*, 18, 126-35, Lin, R.J., Nagy, L., Inoue, S., Shao, W., Miller, W. H. Jr and Evans, R. M. (1998) *Nature* 391, 811-814). Furthermore, histone deacetylase inhibitors, phenylbutyrate and trichostatin A have shown promise in the treatment of promyelocytic leukemia and several other
- 5 HDAC inhibitors are being studied and are nearing the clinic (Byrd, J.C., Shinn, C., Ravi, R., Willis, C.R., Waselenko, J.K., Flinn, I.W., Dawson, N.A., Grever, M.R. (1999) *Blood* 94, 1401-1408, Kim, Y.B., Lee, K.H., Sugita, K., Yoshida, M., Horinouchi, S. (1999) *Oncogene* 18, 2461-2470, Cohen, L.A., Amin, S., Marks, P.A., Rifkind, R.A., Desai, D., Richon, V.M. (1999) *Anticancer Res.* 19, 4999-5005). In addition, the HDAC inhibitor, butyrate was found to decrease
- 10 expression of pro-inflammatory cytokines TNF- α , TNF- β , IL-6, and IL-1- β . These effects are thought to result from inhibition of NF κ B activation (Segain JP, Raingeard de la Bletiere D, Bourreille, A., Leray V., Gérois, N., Rosales, C., Ferrier, L., Bonnet, C., Blottiere, H.M., Galmiche, J.P. (2000) Butyrate inhibits inflammatory responses through NF κ B inhibition: implications for Crohn's disease. *Gut* 47, 397-403) and its ability to inhibit histone deacetylases
- 15 (Jnan M.S., Rasoulpour, R.J., Yin, L., Hubbard, A.K., Rosenberg, D.W., Giardina, C. (2000). The luminal short-chain fatty acid butyrate modulates NF- κ B activity in a human colonic epithelial cell line. *Gastroenterology* 118, 724-34).
- 20 The discovery of the HDAC inhibitor trapoxin, made it possible to isolate the first human histone deacetylase, HDAC1, using an affinity matrix column to which a trapoxin-like molecule was bound (Taunton, J., Collins, J. L., and Schreiber, S. (1996) *J. Am. Chem. Soc.* 118, 10412-10422). Subsequently, seven other human HDAC enzyme isoforms were reported (Taunton, J., Hassig, C. A. and Schreiber, S.L. (1996). *Science* 272, 408-411, Yang, W. m., Inouye, C., Zeng,
- 25 Y., Bearss, D., and Seto, D. (1996) *Proc. Natl. Acad. Sci. U.S.A.* 93, 12845-12850, Yang, W. M., Yao, y. L., Sun, J. M., Davie, J. R., and Seto, E. (1997). *J. Biol Chem.* 272, 28001-28007, Emiliani, S., Fischle, W., Van Lint, C., Al-Abed, Y., and Verdin, E. (1998). *Proc. Natl. Acad. Sci. U.S.A.* 95, 2795-27800). These 8 HDACs have been divided into class I (HDACs 1-3 and 8 similar to the yeast gene Rpd3) and class II HDACs (4-7 similar to yeast gene hda1 (Grozinger,
- 30 C. M., Hassig, C.A., and Schreiber, S. L. (1999). *Proc. Natl. Acad. Sci. U.S.A.* 96, 4983-4988.)

WO 02/50285

PCT/EP01/14928

based on sequence homology. Here we report the isolation and characterization of a potential new HDAC, referred to herein as HDAC9, which displays sequence similarity to the hda1 class II HDACs. HDAC9 has characteristics that bridge HDAC class I and class II.

5

SUMMARY OF THE INVENTION

The present invention relates to histone deacetylases, in particular to a novel histone deacetylase HDAC9.

In a first aspect, the invention provides an isolated polypeptide comprising an amino acid sequence as set forth in SEQ ID NO:1, SEQ ID NO 5 or SEQ ID NO 6. Furthermore, the invention provides an isolated polypeptide consisting of an amino acid sequence as set forth in SEQ ID NO:1, SEQ ID NO 5 or SEQ ID NO 6. The amino acid sequence as set forth in SEQ ID NO:1, SEQ ID NO 5 or SEQ ID NO 6 shows a considerable degree of homology to that of known members of the family of HDACs. For convenience, the polypeptide consisting of the amino acid sequence as set forth in SEQ ID NO:1, SEQ ID NO 5 or SEQ ID NO 6 will be designated as histone deacetylase 9 or HDAC9. Such a polypeptide, or a fragment thereof, is expressed in various normal tissues, for example, HDAC9 was present in normal testes, stomach, spleen, small intestine, placenta, liver, kidney, colon, lung, heart, and brain, as an approximately 3 kb transcript. HDAC9 was not detected in muscle, but this lane also did not hybridize GAPDH (Figure 7). Fragments of the isolated polypeptide having an amino acid sequence as set forth in SEQ ID NO:1, SEQ ID NO 5 or SEQ ID NO 6 will comprise polypeptides comprising from about 5 to 148 amino acids, preferably from about 10 to about 143 amino acids, more preferably from about 20 to about 100 amino acids, and most preferably from about 20 to about 50 amino acids. Such fragments also form a part of the present invention. Preferably, fragments will encompass the catalytic domain, which is predicted to exist between amino acid number 1 to 390. In accordance with this aspect of the invention there are provided novel polypeptides of human origin as well as biologically, diagnostically or therapeutically useful fragments, variants and derivatives thereof, variants and derivatives of the fragments, and analogs of the foregoing.

In a second aspect, the invention provides an isolated DNA comprising a nucleotide sequence that encodes a polypeptide as mentioned above. In particular, the invention provides

WO 02/50285

PCT/EP01/14928

(1) an isolated DNA comprising the nucleotide sequence as set forth in SEQ ID NO:2; SEQ ID NO 7 or SEQ ID NO 8 (2) an isolated DNA comprising the nucleotide sequence set forth in SEQ ID NO:3; (3) an isolated DNA capable of hybridizing under high stringency conditions to the nucleotide sequence set forth in SEQ ID NO:3; and (4) an isolated DNA comprising the nucleotide sequence set forth in SEQ ID NO:4. Also provided are nucleic acid sequences comprising at least about 15 bases, preferably at least about 20 bases, more preferably a nucleic acid sequence comprising about 30 contiguous bases of SEQ ID NO:2, SEQ ID NO 7 or SEQ ID NO 8 or SEQ ID NO:3. Also within the scope of the present invention are nucleic acids that are substantially similar to the nucleic acid with the nucleotide sequence as set forth in SEQ ID NO:2, SEQ ID NO 7 or SEQ ID NO 8 or SEQ ID NO:3. In a preferred embodiment, the isolated DNA takes the form of a vector molecule comprising at least a fragment of a DNA of the present invention, in particular comprising the DNA consisting of a nucleotide sequence as set forth in SEQ ID NO:2, SEQ ID NO 7 or SEQ ID NO 8 or SEQ ID NO:3.

A third aspect of the present invention encompasses a method for the diagnosis of conditions associated with abnormal regulation of gene expression which includes, but is not limited to, conditions associated with abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, or psoriasis in a human which comprises detecting abnormal transcription of messenger RNA transcribed from the natural endogenous human gene encoding the novel polypeptide consisting of the amino acid sequence set forth in SEQ ID NO:1, SEQ ID NO 5 or SEQ ID NO 6 in an appropriate tissue or cell from a human, wherein such abnormal transcription is diagnostic of the human's affliction with such a condition. In particular, the said natural endogenous human gene encoding the novel polypeptide consisting of the amino acid sequence set forth in SEQ ID NO:1, SEQ ID NO 5 or SEQ ID NO 6 comprises the genomic nucleotide sequence set forth in SEQ ID NO:4. In one embodiment of the present invention, the diagnostic method comprises contacting a sample of said appropriate tissue or cell or contacting an isolated RNA or DNA molecule derived from that tissue or cell with an isolated nucleotide sequence of at least about 15 - 20 nucleotides in length that hybridizes under high stringency conditions with the isolated nucleotide sequence encoding the novel polypeptide having an amino acid sequence set forth in SEQ ID NOs:1, 5 or 6

30

WO 02/50285

PCT/EP01/14928

Another embodiment of the assay aspect of the invention provides a method for the diagnosis of a condition associated with abnormal HDAC9 activity in a human, which comprises measuring the level of deacetylase activity in a certain tissue or cell from a human suffering from such a condition, wherein the presence of an abnormal level of deacetylase activity, relative to
5 the level thereof in the respective tissue or cell of a human not suffering from a condition associated with abnormal HDAC activity, is diagnostic of the human's suffering from said condition.

In accordance with one embodiment of this aspect of the invention there are provided
10 anti-sense polynucleotides that can regulate transcription of the gene encoding the novel HDAC9; in another embodiment, double stranded RNA is provided that can regulate the transcription of the gene encoding the novel HDAC9.

Another aspect of the invention provides a process for producing the aforementioned
15 polypeptides, polypeptide fragments, variants and derivatives, fragments of the variants and derivatives, and analogs of the foregoing. In a preferred embodiment of this aspect of the invention there are provided methods for producing the aforementioned HDAC9 comprising culturing host cells having incorporated therein an expression vector containing an exogenously-derived nucleotide sequence encoding such a polynucleotide under conditions sufficient for
20 expression of the polypeptide in the host cell, thereby causing expression of the polypeptide, and optionally recovering the expressed polypeptide. In a preferred embodiment of this aspect of the present invention, there is provided a method for producing polypeptides comprising or consisting of an amino acid sequence as set forth in SEQ ID NOs:1, 5 or 6 which comprises culturing a host cell having incorporated therein an expression vector containing an exogenously-
25 derived polynucleotide encoding a polypeptide comprising or consisting of an amino acid sequence as set forth in SEQ ID NOs:1, 5 or 6 under conditions sufficient for expression of such a polypeptide in the host cell, thereby causing the production of an expressed polypeptide, and optionally recovering the expressed polypeptide. Preferably, in any of such methods the exogenously derived polynucleotide comprises or consists of the nucleotide sequence set forth in
30 SEQ ID NOs:2, 7 or 8 the nucleotide sequence set forth in SEQ ID NO:3, or the nucleotide

WO 02/50285

PCT/EP01/14928

sequence set forth in SEQ ID NO:4. In accordance with another aspect of the invention there are provided products, compositions, processes and methods that utilize the aforementioned polypeptides and polynucleotides for, *inter alia*, research, biological, clinical and therapeutic purposes.

5

In certain additional preferred embodiments of this aspect of the invention there is provided an antibody or a fragment thereof which specifically binds to a polypeptide that comprises the amino acid sequence set forth in SEQ ID NOs:1, 5 or 6 i.e., all HDAC9 variants. In certain particularly preferred embodiments in this regard, the antibodies are highly selective for human HDAC9 polypeptides or portions of human HDAC9 polypeptides.

10

In a further aspect, an antibody or fragment thereof is provided that binds to a fragment or portion of the amino acid sequence set forth in SEQ ID NOs:1, 5 or 6.

15

In another aspect, methods of treating a condition in a subject, wherein the condition is associated with abnormal HDAC9 gene expression, an increase or decrease in the presence of HDAC9 polypeptide in a subject, or an increase or decrease in the activity of HDAC 9 polypeptide, by the administration of an effective amount of an antibody that binds to a polypeptide with the amino acid sequence set out in SEQ ID NOs:1, 5 or 6., or a fragment or portion thereof to the subject are provided. Also provided are methods for the diagnosis of a disease or condition associated with abnormal HDAC9 gene expression or an increase or decrease in the presence of the HDAC9 in a subject, or an increase or decrease in the activity of HDAC 9 polypeptide, which comprises utilizing conventional methodologies, including, for example, the H4 histone assay that was previously described (Inokoshi, J., Katagiri, M., Arima, S., Tanaka, H., Hayashi, M., Kim, Y.-B., Furumai, R., Yoshida, M., Horinouchi, S., Omura, S. (1999) *Biochem. Biophys. Res. Com.* **256**, 372-376.).

20

In yet another aspect, the invention provides host cells which can be propagated in vitro, preferably vertebrate cells, in particular mammalian cells, or bacterial cells, which are capable upon growth in culture of producing a polypeptide that comprises the amino acid sequence set

25

30

WO 02/50285

PCT/EP01/14928

forth in SEQ ID NOs:1, 5 or 6 or fragments thereof, where the cells contain transcriptional control DNA sequences, where the transcriptional control sequences control transcription of RNA encoding a polypeptide with the amino acid sequence according to SEQ ID NOs:1, 5 or 6. or fragments thereof. This includes, but is not limited to, the propagation of HDAC9 in a
5 plasmid and the production of DNA, RNA or protein in human or insect cells or bacteria using the endogenous HDAC9 promoter or any other transcriptional control sequence.

In yet another aspect of the present invention there are provided assay methods and kits comprising the components necessary to detect above-normal expression of polynucleotides encoding a polypeptide comprising an amino acid sequence as set forth in SEQ ID NOs:1, 5 or 6.
10 , or polypeptides comprising an amino acid sequence set forth in SEQ ID NOs:1, 5 or 6. , or fragments thereof, in body tissue samples derived from a patient, such kits comprising e.g., antibodies that bind to a polypeptide comprising an amino acid sequence set forth in SEQ ID NOs:1, 5 or 6 or to fragments thereof, or oligonucleotide probes that hybridize with
15 polynucleotides of the invention. In a preferred embodiment, such kits also comprise instructions detailing the procedures by which the kit components are to be used.

In another aspect, the invention is directed to use of a polypeptide comprising an amino acid sequence set forth in SEQ ID NOs:1, 5 or 6. or fragment thereof, polynucleotide encoding
20 such a polypeptide or a fragment thereof, or antibody that binds to said polypeptide comprising an amino acid sequence set forth in SEQ ID NOs:1, 5 or 6. or a fragment thereof in the manufacture of a medicament to treat diseases associated with abnormal HDAC activity or gene expression.

25 Another aspect is directed to pharmaceutical compositions comprising a polypeptide comprising or consisting of an amino acid sequence set forth in SEQ ID NOs:1, 5 or 6. or fragment thereof, a polynucleotide encoding such a polypeptide or a fragment thereof, or antibody that binds to such a polypeptide or a fragment thereof, in conjunction with a suitable pharmaceutical carrier, excipient or diluent, for the treatment of diseases associated with
30 abnormal HDAC activity or gene expression.

WO 02/50285

PCT/EP01/14928

In another aspect, the invention is directed to methods for the identification of molecules that can bind to a polypeptide comprising an amino acid sequence set forth in SEQ ID NOs:1, 5 or 6. and/or modulate the activity of a polypeptide comprising an amino acid sequence set forth in SEQ ID NOs:1, 5 or 6. or molecules that can bind to nucleic acid sequences that modulate the transcription or translation of a polynucleotide encoding a polypeptide comprising an amino acid sequence set forth in SEQ ID NOs:1, 5 or 6. Such methods are disclosed in, e.g., U.S. Patent Nos. 5,541,070; 5,567,317; 5,593,853; 5,670,326; 5,679,582; 5,856,083; 5,858,657; 5,866,341; 5,876,946; 5,989,814; 6,010,861; 6,020,141; 6,030,779; and 6,043,024, all of which are incorporated by reference herein in their entirety. Molecules identified by such methods also fall within the scope of the present invention.

In a related aspect, the invention is directed to use of the novel HDAC9 to identify associated proteins in HDAC biologically relevant complexes. At present, the proteins that associate with HDAC9 are not known. However, these may be characterized by determining whether HDAC9 associates with proteins that have been previously shown to interact with other HDACs (see Introduction). For example, components of HDAC9 complexes may be determined using conventional methods, including co-immunoprecipitation (see Example 9).

In yet another aspect, the invention is directed to methods for the introduction of nucleic acids of the invention into one or more tissues of a subject in need of treatment with the result that one or more proteins encoded by the nucleic acids are expressed and or secreted by cells within the tissue.

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

WO 02/50285

PCT/EP01/14928

art from reading the following description and from reading the other parts of the present disclosure.

BRIEF DESCRIPTION OF THE DRAWINGS

5 Figure 1 shows the 1156 bp open reading frame that was identified using GENFAM (proprietary software) and used to search databases for the complete HDAC9 cDNA sequence. The respective ORF (SEQ ID NO:3) starts at nucleotide position no. 1 and ends at nucleotide position no. 1156.

10 Figures 2A and 2B show the full length cDNA sequence (SEQ ID NO:2) of HDAC9 and the amino acid sequence (SEQ ID NO:1), respectively. The full length cDNA sequence starts at nucleotide position no. 1 and ends at nucleotide position 2022.

15 Figure 3 shows the genomic DNA sequence in silico (AL022328) (SEQ ID NO:4), aligned with the sequence of clone 198929/HDAC9. The alignment was produced using proprietary software (Novartis Pharmaceuticals, Summit, NJ).

20 Figure 4 is a depiction of the alignment of HDAC9 predicted peptide and *S. pombe* Hda1 peptide. The query is HDAC9 peptide and the subject is *S. pombe* Hda1 peptide. The alignment was produced using Clustalw algorithm (Higgins, D.G., Thompson, J.D., Gibson, T.J. (1996) Using CLUSTAL for multiple sequence alignments. Methods Enzymol 266, 383-402).

25 Figure 5 shows the alignment of HDAC1 and HDAC9v1 and locations of the putative catalytic domain amino acids and Rb-binding domain. Catalytic domain amino acids are boxed and putative Rb domain amino acids are contained within crosshatched boxes. The alignment was produced using Clustalw algorithm (Higgins, D.G., Thompson, J.D., Gibson, T.J. (1996) Using CLUSTAL for multiple sequence alignments. Methods Enzymol 266, 383-402).

WO 02/50285

PCT/EP01/14928

Figure 6 shows the alignment of HDACs 1-9v1. The alignment was produced using Clustalw algorithm (Higgins, D.G., Thompson, J.D., Gibson, T.J. (1996) Using CLUSTAL for multiple sequence alignments. *Methods Enzymol* 266, 383-402).

5 Figure 7 shows the Northern analysis of HDAC9. (A) Northern blot analysis of the distribution of HDAC9 in normal human tissues. GAPDH was hybridized to the same blot as a control for RNA loading. (B) Northern blot analysis of HDAC9 in matched tumor and normal tissues. GAPDH was hybridized to the same blot as a control for RNA loading.

10 Figure 8 shows Real Time PCR analysis of the distribution of HDAC9 in normal human tissues and cell lines relative to 18S ribosomal RNA. RNA from the human lung carcinoma cell line, A549 was used as an internal control.

Figure 9 shows the alignment of HDAC9v1 with class II HDACs (HDACs 4,5,6, 7). The alignment was produced using Clustalw algorithm (Higgins, D.G., Thompson, J.D., Gibson, T.J. (1996) Using CLUSTAL for multiple sequence alignments. *Methods Enzymol* 266, 383-402). Catalytic domain amino acids are boxed.

Figure 10 shows the alignment of HDAC9v1 with class I HDACs (HDACs 1,2,3,8). The alignment was produced using Clustalw algorithm (Higgins, D.G., Thompson, J.D., Gibson, T.J. (1996) Using CLUSTAL for multiple sequence alignments. *Methods Enzymol* 266, 383-402). Catalytic domain amino acids are boxed.

Figure 11 There are three HDAC9 sequence variants (HDAC9v1, HDAC9v2, and HDAC9v3). HDAC9v1 and HDAC9v2 were found by searching the human EST database and HDAC9v3 was found as a predicted transcript in the Celera Sequence database. (A) shows an alignment of the 3 HDAC9 variant peptide sequences. (B) shows a schematic of class I and class II HDAC peptide sequences. Catalytic domains are in filled boxes and putative LXCXE motifs are in open boxes (C) is a schematic of the genomic structures of HDAC9v1 and HDAC9v2.

WO 02/50285

PCT/EP01/14928

Exons are shown as filled boxes and introns are shown as lines between the filled boxes. Lengths of boxes and lines represent the lengths of exons and introns.

Figure 12 shows that HDAC9 is an enzymatically active histone deacetylase. (A) HDAC9 catalytic activity is comparable to the activity of HDAC3 and HDAC4. (B) shows that HDAC1 was more efficient than HDAC3, HDAC4, and HDAC9 at deacetylating the histone substrate in this assay.

Figure 13 shows that HDAC9 is a nuclear protein and shows that HDAC9-flag is in vitro translated.

Figure 14 shows DNA and peptide sequences for HDAC9v3 and HDAC9v2.

DETAILED DESCRIPTION OF THE INVENTION

All patent applications, patents and literature references cited herein are hereby incorporated by reference in their entirety.

In practicing the present invention, many conventional techniques in molecular biology, microbiology, and recombinant DNA are used. These techniques are well known and are explained in, for example, Current Protocols in Molecular Biology, Volumes I, II, and III, 1997 (F. M. Ausubel ed.); Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.; DNA Cloning: A Practical Approach, Volumes I and II, 1985 (D. N. Glover ed.); Oligonucleotide Synthesis, 1984 (M. L. Gait ed.); Nucleic Acid Hybridization, 1985, (Hames and Higgins); Transcription and Translation, 1984 (Hames and Higgins eds.); Animal Cell Culture, 1986 (R. I. Freshney ed.); Immobilized Cells and Enzymes, 1986 (IRL Press); Perbal, 1984, A Practical Guide to Molecular Cloning; the series, Methods in Enzymology (Academic Press, Inc.); Gene Transfer Vectors for Mammalian Cells, 1987 (J. H. Miller and M. P. Calos eds., Cold Spring Harbor Laboratory); and Methods in Enzymology Vol. 154 and Vol. 155 (Wu and Grossman, and Wu, eds., respectively).

WO 02/50285

PCT/EP01/14928

The following abbreviations used throughout the disclosure are listed herein below:
histone deacetylase (HDAC), histone deacetylase-like protein (HDLP)

In its broadest sense, the term "substantially similar", when used herein with respect to a
5 nucleotide sequence, means a nucleotide sequence corresponding to a reference nucleotide
sequence, wherein the corresponding sequence encodes a polypeptide having substantially the
same structure and function as the polypeptide encoded by the reference nucleotide sequence,
e.g. where only changes in amino acids not affecting the polypeptide function occur. Desirably
the substantially similar nucleotide sequence encodes the polypeptide encoded by the reference
10 nucleotide sequence. The percentage of identity between the substantially similar nucleotide
sequence and the reference nucleotide sequence desirably is at least 80%, more desirably at least
85%, preferably at least 90%, more preferably at least 95%, still more preferably at least 99%.
Sequence comparisons are carried out using Clustalw (see, for example, Higgins, D.G. et al.
Methods Enzymol. 266:383-402 (1996)). Clustalw alignments were performed using default
15 parameters.

A nucleotide sequence "substantially similar" to reference nucleotide sequence
hybridizes to the reference nucleotide sequence in 7% sodium dodecyl sulfate (SDS), 0.5 M
NaPO₄, 1 mM EDTA at 50°C with washing in 2X SSC, 0.1% SDS at 50°C, more desirably in
20 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄, 1 mM EDTA at 50°C with washing in 1X
SSC, 0.1% SDS at 50°C, more desirably still in 7% sodium dodecyl sulfate (SDS), 0.5 M
NaPO₄, 1 mM EDTA at 50°C with washing in 0.5X SSC, 0.1% SDS at 50°C, preferably in 7%
sodium dodecyl sulfate (SDS), 0.5 M NaPO₄, 1 mM EDTA at 50°C with washing in 0.1X SSC,
0.1% SDS at 50°C, more preferably in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄, 1 mM
25 EDTA at 50°C with washing in 0.1X SSC, 0.1% SDS at 65°C, yet still encodes a functionally
equivalent gene product.

"Elevated transcription of mRNA" refers to a greater amount of messenger RNA
transcribed from the natural endogenous human gene encoding the novel polypeptide of the
30 present invention present in an appropriate tissue or cell of an individual suffering from a

WO 02/50285

PCT/EP01/14928

condition associated with abnormal HDAC9 activity than in a subject not suffering from such a disease or condition; in particular at least about twice, preferably at least about five times, more preferably at least about ten times, most preferably at least about 100 times the amount of mRNA found in corresponding tissues in humans who do not suffer from such a condition. Such

5 elevated level of mRNA may eventually lead to increased levels of protein translated from such mRNA in an individual suffering from a condition associated with abnormal cellular proliferation as compared with a healthy individual. It is also understood that "elevated transcription of mRNA" may refer to a greater amount of messenger RNA transcribed from genes the expression of which is modulated by HDAC9 either alone or in combination with other

10 molecules.

A "host cell," as used herein, refers to a prokaryotic or eukaryotic cell that contains heterologous DNA that has been introduced into the cell by any means, e.g., electroporation, calcium phosphate precipitation, microinjection, transformation, viral infection, and the like.

"Heterologous" as used herein means "of different natural origin" or represent a non-

15 natural state. For example, if a host cell is transformed with a DNA or gene derived from another organism, particularly from another species, that gene is heterologous with respect to that host cell and also with respect to descendants of the host cell which carry that gene. Similarly, heterologous refers to a nucleotide sequence derived from and inserted into the same natural, original cell type, but which is present in a non-natural state, e.g. a different copy

20 number, or under the control of different regulatory elements.

A "vector" molecule is a nucleic acid molecule into which heterologous nucleic acid may be inserted which can then be introduced into an appropriate host cell. Vectors preferably have one or more origin of replication, and one or more site into which the recombinant DNA can be inserted. Vectors often have convenient means by which cells with vectors can be selected from

25 those without, e.g., they encode drug resistance genes. Common vectors include plasmids, viral genomes, and (primarily in yeast and bacteria) "artificial chromosomes."

"Plasmids" generally are designated herein by a lower case p preceded and/or followed by capital letters and/or numbers, in accordance with standard naming conventions that are familiar to those of skill in the art. Starting plasmids disclosed herein are either commercially

30 available, publicly available on an unrestricted basis, or can be constructed from available

WO 02/50285

PCT/EP01/14928

plasmids by routine application of well known, published procedures. Many plasmids and other cloning and expression vectors that can be used in accordance with the present invention are well known and readily available to those of skill in the art. Moreover, those of skill readily may construct any number of other plasmids suitable for use in the invention. The properties,
5 construction and use of such plasmids, as well as other vectors, in the present invention will be readily apparent to those of skill from the present disclosure.

The term "isolated" means that the material is removed from its original environment (e.g., the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide or polypeptide present in a living animal is not isolated, but the same
10 polynucleotide or polypeptide, separated from some or all of the coexisting materials in the natural system, is isolated, even if subsequently reintroduced into the natural system. Such polynucleotides could be part of a vector and/or such polynucleotides or polypeptides could be part of a composition, and still be isolated in that such vector or composition is not part of its natural environment.

As used herein, the term "transcriptional control sequence" refers to DNA sequences, such as initiator sequences, enhancer sequences, and promoter sequences, which induce, repress, or otherwise control the transcription of protein encoding nucleic acid sequences to which they are operably linked.

As used herein, "human transcriptional control sequences" are any of those
20 transcriptional control sequences normally found associated with the human gene encoding the novel HDAC9 polypeptide of the present invention as it is found in the respective human chromosome. It is understood that the term may also refer to transcriptional control sequences normally found associated with human genes the expression of which is modulated by HDAC9 either alone or in combination with other molecules.

As used herein, "non-human transcriptional control sequence" is any transcriptional control sequence not found in the human genome.

The term "polypeptide" is used interchangeably herein with the terms "polypeptides" and "protein(s)".

As used herein, a "chemical derivative" of a polypeptide of the invention is a polypeptide
30 of the invention that contains additional chemical moieties not normally a part of the molecule.

WO 02/50285

PCT/EP01/14928

Such moieties may improve the molecule's solubility, absorption, biological half life, etc. The moieties may alternatively decrease the toxicity of the molecule, eliminate or attenuate any undesirable side effect of the molecule, etc. Moieties capable of mediating such effects are disclosed, for example, in Remington's Pharmaceutical Sciences, 16th ed., Mack Publishing Co., Easton, Pa. (1980).

As used herein, "HDAC9" refers to the amino acid sequences of substantially purified HDAC9 obtained from any species, particularly mammalian, including bovine, ovine, porcine, murine, equine, and preferably human, from any source, whether natural, synthetic, semi-synthetic, or recombinant.

As used herein, "HDAC activity", including "HDAC9 activity" refers to the ability of an HDAC polypeptide to deacetylate histone proteins, including ³H-labeled H4 histone peptide. Such activity may be measured according to conventional methods, for example as described in Inokoshi, J., Katagiri, M., Arima, S., Tanaka, H., Hayashi, M., Kim, Y.-B., Furumai, R., Yoshida, M., Horinouchi, S., and Omura, S. (1999) Biochem. Biophys. Res. Com. 256, 372-376. A biologically "active" protein refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule.

The term "agonist", as used herein, refers to a molecule which when bound to HDAC9, causes a change in HDAC9 which modulates the activity of HDAC9.. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules that bind to HDAC9.

The terms "antagonist" or "inhibitor" as used herein, refer to a molecule which when bound to HDAC9, blocks or modulates the biological activity of HDAC9. Antagonists and inhibitors may include proteins, nucleic acids, carbohydrates, or any other molecules, natural or synthetic that bind to HDAC9.

HDAC9 was identified using proprietary computer software called GENFAM to search for new human sequences that are related to histone deacetylases in the Celera Human Genome Database, Incyte LIFESEQ® database and the public High Throughput Genomic database. An

WO 02/50285

PCT/EP01/14928

1156 bp open reading frame (ORF) was identified and used to search a database of sequenced clones from pan-tissue and dorsal root ganglion cDNA libraries. Four clones were found to contain the ORF (M6, K10, P3, F23), two from each library. Of these clones, M6, from the pan-tissue library was determined to be the most complete cDNA as a result of sequence analysis and *in vitro* translation. BLAST (Altschul S.F. et al Nucleic Acid Res 25:3389-402 (1997)) was used to search the Genbank database using cDNA clone M6. Genomic sequence AL022328 was found to contain exons that were identical in sequence to the M6 cDNA. A Clustalw alignment of the antisense sequence of HDAC9 (2022 to 8) with genomic sequence AL022328 is shown in Figure 3. The first 7 bases of the HDAC9 predicted cDNA are not aligned, presumably because they occur following the next intron and this sequence was probably too short for the software to determine an alignment. The sequence of cDNA clone M6 was confirmed by automated DNA sequencing (ACGT, Inc., Northbrook, IL). Based upon the predicted cDNA sequence from genomic sequence AL022328, 44 bases were missing from the N-terminus of M6. This sequence was subsequently added by PCR.

The full length cDNA for HDAC9 predicts a protein of 673 amino acids. The HDAC9 cDNA sequence is 2022 base pairs in length. In order to determine the percent similarity of HDAC9 to other known HDACs, a Clustalw multiple sequence alignment was performed using complete peptide sequences for HDACs 1-9. HDAC9 is most similar in peptide sequence to human HDAC6 at 37%. The Clustalw alignment of HDAC9 with class II HDACs is shown in Figure 9. HDAC9 was also 40% similar to a yeast class II sequence hda1 from *S. pombe*. The Clustalw alignment of human HDAC9 and *S. pombe* is shown in Figure 4. HDAC9 was less similar to class I HDACs ($\leq 18\%$). The Clustalw alignment of HDAC9 to class I HDACs is shown in Figure 10. HDAC9 possesses a putative catalytic domain which encompasses approximately 317 aa (~6 to 323) based upon alignments of HDAC9 with the putative catalytic domains of all of the other known HDACs. To identify the catalytic domain of HDAC9, Clustalw alignments were performed separately using HDAC9 complete peptide and catalytic domain sequences from class I HDACs (1-3 and 8) or class II HDACs (4-7). 13 amino acids were previously shown to confer deacetylase activity, based upon inactivation by single amino acid mutations and the three dimensional structure formed by a complex of HDAC-like protein (HDLp), Zn²⁺ and HDAC inhibitors (Finnin, M. S., Donigian, J. R., Cohen, A., Richon, V. M.,

WO 02/50285

PCT/EP01/14928

Rifkind, R. a., Marks, P. A., Breslow, R., and Pavletich, N. P. (1999) Structures of a histone deacetylase homologue bound to TSA and SAHA inhibitors. *Nature* 401, 188-193). These 13 amino acids include Pro 22, His 131, His 132, Gly 140, Phe 141, Asp 166, Asp 168, His 170, Asp 173, Phe 198, Asp 258, Leu 265, and Tyr 297. 12 out of 13 of these amino acids are conserved in HDAC9. The amino acid that is not conserved is Leu 265. This hydrophobic residue forms part of the TS binding pocket and is replaced in HDAC9 with Glu at amino acid 272. Leu 265 is replaced with Met in HDAC8 and Lys in HDAC 6 domain 1. This suggests that this residue is not highly conserved and need not be identical to other HDACs. The second residue that differs from HDLP, HDAC1, and HDAC2, Asp 173 is substituted with Gln at position 177 in HDAC9, a difference that is also present in the HDAC6 catalytic domain 1. Furthermore, Asp 173 is substituted with Asn in HDACs 4,5, 6 (domain 2), and 7. This evidence suggests that these Asp173 substitutions do not affect HDAC activity.

An amino acid sequence motif was previously found to be important for the binding of HDACs 1 and 2 to retinoblastoma protein (Rb). Complexes of HDACs 1 and 2 and Rb induce repression of E2F responsive promoters (Brehm, A., Miska, E. A., McCance, D. J., Reid, J. L., Bannister, A. J., and Kouzarides, T. (1998) *Nature* 391, 597-601). An Rb-binding motif fits the sequence model LXCXE, where "X" can be any amino acid. The LXCXE domain has been found to be dispensible for growth suppression function of Rb, but is necessary for HDAC binding (Chen, T.-T. and Wang, J. Y. J. (2000) *Mol. Cell Biol.* 20, 5571-5580). The Rb-binding domain that was previously determined in HDAC1 is located from amino acid 414 to amino acid 419 and is the sequence IACEE. So far, it has not been determined whether other HDACs are capable of binding to Rb. However, HDAC 9 contains a putative Rb-binding motif, LSCIL, that aligned with HDAC1 IACEE and is located between amino acids 560 and 564. Co-immunoprecipitation of HDAC9 with Rb is one strategy that may be used to validate the function of this motif in HDAC9.

As a member of the HDAC family, HDAC9 could form biologically relevant complexes with proteins and display functions that have been described for other HDACs. For example, it is likely to be involved in the regulation of transcription as a component of complexes that are involved in transcriptional repression that is mediated through interactions of HDACs with multi-protein complexes and which requires deacetylase activity. Thus, increased activity or

WO 02/50285

PCT/EP01/14928

expression of HDAC9 may be associated with numerous pathological conditions, including but not limited to, abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, host inflammatory or immune response, or psoriasis.

5 Thus, the DNA/amino acid sequence and predicted structure of HDAC9 will be useful for designing agents (e.g. antagonists or inhibitors) useful to ameliorate conditions associated with abnormal HDAC activity. These may include, for example, antiproliferative or antiinflammatory agents either through the use of small molecules or proteins (e.g. antibodies) directed against it or associated proteins in HDAC transcription repressor complexes. In
10 addition, protein derived from the HDAC9 sequence may also be used as a therapeutic to modify host cell proliferative or inflammatory responses.

To determine the expression pattern of the novel polypeptide, a panel of mRNAs from a variety of human tissues is subjected to Northern analysis. Data indicate that HDAC9 is
15 expressed in human tissues, being detectable in brain, colon, heart, kidney, liver, placenta, small intestine, spleen, stomach and testes. Thus, HDAC9 represents a transcribed gene.

Therefore, in one aspect, the present invention relates to a novel histone deacetylase (HDAC). As outlined above, HDAC9 is clearly a member of the HDAC family since it is highly similar to other HDAC proteins in the hda1 class II HDACs. It also shares many similarities
20 with the HDAC family.

The present invention relates to an isolated polypeptide comprising the amino acid sequence set forth in SEQ ID NO:1. For example, such a polypeptide may be a fusion protein including the amino acid sequence of the novel HDAC9. In another aspect the present invention relates to an isolated polypeptide consisting of the amino acid sequence set forth in SEQ ID
25 NO:1, which is, in particular, the novel HDAC9.

The invention includes nucleic acid or nucleotide molecules, preferably DNA molecules, in particular encoding the novel HDAC9. Preferably, an isolated nucleic acid molecule, preferably a DNA molecule, of the present invention encodes a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:1 SEQ ID NO 5 or SEQ ID NO 6. Likewise
30 preferred is an isolated nucleic acid molecule, preferably a DNA molecule, encoding a

WO 02/50285

PCT/EP01/14928

polypeptide consisting of the amino acid sequence set forth in SEQ ID NO:1, SEQ ID NO 5 or SEQ ID NO 6. Such a nucleic acid or nucleotide, in particular such a DNA molecule, preferably comprises a nucleotide sequence selected from the group consisting of (1) the nucleotide sequence as set forth in SEQ ID NO:2, 7 or 8 which is the complete cDNA sequence encoding the polypeptide consisting of the amino acid sequence set forth in SEQ ID NO:1, 5 and 6, respectively, (2) the nucleotide sequence set forth in SEQ ID NO:3, which corresponds to the open reading frame of the cDNA sequence set forth in SEQ ID NO:2; (3) a nucleotide sequence capable of hybridizing under high stringency conditions to a nucleotide sequence set forth in SEQ ID NO:3; and (4) the nucleotide sequence set forth in SEQ ID NO:4, which corresponds to the endogenous genomic human DNA encoding the polypeptide consisting of the amino acid sequence set forth in SEQ ID NO:1. Such hybridization conditions may be highly stringent or less highly stringent, as described above. In instances wherein the nucleic acid molecules are deoxyoligonucleotides ("oligos"), highly stringent conditions may refer, e.g., to washing in 6X SSC/0.05% sodium pyrophosphate at 37 °C (for 14-base oligos), 48 °C (for 17-base oligos), 55 °C (for 20-base oligos), and 60 °C (for 23-base oligos). Suitable ranges of such stringency conditions for nucleic acids of varying compositions are described in Krause and Aaronson (1991), *Methods in Enzymology*, 200:546-556 in addition to Maniatis et al., cited above.

These nucleic acid molecules may act as target gene antisense molecules, useful, for example, in target gene regulation and/or as antisense primers in amplification reactions of target gene nucleic acid sequences. Further, such sequences may be used as part of ribozyme and/or triple helix sequences, also useful for target gene regulation. Still further, such molecules may be used as components of diagnostic methods whereby the presence of an allele causing a disease associated with abnormal HDAC9 expression or activity, for example, abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, host inflammatory or immune response, or psoriasis, may be detected.

The invention also encompasses (a) vectors that contain at least a fragment of any of the foregoing nucleotide sequences and/or their complements (i.e., antisense); (b) vector molecules, preferably vector molecules comprising transcriptional control sequences, in particular expression vectors, that contain any of the foregoing coding sequences operatively associated

WO 02/50285

PCT/EP01/14928

with a regulatory element that directs the expression of the coding sequences; and (c) genetically engineered host cells that contain a vector molecule as mentioned herein or at least a fragment of any of the foregoing nucleotide sequences operatively associated with a regulatory element that directs the expression of the coding sequences in the host cell. As used herein, regulatory elements include, but are not limited to, inducible and non-inducible promoters, enhancers, operators and other elements known to those skilled in the art that drive and regulate expression. Preferably, host cells can be vertebrate host cells, preferably mammalian host cells, such as human cells or rodent cells, such as CHO or BHK cells. Likewise preferred, host cells can be bacterial host cells, in particular *E.coli* cells.

Particularly preferred is a host cell, in particular of the above described type, which can be propagated in vitro and which is capable upon growth in culture of producing an HDAC9 polypeptide, in particular a polypeptide comprising or consisting of an amino acid sequence set forth in SEQ ID NO:1, wherein said cell contains some fragment or complete sequence of HDAC9 coding sequence in a construct that is controlled by one or more transcriptional control sequences that is not a transcriptional control sequence of the natural endogeneous human gene encoding said polypeptide, wherein said one or more transcriptional control sequences control transcription of a DNA encoding said polypeptide. Possible transcriptional control sequences include, but are not limited to, bacterial or viral promoter sequences.

The invention includes the complete sequence of the gene as well as fragments of any of the nucleic acid sequences disclosed herein. Fragments of the nucleic acid sequences encoding the novel HDAC9 polypeptide may be used as a hybridization probe for a cDNA library to isolate other genes which have a high sequence similarity to the HDAC9 gene or similar biological activity. Probes of this type preferably have at least about 30 bases and may contain, for example, from about 30 to about 50 bases, about 50 to about 100 bases, about 100 to about 200 bases, or more than 200 bases. The probe may also be used to identify a cDNA clone corresponding to a full length transcript and a genomic clone or clones that contain the complete HDAC9 gene including regulatory and promoter regions, exons, and introns. An example of a screen comprises isolating the coding region of the HDAC9 gene by using the known DNA sequence to synthesize an oligonucleotide probe. Labeled oligonucleotides having a sequence complementary to that of the gene of the present invention may be used to screen a library of

WO 02/50285

PCT/EP01/14928

human cDNA, genomic DNA or mRNA to determine which members of the library to which the probe hybridizes.

In addition to the gene sequences described above, homologs of such sequences, as may, for example, be present in other species, may be identified and may be readily isolated, without undue experimentation, by molecular biological techniques well known in the art. Further, there may exist genes at other genetic loci within the genome that encode proteins which have homology to one or more domains of such gene products. These genes may also be identified via similar techniques. For example, the isolated nucleotide sequence of the present invention encoding the novel HDAC9 polypeptide may be labeled and used to screen a cDNA library constructed from mRNA obtained from the organism of interest. Hybridization conditions will be of a lower stringency when the cDNA library is derived from an organism different from the type of organism from which the labeled sequence was derived. Alternatively, the labeled fragment may be used to screen a genomic library derived from the organism of interest, again, using appropriately stringent conditions. Such low stringency conditions will be well known to those of skill in the art, and will vary predictably depending on the specific organisms from which the library and the labeled sequences are derived. For guidance regarding such conditions see, for example, Sambrook et al. cited above.

Further, a previously unknown differentially expressed gene-type sequence may be isolated by performing PCR using two degenerate oligonucleotide primer pools designed on the basis of amino acid sequences within the gene of interest. The template for the reaction may be cDNA obtained by reverse transcription of mRNA prepared from human or non-human cell lines or tissue known or suspected to express a differentially expressed gene allele. The PCR product may be subcloned and sequenced to ensure that the amplified sequences represent the sequences of a differentially expressed gene-like nucleic acid sequence. The PCR fragment may then be used to isolate a complete cDNA clone by a variety of conventional methods. For example, the amplified fragment may be labeled and used to screen a bacteriophage cDNA library. Alternatively, the labeled fragment may be used to screen a genomic library.

PCR technology may also be utilized to isolate full length cDNA sequences. For example, RNA may be isolated, following standard procedures, from an appropriate cellular or tissue source. A reverse transcription reaction may be performed on the RNA using an

WO 02/50285

PCT/EP01/14928

oligonucleotide primer specific for the most 5' end of the amplified fragment for the priming of first strand synthesis. The resulting RNA/DNA hybrid may then be "tailed" with guanines using a standard terminal transferase reaction, the hybrid may be digested with RNAase H, and second strand synthesis may then be primed with a poly-C primer. Thus, cDNA sequences upstream of the amplified fragment may easily be isolated. For a review of cloning strategies which may be used, see e.g., Sambrook et al., 1989, supra.

In cases where the gene identified is the normal, or wild type, gene, this gene may be used to isolate mutant alleles of the gene. Such an isolation is preferable in processes and disorders which are known or suspected to have a genetic basis. Mutant alleles may be isolated from individuals either known or suspected to have a genotype which contributes to disease symptoms related to abnormal HDAC activity, including, but not limited to, conditions such as abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, host inflammatory or immune response, or psoriasis. Mutant alleles and mutant allele products may then be utilized in the diagnostic assay systems described below.

A cDNA of the mutant gene may be isolated, for example, by using PCR, a technique which is well known to those of skill in the art. In this case, the first cDNA strand may be synthesized by hybridizing an oligo-dT oligonucleotide to mRNA isolated from tissue known or suspected to be expressed in an individual putatively carrying the mutant allele, and by extending the new strand with reverse transcriptase. The second strand of the cDNA is then synthesized using an oligonucleotide that hybridizes specifically to the 5' end of the normal gene. Using these two primers, the product is then amplified via PCR, cloned into a suitable vector, and subjected to DNA sequence analysis through methods well known to those of skill in the art. By comparing the DNA sequence of the mutant gene to that of the normal gene, the mutation(s) responsible for the loss or alteration of function of the mutant gene product can be ascertained.

Alternatively, a genomic or cDNA library can be constructed and screened using DNA or RNA, respectively, from a tissue known to or suspected of expressing the gene of interest in an individual suspected of or known to carry the mutant allele. The normal gene or any suitable fragment thereof may then be labeled and used as a probe to identify the corresponding mutant allele in the library. The clone containing this gene may then be purified through methods routinely practiced in the art, and subjected to sequence analysis as described above.

Additionally, an expression library can be constructed utilizing DNA isolated from or cDNA synthesized from a tissue known to or suspected of expressing the gene of interest in an individual suspected of or known to carry the mutant allele. In this manner, gene products made by the putatively mutant tissue may be expressed and screened using standard antibody screening techniques in conjunction with antibodies raised against the normal gene product, as described below. (For screening techniques, see, for example, Harlow, E. and Lane, eds., 1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor Press, Cold Spring Harbor.) In cases where the mutation results in an expressed gene product with altered function (e.g., as a result of a missense mutation), a polyclonal set of antibodies are likely to cross-react with the mutant gene product. Library clones detected via their reaction with such labeled antibodies can be purified and subjected to sequence analysis as described above.

The present invention includes those proteins encoded by nucleotide sequences set forth in any of SEQ ID NOs:2, 3, 4, 7 or 8 in particular, a polypeptide that is or includes the amino acid sequence set out in SEQ ID NO:1, 5 or 6 or fragments thereof.

Furthermore, the present invention includes proteins that represent functionally equivalent gene products. Such an equivalent differentially expressed gene product may contain deletions, additions or substitutions of amino acid residues within the amino acid sequence encoded by the differentially expressed gene sequences described, above, but which result in a silent change, thus producing a functionally equivalent differentially expressed gene product. 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 involved.

For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Functionally equivalent," as utilized herein, may refer to a protein or polypeptide capable of exhibiting a substantially similar *in vivo* or *in vitro* activity as the endogenous differentially expressed gene products encoded by the differentially expressed gene sequences described above. "Functionally equivalent" may also refer to proteins or polypeptides capable of interacting with other cellular or extracellular molecules in a manner

WO 02/50285

PCT/EP01/14928

substantially similar to the way in which the corresponding portion of the endogenous differentially expressed gene product would. For example, a "functionally equivalent" peptide would be able, in an immunoassay, to diminish the binding of an antibody to the corresponding peptide (i.e., the peptide the amino acid sequence of which was modified to achieve the

5 "functionally equivalent" peptide) of the endogenous protein, or to the endogenous protein itself, where the antibody was raised against the corresponding peptide of the endogenous protein. An equimolar concentration of the functionally equivalent peptide will diminish the aforesaid binding of the corresponding peptide by at least about 5%, preferably between about 5% and 10%, more preferably between about 10% and 25%, even more preferably between about 25% and 50%, and most preferably between about 40% and 50%.

The polypeptides of the present invention may be produced by recombinant DNA technology using techniques well known in the art. Therefore, there is provided a method of producing a polypeptide of the present invention, which method comprises culturing a host cell having incorporated therein an expression vector containing an exogenously-derived

15 polynucleotide encoding a polypeptide comprising an amino acid sequence as set forth in SEQ ID NOs:1, 5 or 6 under conditions sufficient for expression of the polypeptide in the host cell, thereby causing the production of the expressed polypeptide. Optionally, said method further comprises recovering the polypeptide produced by said cell. In a preferred embodiment of such a method, said exogenously-derived polynucleotide encodes a polypeptide consisting of an amino

20 acid sequence set forth in SEQ ID NOs:1, 5 or 6. Preferably, said exogenously-derived polynucleotide comprises the nucleotide sequence as set forth in any of SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO: 7 or SEQ ID NO:8. In case of using the nucleotide sequence set forth in SEQ ID NO:3, i.e. the open reading frame, the sequence, when inserted into a vector, may be followed by one or more appropriate translation stop codons, preferably by the

25 natural endogenous stop codon TGA beginning at nucleotide 2021 in the cDNA sequence.

Thus, methods for preparing the polypeptides and peptides of the invention by expressing nucleic acid encoding respective nucleotide sequences are described herein. Methods which are well known to those skilled in the art can be used to construct expression vectors containing protein coding sequences and appropriate transcriptional/translational control signals. These

30 methods include, for example, in vitro recombinant DNA techniques, synthetic techniques and in

WO 02/50285

PCT/EP01/14928

vivo recombination/genetic recombination. See, for example, the techniques described in Sambrook et al., 1989, supra, and Ausubel et al., 1989, supra. Alternatively, RNA capable of encoding differentially expressed gene protein sequences may be chemically synthesized using, for example, synthesizers. See, for example, the techniques described in "Oligonucleotide
5 Synthesis", 1984, Gait, M. J. ed., IRL Press, Oxford, which is incorporated by reference herein in its entirety.

A variety of host-expression vector systems may be utilized to express the HDAC9 gene coding sequences of the invention. Such host-expression systems represent vehicles by which the coding sequences of interest may be produced and subsequently purified, but also represent cells
10 which may, when transformed or transfected with the appropriate nucleotide coding sequences, exhibit the HDAC9 gene protein of the invention *in situ*. These include but are not limited to microorganisms such as bacteria (e.g., *E. coli*, *B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing differentially expressed gene protein coding sequences; yeast (e.g. *Saccharomyces*, *Pichia*) transformed with
15 recombinant yeast expression vectors containing the differentially expressed gene protein coding sequences; insect cell systems infected or transfected with recombinant virus expression vectors (e.g., baculovirus) containing the differentially expressed gene protein coding sequences; plant cell systems infected with recombinant virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant vectors, including
20 plasmids, (e.g., Ti plasmid) containing protein coding sequences; or mammalian cell systems (e.g. COS, CHO, BHK, 293, 3T3) harboring recombinant expression constructs containing promoters derived from the genome of mammalian cells (e.g., metallothioneine promoter) or from mammalian viruses (e.g., the adenovirus late promoter; the vaccinia virus 7.5K promoter, or the CMV promoter).

25 Expression of the HDAC9 of the present invention by a cell from an HDAC9 encoding gene that is native to the cell can also be performed. Methods for such expression are detailed in, e.g., U.S. Patents 5,641,670; 5,733,761; 5,968,502; and 5,994,127, all of which are expressly incorporated by reference herein in their entirety. Cells that have been induced to express HDAC9 by the methods of any of U.S. Patents 5,641,670; 5,733,761; 5,968,502; and 5,994,127

WO 02/50285

PCT/EP01/14928

can be implanted into a desired tissue in a living animal in order to increase the local concentration of HDAC9 in the tissue.

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the protein being expressed. For example, when a large quantity of such a protein is to be produced, for the generation of antibodies or to screen peptide libraries, for example, vectors which direct the expression of high levels of fusion protein products that are readily purified may be desirable. In this respect, fusion proteins comprising hexahistidine tags may be used, such as EpiTag vectors including pCDNA3.1/His (Invitrogen, Carlsbad, CA). Other vectors include, but are not limited, to the *E. coli* expression vector pUR278 (Ruther et al., 1983, EMBO J. 2:1791), in which the protein coding sequence may be ligated individually into the vector in frame with the lac Z coding region so that a fusion protein is produced; pIN vectors (Inouye & Inouye, 1985, Nucleic Acids Res. 13:3101-3109; Van Heeke & Schuster, 1989, J. Biol. Chem. 264:5503-5509); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. The pGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned target gene protein can be released from the GST moiety. Fusion proteins containing Flag tags, such as 3X Flag (Sigma, St. Louis, MO) or myc tags, for example pCDNA3.1/myc-His (Invitrogen, Carlsbad, CA) may be used. These fusions allow coimmunoprecipitation and Western detection of proteins for which antibodies are not yet available.

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

WO 02/50285

PCT/EP01/14928

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

Among known bacterial promoters suitable for expression of polynucleotides and polypeptides in accordance with the present invention are the *E. coli* lacI and lacZ promoters, the T3 and T7 promoters, the T5 tac promoter, the lambda PR, PL promoters and the trp promoter. Among known eukaryotic promoters suitable in this regard are the CMV immediate early promoter, the HSV thymidine kinase promoter, the early and late SV40 promoters, the promoters of retroviral LTRs, such as those of the Rous sarcoma virus ("RSV"), and metallothionein promoters, such as the mouse metallothionein-I promoter. For example, a plasmid construct that encodes the coding region of β -Galactosidase, chloramphenicol acetyltransferase, green fluorescent protein or luciferase. This construct could be used to screen for small molecules that modulate HDAC9 transcription. Such molecules are potential therapeutics. Furthermore, an HDAC9 reporter gene could be used to examine the effects of an HDAC9 therapeutic in mammalian cells or xenografts using fluorescent reporters and imaging techniques, such as fluorescence microscopy or Biophotonic *in vivo* imaging, a technology that produces visual and quantitative measurements in real time (Xenogen, Palo Alto, CA). Changes in these reporters in normal, diseased or drug-treated tissue or cells would be indicators of changes in HDAC9 expression or activity.

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

WO 02/50285

PCT/EP01/14928

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the coding sequence of interest may be ligated to an adenovirus transcription/translation control complex, e.g., the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by in vitro or in vivo recombination. Insertion in a non-essential region of the viral genome (e.g., region E1 or E3) will result in a recombinant virus that is viable and capable of expressing the desired protein in infected hosts (e.g., See Logan & Shenk, 1984, Proc. Natl. Acad. Sci. USA 81:3655-3659). Specific initiation signals may also be required for efficient translation of inserted gene coding sequences. These signals include the ATG initiation codon and adjacent sequences. In cases where an entire gene, including its own initiation codon and adjacent sequences, is inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only a portion of the gene coding sequence is inserted, exogenous translational control signals, including, perhaps, the ATG initiation codon, must be provided. Furthermore, the initiation codon must be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements, transcription terminators, etc. (see Bittner et al., 1987, Methods in Enzymol. 153:516-544). Other common systems are based on SV40, retrovirus or adeno-associated virus. Selection of appropriate vectors and promoters for expression in a host cell is a well known procedure and the requisite techniques for expression vector construction, introduction of the vector into the host and expression in the host per se are routine skills in the art. Generally, recombinant expression vectors will include origins of replication, a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence, and a selectable marker to permit isolation of vector containing cells after exposure to the vector.

In addition, a host cell strain may be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may be important for the function of the protein. Different host cells have characteristic and specific

WO 02/50285

PCT/EP01/14928

mechanisms for the post-translational processing and modification of proteins. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells which possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. Such mammalian host cells include but are not limited to CHO, VERO, BHK, HeLa, COS, MDCK, 293, 3T3, WI38, etc. and are well known to one of skill in the art.

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

A number of selection systems may be used, including but not limited to, the herpes simplex virus thymidine kinase (Wigler, et al., 1977, Cell 11:223), hypoxanthine-guanine phosphoribosyltransferase (Szybalska & Szybalski, 1962, Proc. Natl. Acad. Sci. USA 48:2026), and adenine phosphoribosyltransferase (Lowy, et al., 1980, Cell 22:817) genes can be employed in tk⁻, hprt⁻ or aprt⁻ cells, respectively. Also, antimetabolite resistance can be used as the basis of selection for dhfr, which confers resistance to methotrexate (Wigler, et al., 1980, Natl. Acad. Sci. USA 77:3567; O'Hare, et al., 1981, Proc. Natl. Acad. Sci. USA 78:1527); gpt, which confers resistance to mycophenolic acid (Mulligan & Berg, 1981, Proc. Natl. Acad. Sci. USA 78:2072); neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin, et al., 1981, J.

WO 02/50285

PCT/EP01/14928

Mol. Biol. 150:1); and hygromycin, which confers resistance to hygromycin (Santerre, et al., 1984, Gene 30:147) genes.

An alternative fusion protein system allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht, et al., 1991, Proc. Natl. Acad. Sci. USA 88: 8972-8976). In this system, the gene of interest is subcloned into a vaccinia recombination plasmid such that the gene's open reading frame is translationally fused to an amino-terminal tag consisting of six histidine residues. Extracts from cells infected with recombinant vaccinia virus are loaded onto Ni^{2+} nitriloacetic acid-agarose columns and histidine-tagged proteins are selectively eluted with imidazole-containing buffers.

When used as a component in assay systems such as those described below, a protein of the present invention may be labeled, either directly or indirectly, to facilitate detection of a complex formed between the protein and a test substance. Any of a variety of suitable labeling systems may be used including, but not limited to, radioisotopes such as ^{125}I ; enzyme labeling systems that generate a detectable calorimetric signal or light when exposed to substrate; and fluorescent labels.

Where recombinant DNA technology is used to produce a protein of the present invention for such assay systems, it may be advantageous to engineer fusion proteins that can facilitate labeling, immobilization, detection and/or isolation

Indirect labeling involves the use of a protein, such as a labeled antibody, which specifically binds to a polypeptide of the present invention. Such antibodies include but are not limited to polyclonal, monoclonal, chimeric, single chain, Fab fragments and fragments produced by an Fab expression library.

In another embodiment, nucleic acids comprising a sequence encoding HDAC9 protein or functional derivative thereof, may be administered to promote normal biological function, for example, normal transcriptional regulation, by way of gene therapy. Gene therapy refers to therapy performed by the administration of a nucleic acid to a subject. In this embodiment of the invention, the nucleic acid produces its encoded protein that mediates a therapeutic effect by promoting normal transcriptional regulation..

WO 02/50285

PCT/EP01/14928

Any of the methods for gene therapy available in the art can be used according to the present invention. Exemplary methods are described below.

In a preferred aspect, the therapeutic comprises a HDAC9 nucleic acid that is part of an expression vector that expresses a HDAC9 protein or fragment or chimeric protein thereof in a suitable host. In particular, such a nucleic acid has a promoter operably linked to the HDAC9 coding region, said promoter being inducible or constitutive, and, optionally, tissue-specific. In another particular embodiment, a nucleic acid molecule is used in which the HDAC9 coding sequences and any other desired sequences are flanked by regions that promote homologous recombination at a desired site in the genome, thus providing for intrachromosomal expression of the HDAC9 nucleic acid (Koller and Smithies, 1989, Proc. Natl. Acad. Sci. USA 86:8932-8935; Zijlstra et al., 1989, Nature 342:435-438).

Delivery of the nucleic acid into a patient may be either direct, in which case the patient is directly exposed to the nucleic acid or nucleic acid-carrying vector, or indirect, in which case, cells are first transformed with the nucleic acid *in vitro*, then transplanted into the patient. These two approaches are known, respectively, as *in vivo* or *ex vivo* gene therapy.

In a specific embodiment, the nucleic acid is directly administered *in vivo*, where it is expressed to produce the encoded product. This can be accomplished by any of numerous methods known in the art, e.g., by constructing it as part of an appropriate nucleic acid expression vector and administering it so that it becomes intracellular, e.g., by infection using a defective or attenuated retroviral or other viral vector (see, e.g., U.S. Pat. No. 4,980,286 and others mentioned *infra*), or by direct injection of naked DNA, or by use of microparticle bombardment (e.g., a gene gun; Biolistic, Dupont), or coating with lipids or cell-surface receptors or transfecting agents, encapsulation in liposomes, microparticles, or microcapsules, or by administering it in linkage to a peptide which is known to enter the nucleus, by administering it in linkage to a ligand subject to receptor-mediated endocytosis (see e.g., U.S. Patents 5,166,320; 5,728,399; 5,874,297; and 6,030,954, all of which are incorporated by reference herein in their entirety) (which can be used to target cell types specifically expressing the receptors), etc. In another embodiment, a nucleic acid-ligand complex can be formed in which the ligand comprises a fusogenic viral peptide to disrupt endosomes, allowing the nucleic acid to avoid lysosomal degradation. In yet another embodiment, the nucleic acid can be targeted *in vivo*

WO 02/50285

PCT/EP01/14928

for cell specific uptake and expression, by targeting a specific receptor (see, e.g., PCT Publications WO 92/06180; WO 92/22635; WO92/20316; WO93/14188; and WO 93/20221). Alternatively, the nucleic acid can be introduced intracellularly and incorporated within host cell DNA for expression, by homologous recombination (see, e.g., U.S. Patents 5,413,923; 5,416,260; and 5,574,205; and Zijlstra et al., 1989, Nature 342:435-438).

5 In a specific embodiment, a viral vector that contains the HDAC9 nucleic acid is used. For example, a retroviral vector can be used (see, e.g., U.S. Patents 5,219,740; 5,604,090; and 5,834,182). These retroviral vectors have been modified to delete retroviral sequences that are not necessary for packaging of the viral genome and integration into host cell DNA. The HDAC9
10 nucleic acid to be used in gene therapy is cloned into the vector, which facilitates delivery of the gene into a patient.

Adenoviruses are other viral vectors that can be used in gene therapy. Adenoviruses are especially attractive vehicles for delivering genes to respiratory epithelia. Adenoviruses naturally infect respiratory epithelia where they cause a mild disease. Other targets for adenovirus-based
15 delivery systems are liver, the central nervous system, endothelial cells, and muscle. Adenoviruses have the advantage of being capable of infecting non-dividing cells. Methods for conducting adenovirus-based gene therapy are described in, e.g., U.S. Patents 5,824,544; 5,868,040; 5,871,722; 5,880,102; 5,882,877; 5,885,808; 5,932,210; 5,981,225; 5,994,106; 5,994,132; 5,994,134; 6,001,557; and 6,033,8843, all of which are incorporated by reference
20 herein in their entirety.

Adeno-associated virus (AAV) has also been proposed for use in gene therapy. Methods for producing and utilizing AAV are described, e.g., in U.S. Patents 5,173,414; 5,252,479; 5,552,311; 5,658,785; 5,763,416; 5,773,289; 5,843,742; 5,869,040; 5,942,496; and 5,948,675, all of which are incorporated by reference herein in their entirety.

25 Another approach to gene therapy involves transferring a gene to cells in tissue culture by such methods as electroporation, lipofection, calcium phosphate mediated transfection, or viral infection. Usually, the method of transfer includes the transfer of a selectable marker to the cells. The cells are then placed under selection to isolate those cells that have taken up and are expressing the transferred gene. Those cells are then delivered to a patient.

WO 02/50285

PCT/EP01/14928

In this embodiment, the nucleic acid is introduced into a cell prior to administration in vivo of the resulting recombinant cell. Such introduction can be carried out by any method known in the art, including but not limited to transfection, electroporation, microinjection, infection with a viral or bacteriophage vector containing the nucleic acid sequences, cell fusion, chromosome-mediated gene transfer, microcell-mediated gene transfer, spheroplast fusion, etc. Numerous techniques are known in the art for the introduction of foreign genes into cells and may be used in accordance with the present invention, provided that the necessary developmental and physiological functions of the recipient cells are not disrupted. The technique should provide for the stable transfer of the nucleic acid to the cell, so that the nucleic acid is expressible by the cell and preferably heritable and expressible by its cell progeny.

The resulting recombinant cells can be delivered to a patient by various methods known in the art. In a preferred embodiment, epithelial cells are injected, e.g., subcutaneously. In another embodiment, recombinant skin cells may be applied as a skin graft onto the patient. Recombinant blood cells (e.g., hematopoietic stem or progenitor cells) are preferably administered intravenously. The amount of cells envisioned for use depends on the desired effect, patient state, etc., and can be determined by one skilled in the art.

Cells into which a nucleic acid can be introduced for purposes of gene therapy encompass any desired, available cell type, and include but are not limited to epithelial cells, endothelial cells, keratinocytes, fibroblasts, muscle cells, hepatocytes; blood cells such as T lymphocytes, B lymphocytes, monocytes, macrophages, neutrophils, eosinophils, megakaryocytes, granulocytes; various stem or progenitor cells, in particular hematopoietic stem or progenitor cells, e.g., as obtained from bone marrow, umbilical cord blood, peripheral blood, fetal liver, etc.

In a preferred embodiment, the cell used for gene therapy is autologous to the patient.

In an embodiment in which recombinant cells are used in gene therapy, a HDAC9 nucleic acid is introduced into the cells such that it is expressible by the cells or their progeny, and the recombinant cells are then administered in vivo for therapeutic effect. In a specific embodiment, stem or progenitor cells are used. Any stem-and/or progenitor cells that can be isolated and maintained in vitro can potentially be used in accordance with this embodiment of the present invention. Such stem cells include but are not limited to hematopoietic stem cells (HSC), stem cells of epithelial tissues such as the skin and the lining of the gut, embryonic heart muscle cells,

WO 02/50285

PCT/EP01/14928

liver stem cells (see, e.g., WO 94/08598), and neural stem cells (Stemple and Anderson, 1992, Cell 71:973-985).

Epithelial stem cells (ESCs) or keratinocytes can be obtained from tissues such as the skin and the lining of the gut by known procedures (Rheinwald, 1980, Meth. Cell Bio. 21A:229).
5 In stratified epithelial tissue such as the skin, renewal occurs by mitosis of stem cells within the germinal layer, the layer closest to the basal lamina. Stem cells within the lining of the gut provide for a rapid renewal rate of this tissue. ESCs or keratinocytes obtained from the skin or lining of the gut of a patient or donor can be grown in tissue culture (Pittelkow and Scott, 1986, Mayo Clinic Proc. 61:771). If the ESCs are provided by a donor, a method for suppression of
10 host versus graft reactivity (e.g., irradiation, drug or antibody administration to promote moderate immunosuppression) can also be used.

With respect to hematopoietic stem cells (HSC), any technique which provides for the isolation, propagation, and maintenance in vitro of HSC can be used in this embodiment of the invention. Techniques by which this may be accomplished include (a) the isolation and
15 establishment of HSC cultures from bone marrow cells isolated from the future host, or a donor, or (b) the use of previously established long-term HSC cultures, which may be allogeneic or xenogeneic. Non-autologous HSC are used preferably in conjunction with a method of suppressing transplantation immune reactions of the future host/patient. In a particular embodiment of the present invention, human bone marrow cells can be obtained from the
20 posterior iliac crest by needle aspiration (see, e.g., Kodo et al., 1984, J. Clin. Invest. 73:1377-1384). In a preferred embodiment of the present invention, the HSCs can be made highly enriched or in substantially pure form. This enrichment can be accomplished before, during, or after long-term culturing, and can be done by any techniques known in the art. Long-term cultures of bone marrow cells can be established and maintained by using, for example, modified
25 Dexter cell culture techniques (Dexter et al., 1977, J. Cell Physiol. 91:335) or Witlock-Witte culture techniques (Witlock and Witte, 1982, Proc. Natl. Acad. Sci. USA 79:3608-3612).

In a specific embodiment, the nucleic acid to be introduced for purposes of gene therapy comprises an inducible promoter operably linked to the coding region, such that expression of the nucleic acid is controllable by controlling the presence or absence of the appropriate inducer
30 of transcription.

WO 02/50285

PCT/EP01/14928

A further embodiment of the present invention relates to a purified antibody or a fragment thereof which specifically binds to a polypeptide that comprises the amino acid sequence set forth in SEQ ID NOs:1, 5 or 6 or to a fragment of said polypeptide. A preferred embodiment relates to a fragment of such an antibody, which fragment is an Fab or F(ab')₂ fragment. In particular, the antibody can be a polyclonal antibody or a monoclonal antibody.

Described herein are methods for the production of antibodies capable of specifically recognizing one or more differentially expressed gene epitopes. Such antibodies may include, but are not limited to polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab')₂ fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above. Such antibodies may be used, for example, in the detection of a fingerprint, target, gene in a biological sample, or, alternatively, as a method for the inhibition of abnormal target gene activity. Thus, such antibodies may be utilized as part of disease treatment methods, and/or may be used as part of diagnostic techniques whereby patients may be tested for abnormal levels of the HDAC9 polypeptide, or for the presence of abnormal forms of the HDAC9 polypeptide.

For the production of antibodies to the HDAC9 polypeptide, various host animals may be immunized by injection with the HDAC9 polypeptide, or a portion thereof. Such host animals may include but are not limited to rabbits, mice, and rats, to name but a few. Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and Corynebacterium parvum.

Polyclonal antibodies are heterogeneous populations of antibody molecules derived from the sera of animals immunized with an antigen, such as target gene product, or an antigenic functional derivative thereof. For the production of polyclonal antibodies, host animals such as those described above, may be immunized by injection with the HDAC9 polypeptide, or a portion thereof, supplemented with adjuvants as also described above.

WO 02/50285

PCT/EP01/14928

Monoclonal antibodies, which are homogeneous populations of antibodies to a particular antigen, may be obtained by 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 of Kohler and Milstein, (1975, Nature 256:495-497; and U.S. Pat. No. 4,376,110), the human B-cell hybridoma technique (Kosbor et al., 1983, Immunology Today 4:72; Cole et al., 1983, Proc. Natl. Acad. Sci. USA 80:2026-2030), and the EBV-hybridoma technique (Cole et al., 1985, Monoclonal Antibodies And Cancer Therapy, Alan R. Liss, Inc., pp. 77-96). Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the mAb of this invention may be cultivated in vitro or in vivo. Production of high titers of mAbs in vivo makes this the presently preferred method of production.

In addition, techniques developed for the production of "chimeric antibodies" (Morrison et al., 1984, Proc. Natl. Acad. Sci., 81:6851-6855; Neuberger et al., 1984, Nature, 312:604-608; Takeda et al., 1985, Nature, 314:452-454) by splicing the genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable or hypervariable region derived from a murine mAb and a human immunoglobulin constant region.

Alternatively, techniques described for the production of single chain antibodies (U.S. Pat. No. 4,946,778; Bird, 1988, Science 242:423-426; Huston et al., 1988, Proc. Natl. Acad. Sci. USA 85:5879-5883; and Ward et al., 1989, Nature 334:544-546) can be adapted to produce differentially expressed gene-single chain antibodies. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide.

Most preferably, techniques useful for the production of "humanized antibodies" can be adapted to produce antibodies to the polypeptides, fragments, derivatives, and functional equivalents disclosed herein. Such techniques are disclosed in U.S. Patent Nos. 5,932,448; 5,693,762; 5,693,761; 5,585,089; 5,530,101; 5,910,771; 5,569,825; 5,625,126; 5,633,425; 5,789,650; 5,545,580; 5,661,016; and 5,770,429, the disclosures of all of which are incorporated by reference herein in their entirety.

WO 02/50285

PCT/EP01/14928

Antibody fragments that recognize specific epitopes may be generated by known techniques. For example, such fragments include but are not limited to: the F(ab')₂ fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the F(ab')₂ fragments. Alternatively, Fab expression libraries may be constructed (Huse et al., 1989, Science, 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity.

An antibody of the present invention can be preferably used in a method for the diagnosis of a condition associated with abnormal HDAC9 expression or activity, for example, abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, host inflammatory or immune response, or psoriasis, in a human which comprises: measuring the amount of a polypeptide comprising the amino acid sequence set forth in SEQ ID NOs:1, 5 or 6, or fragments thereof, in an appropriate tissue or cell from a human suffering from a condition associated with abnormal HDAC9 activity, wherein the presence of an elevated amount of said polypeptide or fragments thereof, relative to the amount of said polypeptide or fragments thereof in the respective tissue from a human not suffering from a condition associated with abnormal HDAC9 activity is diagnostic of said human's suffering from such condition. Such a method forms a further embodiment of the present invention. Preferably, said detecting step comprises contacting said appropriate tissue or cell with an antibody which specifically binds to a polypeptide that comprises the amino acid sequence set forth in SEQ ID NOs:1, 5 or 6 or a fragment thereof and detecting specific binding of said antibody with a polypeptide in said appropriate tissue or cell, wherein detection of specific binding to a polypeptide indicates the presence of a polypeptide that comprises the amino acid sequence set forth in SEQ ID NOs:1, 5 or 6 or a fragment thereof.

Particularly preferred, for ease of detection, is the sandwich assay, of which a number of variations exist, all of which are intended to be encompassed by the present invention.

For example, in a typical forward assay, unlabeled antibody is immobilized on a solid substrate and the sample to be tested brought into contact with the bound molecule. After a suitable period of incubation, for a period of time sufficient to allow formation of an antibody-antigen binary complex. At this point, a second antibody, labeled with a reporter molecule capable of inducing a detectable signal, is then added and incubated, allowing time sufficient for the formation of a ternary complex of antibody-antigen-labeled antibody. Any unreacted material

WO 02/50285

PCT/EP01/14928

is washed away, and the presence of the antigen is determined by observation of a signal, or may be quantitated by comparing with a control sample containing known amounts of antigen.

Variations on the forward assay include the simultaneous assay, in which both sample and antibody are added simultaneously to the bound antibody, or a reverse assay in which the labeled antibody and sample to be tested are first combined, incubated and added to the unlabeled surface bound antibody. These techniques are well known to those skilled in the art, and the possibility of minor variations will be readily apparent. As used herein, "sandwich assay" is intended to encompass all variations on the basic two-site technique. For the immunoassays of the present invention, the only limiting factor is that the labeled antibody be an antibody which is specific for the HDAC9 polypeptide or a fragment thereof.

The most commonly used reporter molecules in this type of assay are either enzymes, fluorophore- or radionuclide-containing molecules. In the case of an enzyme immunoassay an enzyme is conjugated to the second antibody, usually by means of glutaraldehyde or periodate. As will be readily recognized, however, a wide variety of different ligation techniques exist, which are well-known to the skilled artisan. Commonly used enzymes include horseradish peroxidase, glucose oxidase, beta-galactosidase and alkaline phosphatase, among others. The substrates to be used with the specific enzymes are generally chosen for the production, upon hydrolysis by the corresponding enzyme, of a detectable color change. For example, p-nitrophenyl phosphate is suitable for use with alkaline phosphatase conjugates; for peroxidase conjugates, 1,2-phenylenediamine or toluidine are commonly used. It is also possible to employ fluorogenic substrates, which yield a fluorescent product rather than the chromogenic substrates noted above. A solution containing the appropriate substrate is then added to the tertiary complex. The substrate reacts with the enzyme linked to the second antibody, giving a qualitative visual signal, which may be further quantitated, usually spectrophotometrically, to give an evaluation of the amount of HDAC9 which is present in the serum sample.

Alternately, fluorescent compounds, such as fluorescein and rhodamine, may be chemically coupled to antibodies without altering their binding capacity. When activated by illumination with light of a particular wavelength, the fluorochrome-labeled antibody absorbs the light energy, inducing a state of excitability in the molecule, followed by emission of the light at a characteristic longer wavelength. The emission appears as a characteristic color visually

WO 02/50285

PCT/EP01/14928

detectable with a light microscope. Immunofluorescence and EIA techniques are both very well established in the art and are particularly preferred for the present method. However, other reporter molecules, such as radioisotopes, chemiluminescent or bioluminescent molecules may also be employed. It will be readily apparent to the skilled artisan how to vary the procedure to suit the required use.

This invention also relates to the use of polynucleotides of the present invention as diagnostic reagents. In particular, the invention relates to a method for the diagnosis of a condition associated with abnormal HDAC9 expression or activity, for example, abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, host inflammatory or immune response, or psoriasis in a human which comprises: detecting elevated transcription of messenger RNA transcribed from the natural endogenous human gene encoding the polypeptide consisting of an amino acid sequence set forth in SEQ ID NOs:1, 5 or 6 in an appropriate tissue or cell from a human, wherein said elevated transcription is diagnostic of said human's suffering from the condition associated with abnormal HDAC9 expression or activity. In particular, said natural endogenous human gene comprises the nucleotide sequence set forth in SEQ ID NO:4, 7 or 8. In a preferred embodiment such a method comprises contacting a sample of said appropriate tissue or cell or contacting an isolated RNA or DNA molecule derived from that tissue or cell with an isolated nucleotide sequence of at least about 20 nucleotides in length that hybridizes under high stringency conditions with the isolated nucleotide sequence encoding a polypeptide consisting of an amino acid sequence set forth in SEQ ID NOs:1, 5 or 6.

Detection of a mutated form of the gene characterized by the polynucleotide of SEQ ID NO:4 7 or 8 which is associated with a dysfunction will provide a diagnostic tool that can add to, or define, a diagnosis of a disease, or susceptibility to a disease, which results from under-expression, over-expression or altered spatial or temporal expression of the gene. Individuals carrying mutations in the gene may be detected at the DNA level by a variety of techniques.

Nucleic acids, in particular mRNA, for diagnosis may be obtained from a subject's cells, such as from blood, urine, saliva, tissue biopsy or autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR or other amplification techniques prior to analysis. RNA or cDNA may also be used in similar fashion. Deletions and insertions can be detected by a change in size of the amplified product in

WO 02/50285

PCT/EP01/14928

comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to labeled nucleotide sequences encoding the HDAC9 polypeptide of the present invention. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase digestion or by differences in melting temperatures. DNA sequence differences may also be detected by alterations in electrophoretic mobility of DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing (e.g., Myers et al., Science (1985) 230:1242). Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (see Cotton et al., Proc Natl Acad Sci USA (1985) 85: 4397-4401). In another embodiment, an array of oligonucleotides probes comprising nucleotide sequence encoding the HDAC9 polypeptide of the present invention or fragments of such a nucleotide sequence can be constructed to conduct efficient screening of e.g., genetic mutations. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability (see for example: M. Chee et al., Science, Vol 274, pp 610-613 (1996)).

The diagnostic assays offer a process for diagnosing or determining a susceptibility to disease through detection of mutation in the HDAC9 gene by the methods described. In addition, such diseases may be diagnosed by methods comprising determining from a sample derived from a subject an abnormally decreased or increased level of polypeptide or mRNA. Decreased or increased expression can be measured at the RNA level using any of the methods well known in the art for the quantitation of polynucleotides, such as, for example, nucleic acid amplification, for instance PCR, RT-PCR, RNase protection, Northern blotting and other hybridization methods. Assay techniques that can be used to determine levels of a protein, such as a polypeptide of the present invention, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays.

Thus in another aspect, the present invention relates to a diagnostic kit which comprises:
(a) a polynucleotide of the present invention, preferably the nucleotide sequence of SEQ ID NO:2, 3, 4, 7 or 8 or a fragment thereof;
(b) a nucleotide sequence complementary to that of (a);

WO 02/50285

PCT/EP01/14928

(c) a polypeptide of the present invention, preferably the polypeptide of SEQ ID NOs:1, 5 or 6 or a fragment thereof; or

(d) an antibody to a polypeptide of the present invention, preferably to the polypeptide of SEQ ID NOs:1, 5 or 6.

5 It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component. Such a kit will be of use in diagnosing a disease or susceptibility to a disease, particularly to a disease or condition associated with abnormal HDAC9 expression or activity, for example, abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, host inflammatory or immune response, or psoriasis.

10 The nucleotide sequences of the present invention are also valuable for chromosome localization. The sequence is specifically targeted to, and can hybridize with, a particular location on an individual human chromosome. The mapping of relevant sequences to chromosomes according to the present invention is an important first step in correlating those
15 sequences with gene associated disease. Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found in, for example, V. McKusick, Mendelian Inheritance in Man (available on-line through Johns Hopkins University Welch Medical Library). The relationship between genes and diseases that have been mapped to the same
20 chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes).

The differences in the cDNA or genomic sequence between affected and unaffected individuals can also be determined. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative
25 agent of the disease.

An additional embodiment of the invention relates to the administration of a pharmaceutical composition, in conjunction with a pharmaceutically acceptable carrier, excipient or diluent, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of HDAC9, antibodies to that polypeptide, mimetics, agonists, antagonists, or
30 inhibitors of HDAC9 function. The compositions may be administered alone or in combination

WO 02/50285

PCT/EP01/14928

with at least one other agent, such as stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs or hormones.

5 In addition, any of the therapeutic proteins, antagonists, antibodies, agonists, antisense sequences or vectors described above 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
10 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. Antagonists and agonists of HDAC9 may be made using methods which are generally known in the art.

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

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which
20 facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, Pa.).

Pharmaceutical compositions for oral administration can be formulated using
25 pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combination of active
30 compounds with solid excipient, optionally grinding a resulting mixture, and processing the

WO 02/50285

PCT/EP01/14928

mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums
5 including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel,
10 polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or
15 sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in
20 aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles
25 include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

WO 02/50285

PCT/EP01/14928

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, 5 dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized 10 powder which may contain any or all of the following: 1-50 mM histidine, 0.1%-2% sucrose, and 2-7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of the HDAC9, such labeling would include amount, frequency, and method of administration.

15 Pharmaceutical 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.

For any compound, the therapeutically effective dose can be estimated initially either in 20 cell culture assays, e.g., of neoplastic cells, or in animal models, usually mice, rabbits, dogs, or pigs. The 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 25 HDAC9 or fragments thereof, antibodies of HDAC9, agonists, antagonists or inhibitors of HDAC9, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic 30 index, and it can be expressed as the ratio, LD50/ED50. Pharmaceutical compositions which

WO 02/50285

PCT/EP01/14928

exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, 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 that requires 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, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or once every two weeks depending on half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, 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. Pharmaceutical formulations suitable for oral administration of proteins are described, e.g., in U.S. Patents 5,008,114; 5,505,962; 5,641,515; 5,681,811; 5,700,486; 5,766,633; 5,792,451; 5,853,748; 5,972,387; 5,976,569; and 6,051,561.

The following Examples illustrate the present invention, without in any way limiting the scope thereof.

EXAMPLES

Example 1: Identification of a novel HDAC related human DNA sequence using bioinformatics
HDAC9 was identified using computer software for the identification of new members of gene families based on a strategy to find maximal evolutionary links among known HDAC family members by first searching the non-redundant amino acid database, followed by searching less

WO 02/50285

PCT/EP01/14928

diverse databases such as the Celera Human Genome Database (CHGD), public High Throughput Genomic (HTG) database and the Incyte LIFESEQ™ database. Smith-Waterman (Pearson W. R. Comparison of methods for searching protein sequence databases. *Protein Sci* (1995) 4,1145-60) and Hidden Markov Models (probability models derived from diversity of amino acids at every position (Eddy S. R. Hidden Markov models. *Curr Opin Struct Biol* (1996) 6, 361-5) were performed. An 1156 bp open reading frame (ORF) was identified and used to search a database of sequenced clones from pan-tissue and dorsal root ganglion cDNA libraries.

Example 2: Construction of pan-tissue and dorsal root ganglion cDNA libraries

- 10 Pan-tissue and dorsal root ganglion cDNA libraries are prepared from polyA+ RNA. Total RNA was extracted from a pooled sample of 31 human tissues or dorsal root ganglia and isolated using TRIZOL reagent according to manufacturer's instructions (Life Technologies, Rockville, MD). mRNA is isolated using Polytract mRNA Isolation System III according to manufacturer's instructions (Promega, Madison, WI). Total RNA is hybridized to a biotinylated-oligo (dt) probe.
- 15 The oligo (dt)-mRNA hybrids are captured on streptavidin magnetosphere particles and eluted in RNase-free H₂O. 3 ul of biotinylated-oligo(dt) probe (50 pmol/ul) and 13 ul of 20X SSC is added to 60-150 ug of RNA that is heated to 65°C in RNase free water. This mixture is incubated at room temperature until it is completely cooled. Streptavidin-paramagnetic particles (beads) are resuspended and washed 3 times in 0.5X SSC and then resuspended in 0.5X SSC. The RNA-oligo(dt) hybrids from the previous step are added to these beads. To release the poly-A RNA from the beads, the beads are resuspended in RNase-free water and magnetically captured and then the eluate from the beads is ethanol precipitated. First and second strand cDNA synthesis is performed using a modified procedure from Life Technologies (D'Alessio, J. M., Gruber, C.E., Cain, C. R., and Noon, M. C. (1990) *Focus* 12, 47). First strand synthesis is performed by
- 25 incubating 1-5 ug of RNA that is heated to 60°C in 1X 1st strand buffer (Life Technologies)/6 mM DTT/600 nM dNTPs/2 units anti-Rnase. This mixture is incubated at 40°C for 2 min, then Superscript II reverse transcriptase (RT) and 1 ul of Display Thermo RT terminator mix is added and the mixture is incubated at 40°C for 1 h, followed by incubation at 60°C for 10 min. Second strand synthesis is performed in 1x second strand buffer (Life Technologies) in DEPC-H₂O/66
- 30 nM/1 ul E.coli DNA ligase/4 ul E. coli DNA polymerase I/1 ul E. coli Rnase H. This mixture is

WO 02/50285

PCT/EP01/14928

incubated at 10⁰C for 10 min and then at 16⁰C for 2h. To this mixture, 2 ul of T4 DNA polymerase is added and incubation is continued at 16⁰C for 5 min. The reaction is stopped with 10 ul of 0.5M EDTA, extracted with phenol/chloroform/isoamyl alcohol and then ethanol precipitated. Sal I and Not I adaptors are added to the 5' ends of the cDNAs by ligation for directional cloning using conventional methodology. The cDNAs are then passed through a size fractionation column to retrieve cDNAs that are >500 bp in length according to manufacturers instructions (Life Technologies, Rockville, MD). cDNAs are ligated to Sal I/Not I digested Gateway compatible pCMV-Sport6 vector (Life Technologies, Rockville, MD) using conventional methods. Competent DH10B cells (Life Technologies, Rockville, MD) are transformed with the resulting library using conventional methods. Semi-solid amplification of the libraries is performed according to the manufacturer's instructions (Life Technologies, Rockville, MD).

Example 3: Preparation of full length cDNA encoding the novel HDAC9 consisting of SEQ ID NO.1, 5 or 6: An 1156 base pair ORF was used to search a database of sequenced clones from pan-tissue and dorsal root ganglion cDNA libraries using BLAST. Four clones were found to contain the ORF (M6, K10, P3, F23), two from each library. Of these clones M6 from the pan-tissue library was determined to be the most complete, but missing approximately 44 bp from the N-terminus. A protein slightly smaller than that predicted for the complete cDNA was observed by *in vitro* translation. The result that proteins were observed by *in vitro* translation of the incomplete cDNA, suggests possibility of alternate translation initiation sites within HDAC9. Specifically, sequencing of HDAC9 in pCMVSPORT6 was performed using an automated ABI Sequencer (ACGT, Northbrook, IL). PCR was performed using conditions listed in the ABI Prism BigDye™ Terminator Cycle Sequencing Ready Reaction Kit manual and are as follows: denaturation at 96°C for 30 seconds, annealing at 50° C for 15 seconds, extension at 60°C for 4 minutes, for a total of 25 cycles. Each round of sequencing provided between 200 and 600 bp of sequence. PCR primers for 1st round sequencing were 5'-ATTTAGGTGACACTATAG -3' (Sp6, sense) and 5'-TAATACGACTCACTATAGGG -3' (T7, antisense). Results of sequencing using Sp6 primer are as follows. Bolded sequence is pCMVSPORT6 vector sequence.

CT**gg**tACCGGTCGGGAATCCCGGGATATCGTCGACCCACGCGTCCG/GGCTGCT

WO 02/50285

PCT/EP01/14928

CCCGGCCGAAGCCCCGAGTGCAGATCGAGCGTCCTGAGCGCCTGACCGCAGCCCT
 GGATCGCCTGCGGCAGCGCGGCCTGGAACAGAGGTGTCTGCGGTTGTACGCCGCG
 AGGCCTCGGAAGAGGAGCTGGGCTGGTGCACAGCCCAGAGTATGTATCCCTGGTC
 AGGGAGACCCAGTCTCTAGGCAAGGAGGAGCTGAGGCGCTGTCCGGACAGTTCGA
 5 CGCCATCTACTTCCACCCGAGTACCTTTACTGCGCGCGGCTGGCCGAGGGGCTGG
 ACTGCAGCTGGTGGACGCTGTGCTCACTGGAGCTGTGCA:AAATGGGCTTGGCCTGG
 TGAGGCTCCCGGCACCATGGCCAGAGGCGGCTGCAACGGGTTCTGCGTGTTC
 ACAACGTGGCCATAGCAGCTGCACATGCCAAGCAGAAACACGGGCTACACAGGATC
 CTCGCTGTGGACTGGGGATGTGCACCATGGCAGGGGGATCCAGTATCTCTTTGAAG
 10 GATGACCCACAGCTCCTTTACTTCTCCTGGCACCGCTATGAGCATTGGGCGCCTTCT
 GGCCTTTCTGCGAGAGTCAAGATGAgACGCATGGGGGGCGGGGACAGGGCCTCGGC
 TTCCTGTCACTGCCCTGACCAAGTTgGGGGAATGGGGAACGCTGACTTACGTG
 GCTGGCCTTCTTGACCTTGCTGGTTCCA_cTGGCCTTTTGGAGTTTGACCTGA_gCTGG
 GTGCTTGGT_cTC_gGCAGGGATTTGACT_{cage}CaT_cgGGACCCTGAgGGGGCAA. Results
 15 of sequencing using the T7 primer were:
 TCAAGCCACCAGGTGAGGATGGCACTGCAACATCTTCCACTGAGGCTCCAGCTGCCC
 TCTCAGGTACATCAGGGCCTGGACGTCTCTGGGAGGCCACAGAGGAAGGGCCTA
 GGCTAGGAGGTGCCTCTCCATTACACACCGGGCCAGGATCCCTGCTAGCTGGGGTG
 TGGAGTTCTCTCCAGGAGGGCCAGGACTCGGCCCTGCCAGCCCCGAAGCATTG
 20 CAGCCAGGAGTGCAGCGTGGGGGCCCTGCAGGCCATGGCCAGGCCCCAGCGCCACC
 AGCACCAGGTACAGGCTGGAAGCCATAGGCCAGGGGCAGC_aCCAAGCCCAAGATGCA
 GCTCAGGAAACCACCGTCACTACTGGCAGTGGCGTGGAGACATGGAACATGGA[T
 AGGGCAG_cCGCCTCCTTGCCCTGATGTTACGCCACAGACT_cCTCCCCTCATGGGCGA
 AGTCTGGAGGCCGGTCCA_gCTGT_{ta}GGCCACGCACAG_{gt}CTCTGGGCTCC_{gt}GGGACA
 25 gGCCT:TTTNGAAAAGAtATTaAGGGTGGGTGTGA_{acagg}GCTGGAATGGCTGGTATA_c
 AcTGTTAcCTGCCATT. 2nd and 3rd round sequencing primers are designed to prime sequence
 obtained from the previous round of sequencing. 2nd round primers are 5'-GTCATCA
 CTGGCAGTGGCGTG -3' (HUF7392, antisense) and 5'-TGGACTGCAGCTGGTGG -3' (DF-2,
 sense). Results of sequencing using the DF-2 primer were: CTGGcAAATGGGCTTGCCCTGG
 30 TGAGGCCTCCCGGCACCATGGCCAGAGGGCGGCTGCCAACGGGTTCTGCGTGTTC

WO 02/50285

PCT/EP01/14928

AACAACTGGCCATAGCAGCTGCACATGCCAAGCAGAAACACGGGCTACACAGGAT
 CCTCGTCGTGGACTGGGATGTGCACCATGGCCAGGGGATCCAGTATCTCTTTGAGGA
 TGACCCACAGCGTCCTTTACTTCTCTGGCACCGCTATGAGCATGGGCGCTTCTGGCCT
 TTCTTGCAGAGTCAGATGCAGACGCAGTGGGGCGGGACAGGGCCTCGGCTTCAC
 5 TGTCAACTGCCCTGGAACAGGTTGGGATGGGAAACGCTGACTACGTGGCTGCCTT
 cTGACCTGCTGCTCCCACTGGCCTTTGAGTTTGACCTGAGCTGGTGTGCTCTCG
 GCAGGATTGACTCAGCCATCGGGACCTGAGGGGCAAATGCAGGCCACGCCAGA
 GTGCTTCGCCACCTCACACAGCTGCTGCAGGTGCTGGCCGGCGCCGGGTCTGTGC
 CGTGTGGAGGGCGGCTACCACTGGAGTCACTGGCGGAGTCAGTGTGCATGACAG
 10 TACAGACGCTGCTGGGTGACCCGGcCCCACCCCTGTCAGGGCCAATGGCGCCATGTC
 AGAGTGCCCTAgAgTCATTCAgAGTGCCCGTGTGCCAGGcCCCGCACTGGAAAgAgG
 CTTCAgCAGCAAgATGTGACCGcGTGCGCATGAACCCCA. Sequencing results for the
 HUF7392 primer were: TGtaTAGGGcAGCCGCCTCCTTGCC
 CCTGATGTTCAGCCACAGACTCCTCCCGTCATGGGCGAGG
 15 TCTGAGGCGCTCCAGCTGTCCAGGGCCACGCACAGCAGCCTCTGGGCTCCGTG
 GGACAGGCCTCTCCGAACAGCCACATCCAGGGTGGCTGCTGCAGCAGAGGCTGGAG
 TGGCTGCTATACCACTGTTACCTGCCCATCCAGCATCCCATCTAAGAGGTACAGGA
 GCTTCCCAAGTGCAGTGAGGGCCTCCTCCGGGCCAGGACTCGTGTGGCTGGCCC
 AGGCTTCTGTCTCCTCCCTCAGGGCTGACGCTTCTGTTGGATGACGTCAGGGGGCAG
 20 AACCAATGTGATATCCGGCGTTGTCAAGGGCAACAGCGGTGCGGACAGAGGGTGCG
 GGGCAGAGGCACgGCTGGTCCAgGAGGGAGCTCGGTGCAGATGCAGcTGCCTTACAC
 ACTGgACCCCCAGGCAGCAGAGGTGGAGGCCTCCCTCTGGGGAGTG. 3rd round
 sequencing primers were 5'-AACAGCGGTG C GGACAGA -3' (HUF2A, antisense) and 5'-
 CTGGAGTCACTGGCGGAG -3' (DF3A, sense). Results of sequencing using DF3A primer
 25 were: AgcaCAGA cGCTgCTGGGTGACCCGGCCACCCCTG
 TCAGGGCCAATGGCGCCATGTCAGAGTGCCCTAGAGTCCATCCAGAGTGcCCGTGCT
 GCCCAGGCCCGCACTGGAAGAGCCTCCAGCAGCAAGATGTGACCGCTGTGCCGAT
 GAGCCCCAGCAGCCACTCCCCAGAGGGGAGGCCCTCCACCTCTGCTGCTGGGGGTC
 CAGTGTGTAAAGCAGCTGCATCTGCACCGAGCTCCCTCCTGGACCAGCCGTGCCTCT
 30 GCCCCGACCCCTCTGTCCGCACCGCTGTTGCCCTGACAACGCCGGATATCACATTGG

WO 02/50285

PCT/EP01/14928

TTCTGCCCCCTGACGTCATCCAACAGGAAGCGTCAGCCCTGAGGGAGGAGACAGAA
GCCTGGGCCAGGCCACACGAGTCCCTGGCCCGGGAGGAGGCCCTcACTGcACTTGGG
AAGCTCCTGTACCTcTTAgATGGGATGCTGGATGGGCAGGTGAACAgTGGTATA.

Results of sequencing using HUF2A primer were: TgcaCGGATGGTCCAGGAGGGAGCTCG

5 GTGCAAATGCAGCTGCCTTACACACTGGACCCCAAGGCAGCAgAGGTGGAGGCCTC
CCCTcTGGGGAGTGGCTGCTGGGGCTCATCGGCACAGCGGTACATCTTGCTGCTGG
AGGCTCTTCCAGTGCGGGGCCTGGGCAGCACGGGCACCTCTGGATGGACTCTAGGGC
ACTCTGACATGGCGCCATTGGCCCTGACAGGGGTGGGGCCGGGTACCCAGCAGCG
TCTGTACTGTATGCACACTGACTCCGCCAGTGACTCCAGGTGGTAGCCGCCCTCCA
10 GCACGGCACAgACCCGGCCGGCCGACACCTGCAGCAGCTGTGTGAGGTGGGCG
AAGCACTCTGGCGTGGCCTGCATTTGCCCTCAGGGTCCCGATGGCTTGAGTCAAA
TCCTGCCGAGACCAGCACCAAGCTCAGGGTCAAACCTCAAAGGCCAGTGGGAGCAGCA
GGTGCAAGGAAGGCAGCCACgTATCAGCGTTTCCATCCCAACCTGgTCCAGGGGCA
GGTTGAACAGTGAAGCCGAGGGCCCTTGTCCCgCCCCACCTTGCGTCTGCATcGA
15 CTCTCGCAGGAAAGGCCAAgAAGCgCCCATgCTATTTT. The overlapping sequence from
the combined sense and antisense sequencing was reconstructed to give the complete cDNA
sequence of HDAC9. See Figure 2A.

BLAST is used to search the Genbank database using cDNA clone M6 as the query to
identify a genomic sequence containing M6 cDNA sequence. The results of this search identified
20 a genomic sequence AL022328 that was found to contain exons that were identical in sequence
to the M6 cDNA. The sequence of cDNA clone M6 was confirmed by automated DNA
sequencing (ACGT, Inc. Northbrook, IL). See Figure 2A.

The remaining 44 bp of N-terminal sequence was added by PCR using the nested sense
strand primers 5'-GCGGTCGACGCCACCATGGGGACCGCGCTTGTGTACCATGAGGAC
25 ATG-3' and 5'-GTGTACCATGAGGACATGACGGCCACCGGCTGCTCTGGGACGACC
CCGAGTGC-3' and the 3' primer 5'-GAACCAATGTGATATCCGGCGTTG-3'. The 5' primer
added a kozak sequence and a SalI site for cloning and the 3' primer sequence overlaps the
EcoRV site in HDAC9. PCR was performed using a step-cycle file for amplification using 1
cycle of 94°C for 30 seconds, 68°C for 30 seconds, and 72°C for 1 minute, followed by 20 cycles
30 of 94°C for 30 seconds and 72°C for 1 minute.

WO 02/50285

PCT/EP01/14928

Example 3 HDAC9 sequence variants

Three variants of the HDAC9 sequence, HDAC9v1, HDAC9v2, and HDAC9v3 were found. HDAC9v1 is the original sequence found and described above. HDAC9v2 was found in the human dorsal root ganglion cDNA library and in AL022328 genomic sequence. HDAC9v3 is a predicted transcript that lacks a stop codon that was found in the Celera human genomic database. HDAC9v1 contains 20 exons and HDAC9v2 has 20 exons. Comparison of the peptide sequences of HDAC9 variants demonstrated that HDAC9v1 and HDAC9v2 were identical up to exon 17, but diverge after this exon. HDAC9v2 has an extended intron between exon 17 and 18 and an extended exon 18 that contains HDAC9v1 exon 19, but lacks 20, as a result of a single nucleotide insertion at nucleotide 446. This insertion frame shifts the sequence and shortens the peptide by 11 amino acids (Fig 11A). Compared to HDAC9v1 and HDAC9v2, HDAC9v3 has an internal deletion of amino acids 219 through 240 and diverges in its C-terminal beginning at amino acid 486. HDAC9 is the first HDAC enzyme for which sequence variants have been reported. HDAC9v1 is the sequence variant that is characterized, unless otherwise noted.

Example: 4 Identification of HDAC-associated sequence motifs.

The M6 clone was analyzed for the presence of motifs that would indicate an HDAC catalytic domain and a binding site for Rb and Rb-like proteins. HDACs are characterized by the presence of a catalytic domain with conserved amino acids. Most of the HDACs that have been identified to date have one catalytic domain, with the exception of HDAC6 that has two domains. N-terminal catalytic domains have been associated with class I HDACs, while C-terminal catalytic domains are associated with class II HDACs. An N-terminal catalytic domain was found in HDAC9 based upon PFAM prediction and alignment with the catalytic domains of

WO 02/50285

PCT/EP01/14928

other HDACs. A set of conserved amino acids were previously shown to be critical for HDAC activity and provide the critical contacts for HDAC inhibitor, TSA, based upon single amino acid mutations in HDAC1 and the three dimensional structure formed by a complex of an HDAC-like protein (HDLP), Zn^{2+} and HDAC inhibitor TSA (Hassig CA, Tong JK, Fleischer TC, Owa T, Grable PG, Ayer DE, Schreiber SL. (1998) *Proc Natl Acad Sci U S A*. 95, 3519-3524; Finnin, M. S., Donigian, J. R., Cohen, A., Richon, V. M., Rifkind, R. a., Marks, P. A., Breslow, R., and Pavletich, N. P. (1999) Structures of a histone deacetylase homologue bound to TSA and SAHA inhibitors. *Nature* 401, 188-193). A bacterial protein with similarities in sequence and enzymatic activity to human HDACs and the only class I HDAC-like structure elucidated, HDLP was used as an HDAC template. Many of these conserved amino acids with a few exceptions were found in HDAC9 (Table 4). Alignments of HDAC peptide sequences indicated that the hydrophobic residue Leu 265 that forms part of the binding pocket in HDLP is replaced with Glu at amino acid 272 in HDAC9. Similarly, Leu 265 is also replaced with Met in HDAC8 and with Lys in HDAC6 domain 1. Furthermore, Asp 173 in HDLP is substituted with Gln at position 177 in HDAC9, a difference that was also found in the HDAC6 catalytic domain 1. This Asp is substituted with Asn in HDAC4, HDAC5, HDAC6 domain 2, and HDAC7. HDAC1-8 have been shown to be catalytically active, hence the amino acid substitutions in these proteins have no enzymatic consequences.

HDAC9 is similar in sequence to class I and class II HDACs. HDACs have been classified by their sequence similarity with yeast HDACs Rpd3, Hda1, and Sir2 and by catalytic domain location. Alignment of the peptide sequences of HDAC9, yeast HDACs Rpd3, Hda1, Hda1 subfamily member from fission yeast, cryptic loci regulator 3 (Clr3), and Sir2 determined

WO 02/50285

PCT/EP01/14928

that HDAC9 had the highest sequence similarity with Ctr3 (Table 1). However, the sequence similarity is not high enough to categorize HDAC9.

Alignment of human HDACs 1-9 and Sir 1-7 peptide sequences demonstrated that HDAC9 was most similar to class II human HDAC6 (Table 2). Alignment of class I and class II HDAC catalytic domains with HDAC9 catalytic domains demonstrated that HDAC6 catalytic domain 1 has the most sequence similarity with HDAC9 (Table 3).

In order to compare the locations of catalytic domains in HDACs, PFAM predictions were made of the catalytic domains in HDAC peptides (Fig. 11B). The location of HDAC9 catalytic domain was at the N-terminus, similar to class I HDACs, and was estimated as spanning the amino acid sequence from amino acid 4 to 323. In addition, the average length of class I HDACs is 443 amino acids, while the average length of class II HDACs is 1069 amino acids. The 673 amino acid HDAC9 peptide is between the average sizes of class I and class II HDACs (Fig. 11B).

Table 1.

HDAC Class	HDAC Isoform	%Similarity to HDAC9
Class I	Rpd3	16
Class II	Hda1	18
	Ctr3	23
Class III	Sir2	5

Table 2.

HDAC Class	HDAC Isoform	% Similarity to HDAC9
Class I	HDAC1	14
	HDAC2	15
	HDAC3	15
	HDAC8	22
Class II	HDAC4	21
	HDAC5	19
	HDAC6	37
	HDAC7	20

WO 02/50285

PCT/EP01/14928

Class III	Sir1	5
	Sir2	7
	Sir3	11
	Sir4	4
	Sir5	8
	Sir6	10
	Sir7	15

Table 3.

HDAC Class	HDAC Isoform	% Similarity to HDAC9
Class I	HDAC1	20
	HDAC2	20
	HDAC3	20
	HDAC8	19
Class II	HDAC4	39
	HDAC5	38
	HDAC6-1	55
	HDAC6-2	53
	HDAC7	40

- 5 The protein product of the retinoblastoma protein (Rb) gene is a transcriptional regulator that controls DNA synthesis, the cell cycle, differentiation and apoptosis and plays a tissue-specific role normal development. Rb complexes with the transcription factor E2F, an interaction that is regulated by phosphorylation. Mutations in Rb lead to a hereditary form of cancer of the retina, retinoblastoma. Mutations have also been found in a number of mesenchymal and
- 10 epithelial cancers. Mutations that affect regulators of Rb phosphorylation including, cyclin D1, cdk4, and p16 have been found in many cancers. Therefore, Rb function is thought to play a critical role in tumorigenesis (Sellers, W.R., Kaelin, W.G. Jr. (1997) *J. Clin. Oncol.* 15, 3301-3312, DiCiommo, D., Gallie, B.L., Bremner, R.(2000) *Semin. Cancer Biol.* 10, 255-269). An Rb-binding motif was previously defined as the amino acid sequence LXCXE, where "X" can be
- 15 any amino acid (Chen, T.-T. and Wang, J. Y. J. (2000) *Mol. Cell Biol.* 20, 5571-5580). The LXCXE domain in HDAC1 was found to be dispensible for growth suppression function of Rb, but necessary for HDAC binding to Rb. Two putative Rb-binding motifs were found in HDAC9 (Fig. 11A, green boxes). LLCVA is located between amino acids 510 and 515, and LSCIL located between amino acids 560 and 564. Both are present in HDAC9v1 and HDAC9v2.

WO 02/50285

PCT/EP01/14928

Table 4.

HDAC Isoform	Amino acids in the catalytic domains of HDAC isoforms															
	Pro	Tyr	His	His	Gly	Phe	Asp	Asp	His	Asp	Phe	Asp	Leu	Asp	Leu	Tyr
HDLP	22	91	131	132	140	141	166	168	170	173	198	258	265	297		
HDAC1	Pro	Glu	His	His	Gly	Phe	Asp	Asp	His	Asp	Phe	Asp	Leu	Asp	Leu	Tyr
	29	98	139	140	148	149	174	176	178	181	205	264	271	303		
HDAC2	Pro	Glu	His	His	Gly	Phe	Asp	Asp	His	Asp	Phe	Asp	Leu	Asp	Leu	Tyr
	30	99	140	141	149	150	175	177	179	182	206	265	272	304		
HDAC3	Pro	Asp	His	His	Gly	Phe	Asp	Asp	His	Asp	Phe	Asp	Leu	Asp	Leu	Tyr
	23	92	134	135	143	144	167	168	171	174	199	259	266	298		
HDAC4	Pro	Trp	His	His	Gly	Phe	Asp	Asp	His	Asn	Phe	Asp	Leu	Asp	Leu	His
	592	762	802	803	900	901	838	839	843	846	870	934	943	976		
HDAC5	Pro	Trp	His	His	Gly	Phe	Asp	Asp	His	Asn	Phe	Asp	Leu	Asp	Leu	His
	705	793	832	833	841	842	868	869	873	876	900	964	973	1006		
HDAC6-1	Pro	Tyr	His	His	Gly	Tyr	Asp	Asp	His	Gln	Phe	Asp	Lys	Asp	Lys	Tyr
	106	175	215	216	224	225	251	252	255	258	283	346	353	386		
HDAC6-2	Pro	Tyr	His	His	Gly	Phe	Asp	Asp	His	Asn	Phe	Asp	Leu	Asp	Leu	Tyr
	501	570	593	594	602	603	647	648	651	654	679	742	749	782		
HDAC7	Pro	Tyr	His	His	Gly	Phe	Asp	Asp	His	Asn	Phe	Asp	Leu	Asp	Leu	His
	502	589	629	630	638	639	668	669	673	676	700	764	773	806		
HDAC8	N/A	Tyr	His	His	Gly	Phe	Asp	Asp	His	Asp	Phe	Asp	Met	Asp	Met	Tyr
	100	141	142	150	151	176	177	180	183	208	267	274	306			
HDAC9	Pro	Tyr	His	His	Gly	Phe	Asp	Asp	His	Gln	Phe	Asp	Glu	Asp	Glu	Tyr
	21	94	134	135	143	144	170	173	174	177	205	265	272	305		

Non-conserved amino acids (bold text). No alignment (N/A).

WO 02/50285

PCT/EP01/14928

Example 5: mRNA distribution of HDAC9 in normal tissues

mRNA distribution of HDAC9 in normal tissues is investigated using Northern analysis. Probes are prepared by ³²P-labeling a 750 bp EcorV/NotI HDAC9 fragment using Redi-Prime random nucleotide labelling kit according to manufacturer's instructions (Amersham, Piscataway, NJ). A Northern blot containing polyA+ RNA from 12 normal tissues (Origene Technologies, Rockville, MD) and an array of matched tumor *versus* normal cDNAs (Clontech, Palo Alto, CA) are probed with the [³²P]-labeled 750 bp EcorV/NotI HDAC9 fragment and washed under high stringency conditions (68°C). Hybridized blots are washed two times for 15 min at 68°C in 2 X SSC /0.1% SDS followed by two 30 min washes in 0.1 X SSC/0.1% SDS at 68°C. The blot is exposed to film with an intensifying screen for 18 hr. Results indicate that an approximately 3.0 Kb HDAC9 mRNA is detected in brain, colon, heart kidney, liver, lung, placenta, small intestine, spleen, stomach and testes. HDAC9 message was not detected in muscle, but GAPDH was also not detected. See Figure 7.

15 Analogous computer techniques using BLAST (Altschul, S.F. 1993, 1990 refs) are used to search for identical or related molecules in nucleotide databases such as GenBank or the LIFESEQ™ database. The basis of the search is the product score which is defined as:

$$\% \text{ sequence identity} \times \% \text{ maximum BLAST score}$$

20 100

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1-2% error; and at 70, the match will be exact. Homologous molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

25 The results of Northern analysis are reported as a list of libraries in which the transcript encoding HDAC9 occurs. Abundance and percent abundance are also reported. Abundance directly reflects the number of times a particular transcript is represented in a cDNA library, and

WO 02/50285

PCT/EP01/14928

percent abundance is abundance divided by the total number of sequences examined in the cDNA library.

In this case, electronic Northern analysis of LIFESEQ™ database (Incyte Pharmaceuticals, Inc. Palo Alto, Calif) indicates tissue distribution of the HDAC9 sequence as seen in Table 5. These results are reported as a list of cDNA libraries in which the transcript encoding HDAC9 occurs. The presence of HDAC9 in 20 libraries from different tissue-specific and mixed tissue sources indicates that HDAC9, like other HDAC family members may be found as an expressed gene in a wide range of tissues. This result is supported by the Northern hybridization of an HDAC9 probe to mRNAs from 12 normal tissues (see Figure 7).

Table 5. Tissue distribution determined electronically from LIFESEQ™ database.

Tissue Category
Cardiovascular System
Connective Tissue
Digestive System
Embryonic Structures
Endocrine System
Exocrine Glands
Genitalia, Female
Genitalia, Male
Germ Cells
Hemic and Immune System
Liver
Musculoskeletal System
Nervous System
Pancreas
Respiratory System
Sense Organs
Skin
Stomatognathic System
Unclassified/Mixed
Urinary Tract

Example 6: Real time PCR survey of HDAC9 distribution in human normal tissues and cell lines.

Real Time PCR. Total RNA from cultured cell lines was isolated with the Rneasy 96 kit according to the manufacturers protocol (Qiagen, Valencia CA). RNA from human tissues was purchased (Clontech Inc, Palo Alto, CA) and the tissue sources are listed in table 6 below.

WO 02/50285

PCT/EP01/14928

Table 6. Tissue sources of RNA for real time PCR analysis

Tissue	Sex of donor	Age range of donor (yrs.)	Number of samples pooled
Brain 1	M	57	1
Brain 2	F	16&36	2
Cerebellum	M	64	1
Spinal cord	M/F	17-72	31
Fetal brain	M/F	20-23 wks	8
Trachea	M/F	17-70	84
Liver 1	M	27	1
Liver 2	M/F	15&35	2
Fetal liver	?	15-24 wks	?
Stomach	M/F	23-61	15
Pancreas	M/F	17-69	18
Colon	M	35&50	2
Intestine	M/F	25&30	2
Kidney	M/F	24-55	8
Bone marrow	M/F	18-68	24
Spleen	M	22-60	7
Thymus	M	6-45	9
Thyroid	M/F	10-46	4
Adrenal gland	M	32-50	6
Salivary gland	M/F	13-78	43
Mammary gland	F	23-47	8
Skeletal muscle	M/F	23-56	10
Testis	M	28-64	25
Prostate 1	M	26-64	23
Prostate 2	M	14-60	10
Placenta	F	22-41	15

Numbers following tissues represent separate samples from the same tissue type: Male (M), Female (F)

- 5 Human cell lines, H1299 human lung carcinoma, T24 bladder carcinoma, SJRH30 muscle rhabdomyosarcoma, SJSA-1 osteosarcoma, human fibroblasts, and A549 human lung carcinoma, were obtained from American Type Tissue Culture Collection. Total RNA was isolated from

WO 02/50285

PCT/EP01/14928

human cell lines using RNA easy kit according to the manufacturers instructions (Qiagen, Valencia, CA). RNAs were quantified using RT-PCR on an ABI Prism Sequence Detection System. The primers used for detection of HDAC9 were forward primer 5'-

GGATCCAGTATCTCTT TGAGGATGAC-3', reverse primer 5'-

- 5 AGAAGCGCCCATGCTCATA-3', and Taqman probe 5'-AGCGTCCTTTACT
TCTCCTGGCACCG-3'. The Taqman Reaction System (Eorogentec, Belgium) was used with 10
ng total RNA in a 25 µl reaction in the proportions indicated by the manufacturer but
supplemented with 0.25 U/µl reverse transcriptase (MultiScribe ABI, Perkin Elmer, Branchburg
NJ) and 0.08 U/µl RNaseOUT RNase inhibitor (Life Technologies, Gaithersburg, MD). The
10 reverse reaction was initiated with a 5 min incubation at 48 °C for the reverse transcription of the
mRNA followed by a 10 min incubation at 95 °C to inactivate the reverse transcriptase and
simultaneously activate the 'hot-start' thermostable DNA polymerase. This was followed by 50
cycles of a two-step PCR reaction with alternating 15 sec at 95 °C and 60 sec at 60 °C.
Computations were performed using ABI sequence detection software (version 1.6.3). The RT-
15 PCR assays were standardized with cRNAs transcribed in vitro with the T7 RNA polymerase
reaction using the Maxiscript kit (AMBION Inc., Austin, TX) according to the manufacturers
protocol. The RT-PCR assays were standardized with a dilution series of total RNA isolated
from A549 lung tumor cells. Parallel to the RT-PCR, the total amount of RNA in each reaction
was quantitated in a fluorometric assay using the RiboGreen kit (Molecular Probes Inc., address)
20 according to the manufacturers instructions, using mammalian ribosomal RNA provided with the
kit as standard.

Real time PCR was also used to survey the distribution and levels of HDAC9 in tissues
and tumor cell lines, relative to the levels of 18S ribosomal RNA . RNA from the human A549

WO 02/50285

PCT/EP01/14928

lung carcinoma cell line was arbitrarily chosen as an internal control for the levels of total RNA in the samples. The levels of HDAC9 and 18S rRNA in A549 cells were set at 100 % and the levels of HDAC9 and 18S rRNA in other tissues and cell lines were measured as a percent of the level of these genes in A549 RNA. The levels of 18S ribosomal RNA ranged between 82% and 126% of the A549 internal control in all of the RNA samples, suggesting that there were similar amounts of RNA in the analyzed tissue samples. HDAC9 was detected at varying levels by real time PCR in a wide range of tissues (Fig. 8), confirming the Northern blot analysis (Fig. 7). In normal tissues, HDAC9 was detected at the highest levels in fetal brain (894%), cerebellum (538%), and thymus (589%). In tumor cell lines, HDAC9 was detected at the highest levels in SJRH30 cells (850%) (Fig. 8). These results suggest that HDAC9 is differentially expressed in some tissues at the RNA level.

Example 7: HDAC Enzyme Assay

Preparation of HDAC9-flag. A flag epitope tag sequence was added to the 3' end of HDAC9v1 by PCR. The PCR primers were 5'-ACGCCGGATATCACATTGGT TCTGC-3' and 5'-GCGGAATTCTTATTATTATCATCATCATCTTTATAATCCCC GTCGACAGCCACCAGGTGAGGATGGCA -3'. The flag-tagged HDAC9v1 was reconstructed using the EcoRV site in the 1st primer and subcloned into the XbaI and EcoRI sites of human expression vector pCDNA3.1(-) (Invitrogen, Carlsbad, CA).

HDAC activity assay. HDAC activity assays are performed as previously described (Emiliani, S., Fischle, W., Van Lint, C., Al-Abed, Y., and Verdin, E. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2795-2800). 5x10⁶ 293 cells grown to 50% confluency in 100 mm dishes are transfected with 30 ug of C-terminally flag-tagged HDAC1, HDAC3, HDAC4, HDAC6, HDAC7, or HDAC9 using Geneporter transfection kit according to the manufacturers

WO 02/50285

PCT/EP01/14928

instructions. The cell culture medium is changed 5 h after transfection. 48 h after transfection cells are washed in cold PBS and scraped into 1 ml of IP buffer (50mM Tris-HCl pH 7.5, 120mM NaCl, 0.5mM EDTA, 0.5% NP-40) and incubated on a rocker for 20 min. Cellular debris is pelleted in a centrifuge at 14K for 20 min. The supernatant is precleared for 1 h with protein G beads (Pharmacia Biotech) in IP buffer. Immunoprecipitations are performed by incubating the precleared supernatant with either α -FLAG M2 agarose affinity gel (Sigma) for 2 h at 4°C or anti-HDAC2 (Santa Cruz) for 1 h followed by incubation with protein G beads for 1 h at 4°C. The beads are then washed three times for 5 min in IP buffer and then washed three times in high salt IP buffer (50mM Tris-HCl pH 7.5, 1000 mM NaCl, 0.5mM EDTA, 0.5% NP-40) at 4°C. IPs are then washed two times for 2 min in 1ml of HD-buffer (10mM Tris-HCl pH 8.0, 10mM NaCl, 10% glycerol). When trapoxin inhibition is determined Ips are incubated with 0.3, 3, 30 and 300 nM TPX in HD-buffer for 20 min. Supernatants are incubated with 100000 cpm substrate ($[^3\text{H}]\text{-Ac(H41-24) SGRGKGGKGLGKGGAKRHRKVLRD}$, in vitro/chemically acetylated using BOP-chemistry) in 30 μl HD-buffer or TPX in HD-buffer, resuspending the sepharose by gently tapping the tube and shaking in an Eppendorf 5436 Thermomixer at full speed at 37°C for 2h. 170 μl HD-buffer and 50 μl stop-mix (1M HCl, 0.16M HAc) are added, vortexed for 15' min, 600 μl ethylacetate is then added and vortexed for 45 minutes, then centrifuged at 14000g for 7 minutes. 540 μl of the organic (upper) phase is then counted in 5 ml scintillation liquid using conventional techniques.

HDAC9 is catalytically active. *In vitro* histone deacetylase assays using immunoprecipitated HDAC9 and an ^3H -acetylated histone H4 peptide as substrate were performed to determine whether HDAC9 was catalytically active and to compare the activity of HDAC9 to known catalytically active HDAC1, HDAC3, and HDAC4. An HDAC-related protein

WO 02/50285

PCT/EP01/14928

that lacks catalytic activity, HDRP/MITR/HDACC was used as a negative control (Zhou, X., Richon, V.M., Rifkind, R.A., Marks, P.A. (2000) Identification of a transcriptional repressor related to the noncatalytic domain of histone deacetylases 4 and 5. *Proc Natl Acad Sci U S A* 97, 1056-61). These results demonstrated that HDAC9 could deacetylate the histone peptide
5 substrate at a level that was equivalent to HDAC3 and HDAC4 (Fig. 12A), while HDAC1 was more effective in this assay (Fig. 12B).

Example 8 HDAC9 expression and cellular localization

HDAC9 is expressed *in vitro* using 1 ug of the M6 clone, 2 ul of ³⁵S-Methionine and Sp6 TNT Quick Coupled Transcription/Translation System according to manufacturer instructions.

10 (Promega, Madison, WI). Proteins are electrophoresed on a SDS-PAGE gel according to conventional methods and visualized by a Storm phosphorimager. The complete HDAC9 sequence molecular weight is estimated *in silico* as 72 kda using VectorNTI Suite software (Informax, North Bethesda, MD). A doublet was observed on a 10% SDS-PAGE gel. Doublets have also been observed when HDAC1 is translated *in vitro*. These doublets suggest that there is
15 potentially a second translation initiation site. Furthermore, these results suggest that HDAC9 is an expressed gene. See Figure 13.

1X10⁵ Cos7 cells are plated onto chamber slides. Cells are transfected on the slides with 2 ug of flag epitope-tagged HDAC9 or a cytoplasmically expressed protein (Ena-flag) using Geneporter2 in serum free medium according to the manufacturers instructions. The cell culture
20 medium is changed 24 h after transfection. 48 h after transfection, cells are washed three times with PBS, fixed for 15 min. in 5% formaldehyde, washed two times in PBS, and blocked for 30 minutes at room temperature in 10% fetal calf serum (Sigma) in PBS with 0.5% Triton-X-100 to permeablize the cells. The cells are washed again two times in PBS and then incubated with 25

WO 02/50285

PCT/EP01/14928

mg/ml anti-Flag-FITC conjugate for 1 hour. The stained cells are washed with PBS and photographed using fluorescence microscopy.

HDAC9 is a nuclear protein. The translated HDAC9 peptide sequence predicts a 72 Kda protein and this was confirmed by *in vitro* translation (Fig. 13A). In order to determine the cellular localization of HDAC9, flag epitope-tagged HDAC9, Enabled (Ena) or pCMV4flag were transfected into Cos7 and 293 cells or cells were mock transfected without plasmid. The flag epitope was detected by fluorescence immunocytochemistry 48 h after transfection (Fig 13B). Ena is a cytoskeleton-associated cytoplasmic protein substrate of Abl tyrosine kinase that transduces the axon-repulsive function of the Roundabout receptor during axon guidance (Gertler FB, Comer AR, Juang JL, Ahern SM, Clark MJ, Liebl EC, Hoffmann FM. (1995) enabled, a dosage-sensitive suppressor of mutations in the Drosophila Abl tyrosine kinase, encodes an Abl substrate with SH3 domain-binding properties. Genes Dev. 9, 521-533. Bashaw GJ, Kidd T, Murray D, Pawson T, Goodman CS. (2000) Repulsive axon guidance: Abelson and Enabled play opposing roles downstream of the roundabout receptor. Cell. 101, 703-715). As expected, Ena was detected in the cytoplasm, whereas HDAC9 was detected in the nuclei of these cells. The detection of HDAC9 in the nuclei of both Cos7 and 293 cells suggested that HDAC9 was predominantly a nuclear protein.

Example 9: Identification of associated proteins in HDAC complexes

Transfection. 1×10^7 Cos7 cells are transfected with 10 ug of either C-terminally flag epitope-tagged HDAC1, HDAC2, HDAC3, HDAC4, HDAC6, HDAC7, or HDAC9 in pCDNA3.1 expression vector or Flag vector or buffer (Mock) as transfection controls. by electroporation using a Gene Pulser II instrument (Biorad, Hercules CA) set at 0.3Kv/ 500 uF.

WO 02/50285

PCT/EP01/14928

Immunoprecipitation. Immunoprecipitations are performed as described (Grozinger, C. M., Hassig, C. A., and Schreiber, S. L. 1999. Proc. Natl. Acad. Sci. USA 96, 4868-4873). Whole cell extracts are prepared 48h after transfection by scraping cells into JLB buffer (50 mM Tris-HCL, pH 8, 150 mM NaCl, 10% glycerol, 0.5% Triton-X-100) containing complete protease inhibitor cocktail (Boehringer-Mannheim). Lysis is continued at 4°C for 10 min. and then cellular debris is pelleted by centrifugation at 14K for 5 minutes. Supernatants are pre-cleared with Sepharose A/G-plus agarose beads (Santa Cruz). Recombinant proteins are immunoprecipitated from pre-cleared supernatant by incubation with α -FLAG M2 agarose affinity gel (Sigma) for 2 h at 4°C or anti-HDAC1 (Santa Cruz, Santa Cruz, CA) for 1 h at 4°C, followed by incubation with Sepharose A/G beads. For Western blot analysis, the beads are washed with MSWB buffer (50 mM Tris-HCl, pH 8, 150 mM NaCl, 1 mM EDTA, 0.1% NP-40) and the proteins are separated by SDS/PAGE. Western blots are probed with anti-flag M2 (Sigma), HDAC1 (Santa Cruz), anti-HDAC2 (Santa Cruz), anti-HDAC6 (Santa Cruz), anti-Rb (PharMingen), or anti-mSin3A (Transduction Labs, Lexington, KY)

HDAC9 associates with proteins in the mSin3A complex. Class I HDACs, but not class II HDACs were previously found to be associated with the mSin3A complexes. The core HDAC1 complex consists of HDAC1, HDAC2, RbAp46, RbAp48. This core complex has been found to associate with an mSin3A complex that is involved in transcriptional repression through an Rb and E2F complex (Luo RX, Postigo AA, Dean DC. (1998) Rb interacts with histone deacetylase to repress transcription. Cell. 92, 463-473; Magnaghi-Jaulin L, Groisman R, Naguibneva I, Robin P, Lorain S, Le Villain JP, Troalen F, Trouche D, Harel-Bellan A. (1998) Retinoblastoma protein represses transcription by recruiting a histone deacetylase. Nature. 391, 601-605; Brehm A, Miska EA, McCance DJ, Reid JL, Bannister AJ, Kouzarides T. (1998)

WO 02/50285

PCT/EP01/14928

Retinoblastoma protein recruits histone deacetylase to repress transcription. Nature. 391, 597-601). In order to determine whether HDAC9 was a part of this complex, endogenous HDAC1, HDAC2, Rb, and mSin3 proteins were co-immunoprecipitated from cells transfected with flag-epitope tagged HDAC1, HDAC3, HDAC4, HDAC6, HDAC7 or HDAC9. To assure that

5 transfected flag epitope-tagged HDACs could be detected in cells, the levels of HDAC expression were detected by immunoprecipitation and Western blotting with antiserum to the flag epitope. To determine which HDACs associated with components of the Sin3 complex, endogenous proteins in the Sin3 complex were immunoprecipitated and the associated HDACs were detected by Western blotting flag epitope-specific antibody HDAC9 was found to associate
10 with HDAC1, HDAC2, Rb, and mSin3A, suggesting that HDAC9 is a component of an mSin3A complex.

HDAC9 associates with SMRT and NCoR. Since corepressors SMRT and NCoR associate with the mSin3 core complex, experiments were performed to co-immunoprecipitate HDACs with NCoR and SMRT (Fig. 15). HDAC9 co-immunoprecipitated with both of these
15 proteins, suggesting that HDAC9 associates with SMRT, and NCoR. Western analysis of the flag-detected blots with anti-NCoR indicated that NCoR was immunoprecipitated. As previously reported, SMRT co-immunoprecipitated with HDAC4 and HDAC6, and HDAC6 and HDAC7 did not associate with the Sin3A complex.

HDAC9 associates with 14-3-3 and Erk proteins. HDAC4 was previously found to
20 associate with 14-3-3- β , 14-3-3- ϵ , CamK, Erk1, and Erk 2 proteins, which sequester HDAC4 in the cytoplasm and prevent phosphorylated HDAC4 and HDAC5 from entering the nucleus and repressing MEF2 activated transcription. In order to determine whether HDAC9 associate with these proteins, experiments were performed to co-immunoprecipitate HDACs with 14-3-3 and

WO 02/50285

PCT/EP01/14928

Erk proteins. All of the HDACs tested associated with 14-3-3s and Erks. These results suggest that the association of HDACs with 14-3-3 and Erks might be a general mechanism of sequestering HDACs in the cytoplasm.

Classification of HDAC9. HDACs have been classified by sequence similarity to yeast HDACs, sequence length, location of catalytic domain, cellular localization, associating proteins, and sensitivity to HDAC inhibitors. The data in this study suggests that HDAC9 has characteristics of both class I and class II HDACs. HDAC9 had sequence similarity with class II yeast hda1 subfamily member Clr3 and HDAC6 catalytic domain 1. In addition, the 3 Kb HDAC9 transcript was only detected in kidney and testis, suggesting that it might have a limited tissue distribution like class II HDACs. HDAC9 was between class I and class II HDACs in length. Class I HDACs average 443 bp in length, whereas class II HDACs average 1069 bp in length. However, HDAC9 was found to have an N-terminal catalytic domain, as opposed to the C-terminal domains that have been found in class II HDACs. HDAC6 is an exception that has both N-terminal and C-terminal catalytic domains. Furthermore, class I HDACs are nuclear proteins, while class II HDACs are nucleo-cytoplasmic. Immunocytochemistry indicated that HDAC9 was predominantly nuclear and was detected in a different subcellular compartment in comparison to the Ena protein that is expressed in the cytoplasm. In contrast to the 3 Kb HDAC9 transcript that might be differentially expressed, a 3.5 Kb HDAC9 transcript that was identified by Northern analysis was expressed ubiquitously in normal tissues, tumor tissues and cell lines, similar to class I HDACs. In addition, HDAC9 was found to co-immunoprecipitate with proteins that were previously only associated with class I HDAC complexes, including HDAC1, HDAC2, mSin3A, and Rb. HDAC9 also has putative C-terminal LXCXE motifs that so far have only been

WO 02/50285

PCT/EP01/14928

found in HDAC1. HDAC9 was also found to associate with NCoR and SMRT. This evidence suggests HDAC9 had characteristics that bridged those of class I and class II HDACs.

WO 02/50285

PCT/EP01/14928

What is claimed is:

- 5 1. An isolated polypeptide comprising the amino acid sequence set forth in
SEQ ID NO:1, SEQ ID NO 5 or SEQ ID NO 6.
2. An isolated polypeptide consisting of the amino acid sequence set forth in
SEQ ID NO:1, SEQ ID NO 5 or SEQ ID NO 6.
- 10 3. An isolated DNA comprising a nucleic acid sequence that encodes the
polypeptide of claim 1 or 2.
4. A vector molecule comprising at least a fragment of the isolated DNA
according to claim 3.
- 15 5. The vector molecule according to claim 4 comprising transcriptional control
sequences.
6. A host cell comprising the vector molecule according to claim 5.
- 20 7. The isolated DNA according to claim 3, comprising a nucleotide sequence
selected from the group consisting of (1) the nucleotide sequence set forth in SEQ ID NO:2, 7 or
8, being the complete cDNA sequence encoding the polypeptide as defined in claim 2; (2) the
nucleotide sequence set forth in SEQ ID NO:3, being the open reading frame of the cDNA
25 sequence encoding the polypeptide as defined in claim 2; (3) a nucleotide sequence capable of
hybridizing under high stringency conditions to a nucleotide sequence set forth in SEQ ID NO:3;
and (4) the nucleotide sequence set forth in SEQ ID NO:4, being the endogenous genomic
human DNA encoding the polypeptide as defined in claim 2.

WO 02/50285

PCT/EP01/14928

8. A vector molecule comprising at least a fragment of an isolated DNA molecule according to claim 7.
9. The vector molecule according to claim 8 comprising transcriptional control sequences.
10. A host cell comprising the vector molecule according to claim 9.
11. A host cell which can be propagated in vitro and which is capable upon growth in culture of producing a polypeptide according to claim 1 or 2, wherein said cell comprises at least one transcriptional control sequence that is not a transcriptional control sequence of the natural endogeneous human gene encoding the polypeptide of claim 2, wherein said one or more transcriptional control sequences control transcription of a DNA encoding a polypeptide according to claim 1 or 2.
12. A method for the diagnosis of a condition associated with abnormal regulation of gene expression which includes, abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, host inflammatory or immune response, or psoriasis in a human which comprises: detecting abnormal transcription of messenger RNA transcribed from the natural endogeneous human gene encoding the polypeptide as defined in claim 2 in an appropriate tissue or cell from a human, wherein said abnormal transcription is diagnostic of said condition.
13. The method of claim 12, wherein said natural endogeneous human gene comprises the nucleotide sequence set forth in SEQ ID NO:4, 7 or 8.
14. The method of claim 12, comprising contacting a sample of said appropriate tissue or cell or contacting an isolated RNA or DNA molecule derived from said tissue or cell with an isolated nucleotide sequence of at least about 15-20 nucleotides in length

WO 02/50285

PCT/EP01/14928

that hybridizes under high stringency conditions with the isolated nucleotide sequence as defined in claim 3.

15. A method for the diagnosis of a condition associated with abnormal
5 HDAC9 expression or activity in a human which comprises:
measuring the amount of a polypeptide comprising the amino acid
sequence set forth in SEQ ID NO:1, 5 or 6 or fragments thereof, in an appropriate tissue or cell
from a human suffering from said condition wherein the presence of an abnormal amount of said
polypeptide or fragments thereof, relative to the amount of said polypeptide or fragments thereof
10 in the respective tissue from a human not suffering from said condition associated with abnormal
HDAC9 expression or activity is diagnostic of said human's suffering from a condition
16. The method of claim 15, wherein said detecting step comprises contacting
said appropriate tissue or cell with an antibody which specifically binds to a polypeptide that
15 comprises the amino acid sequence set forth in SEQ ID NO:1, 5 or 6 or a fragment thereof and
detecting specific binding of said antibody with a polypeptide in said appropriate tissue or cell,
wherein detection of specific binding to a polypeptide indicates the presence of a polypeptide
that comprises the amino acid sequence set forth in SEQ ID NO:1, 5 or 6 or a fragment thereof.
- 20 17. An antibody or a fragment thereof which specifically binds to a
polypeptide that comprises the amino acid sequence set forth in SEQ ID NO:1, 5 or 6 or to a
fragment of said polypeptides.
- 25 18. An antibody fragment according to claim 17 which is an Fab or F(ab')₂
fragment.
19. An antibody according to claim 17 which is a polyclonal antibody.
20. An antibody according to claim 17 which is a monoclonal antibody.

30

WO 02/50285

PCT/EP01/14928

21. A method for producing a polypeptide as defined in claim 1 or 2, which method comprises:

5 culturing a host cell having incorporated therein an expression vector comprising an exogenously-derived polynucleotide encoding a polypeptide comprising an amino acid sequence as set forth in SEQ ID NO:1, 5 or 6 under conditions sufficient for expression of the polypeptide in the host cell, thereby causing the production of the expressed polypeptide.

10 22. The method according to claim 21, said method further comprising recovering the polypeptide produced by said cell.

23. The method according to claim 21, wherein said exogenously-derived polynucleotide encodes a polypeptide consisting of an amino acid sequence set forth in SEQ ID NO:1, 5 or 6.

15 24. The method according to claim 21, wherein said exogenously-derived polynucleotide comprises the nucleotide sequence as set forth in SEQ ID NO:2, 7 or 8.

20 25. The method according to claim 21, wherein said exogenously-derived polynucleotide comprises the nucleotide sequence as set forth in SEQ ID NO:3.

26. The method according to claim 21, wherein said exogenously-derived polynucleotide consists of the nucleotide sequence as set forth in SEQ ID NO:3.

25 27. The method according to claim 24, wherein said exogenously-derived polynucleotide comprises the nucleotide sequence as set forth in SEQ ID NO:4.

Figure 1.

1 GGGCCCGAGG CTTCTGCGTC CGTCGTGGTT CCTCGCTCCG
41 GGGCGGAGTT CGCGATAGCG ATCGGGGAGC AGGACGCGGG
82 GCGTGGACCC AGGTCCGAGG CGAGGAAGCC GTAACCATG
123 CGCGGGGAGC CTCCCCCTTC GACTGCAGCC TCGCTCCGTG
164 CTTTCTGCGC GCCTGGGATC CCGGAGCCTG OCTAGGTTCT
205 GTGCGCTCCC GCGCAGGCCG GTGCCCGCCG CCCGCTGCG
246 CCCCAGGCAG GTCCCAGGCC TCCGGCTGCT CCCGGCCGAA
287 GCCCCGAGTG CGAGATCGAG CGTCTGAGC GCCTGACCGC
328 AGCCCTGGAT CGCCTGCCGC AGCGCGGCT GGAACAGAGG
369 TGTCTGCGGT TGTAGCCCG CGAGGCCCTCG GAAAGGAGC
410 TGGGCTGGT GCACAGAGTA CCTTCACTG CGCGCGGCTG
451 GCGCAGGGG CTGGACTGCA GCTGGTGGAC GCTGTGCTCA
492 CTGGAGCTGT GCAAAATGGG CTTGCCCTGG TGAGGCCTCC
533 CGGGCACCAT GGCCAGAGGG CGGCTGCCAA CGGGTTCTGT
574 GTGTTCAACA ACGTGGCCAT AGCAGCTGCA CATGCCAAGC
615 AGAAACACGG GCTACACAGG ATCCTCGTCG TGGACTGGGA
655 TGTGCCACCAT GGCCAGGGGA TCCAGTATCT CTTGAGGAT
696 GACCCACAGG TCCTTTACTT CTCTGGCAC CGCTATGAGC
737 ATGGCGGCTT CTGGCCTTTC CTGCGAGAGT CAGATGCAGA
778 CGCAGTGGGG CGGGGACAGG GCCTCGGCTT CACTGTCAAC
819 CTGCCCTGGA ACCAGGTTGG GATGGGAAAC GCTGACTACG
860 TGGCTGCCCTT CCTGCACCTG CTGCTCCAC TGGCCTTGA
901 GTTTGACCCT GAGCTGGTGC TGGTCTCGGC AGGATTGAC
942 TCAGCCATCG GGGACCTGA GGGGCAAATG CAGGCCACGC
983 CAGAGTGCTT CGCCACCTC ACACAGCTGC TGAGGTGCT
1024 GGCCGGCGGC CGGGTCTGTG CCGTGCTGGA GGCGGCTAC
1065 CACCTGGAGT CACTGGCGGA GTCAGTGTGC ATGACAGTAC
1106 AGACGCTGCT GGGTGACCCG GCCCCACCC TGTCAGGGCC
1147 AATGGCGCC

WO 02/50285

2/29

PCT/EP01/14928

Figure 2.

A.

```

1 atgggggaccg cgcttggtga ccatgaggac atgacggcca cccggctgct ctgggacgac
61 cccgagtgccg agatcgagcg tcttgagcgc ctgaccgcag cccctggatcg cctgcggcag
121 cgcggcctgg aacagaggtg tctgcggttg tccgcccgcg aggcctcgga agaggagctg
181 ggcctggtgc acagcccaga gtatgtatcc ctggtcaggg agaccaggt cctagccaaag
241 gaggagctgc aggcgctgct cggacagttc gacgccaatc acttccacc gagtaccttt
301 cactgcgcgc ggctggccgc aggggctgga cctcccgggc accatggcca gaggcggtct
361 gctgtgcaaa atgggcttgc cctggtgagg cctcccgggc accatggcca gaggcggtct
421 gccaaagggt tctgtgtgtt caacaacgtg gccatagcag ctgcacatgc caagcagaaa
481 caggggctac acaggatcct cgtcgtggac tgggatgtgc accatggcca ggggatccag
541 tatctctttg aggatgaccc cagcgtcctt tactttcctt ggcaaccgta tgagcatggg
601 cgcttctggc ctttctctgc agagtcagat gcagacgcag tggggcgggg acagggctct
661 ggcttctactg tcaacctgcc ctggaaccag gttgggatgg gaaacgctga ctactgggtc
721 gccttctctgc acctgctgct cccactggcc tttgagtttg accctgagct ggtgctggtc
781 tggcgaggat ttgactcagc catcggggac cctgaggggc aaatgcagtc cacgccagag
841 tgcttcgccc acctcacaca gctgctgcag gtgctggccg gcggccgggt cgtgcccgtg
901 ctggaggcg gctaccacct ggagtcactg gcggagtcag tgtgcatgac agtacagacg
961 ctgctgggtg acccgggccc acccctgtca gggccaatgg cgccatgtca gagggtcgag
1021 gggagtgccc tagagtccat ccagagtgcc cgtgctgccc agggcccgca ctgggaagagc
1081 ctccagcagc aagatgtgac cgctgtgcgc atgagcccca gcggccactc cccagagggg
1141 agggctccac ctctgctgcc tgggggtcca gtgtgttaagg cagctgcatc tgcaccgagc
1201 tccctcctgg accagccgtg cctctgcccc gcacccctctg tccgcaccgc tgttgcctctg
1261 acaacgcccg atatcacatt ggttctgccc cctgacgtca tccaacagga agcgtcagcc
1321 ctgagggagg agacagagag ctgggccagg ccacacaggt ccttggcccg gggagggccc
1381 ctacatgcac ttgggaagct cctgtacetc ttegatggga tgcggatgg gcaggtgaac
1441 agtggtatag cagccactcc agcctctgct gcagcagcca cccctggatgt ggtgttctcg
1501 agaggcctgt cccacggagc ccagaggctg ctgtgcgtgg cccctgggaca gctgtaccgg
1561 cctccagacc tgcacctga cgggaggagt ctgtggctga acatcagggg caaggaggcg
1621 gctgccctat ccatgttcca tgtctccag ccactgccag tgatgacggg tggtttctcg
1681 agctgcatct tgggcttggg gctgccctgt gcctatggct tccagcctga cctgtgctg
1741 gtggcgctgg ggcctggcca tggcctgcag gggcccccag ctgcactcct ggtgcaatg
1801 ctccgggggg tggcaggggg ccgagtcctg gccctcctgg aggaggaactc cacaccccaa
1861 ctacagggga tcttggcccg ggtgctgaat ggagaggcac ctctagcct agggccttcc
1921 tctgtggcct cccagaggga cgtccaggcc ctgatgtacc tggaggggca gctggagcct
1981 cagtggaaga tgttgagtg ccatcctcac ctggtggctt ga

```

B.

```

MGVALVYHED NTATRLWDD PECEIERFER LTAALDLRQ RGLQRCRLR SAREASEEEL
GLVHSPETVS LVRETQVLGK EELQALSGQF DAIFYHPSTP HCARLAAGAG LQLVDAVLTG
AVQNGALAVR PPGHHGQRAA ANGFCVFMV AIAAAHAKQK HGLHRLVVD WDVHHCQGIQ
YLFEDDFSVL YFSWHRYEHG RFWPFLRESD ADAGVRGQGL GFTVNLPMWQ VMGMNADIYA
AFILHLLPLA FEFDPPELVLV SAGFDSAIGD PEGQMQATPE CFAHLTQLLQ VLAGGRVCAV
LEGQYHLES L AEFVCMVTQT LLGDPAPFLS GPMAPCORCE GSALESIQSA RAAQAPHWKS
LQQQDVTAVP MSPSSHSEFEG RPPPLLPGGP VKAAASAPS SLLDQPCICP APSVRTAVAL
TTPDITLVLP PDVIOQEASA LREETEAWAR PHESLAREEA LTRALGXLLYL LDGMLDGGQVN
SGIAATPASA AANTLDVAVR RGLSHGAQRL LCVALGQLDR PFDLHNDGRS LMLNIRGKEA
AALSMFVST PLFVMTGGFL SCILGLVLPL AYGFOPDLVL VALGPGHGLQ GPHALLAAM
LRGLAGGRVL ALLEENSTPQ LAGILARVLN GEAPPSLGFPS SVASPEDVQA LMYLRGQLEP
QHWMLQCPHF LWA

```

WO 02/50285

3/29

PCT/EP01/14928

Figure 3.

AL022328 vs HDAC9:

AL022328	2	tcaagccaccagggtgaggatggcactaca.....ctcacctgcaacatct	180
HDAC9	1	tcaagccaccagggtgaggatggca.....ctgcaacatct	35
AL022328	181	tccactgaggctccagctgccctctcaggtacatcagggcctggacgtcc	230
HDAC9	36	tccactgaggctccagctgccctctcaggtacatcagggcctggacgtcc	85
AL022328	231	tctggggaggccacagaggagggcctaggctaggaggtgctctccatt	280
HDAC9	86	tctggggaggccacagaggagggcctaggctaggaggtgctctccatt	135
AL022328	281	cagcaccggggtcaggatccctgctagctggggtgtggagttctgga...	322
HDAC9	136	cagcaccggggtcaggatccctgctagctggggtgtggagtt.....	177
AL022328	322	.cttacctcctccaggagggtcaggactcggccctgcagccccgaa	448
		<<<<	
HDAC9	177	ctctccaggagggtcaggactcggccctgcagccccgaa	221
AL022328	449	gcattgcagccagggtgcagcgtgggggcctgcaggccatggccaggc	498
HDAC9	222	gcattgcagccagggtgcagcgtgggggcctgcaggccatggccaggc	271
AL022328	499	cccagcgccaccagcaccgggtcaggctggaagccataggccaggggcag	548
HDAC9	272	cccagcgccaccagcaccgggtcaggctggaagccataggccaggggcag	321
AL022328	549	caccaagcccaagatgcagctcagggaaccacgggtcatctgtg....c	587
HDAC9	322	caccaagcccaagatgcagctcagggaaccacgggtcat.....	360
AL022328	587	tcaccactggcagtgccgtggagacatggaacatggataggccagccgc	837
		<<<<	
HDAC9	360	cactggcagtgccgtggagacatggaacatggataggccagccgc	406
AL022328	838	tccttgccctgatgttcagccacagactcctctcc.....cctaccc	1174
HDAC9	407	tccttgccctgatgttcagccacagactcctcc.....cc	441
AL022328	1175	gtcatggggcagggtctggggccgggtccagctgtcccaggccacgcaca	1224
HDAC9	442	gtcatggggcagggtctggggccgggtccagctgtcccaggccacgcaca	491
AL022328	1225	gcagcctgga.....cttacctctgggtccgtgggacaggcctctccga	1398
		<<<< 139 <<<<	
HDAC9	492	gcagc.....ctctgggtccgtgggacaggcctctccga	526
AL022328	1399	acagccacatccagggtggctgctgcagcagggtggagtggtgctat	1448
HDAC9	527	acagccacatccagggtggctgctgcagcagggtggagtggtgctat	576
AL022328	1449	accactgttcacctgtg....cccactgcccacccagcctccatcta	2208
		<<<< 725 <<<<	

WO 02/50285

4/29

PCT/EP01/14928

HDAC9 577 accactgttcac.....ctgccatccagcatcccatcta 611

AL022328 2209 agaggtacaggagcttcccaagtgcagtgaaggcctctcccgggccaagg 2258
|||||

HDAC9 612 agaggtacaggagcttcccaagtgcagtgaaggcctctcccgggccaagg 661

AL022328 2259 gactcgtgtggcctgtg....cccactggcccaggcttctgtctcctc 2406
|||||<<<< 113 <<<<|

HDAC9 662 gactcgtgtggc.....ctggcccaggcttctgtctcctc 696

AL022328 2407 cctcagggtctgacgtctcctgttgatgacgtcagggggcagaaccaatg 2456
|||||

HDAC9 697 cctcagggtctgacgtctcctgttgatgacgtcagggggcagaaccaatg 746

AL022328 2457 tgatattcggcggtgtcagggtcagcggtgcggacaggggtgcgggg 2506
|||||

HDAC9 747 tgatattcggcggtgtcagggtcagcggtgcggacaggggtgcgggg 796

AL022328 2507 cagaggcaggggtgtccaggaggagctcgggtgcagatgcagctgcctt 2556
|||||

HDAC9 797 cagaggcaggggtgtccaggaggagctcgggtgcagatgcagctgcctt 846

AL022328 2557 acacactggaccccccaggcagcagaggtggaggcctccctctggggagt 2606
|||||

HDAC9 847 acacactggaccccccaggcagcagaggtggaggcctccctctggggagt 896

AL022328 2607 ggctgctggggctcatcggcacagcggtcacatctagg....ctgacctt 2721
|||||<<<< 79 <<<<|

HDAC9 897 ggctgctggggctcatcggcacagcggtcacat.....ctt 932

AL022328 2722 gctgctggaggctcttcagtcggggcctgggtagcagggcactctgg 2771
|||||

HDAC9 933 gctgctggaggctcttcagtcggggcctgggtagcagggcactctgg 982

AL022328 2772 atggactctagggcactgtg....cctacctccctcgacactctgacat 2868
|||||<<<< 61 <<<<|

HDAC9 983 atggactctagggca.....ctccctcgacactctgacat 1018

AL022328 2869 ggcgccattggccctgacaggggtggggcgggtcaccagcagcgtctg 2918
|||||

HDAC9 1019 ggcgccattggccctgacaggggtggggcgggtcaccagcagcgtctg 1068

AL022328 2919 tactgtcatgcacactgactccgcagtgactccaggtggtagccgcct 2966
|||||<<

HDAC9 1069 tactgtcatgcacactgactccgcagtgactccaggtggtagccgc.. 1116

AL022328 2966 ggg....gtcactctcagcacggcacgacccggcggccagcagacc 3192
<<< 189 <<<<|

HDAC9 1116 ctccagcacggcacgacccggcggccagcagacc 1153

AL022328 3193 tgcagcagctgtgtgaggtgggcgaagcactctggcgtggcctgcatctg 3242
|||||

HDAC9 1154 tgcagcagctgtgtgaggtgggcgaagcactctggcgtggcctgcatctg 1203

AL022328 3243 cccctgga....ctacctcagggtcccgatggctgagtcasatcctgc 3357
|||<<<< 79 <<<<|

HDAC9 1204 ccc.....ctcagggtcccgatggctgagtcasatcctgc 1239

AL022328 3358 cgagaccagcagcagctcagggtcacaactaca....gtcactcaaagg 3604

[illegible]

WO 02/50285

6/29

PCT/EP01/14928

```
AL022328 5322 ctctgttccaggccgcgctgccgcaggcgatccagggtgcggtcaggcg 5371
          |||
HDAC9 1886 ctctgttccaggccgcgctgccgcaggcgatccagggtgcggtcaggcg 1935
AL022328 5372 ctcaggacgctcgatctcgcactcggggtggg....cttactcgccca 5475
          |||
HDAC9 1936 ctcaggacgctcgatctcgcactcggggtggg....tcgtccca 1971
AL022328 5476 gagcagccgggtggccgtcatgtctcatggtacacagcgcg 5519
          |||
HDAC9 1972 gagcagccgggtggccgtcatgtctcatggtacacagcgcg 2015
```

WO 02/50285

7/29

PCT/EP01/14928

Figure 4.

Score = 267 bits (676), Expect = 4e-71
 Identities = 143/354 (40%), Positives = 201/354 (56%), Gaps = 14/354 (3%)
 Query: 26 ERPERLTAAADRLRONGLEQRC-----LRISAKKASEELGOLVHSELYVSLVRETOVL 78
 E P R+ + + + G LR+ AREK+ EEL VHS E V T+ +
 Subject: 72 EDPRRVRFEEAIKKAGVSNVPEESDVLELIPAREATLELLQVHSECHYDRVINTEKM 186
 Query: 79 SKELQALSQGFDAIYTHESIHCARLAAGAGLOEVDALVTSVAVONGKALVREPOHSGQIN 138
 E+L L D+ + + + + CARLA G+ + + NVATG V+ N + + VREPEGBH +
 Subject: 139 SHEDLANLEKISDSLYNNESAFECARLACGSALETCTAVVTCOVNLAFAVVRPEGHNALE 198
 Query: 139 AANGCEVNNVATIAAAHAKQHG--LHRILVVVDVVRHGGGICVLTEDDESULVIEVHTR 198
 GEC+ENNV+ + A C+ + R+L+VDP+HNS G O E DE+VLY S HIR
 Subject: 199 HKEGGECINNVSVTARSHLOREPDKIKRVLIVDUEIRHNGTOMAEYDDENVIVSLIIN 258
 Query: 197 YHGRLEPILRESADAVGRCOSLGCTVNLDPNQCCHONADYVXXXXXXXXXXXXXDEE 256
 YE+GRE+P A+ G G G LG TVN+PM+ GNG+ DY+ DE+
 Subject: 259 YHGRFYEGTNYGCAENGCLGGLGRTVNIPDSCAGGCGDYIYAFQVVHNVANTQED 318
 Query: 257 EVLVSAQDSAGDEFGONATETCTAHLTOLLQVLAGGRVCVILEGGYHETSAESVCM 318
 LV+VS GED+ A GD GQ TE +IS+QHL SA GVV LGGYH+ST+ E
 Subject: 319 EVIVSCGDAAGDHIGGLUTFAAYAHNTOMLGLADGVFLSLTCTNDEISTALIA 378
 Query: 317 IVOTILGDFAPPLSGEHAECORCEGSALESICSAKACAPHNSLOOQDVJAVE 370
 O+LIG P L A O A+ + + O+ + + + + + A F
 Subject: 379 VACSLGIGPGRHHTYACD-----AVATINEVTKICQYWRCHREKEDANE 427

Figure 5.

```

10A C1      M A Q T G G T R G V V Y Y D G V G Y T T T G -- G G H O G H O R I H T D B L L I N G L Y R G G E Y T P F H
10A C9      -----M E T A L V D H D T A T R L L W D H E C T E E P E R L T A A L D R L T Q R G L E G R C L R S A R E
               * * * * *      : : * : : * : * : * : :

10A C1      A N A E D T K M S D D Y I K L R S I R P D M S I Y S K Q M Q R V N G E D C P V F D G L E P C O L S T G C --
10A C9      A S E E L I G L V S P E Y V S L V R E T Q V L K E E L Q A L S G D F A -- I Y F H P S T F H C A R A A G A G L
               * * * * *      : : * : * : * : * : * : * : * : * : * : * : * :

10A C1      S V A S K V K L I N Q D T A V A A A G G I P H O K S E A G C C V M D V I A L L I L K H -- G R V L T V D
10A C9      Q L V D A V L G A V G N G A L V R P G - H G G H A A A G I C V T I T A T A A A A K G G L D I V I V D
               ( : * ) * : * : * : * : * : * : * : * : * : * : * : * : * : * : * :

10A C1      I I I I G I G V L L A Y T T R D V H T V S F K Y G E T -- R P C Y I L R D I G A G K Y T A M T P L R D
10A C9      I I I I G I G Q I L E D D E S V L T F S W A Y I H G R N L L G S I A R A V G R G G G E T D I L P V N Q
               * : * : * : * : * : * : * : * : * : * : * : * : * : * : * :

10A C1      G - I D D E S Y L A I F K P M V S K M H A P Q S A V L Q C G S D S L S C D R L G C R N L Y I K G H A K O V E E V K
10A C9      V G M G R A D V A A R L L L L P L A T E D F E L V Y S A G D S A I G P P G Q H O A T P E C T A G I Q L Q
               : : * : * : * : * : * : * : * : * : * : * : * : * : * : * :

10A C1      S F N - L P M L L G G G G I I R N V A R C H I T A V A L D T E P
10A C9      V L A G G R V C A V L G G G L E S L A E S Y G I T V L L G P A P E L S G P M A P C O R C G S A L E S T O S A
               : : : * * : * : * : * : * : * : * : * : * :

10A C1      -----N E L P Y R-----
10A C9      R A A Q A P M H K S L Q Q Q V Y A V P S F S S H S P D G R P P L L P G G P V C A A A S A P S S I L Q P C L C P
               : * :

10A C1      -----D T F E Y G P D F I K L H S P S I M T I O N----- T H E Y L E X I Q R L F E M L A L P H A
10A C9      A P S V R T A V A L T P D I L V L P P V I Q P A S A L R E T E A W A P H S S A R E T A L T A L G L Y T
               * : * : * : * : * : * : * : * : * : * : * : * : * : * : * :

10A C1      P G V Q M A I P E L A I F E S O D E D D P F K R I S -- I C S D R C H A C E P E -- P S D S E E P E G G H
10A C9      L D G L D G V S G S L A T P A S A A A L L D V A V A N G L S H A G Q U L L V A L G D R P F D I D O R S
               * : * : * : * : * : * : * : * : * : * : * : * : * : * : * :

10A C1      N S S H F K
10A C9      L M L N R K T A A A S I A F H Y S I P L P V M T G E L S C L L V L P L A Y G E Q P D L V I A L G P G H L D
               * : * :

10A C1      -----K A K V A T E D E K E O F L E K-----K E V T L E E K T E E K P E D A V K E E V L
10A C9      G P H A A L A A M R G L A G K V L A L L E S T P L A G T A V I N G E A P E S L G S V A P E D V A
               * * : * : * : * : * : * : * : * : * : * : * : * :

10A C1      A-----
10A C9      L M T A G R O L E P P S K L A C H P H L V A

```

 Catalytic domain residues
 Potential RB-binding pocket residues

WO 02/50285

9/29

PCT/EP01/14928

Alignment results

Figure 6

Sequence format is Pearson

Sequence 1:	HDAC1	482 aa
Sequence 2:	HDAC2	488 aa
Sequence 3:	HDAC3	428 aa
Sequence 4:	HDAC8	377 aa
Sequence 5:	HDAC4	1084 aa
Sequence 6:	HDAC5	1122 aa
Sequence 7:	HDAC6	1122 aa
Sequence 8:	HDAC7	855 aa
Sequence 9:	HDAC9	673 aa

Start of Pairwise alignments

Aligning...

Sequences (1:2)	Aligned. Score: 82
Sequences (1:3)	Aligned. Score: 57
Sequences (1:4)	Aligned. Score: 38
Sequences (1:5)	Aligned. Score: 18
Sequences (1:6)	Aligned. Score: 14
Sequences (1:7)	Aligned. Score: 14
Sequences (1:8)	Aligned. Score: 15
Sequences (1:9)	Aligned. Score: 14
Sequences (2:3)	Aligned. Score: 55
Sequences (2:4)	Aligned. Score: 39
Sequences (2:5)	Aligned. Score: 13
Sequences (2:6)	Aligned. Score: 15
Sequences (2:7)	Aligned. Score: 15
Sequences (2:8)	Aligned. Score: 14
Sequences (2:9)	Aligned. Score: 15
Sequences (3:4)	Aligned. Score: 37
Sequences (3:5)	Aligned. Score: 12
Sequences (3:6)	Aligned. Score: 13
Sequences (3:7)	Aligned. Score: 13
Sequences (3:8)	Aligned. Score: 15
Sequences (3:9)	Aligned. Score: 15
Sequences (4:5)	Aligned. Score: 21
Sequences (4:6)	Aligned. Score: 16
Sequences (4:7)	Aligned. Score: 16
Sequences (4:8)	Aligned. Score: 20
Sequences (4:9)	Aligned. Score: 22
Sequences (5:6)	Aligned. Score: 59
Sequences (5:7)	Aligned. Score: 59
Sequences (5:8)	Aligned. Score: 49
Sequences (5:9)	Aligned. Score: 21
Sequences (6:7)	Aligned. Score: 100
Sequences (6:8)	Aligned. Score: 43
Sequences (6:9)	Aligned. Score: 19
Sequences (7:8)	Aligned. Score: 43

WO 02/50285

10/29

PCT/EP01/14928

Sequences (7:9) Aligned. Score: 19
 Sequences (8:9) Aligned. Score: 20
 Guide tree file created: [/bioinfv/software/biobenchsw/tmp/align/1478.dnd]
 Start of Multiple Alignment
 There are 8 groups
 Aligning...
 Group 1: Sequences: 2 Score:24259
 Group 2: Sequences: 3 Score:18415
 Group 3: Sequences: 4 Score:12882
 Group 4: Delayed
 Group 5: Sequences: 2 Score:9847
 Group 6: Sequences: 3 Score:7569
 Group 7: Sequences: 4 Score:5689
 Group 8: Sequences: 8 Score:2841
 Sequence:9 Score:3452
 Alignment Score 36872
 CLUSTAL-Alignment file created [/bioinfv/software/biobenchsw/tmp/align/1478.out]
 CLUSTAL W (1.81) multiple sequence alignment

```

HDAC5      MNSPNESDGMSGRPSLEILPRTSLHSIPVTVEVFLVLPAMPSSMGGGGGSPSPVFLR
HDAC6      MNSPNESDGMSGRPSLEILPRTSLHSIPVTVEVFLVLPAMPSSMGGGGGSPSPVFLR
HDAC4      MSSQSHFDGLSGRDQFVELLNPARVNHMPETVDVATALLQVAFSAVPMDLKLDHQFSLP
HDAC7      -----MDLRVGGRRFVEPPF-----
HDAC1      -----
HDAC2      -----
HDAC3      -----
HDAC8      -----
HDAC9      -----

HDAC5      GALVGSVDPTLREQQLQCELLALKQCCOLQKLLFAEFQGHDLTRQHEVOLCKHLKQ
HDAC6      GALVGSVDPTLREQQLQCELLALKQCCOLQKLLFAEFQGHDLTRQHEVOLCKHLKQ
HDAC4      -----VAEPALRFQQLQCELLALKQCCOLQKLLFAEFQGHDLTRQHEVOLCKHLKQ
HDAC7      -----EPTLLALQRFQRLNHHFLAGLQ-----
HDAC1      -----
HDAC2      -----
HDAC3      -----
HDAC8      -----
HDAC9      -----

HDAC5      QEMLAAKQCEMLAAKQCELEQCRQREQCELEKQRLQQLLANKKSKESAIAS
HDAC6      QEMLAAKQCEMLAAKQCELEQCRQREQCELEKQRLQQLLANKKSKESAIAS
HDAC4      QEMLAAKQCEMLAAKQCELEQCRQREQCELEKQRLQQLLANKKSKESAIAS
HDAC7      RSVEPHRLSMUTP-----HPELTVGPGQELRLQLLHKDKSKSAVAS
HDAC1      -----
HDAC2      -----
HDAC3      -----
HDAC8      -----
HDAC9      -----

HDAC5      TEVKRLQEFLLSKSEPTPGGLNMSLPQHFACWG--ABHASLDQSSPPQSGPPGTFPSY
HDAC6      TEVKRLQEFLLSKSEPTPGGLNMSLPQHFACWG--ABHASLDQSSPPQSGPPGTFPSY
HDAC4      TEVKRLQEFVFNRRALAHRLNHCISSDPRYWGKTQHSSELDQSSPPQSG--VSTSY
HDAC7      SVVKQKLAEVILAKQQAALERTVHPNSPGIF-----YRTLEP-LETEGATRSMLSP
HDAC1      -----
HDAC2      -----

```

WO 02/50285

11/29

PCT/EP01/14928

```

HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 KLPLPG-PYDSRDDFFLRKTASEPNLKVRSRLKQKVAERSSPLLRKDGTVISTFKKRA
HDAC6 KLPLPG-PYDSRDDFFLRKTASEPNLKVRSRLKQKVAERSSPLLRKDGTVISTFKKRA
HDAC4 NHFVLG-HYDAKDDFFLRKTASEPNLKLRSRLKQKVAERSSPLLRKDGTVVITALKKRP
HDAC7 LPFVSLPSPDPPEHFLEKTVSEPNLKLRYAPR-KSLERRNPLLRKE--SAPPSLRRRP
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 VEITGAGPGASEVCNAPGCGPESP-N-SHSTIAENGFTGCVNIPTEMLPCHRALPLDS
HDAC6 VEITGAGPGASEVCNAPGCGPESP-N-SHSTIAENGFTGCVNIPTEMLPCHRALPLDS
HDAC4 LDVT-----DEACSEAPGCGPESPNDSEGSVAENGIAFAVPSIPATSLAHLVAREG
HDAC7 AETLG---DESPSSSTPAEGCESPNDSEHG-----
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 SPNQFSLYTSPSLPNISLGLQATVTVTNSHLTAEFKLSTQCEAFQALQSLRQGGTLTGK
HDAC6 SPNQFSLYTSPSLPNISLGLQATVTVTNSHLTAEFKLSTQCEAFQALQSLRQGGTLTGK
HDAC4 SAAPLPLYTSPSLPNITLGLPATG-----PSAGTAGQDTERLILPALQDR--LS--
HDAC7 -----FNFIIG-----DEBERTHPTLGFPG-----
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 FMSTSEIPGCLLGVALEGDSFGHNASLLQHVLLLEQARQOSTLIA-----VPLHGQSP
HDAC6 FMSTSEIPGCLLGVALEGDSFGHNASLLQHVLLLEQARQOSTLIA-----VPLHGQSP
HDAC4 LPFGTHLTFYLTSPLEKGG-AAHSPLLQHVLLLEQFPAQAPLVIGL--GALPLHAQS-
HDAC7 FILGSPHTPLFLPHGLEPEAG-GTLPSRLQFILLDFSGSHAPLLTVPGLGFLPHFAQS
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 LVTGERVATSMRTVGRKLPFRHPLERTQSSPLPQSPQALQQLVMQOHOOFLEKQKQ----
HDAC6 LVTGERVATSMRTVGRKLPFRHPLERTQSSPLPQSPQALQQLVMQOHOOFLEKQKQ----
HDAC4 LVGADRVSFSIH--KLQHRPLGRTQSAPLPQNAQALQHLVIQOHOOFLEKHKQOPOQ
HDAC7 LMITERLS-----GSLHWPLERTREPLPPSATAPPPGPMQPRLEQLKTHVQ----
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

```

WO 02/50285

12/29

PCT/EP01/14928

HDAC5 QQLQLGRILITKTGELPQPTTHPEETEELTQQQVLLGEGALTMPEGSTESSTQEDL
 HDAC6 QQLQLGRILITKTGELPQPTTHPEETEELTQQQVLLGEGALTMPEGSTESSTQEDL
 HDAC4 QQLQMWRIJFAPSELPAQFESHPETEELTLEBQ-ALLDEPYLDRLPQCKEAMHAGVGV
 HDAC7 ---VIRKSAFSEKFLAQIPSAEDLETDGGGPGQVVDGSLHRELGHGQPEARGAPL
 HDAC1 -----
 HDAC2 -----
 HDAC3 -----
 HDAC8 -----
 HDAC9 -----

 HDAC5 FEEDDEEDGEEEDCIQVKDEEGSGAEEGPDLEEPGAGYKLFSDAQPLQPLQVYQAPL
 HDAC6 FEEDDEEDGEEEDCIQVKDEEGSGAEEGPDLEEPGAGYKLFSDAQPLQPLQVYQAPL
 HDAC4 KQEP1ESDEEAE-----PPREVEPGQRQ-PSEQELLFRQKALLLEQCRHQCRNYQASH
 HDAC7 QQHPQVLLMEQQR-----LAGRLPRGSTGDTVLLPLAQGGHRLERKQ-----SEPA
 HDAC1 -----
 HDAC2 -----
 HDAC3 -----
 HDAC8 -----
 HDAC9 -----

 HDAC5 SLATVP-----HQALGRTOSSPAAPGGMKSPFPDPVKHL-FTTGVTYDTFMLRHQCMCGN
 HDAC6 SLATVP-----HQALGRTOSSPAAPGGMKSPFPDPVKHL-FTTGVTYDTFMLRHQCMCGN
 HDAC4 YAAQIFVSGGGRPLSRQSPASATFPVSVQEPFTRPR-FTTGLVYDTMLRHQCTCGS
 HDAC7 APASLE-----APEPASQARVLSSSETPARTLFTTGLYDSVMLRHQCSGCD
 HDAC1 -----HAGTQG-TRKVCYTYDGDVGNHYGQ
 HDAC2 -----HAYSQGGGKJKVCTYDGDVGNHYGQ
 HDAC3 -----HAKTVATYDGDVGNHYGQ
 HDAC8 -----MEEPPEADSGQSLVPVYIYSPETVSMCD
 HDAC9 -----MGTAIVTHEDMTATRLMDD

 HDAC5 THVHPERAGRIQSIWRLQETGLSKCERIRGRATLDEIQTVHSEYHLLYGTSPINRQ
 HDAC6 THVHPERAGRIQSIWRLQETGLSKCERIRGRATLDEIQTVHSEYHLLYGTSPINRQ
 HDAC4 SSSHPERAGRIQSIWRLQETGLSKCERIRGRATLDEIQTVHSEYHLLYGTSPINRQ
 HDAC7 NSKHFERAGRIQSIWRLQETGLSKCERIRGRATLDEIQTVHSEYHLLYGTSPINRQ
 HDAC1 G--HPMRPHRIKNTIDLLNYGLYKMEIYEPHRAAEEMTKYHSEDIYKFLRESIRPDM
 HDAC2 G--HPMRPHRIKNTIDLLNYGLYKMEIYEPHRAAEEMTKYHSEDIYKFLRESIRPDM
 HDAC3 G--HPMRPHRIKNTIDLLNYGLYKMEIYEPHRAAEEMTKYHSEDIYKFLRESIRPDM
 HDAC8 S--LAKIPRASHVSLIYAYALHQMRIYKPKVASMEEMATPHTDAYLQHLQKVSQGD
 HDAC9 FECEIERPERLTAAIDRLRQGLRCLRLSAREASEELGLVHSEFYVSLVRETQVLGK

 HDAC5 KLDEKLLGPISQMYAVLPCCGIQVDEPTVWNEHSSAVRMVAGCLLELAFFKVAAGEL
 HDAC6 KLDEKLLGPISQMYAVLPCCGIQVDEPTVWNEHSSAVRMVAGCLLELAFFKVAAGEL
 HDAC4 KLDEKLLGLAS-LVPVRLPCGGVGVDEDTIWNELHSSAAKLAAGCVVELVFKVATGEL
 HDAC7 KLDEKLLAGLLARWFEMLPCCGVGVDTITWELHSSAAKLAAGCVVELVFKVATGEL
 HDAC1 SE-----YKQMRPNVGEDCFVFDGLFEFCQLSTGGSVASAVKLNKQOTDIAYDM
 HDAC2 SE-----YKQMRPNVGEDCFVFDGLFEFCQLSTGGSVASAVKLNKQOTDIAYDM
 HDAC3 QG-----FTRELNAFNVGDCPVFGLFEFCERTGASLQCATQLNNKI CDIAINW
 HDAC8 ED-----HPDSIE-YGLGYDCPATEGIDYAAAGATITAAQCLIDGMCVKAINW
 HDAC9 EE-----LQALSGQFDAIYFHPFTHCARLAAGGLQLVDAVLTGAV

 HDAC5 KNGFAIIRFPCHRAEESTAMGFCFFNSVAITAKILQQR---LNVGKVLIVDWDIHRHNGGT
 HDAC6 KNGFAIIRFPCHRAEESTAMGFCFFNSVAITAKILQQR---LNVGKVLIVDWDIHRHNGGT
 HDAC4 KNGFAVVRFPCHRAEESTPMGFCFFNSVAVAAKILQQR---LSVSKLIVDWDVHRHNGGT

[illegible]

WO 02/50285

14/29

PCT/EP01/14928

HDAC8 -----
HDAC9 RTAVALTTPDITLVLPDVIQQEASALREETEAWARPHESLAREEALTALGKLLYLIDGM

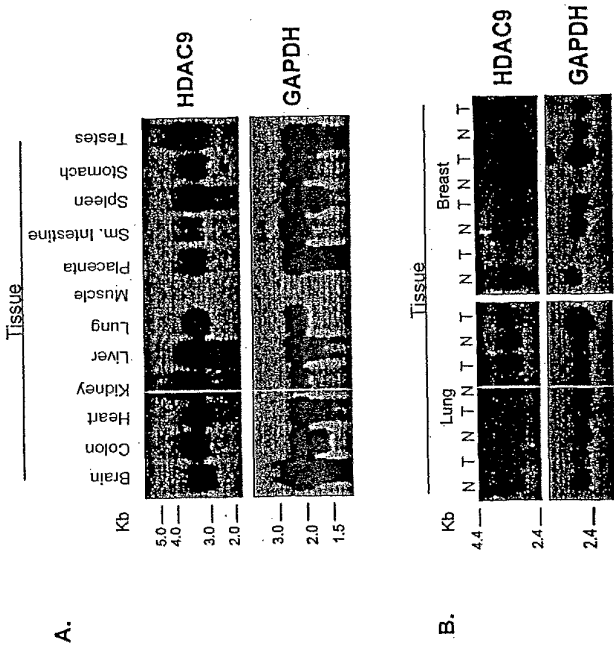
HDAC5 -----
HDAC6 -----
HDAC4 -----
HDAC7 -----
HDAC1 -IKPEARGVHEEYKLA-
HDAC2 GERTDTGTKEPQLGNP-
HDAC3 -RESUVEI-
HDAC8 -----
HDAC9 LDGQVNSGIAATPASAAAATLDVAVRRGLSHGAQELLCVALGQLDRPPDLANDGRSLNLM

HDAC5 -----
HDAC6 -----
HDAC4 -----
HDAC7 -----
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 IRGEAAALSMTFVSTPLPVMTGGFLSCILGLVPLATGFPQDVLVALGPGHGLQGPHA

HDAC5 -----
HDAC6 -----
HDAC4 -----
HDAC7 -----
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 ALLAAMLRGLAGGRVLALLEENSTPQLAGILARVLNGEAPFSLGPSSVASPEDVQALMYL

HDAC5 -----
HDAC6 -----
HDAC4 -----
HDAC7 -----
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 RGGLEFPQWRLQCCHPLVA

Figure 7



WO 02/50285

16/29

PCT/EP01/14928

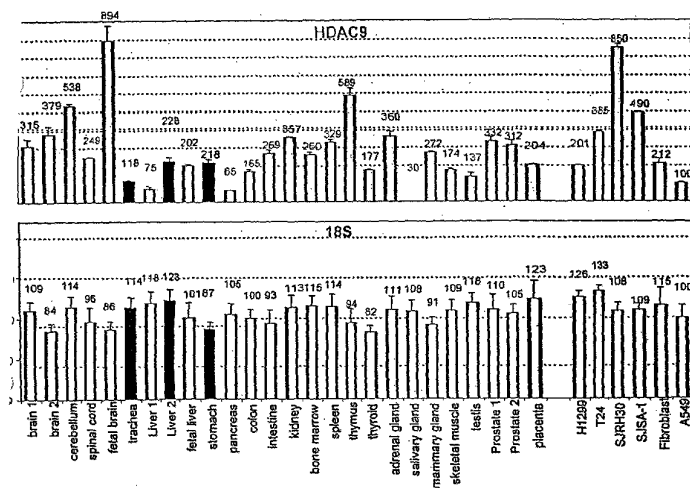


Figure 8

WO 02/50285

17/29

PCT/EP01/14928

Figure 9.

```

Sequence format is Pearson
Sequence 1: HDAC6catalyticdomain      336 aa
Sequence 2: HDAC5catalyticdomain      329 aa
Sequence 3: HDAC6catalyticdomain1     302 aa
Sequence 4: HDAC6catalyticdomain2     481 aa
Sequence 5: HDAC7catalyticdomain      334 aa
Sequence 6: HDAC5completepeptide     673 aa
Start of Pairwise alignments
Aligning...
Sequences (1:2) Aligned. Score: 78
Sequences (1:3) Aligned. Score: 41
Sequences (1:4) Aligned. Score: 45
Sequences (1:5) Aligned. Score: 75
Sequences (1:6) Aligned. Score: 37
Sequences (2:3) Aligned. Score: 42
Sequences (2:4) Aligned. Score: 44
Sequences (2:5) Aligned. Score: 72
Sequences (2:6) Aligned. Score: 37
Sequences (3:4) Aligned. Score: 49
Sequences (3:5) Aligned. Score: 41
Sequences (3:6) Aligned. Score: 55
Sequences (4:5) Aligned. Score: 46
Sequences (4:6) Aligned. Score: 41
Sequences (5:6) Aligned. Score: 38
Guide tree      file created: [/home/nv/software/biobenchw/tmp/align/3664.dnd]
Start of Multiple Alignment
There are 5 groups
Aligning...
Group 1: Sequences: 2      Score: 6517
Group 2: Sequences: 3      Score: 6370
Group 3: Sequences: 4      Score: 4601
Group 4: Sequences: 2      Score: 5205
Group 5: Sequences: 6      Score: 4795
Alignment Score: 35000
CLUSTALW-Alignment: file created: [/home/nv/software/biobenchw/tmp/align/3664.out]
CLUSTAL W (1.81) multiple sequence alignment

```

CYASLGDPPPLLTLPPLFLS---GALASITETIQVDRFWSLRVHDK
FYQILLGDPPPLSGTHAPCQRCEGLAESIQSARAAQHPNKSILQQDPW

WO 02/50285

19/29

PCT/EP01/14928

```

HDA C4 catalytic domain -----
HDA C5 catalytic domain -----
HDA C7 catalytic domain -----
HDA C6 catalytic domain2 CTRSLLGPPPLLLTLPPPLS-----GALASITETIQVIRKPVRSLEFVDRV
HDA C6 catalytic domain3 TVQTLLGDPAPPLSGVWAPYQKCTERSALESTQSAFAAQJYHNSLQQQRV
HDA C9 complete peptide -----

HDA C4 catalytic domain -----
HDA C5 catalytic domain -----
HDA C7 catalytic domain -----
HDA C6 catalytic domain2 EDREQSSSSKLVTNKAQVPAJNRLAERHTLRKQVLEAG-----
HDA C6 catalytic domain3 IAVVHSPSSHSPEGRPPPLFGEPSCAAAAPSSLLDQDCLGKPSVDR
HDA C9 complete peptide -----

HDA C4 catalytic domain -----
HDA C5 catalytic domain -----
HDA C7 catalytic domain -----
HDA C6 catalytic domain2 -----NCKVTSASFQESFPGTSTSEFAVVALTDQPSKATGSAI
HDA C6 catalytic domain3 AVALLTVETLVLFPVJQGLASALRETEAWARPHSELAAREALYALGI
HDA C9 complete peptide -----

HDA C4 catalytic domain -----
HDA C5 catalytic domain -----
HDA C7 catalytic domain -----
HDA C6 catalytic domain2 LAQISLAAJUGAHGQTTSEAVGGAEPQTTSSTVYQGAHL-----
HDA C6 catalytic domain3 ILYLLDGLFGVNSGLATPASAAATLVAVRGLSHGQRLCVALE
HDA C9 complete peptide -----

HDA C4 catalytic domain -----
HDA C5 catalytic domain -----
HDA C7 catalytic domain -----
HDA C6 catalytic domain2 -----
HDA C6 catalytic domain3 QLRPPDLAIDGRSLVLRNGKKAALSMFVNSYPLFVHTGGTSGTLLGL
HDA C9 complete peptide -----

HDA C4 catalytic domain -----
HDA C5 catalytic domain -----
HDA C7 catalytic domain -----
HDA C6 catalytic domain2 -----
HDA C6 catalytic domain3 VLPAYGTQDVLVVALGFGHGLGPPHALLVARRGLAGORVLALEEN
HDA C9 complete peptide -----

HDA C4 catalytic domain -----
HDA C5 catalytic domain -----
HDA C7 catalytic domain -----
HDA C6 catalytic domain2 -----
HDA C6 catalytic domain3 STTQLRGTLAHVLIHGLAPPSLGPSSVASPEPVQALHTLHPLQEPQVGGQ
HDA C9 complete peptide -----

HDA C4 catalytic domain -----
HDA C5 catalytic domain -----
HDA C7 catalytic domain -----
HDA C6 catalytic domain2 -----
HDA C6 catalytic domain3 CHPHLVN
HDA C9 complete peptide -----

```

WO 02/50285

20/29

PCT/EP01/14928

Figure 10.

```

Sequence format is: Pearson
Sequence 1: HDACcatalyticdomain      310 aa
Sequence 2: HDAC2catalyticdomain     310 aa
Sequence 3: HDAC3catalyticdomain     310 aa
Sequence 4: HDAC8catalyticdomain     308 aa
Sequence 5: HDAC9completepeptide    673 aa
Start of Pairwise alignments
Aligning...
Sequences: (1:2) Aligned: Score: 92
Sequences: (1:3) Aligned: Score: 65
Sequences: (1:4) Aligned: Score: 42
Sequences: (1:5) Aligned: Score: 20
Sequences: (2:3) Aligned: Score: 64
Sequences: (2:4) Aligned: Score: 43
Sequences: (2:5) Aligned: Score: 20
Sequences: (3:4) Aligned: Score: 42
Sequences: (3:5) Aligned: Score: 21
Sequences: (4:5) Aligned: Score: 19
Guide tree      file created: [p/bin/mw/software/biobenchsw/tmp/align/3.650.dnd]
Start of Multiple Alignment
There are 4 groups
Aligning...
Group 1: Sequences: 2      Score: 6624
Group 2: Sequences: 3      Score: 5628
Group 3: Sequences: 4      Score: 4878
Group 4:                  Delayed
Sequence 5      Score: 2573
Alignment Score: 8332
CLUSTALW alignment file created: [p/bin/mw/software/biobenchsw/tmp/align/3.650.out]
CLUSTAL W (1.81) multiple sequence alignment

```

[illegible]

WO 02/50285

22/29

PCT/EP01/14928

HDA C1 catalytic domain
HDA C2 catalytic domain
HDA C3 catalytic domain
HDA C4 catalytic domain
HDA C5 complete peptide
KDAALSGTHVSTPLPQHTGGLSTGLQLVFLAYTGQPDLYVALGPGHGLQEPHAAI

HDA C1 catalytic domain
HDA C2 catalytic domain
HDA C3 catalytic domain
HDA C4 catalytic domain
HDA C5 complete peptide
KAMLRGLAGGPGVALLCHSTPQLAGLLAPVLRGAPPSELGTSVASTEPDVALMTHG

HDA C1 catalytic domain
HDA C2 catalytic domain
HDA C3 catalytic domain
HDA C4 catalytic domain
HDA C5 complete peptide
LEFQHTGGLQCHPGLVA

WO 02/50285

23/29

PCT/EP01/14928

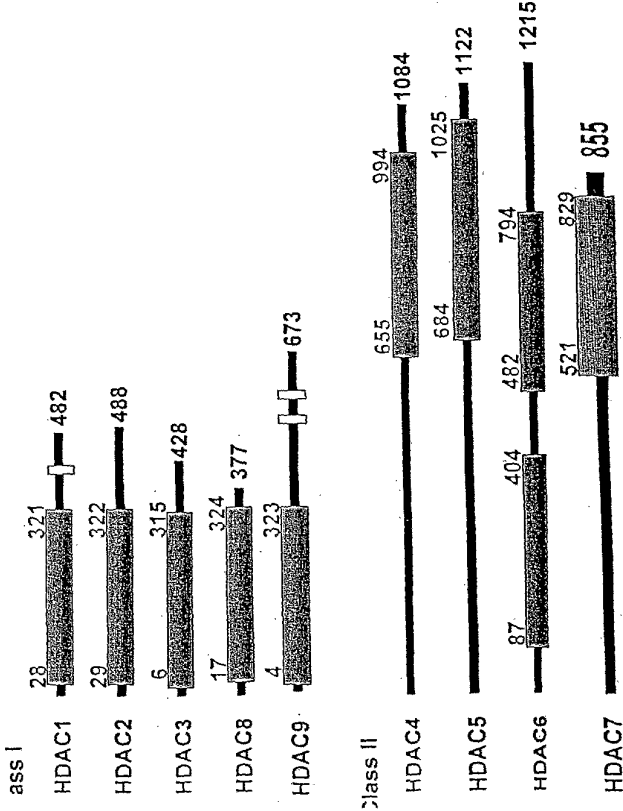
Figure 11a

```

HDAC9v1  MGTALVYHEDMTATRLWDDPECEIERFERLTAALDRLRQGLEQRCRLRSAREASEEEL
HDAC9v2  MGTALVYHEDMTATRLWDDPECEIERFERLTAALDRLRQGLEQRCRLRSAREASEEEL
HDAC9v3  MGTALVYHEDMTATRLWDDPECEIERFERLTAALDRLRQGLEQRCRLRSAREASEEEL
*****
HDAC9v1  GLVHSPEYVSLVRETQVLGKEELQALSGQFDIYFHPSTFHCARLAAGAGLQLVDVLTG
HDAC9v2  GLVHSPEYVSLVRETQVLGKEELQALSGQFDIYFHPSTFHCARLAAGAGLQLVDVLTG
HDAC9v3  GLVHSPEYVSLVRETQVLGKEELQALSGQFDIYFHPSTFHCARLAAGAGLQLVDVLTG
*****
HDAC9v1  AVQNGLALVRPFGHHGQRAAANGFCVFNNAIAAAHAKQKHSGLHRILVVDWDVHHGQGIQ
HDAC9v2  AVQNGLALVRPFGHHGQRAAANGFCVFNNAIAAAHAKQKHSGLHRILVVDWDVHHGQGIQ
HDAC9v3  AVQNGLALVRPFGHHGQRAAANGFCVFNNAIAAAHAKQKHSGLHRILVVDWDVHHGQGIQ
*****
HDAC9v1  YLEDDPSVLYFSWHRYEHGRFWPFLRESADAVGRGQGLGFTVNLPMNQVGMGNADYVA
HDAC9v2  YLEDDPSVLYFSWHRYEHGRFWPFLRESADAVGRGQGLGFTVNLPMNQVGMGNADYVA
HDAC9v3  YLEDDPSVLYFSWHRYEHGRFWPFLRESADAVGRGQGLGFTVNLPMN-----
*****
HDAC9v1  AFLHLLPLAFEFDPPELVLSAGFDSAIGDPEGQMORTPECF AHLTQLLQVLAGGRVCAV
HDAC9v2  AFLHLLPLAFEFDPPELVLSAGFDSAIGDPEGQMORTPECF AHLTQLLQVLAGGRVCAV
HDAC9v3  -----QFDPELVLSAGFDSAIGDPEGQMORTPECF AHLTQLLQVLAGGRVCAV
*****
HDAC9v1  LEGGYHLESIAESVCMVTQTLGDPAPPLSGPMAPCORCEGSALESIQSARAAQAPHWKS
HDAC9v2  LEGGYHLESIAESVCMVTQTLGDPAPPLSGPMAPCORCEGSALESIQSARAAQAPHWKS
HDAC9v3  LEGGYHLESIAESVCMVTQTLGDPAPPLSGPMAPCORCEGSALESIQSARAAQAPHWKS
*****
HDAC9v1  LQQQDVTAVPMSFSSHSPEGRPPPLPGGPVCKAAASAPSSILLQPCCLCPAPSVRTAVAL
HDAC9v2  LQQQDVTAVPMSFSSHSPEGRPPPLPGGPVCKAAASAPSSILLQPCCLCPAPSVRTAVAL
HDAC9v3  LQQQDVTAVPMSFSSHSPEGRPPPLPGGPVCKAAASAPSSILLQPCCLCPAPSVRTAVAL
*****
HDAC9v1  TTPDITLVLPDVIQQA-----
HDAC9v2  TTPDITLVLPDVIQQAASALREETEAWARPHESLAREEALTALGKLLYLLDGMLDQGVN
HDAC9v3  TTPDITLVLPDVIQQAASALREETEAWARPHESLAREEALTALGKLLYLLDGMLDQGVN
*****
HDAC9v1  -----
HDAC9v2  SGIAATFASAAAATLDVAVRRGLSHGAQRLLCVLGLQDLRPPDLAHDGRSLWLNIRGKEA
HDAC9v3  SGIAATFASAAAATLDVAVRRGLSHGAQSWGVGEGLEAMPGGSPAQRLLSSHSTPAHGVP
*****
HDAC9v1  -----CILGLVLPPLAYGFQPDVLVVALGPGHGLQGPHAALLAAM
HDAC9v2  AALSMFHVSTFLPVMTGGFLSCILGLVLPPLAYGFQPDVLVVALGPGHGLQGPHAALLAAM
HDAC9v3  NALPFLFLRFGLRRMTGGFLSCILGLVLPPLAYGFQPDVLVVALGPGHGRAPTLHSLWLC
*****
HDAC9v1  LRGLAGGRVLALLEENSTPQLAGILARVLNGEAPPSLGLSSVASPEDVQALMYLRGQLEP
HDAC9v2  LRGLAGGRVLALLEEVSWAGWR--CCGVGRGKGP--VTASVFAPGPELHTPASRDPGGA
HDAC9v3  FGGWQG-----AESWPSWR-----RGRPGPYVPERAAGASVEDVAVPSSPGGLKSA
*****
HDAC9v1  QWKMLQCHPHLVA
HDAC9v2  EWRGTS-----
HDAC9v3  K-----

```

Figure 11B



WO 02/50285

25/29

PCT/EP01/14928

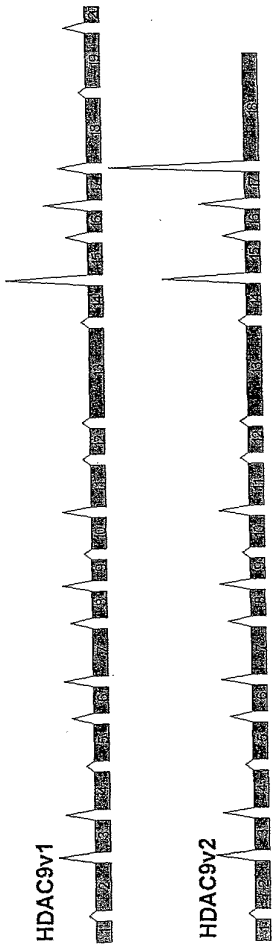


Figure 11C

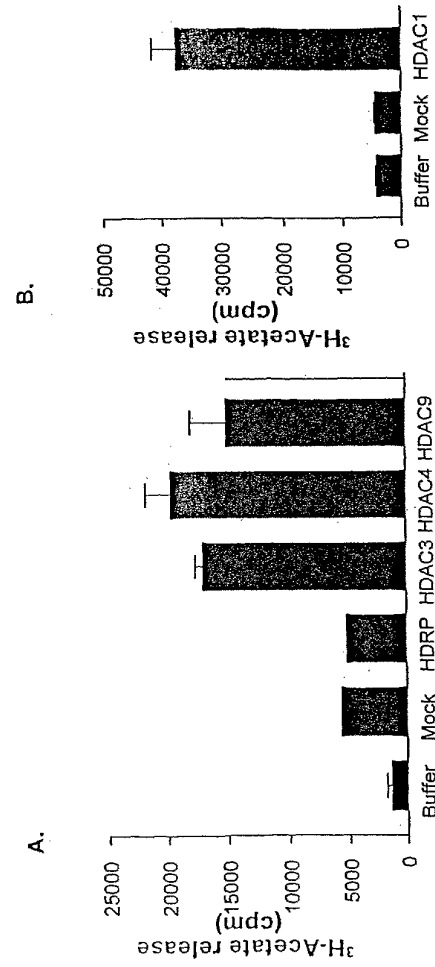


Figure 12B

Figure 12A

WO 02/50285

27/29

PCT/EP01/14928

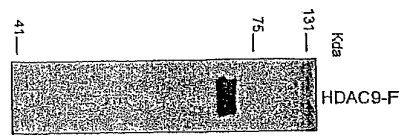


Figure 13

WO 02/50285

28/29

PCT/EP01/14928

Figure 14

SEQ ID NO:7

>HDAC9v2 DNA sequence

```
1 ATGGGGACCGCGCTTGTATCCATGAGGACATGACGGCCACCCGGCTGCTCTGGGACGAC
61 CCCGAGTGCAGATCGAGCGCTCCTGAGCGCCTGACCGCAGCCCTGGATCGCCTGCGGCAG
121 CGCGGCTTGGACAGAGGTGTCTGCGGTTGTGAGCCGCGAGGCTTCGGAAGAGGAGCTG
181 GGCCTGGTGCACAGCCAGAGTATGTATCCTGGTCAGGGAGACCCAGGTCTAGGCAAG
241 GAGGAGTGCAGGCGCTGTCGGGACGTTCGAGCCATCTACTTCCACCGAGTACCTTT
301 CACTGCGCGGCGCTGGCCGCGAGGGCTGGAGTGCAGCTGGTGACGCTGTGCTCACTGGA
361 GCTGTGCAAAATGGGCTTGCCTGGTGGGCTCCCGGCACCATGGCCAGAGGGCGGCT
421 GCCAACGGGTTCTGTGTGTTCAACAGTGGCCATAGCAGCTGCACATGCCAAGCAGAAA
481 CACGGGCTACACAGGATCCTGCTGAGTGGGATGTGACCATGGCCAGGGGATCCAG
541 TATCTCTTTGAGGATGACCCAGCGCTTTACTTCTCCTGGCACCGCTATGAGCATGGG
601 CGCTTCTGGCCTTTCTCGGAGAGTCAGATGCAGACGAGTGGGGCGGGGACAGGGCCTC
661 GGCTTCACTGTCAACCTGCCCTGGAAACAGTTGGGATGGGAAACGCTGACTACGTGGCT
721 GCCTTCTGACATGTCTGCTCCACCTGGCCTTTGAGTTTGACCTTGAGCTGGTGGTGGT
781 TCGGCAGGATTTGACTCAGCCATCGGGGACCTGAGGGGCAATGCAAGGCCACGCCAGAG
841 TGCTTCGCCACCTCACACAGTCTGCTGAGGTGCTGGCGCGCGCGGCTGTGTGCCGTG
901 CTGGAGGGCGGCTACCACTGGAGTCACTGGCGGAGTCAGTGTGCATGACAGTACAGAG
961 CTGCTGGGTGACCCGGCCACCCCTGTGAGGGCCATGGCCCATGTGACAGGTGCGAG
1021 GGGAGTGCCCTAGAGTCCATCCAGAGTGCCTGGTGGTGGCCAGGCCCGCACTGGAAAGAG
1081 CTCACGACAGAGATGTGACCGCTGTGCGGATGAGCCCGAGCAGCCACTCCCGAGAGGG
1141 AGGCCTCCACCTCTGCTGCTGGGGGTCAGTGTGTAAAGGACGCTGCATCTGCACCGAGC
1201 TCCTCCTGGACAGCGCTGCTGCTGCCCCGACCCCTGTGTCCGACCGCTGTTCCTG
1261 ACAAGCGCGGATATCACATTGGTTCTGCCCCCTGACGTCATCCACAGGAGCTGACCC
1321 CTGAGGAGGAGACAGAAAGCCTGGGCGAGGCCACAGAGTCCCTGGCCCGGAGGAGGCC
1381 CTCACTGGACTTGGGAAGCTCCTGTACCTCTTAGATGGGATCCCTGGCCCGGAGGAGGCC
1441 AGTGGTATAGCAGCCACTCCAGCTCTGCTGCTGAGTGGGATGTGCTGTTCGG
1501 AGAGGCTGTGCCCACGGAGCCCGAGGCTGCTGTGCGTGGCCCTGGGACAGCTGGAGCGG
1561 CCTCCAGACTCGCCATGACGGGAGGAGTCTGTGGCTGACATCAGGGGCAAGGAGGCG
1621 GCTGCCCTATCCATGTTCCATGCTGCCACGCACTGCCAGTGATGACCGGTGGTTTCTG
1681 AGTGGCATCTTGGGCTTGGTGTGCTGCCCTGGCCCTATGCTTCCAGCTGACCTGGTGTG
1741 GTGGGCTTGGGCTTGGCCATGGCTGCGAGGCCCGCCACGCTGCACTCTGGTGCAATG
1801 CTTCCGGGGCTGGCAGGGGCGAGTCTTGGCCCTCTGGAGGAGTTAAGTGGGAGGG
1861 TGGAGGTGCTCGGGGCTGGGACGAGGGAGGACCACTGACTGCTTCCGTCTTCCGCCCT
1921 GTTCAGAACTCCACACCCAGCTAGCAGGATCTTGGCCCGGTGCTGAATGGAGAGGC
1981 ACCTCTAGCTAGGCTTCTCTGTGGCTTCCCGAGGAGGCTCAGGCCCTGATGTA
2041 CTTGAGAGGCGAGCTGGAGCCTCAGTGGAGATGTTGAGTGCATCTCTACCTGGTGGC
2101 TTGA
```

SEQ ID NO:5

>HDAC9v2 peptide sequence

```
1 MGTALVYHEDMTATRLWDDPECEIERPERLTAALDRLRQGLRCLRLSAREASEEEL
61 CLVHSPEYVSLVRETOVLKKEELOALSGQFDAYFHPSTFHCARLAAGAGLQVDAVLG
121 AVGNGLALVRPPGHHGQRAAANGPCVFNNVAIAAHAKQKHGLHRLVVDVHGGQGIQ
181 YLEDDPSVLYFSWHRYEHRFWPFLRESADAVGRGGLGFTVNLWNQVGMGNADYVA
241 AFLHLLPLAFDFPRLVLSAFDSATGDEPGMQATPECFHLLTOLLQVLAGGRVCAV
301 LEEGYHLESLAESVCMVTQTLGDPAPPLSGPMAPOQRCEGSALESIQSARAQAAPHWKS
361 LQQQDVTAVPMSFSSHSPEGRFPPLLEGGPVCKAAASAPSSLLDQCLCPAPSVRTAVAL
421 TTPDTTLVLPDPVITQEEASALREETAWARPHESLAREEALTALGRKLYLLDGMLDQVN
```

WO 02/50285

29/29

PCT/EP01/14928

481 SGIAATPASAAAATLDVAVRRGLSHGAQRLLCVLGLQLDORPDLAHDGRSLNLRNGKEA
 541 AALSMFHVSTPLPVMGTGFLSCIILGLVLPLAYGFQPDVLVLVAGPBGHLQGPAAALLAAM
 601 LRGLAGGRVLALLLEEVSWAGWRCGVGRKGFPVTASVFAPGPELHTPASRDPCPGAEWRG
 661 TS

SEQ ID NO:8

>HDAC9v3 DNA sequence

1 ATGGGGACCGCGCTTGTGTACCATGAGGACATGACGGCCACCGCGCTGCTCTGGGACGAC
 61 CCCGAGTGCAGATCGAGCGTCTGAGCGCTGACCGCAGCCCTGGATCGCTGCCGCGAG
 121 CGCGGCTTGGAAACAGAGGTGTCTGCGGTGTTCAGCGCGCAGGCTCGGAAGAGGAGCTG
 181 GGCCTGGTGCACAGCCCAAGTATGTATCCCTGGTCAGGGAGACCCAGGTCTTAGGCAAG
 241 GAGGAGCTGCAGGCGCTGTCCGAGCAGTTCGACGCCATCTACTTCCACCCGAGTACCTTT
 301 CACTGCGCGCGCTGCGCCGAGGGGCTGAGCTGCGAGCTGGTGGACGCTGTGCTCACTGGA
 361 GCTGTGCAAAATGGGCTTGGCTGGTGGAGGCTCCCGGGCACCATGCGCAGAGGGCGGCT
 421 GCCAACGGGTTCTGCGTGTTCACCAACGTGGCCATAGCAGCTGCACATGCCAAGCAGAAA
 481 CACGGGCTACACAGGATCTCTGCTGGTGGACTGGGATGTGCACCATGGCCAGGGGATCCAG
 541 TATCTCTTGGAGGATGACCCAGCGTCTTTACTTCTCTGCGCACCGCTATGAGCATGGG
 601 CGCTTCTGGCTTTCTCGAGAGTCAGATGACGACGAGTGGGGCGGGGACAGGGCCCTC
 661 GGCTTCACTCAACCTGCGCTGGAAACGAGTTTGAACCTGAGCTGGTGTGCTCTCGGCA
 721 GGATTTGACTCAGCCATCGGGGACCTGAGGGGCAAAATGCAAGCCACGCCAGAGTCTTC
 781 GCCACCTCACACAGCTGCTGAGGTGCTGGCGGGCGGGCGGGTCTGTGCCGTGCTGGAG
 841 GCGGCTACCACTGGAGTCACTGGCGGAGTCACTGTCATGACAGTACAGACGCTGCTG
 901 GGTGACCCGCGCCACCCCTGTGAGGCGCAATGGCGCCATGTCAAGGTGCGAGGGGAGT
 961 GCCCTAGAGTCCATCCAGAGTCCCGTGTGCTGCCAGGCGCCGCACTGGAGAGAGCTCCAG
 1021 CAGCAAGATGTACCGCTGTGCCGATGAGCCCGCAGCCCACTCCCGCAGAGGGAGGCT
 1081 CCACCTCTGCTGCGCTGGGGTCCAGTGTGTAAAGGAGCTGCATCTGCACCGAGCTCCCTC
 1141 CTGACCCAGCCGCTGCTGCGCCGACCCCTCTGCGCACCGCTGTGCCCTGACCAAG
 1201 CCGGATATCAATTGGTTCTGCCCTGAGCTCATCCAAAGGAGCGTCAAGCTGAGG
 1261 GAGGAGACAGAGCTGGGCGAGGCCACAGTCTCCGCGCGGAGGAGGCGCTCACT
 1321 GCACCTGGGAGGCTCCTGTACTCTTAGATGGGATGCTGGATGGGAGGCTGACAGTGGT
 1381 ATAGCAGCCACTCCAGCTCTGCTGACAGCAGCCACCTGGATGTGGCTGTTCGAGAGGCT
 1441 CTGTCCACGAGGCCAGAGCTGGGTGTGGGAGAGGGCTGCTGGAGGCAATGCCAGGT
 1501 GGTCTCCAGCACAGAGGCTCAGAGTCAACAGCACCCCTGCCATGGCCCGCTGAATGCT
 1561 TTCTCCACCTCTGCTCTGCGGTTTGGGCTCAGGAGGATGACCGGTGGTTTCTGAGCTGC
 1621 ATCTGGGCTTGGTGTGCGCTGCGCTGAGCTTACGCTTCCAGCTGACCTGGTGTGGTGGG
 1681 CTGGGGCTGGCCATGGCTGCGAGGCCCCACGCTGCCTCTGCTGCAATGCTTCGGG
 1741 GGTGGCAGGGGCGGAGTCTTGGCTCTCTGGAGGAGGAGCGTCCAGGCGCTTATGTA
 1801 CCTGAGAGGGCAGCTGGAGCTCTAGTGAAGATGTTGAGTGCCATCTCACCTGGTGGC
 1861 TTGAAATCGGCCAAG

SEQ ID NO:6

>HDAC9v3 peptide sequence

1 MGTALVYHEDMTATRLWDDPBCRIERPERLTAALDLRQRLGLRCLRLSAREASEEEL
 61 GLVHSPEYSLVRETVLGLKEELQALSGQFDIYFHFSTFHCARLAAGALQLVLDVLTG
 121 AVQNGLALVRPPGHHGQRAAANGFCVFNVAIAAAHAKQKHGLHRLVVDWVHHGGQIQ
 181 YLFEDDPSVLYFVSWHYEHRFWFLRESADAVGRGQGLFTVNLWNQDPFLVLVSA
 241 GFDSAIGDPEGQMQATPECFARLTQLQLVLAGGRVCAVLEGGYHLESIAESVCMVTQTL
 301 GDPAPLSPGPMAPCQRCEGSALESIQSARAAQAPHWKSLLQQDVTAVFMSPSHSPGRRP
 361 PPLPFGSPVCKAASAFSSLLDQPLCPAPSVRTAVALTTPDITLVLFPDVIQDEASALR
 421 ETEANARPHRESLAREEALTALGKLLYLLDGMLDQVNSGIAATPASAAAATLDVAVRRG
 481 LSHGAQSWGVGEGLEAMPGGSPAQRLLSHSTPAHGPNALPLPLRFLRRMTGGFLSC
 541 ILGLVLPLAYGFQPDVLVLVAGPBGHLQGPAAALLAAM
 601 PERAAGASVREDVAVPSSPGGLKSAK

【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)

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
27 June 2002 (27.06.2002)

PCT

(10) International Publication Number
WO 02/050285 A2

(51) International Patent Classification: C12N 15/55, 9/16, C07K 14/47 (74) Agent: BECKER, Konrad; Novartis AG, Corporate Intellectual Property, Patent and Trademark Department, CH-4002 Basel (CH).

(21) International Application Number: PCT/EP01/14928

(22) International Filing Date: 18 December 2001 (18.12.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/256,827 20 December 2000 (20.12.2000) US
60/293,089 23 May 2001 (23.05.2001) US
60/317,984 6 September 2001 (06.09.2001) US

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

(84) Designated States (regional): Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, HU, IT, LU, MC, NL, PT, SE, TR).

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

(71) Applicant (for all designated States except AT, US): NOVARTIS AG [CH/CH]; Lichstrasse 35, CH-4056 Basel (CH).

(71) Applicant (for AT only): NOVARTIS PHARMA GMBH [AT/AT]; Brunner-Strasse 59, A-1230 Vienna (AT).

(48) Date of publication of this corrected version: 10 April 2003

(72) Inventors; and

(75) Inventors/Applicants (for US only): COHEN, Dalia [IR/US]; 236 East Cedar Avenue, Livingston, NJ 07039 (US); BHATTIA, Umesh [US/US]; 5212 Union Avenue, San Jose, CA 95124 (US); CAI, Richard, Lie [CN/US]; 367 Pond Road, Bridgewater, NJ 08807 (US); FISCHER, Denise, Dawn [US/US]; 17 Woodland Road, Bernardsville, NJ 07924 (US).

(15) Information about Correction:
see PCT Gazette No. 15/2003 of 10 April 2003, Section II

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



WO 02/050285 A2

(54) Title: HISTONE DEACETYLASE-RELATED GENE AND PROTEIN

(57) Abstract: Disclosed is an HDAC related genes and gene products. In particular, the invention relates to a protein and variants that is highly homologous to known HDACs and referred to herein as HDAC9, nucleic acid molecules that encode such a protein, antibodies that recognize the protein, and methods for diagnosing conditions related to abnormal HDAC9 activity or gene expression.

WO 02/050285

PCT/EP01/14928

1

HISTONE DEACETYLASE - RELATED GENE AND PROTEIN**FIELD OF THE INVENTION**

This invention relates to a histone deacetylase gene and gene product. In particular, the invention relates to a protein that is highly homologous to known yeast histone deacetylase I (hda1) class II histone deacetylases (HDACs), nucleic acid molecules that encode such a protein, antibodies that recognize the protein, and methods for diagnosing conditions related to abnormal HDAC activity, including, for example, abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, host inflammatory or immune response or psoriasis.

BACKGROUND OF THE INVENTION

Histone acetylation is a major regulatory mechanism that modulates gene expression by altering the accessibility of transcription factors to DNA. Acetylation of histones is a reversible modification of the free Σ -amino group of lysine that occurs during the assembly of nucleosomes and during DNA synthesis. Changes in histone acetylation levels also occur during transcriptional activation and silencing. Acetylation of histones is generally associated with transcriptional activity, whereas deacetylation is associated with transcriptional repression. Histone acetylation levels result from an equilibrium between competing histone acetylases and deacetylases (Emiliani, S., Fischle, W., Van Lindt, C., Al-Abed, Y., and Verdin, E., Proc Nat. Acad. Sci., U. S. A., **95**, 2795-2800 (1998).

HDACs have been shown to play an important role in the regulation of transcription. HDACs function as components of complexes that are involved in transcriptional repression. This is mediated through interactions of HDACs with multi-protein complexes and requires deacetylase activity. HDAC complexes may contain the co-repressor mSin3A (Kasten, M.M., Dorland, S., Stillman, D.J. *Mol. Cell. Biol.* **17**, 4852-4858 (1997)) and mSin3A-associated proteins (Zhang, Y., Iratni, R., Erdjument-Bromage, H., Tempst, P., Reinberg, D. *Cell* **89**, 357-364 (1997); Zhang, Y., Sun, Z.W., Iratni, R., Erdjument-Bromage, H., Tempst, P., Hampsey, M., Reinberg, D. *Mol. Cell.* **1**, 1021-1031(1998)) silencing mediators NcoR (Nagy, L., H.-Y. Kao,

SUBSTITUTE SHEET (RULE 26)

WO 02/050285

2

PCT/EP01/14928

- D. Chakravarti, R. J. Lin, C. A. Hassig, D. E. Ayer, S. L. Schreiber, and R. M. Evans (1997) *Cell* **89**, 373-380 and SMRT (Alland, L. et al., *Nature* 387:49-55 (1997); Heinzel, T. et al., *Nature* 387:43-8 (1997)), transcriptional repressors Rb (Brownell, J. E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D. G., Roth, S.Y., and Allis, C. D. (1996) *Cell* **84**, 843-851), Rb-like proteins p107 (Ferreira, R., Magnaghi-Jaulin, L., Robin, P., Harel-Bellan, A., Trouche, D. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 10493-10498) and p130 (Stiegler, P., De Luca, A. Bagella, L., Giordano, A. (1998) *Cancer Res.* **389**, 187-190), Rb-associated proteins (Nicolas, E., Morales, V., Magnaghi-Jaulin, L., Harel-Bellan, A., Richard-Foy, H., Trouche, D. (2000) *J. Biol. Chem.* **275**, 9797-9804, Lai, A., Lee, J.M., Yang, W.M., DeCaprio, J.A., Kaelin, W.G. Jr., Seto, E., Branton, P.E. (1999) *Mol. Cell. Biol.* **19**, 6632-6641), Mad/Max (Laherty, C., W.-M. Yang, J.-M. Sun, J. R. Davie, E. Seto, and R. N. Eisenman. (1997) *Cell* **89**, 349-456), nuclear hormone receptors (Nagy, L., H.-Y. Kao, D. Chakravarti, R. J. Lin, C. A. Hassig, D. E. Ayer, S. L. Schreiber, and R. M. Evans. (1997) *Cell* **89**, 373-380), nucleosome remodeling factors (Xue, Y., Wong, J., Moreno, G.T., Young, M.K., Cote, J., Wang, W. (1998) *Mol. Cell.* **2**, 851-861), methyl-binding proteins (Fuks, F., Burgers, W.A., Brehm, A., Hughes-Davies, L., Kouzarides, T. (2000) *Nat. Genet.* **24**, 88-91, Nan, X., Ng, H.H., Johnson, C.A., Laherty C.D., Turner, B.M., Eisenman, R.N., Bird, A. (1998) *Nature* **393**, 386-389, Ghosh, A.K., Steele, R., Ray, R.B. (1999) *Biochem. Biophys. Res. Commun.* **260**, 405-409, Ng, H. H., Zhang, Y., Hendrich, B., Johnson, C.A., Turner, B.M., Erdjument-Bromage, H., Tempst, P., Reinberg, D., Bird, A. (1999) *Nat. Genet.* **23**, 58-61), and DNA repair machinery proteins (Yarden, R.I., Brody, L.C. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 4983-4988, Cai, R.L., Yan-Neale, Y., Cueto, M.A., Xu, H., Cohen, D. (2000) *J. Biol. Chem.* **275**, 27909-27916). Furthermore, HDAC1 has been found to bind directly to YY1 (Yang, W.-M., Inouye, C., Zeng, Y., Bearss, D., and Seto, E. (1996) *Proc. Natl. Acad. Sci.* **93**, 122845-12850) and Sp1 (Doetzlhofer, A., Rotheneder, H., Lagger, G., Koranda, M., Kurtev, V., Brosch, G., Wintersberger, E., Seiser, C. (1999) *Mol. Cell. Biol.* **19**, 5504-5511) and HDACs 4 and 5 bind to MEF2 (Grozinger, C. M., and Schreiber, S. L. (2000) *Proc. Natl. Acad. Sci.* **97**, 7835-7840). In addition, HDACs have been found together in complexes (Eilers, A.L., Billin, A.N., Liu, J., Ayer, D.E. (1999) *J Biol Chem* **274**, 32750-32756, Grozinger, C. M., and Schreiber, S. L. (2000) *Proc. Natl. Acad. Sci.* **97**, 7835-7840).

30

SUBSTITUTE SHEET (RULE 26)

WO 02/050285

3

PCT/EP01/14928

- Two distinct classes of yeast histone deacetylases have been identified based upon size and sequence. Yeast class I HDACs include Rpd3, Hos1p, and Hos2p. Class II contains yeast HDA1p. Furthermore, members of these two classes were found to form different complexes. Human HDACs have been classified based upon their similarity to yeast sequences. Class I
- 5 human HDACs include HDACs1-3 and 8. Class II HDACs include HDACs 4-7. The deacetylase core of class I HDACs reside in the first ~390 amino acids. Class II HDAC catalytic domains are located in the C-terminal of these peptides, with the exception of HDAC4 that contains a second catalytic domain in the N-terminus (Grozinger, C. M., Hassig, C. A., and Schreiber, S. L. (1999) *Proc. Natl. Acad. Sci. U. S. A.* 96, 4868-4873).
- 10
- An important approach that has been used to study the function of chromatin acetylation is the use of specific inhibitors of histone deacetylase. Several classes of compounds have been identified that inhibit HDAC. Histone deacetylase inhibitors have been found to have anti-proliferative effects, including induction of G1/S and G2/M cell cycle arrest, differentiation
- 15 (Itazaki, H., K. Nagashima, K. Sugita, H. Yoshida, Y. Kawamura, Y. Yasuda, K. Matsumoto, K. Ishii, N. Uotani, H. Nakai, A. Terui, S. Yoshimatsu, Y. Ikenishi and Y. Nakagawa. (1990) *J. Antibiot.* 12, 1524-1532, Hoshikawa, Y., Kijima, M., Yoshida, M., and Beppu, T. (1991) *Agric. Biol. Chem.* 55, 1491-1497, Hoshikawa, Y., Kwon, H.-J., Yoshida, M., Horinouchi, S., and Beppu, T. (1994) *Exp. Cell Res.* 214, 189-197, Sugita, K., Koizumi, K., and Yoshida, H. (1992)
- 20 *Cancer Res.* 52, 168-172, Yoshida, M., Y. Hoshikawa, K. Koseki, K. Mori and T. Beppu. (1990) *J. of Antibiot.* 43, 1101-106, Yoshida, M., Nomura, S., and Beppu, T. (1987) *Cancer Res.* 47, 3688-3691), and apoptosis (Medina, V., Edmonds, B., Young, G. P., James, R., Appleton, S., Zalewski, P. D. (1997) *Cancer Res.* 57, 3697-3707) of transformed and normal cells and reversal of transformation (Kwon, H. J., Owa, T., Hassig, C. A., Shimada, J., and Schreiber, S. (1998)
- 25 *Proc. Natl. Acad. Sci. U. S. A.* 95, 3356-3361, Kim, M.-S., Son, M.-W., Park, Y. I., and Moon, A. (2000) *Cancer Lett.* 157, 23-30). These effects, along with the presence of HDAC in complexes with fusions of unliganded retinoic acid receptors PML-RAR α and PLZF-RAR α indicate a role for HDACs in tumorigenicity (Grignani, F., De Matteis, S., Nervi, C., Tomassoni, L., Gelmetti, V., Cioce, M., Fanelli, M., Ruthardt, M., Ferrara, F. F., Zamir, I., Seiser, C.,
- 30 Grignani, F., Lazar, M. A., Minucci, S., Pelicci, P. G. (1998) *Nature* 391, 815-818, He, L. Z.,

WO 02/050285

PCT/EP01/14928

4

- Guidez, F., Tribioli, C., Peruzzi, D., Ruthardt, M., Zelent, A., Pandolfi, P. P. (1998) *Nat. Genet.*, **18**, 126-35, Lin, R.J., Nagy, L., Inoue, S., Shao, W., Miller, W. H. Jr and Evans, R. M. (1998) *Nature* **391**, 811-814). Furthermore, histone deacetylase inhibitors, phenylbutyrate and trichostatin A have shown promise in the treatment of promyelocytic leukemia and several other
- 5 HDAC inhibitors are being studied and are nearing the clinic (Byrd, J.C., Shinn, C., Ravi, R., Willis, C.R., Wasilenko, J.K., Flinn, I.W., Dawson, N.A., Grever, M.R. (1999) *Blood* **94**, 1401-1408, Kim, Y.B., Lee, K.H., Sugita, K., Yoshida, M., Horinouchi, S. (1999) *Oncogene* **18**, 2461-2470, Cohen, L.A., Amin, S., Marks, P.A., Rifkind, R.A., Desai, D., Richon, V.M. (1999) *Anticancer Res.* **19**, 4999-5005). In addition, the HDAC inhibitor, butyrate was found to decrease
- 10 expression of pro-inflammatory cytokines TNF- α , TNF- β , IL-6, and IL-1 β . These effects are thought to result from inhibition of NFkB activation (Segain JP, Raingeard de la Bletiere D, Bourreille, A., Leray V., Gervois, N., Rosales, C., Ferrier, L., Bonnet, C., Blottiere, H.M., Galmiche, J.P. (2000) Butyrate inhibits inflammatory responses through NFkappaB inhibition: implications for Crohn's disease. *Gut* **47**, 397-403) and its ability to inhibit histone deacetylases
- 15 (Inan M.S., Rasoulpour, R.J., Yin, L., Hubbard, A.K., Rosenberg, D.W., Giardina, C. (2000). The luminal short-chain fatty acid butyrate modulates NF-kappaB activity in a human colonic epithelial cell line. *Gastroenterology* **118**, 724-34).
- 20 The discovery of the HDAC inhibitor trapoxin, made it possible to isolate the first human histone deacetylase, HDAC1, using an affinity matrix column to which a trapoxin-like molecule was bound (Taunton, J., Collins, J. L., and Schreiber, S. (1996) *J. Am. Chem. Soc.* **118**, 10412-10422). Subsequently, seven other human HDAC enzyme isoforms were reported (Taunton, J., Hassig, C. A. and Schreiber, S.L. (1996). *Science* **272**, 408-411, Yang, W. m., Inouye, C., Zeng,
- 25 Y., Bearss, D., and Seto, D. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 12845-12850, Yang, W. M., Yao, y. L., Sun, J. M., Davie, J. R., and Seto, E. (1997). *J. Biol Chem.* **272**, 28001-28007, Emiliani, S., Fischle, W., Van Lint, C., Al-Abed, Y., and Verdin, E. (1998). *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2795-27800). These 8 HDACs have been divided into class I (HDACs 1-3 and 8 similar to the yeast gene Rpd3) and class II HDACs (4-7 similar to yeast gene hda1 (Grozing,
- 30 C. M., Hassig, C.A., and Schreiber, S. L. (1999). *Proc. Natl. Acad. Sci. U.S.A.* **96**, 4983-4988.)

WO 02/050285

5

PCT/EP01/14928

based on sequence homology. Here we report the isolation and characterization of a potential new HDAC, referred to herein as HDAC9, which displays sequence similarity to the hda1 class II HDACs. HDAC9 has characteristics that bridge HDAC class I and class II.

5

SUMMARY OF THE INVENTION

The present invention relates to histone deacetylases, in particular to a novel histone deacetylase HDAC9.

In a first aspect, the invention provides an isolated polypeptide comprising an amino acid sequence as set forth in SEQ ID NO:1, SEQ ID NO 5 or SEQ ID NO 6. Furthermore, the invention provides an isolated polypeptide consisting of an amino acid sequence as set forth in SEQ ID NO:1, SEQ ID NO 5 or SEQ ID NO 6. The amino acid sequence as set forth in SEQ ID NO:1, SEQ ID NO 5 or SEQ ID NO 6 shows a considerable degree of homology to that of known members of the family of HDACs. For convenience, the polypeptide consisting of the amino acid sequence as set forth in SEQ ID NO:1, SEQ ID NO 5 or SEQ ID NO 6 will be designated as histone deacetylase 9 or HDAC9. Such a polypeptide, or a fragment thereof, is expressed in various normal tissues, for example, HDAC9 was present in normal testes, stomach, spleen, small intestine, placenta, liver, kidney, colon, lung, heart, and brain, as an approximately 3 kb transcript. HDAC9 was not detected in muscle, but this lane also did not hybridize GAPDH (Figure 7). Fragments of the isolated polypeptide having an amino acid sequence as set forth in SEQ ID NO:1, SEQ ID NO 5 or SEQ ID NO 6 will comprise polypeptides comprising from about 5 to 148 amino acids, preferably from about 10 to about 143 amino acids, more preferably from about 20 to about 100 amino acids, and most preferably from about 20 to about 50 amino acids. Such fragments also form a part of the present invention. Preferably, fragments will encompass the catalytic domain, which is predicted to exist between amino acid number 1 to 390. In accordance with this aspect of the invention there are provided novel polypeptides of human origin as well as biologically, diagnostically or therapeutically useful fragments, variants and derivatives thereof, variants and derivatives of the fragments, and analogs of the foregoing.

In a second aspect, the invention provides an isolated DNA comprising a nucleotide sequence that encodes a polypeptide as mentioned above. In particular, the invention provides

WO 02/050285

PCT/EP01/14928

6

(1) an isolated DNA comprising the nucleotide sequence as set forth in SEQ ID NO:2; SEQ ID NO 7 or SEQ ID NO 8 (2) an isolated DNA comprising the nucleotide sequence set forth in SEQ ID NO:3; (3) an isolated DNA capable of hybridizing under high stringency conditions to the nucleotide sequence set forth in SEQ ID NO:3; and (4) an isolated DNA comprising the nucleotide sequence set forth in SEQ ID NO:4. Also provided are nucleic acid sequences comprising at least about 15 bases, preferably at least about 20 bases, more preferably a nucleic acid sequence comprising about 30 contiguous bases of SEQ ID NO:2, SEQ ID NO 7 or SEQ ID NO 8 or SEQ ID NO:3. Also within the scope of the present invention are nucleic acids that are substantially similar to the nucleic acid with the nucleotide sequence as set forth in SEQ ID NO:2, SEQ ID NO 7 or SEQ ID NO 8 or SEQ ID NO:3. In a preferred embodiment, the isolated DNA takes the form of a vector molecule comprising at least a fragment of a DNA of the present invention, in particular comprising the DNA consisting of a nucleotide sequence as set forth in SEQ ID NO:2, SEQ ID NO 7 or SEQ ID NO 8 or SEQ ID NO:3.

A third aspect of the present invention encompasses a method for the diagnosis of conditions associated with abnormal regulation of gene expression which includes, but is not limited to, conditions associated with abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, or psoriasis in a human which comprises detecting abnormal transcription of messenger RNA transcribed from the natural endogenous human gene encoding the novel polypeptide consisting of the amino acid sequence set forth in SEQ ID NO:1, SEQ ID NO 5 or SEQ ID NO 6 in an appropriate tissue or cell from a human, wherein such abnormal transcription is diagnostic of the human's affliction with such a condition. In particular, the said natural endogenous human gene encoding the novel polypeptide consisting of the amino acid sequence set forth in SEQ ID NO:1, SEQ ID NO 5 or SEQ ID NO 6 comprises the genomic nucleotide sequence set forth in SEQ ID NO:4. In one embodiment of the present invention, the diagnostic method comprises contacting a sample of said appropriate tissue or cell or contacting an isolated RNA or DNA molecule derived from that tissue or cell with an isolated nucleotide sequence of at least about 15 - 20 nucleotides in length that hybridizes under high stringency conditions with the isolated nucleotide sequence encoding the novel polypeptide having an amino acid sequence set forth in SEQ ID NOs:1, 5 or 6

30

WO 02/050285

7

PCT/EP01/14928

Another embodiment of the assay aspect of the invention provides a method for the diagnosis of a condition associated with abnormal HDAC9 activity in a human, which comprises measuring the level of deacetylase activity in a certain tissue or cell from a human suffering from such a condition, wherein the presence of an abnormal level of deacetylase activity, relative to
5 the level thereof in the respective tissue or cell of a human not suffering from a condition associated with abnormal HDAC activity, is diagnostic of the human's suffering from said condition.

In accordance with one embodiment of this aspect of the invention there are provided
10 anti-sense polynucleotides that can regulate transcription of the gene encoding the novel HDAC9; in another embodiment, double stranded RNA is provided that can regulate the transcription of the gene encoding the novel HDAC9.

Another aspect of the invention provides a process for producing the aforementioned
15 polypeptides, polypeptide fragments, variants and derivatives, fragments of the variants and derivatives, and analogs of the foregoing. In a preferred embodiment of this aspect of the invention there are provided methods for producing the aforementioned HDAC9 comprising culturing host cells having incorporated therein an expression vector containing an exogenously-derived nucleotide sequence encoding such a polynucleotide under conditions sufficient for
20 expression of the polypeptide in the host cell, thereby causing expression of the polypeptide, and optionally recovering the expressed polypeptide. In a preferred embodiment of this aspect of the present invention, there is provided a method for producing polypeptides comprising or consisting of an amino acid sequence as set forth in SEQ ID NOs:1, 5 or 6 which comprises culturing a host cell having incorporated therein an expression vector containing an exogenously-
25 derived polynucleotide encoding a polypeptide comprising or consisting of an amino acid sequence as set forth in SEQ ID NOs:1, 5 or 6 under conditions sufficient for expression of such a polypeptide in the host cell, thereby causing the production of an expressed polypeptide, and optionally recovering the expressed polypeptide. Preferably, in any of such methods the exogenously derived polynucleotide comprises or consists of the nucleotide sequence set forth in
30 SEQ ID NOs:2, 7 or 8 the nucleotide sequence set forth in SEQ ID NO:3, or the nucleotide

WO 02/050285

8

PCT/EP01/14928

sequence set forth in SEQ ID NO:4. In accordance with another aspect of the invention there are provided products, compositions, processes and methods that utilize the aforementioned polypeptides and polynucleotides for, *inter alia*, research, biological, clinical and therapeutic purposes.

5

In certain additional preferred embodiments of this aspect of the invention there is provided an antibody or a fragment thereof which specifically binds to a polypeptide that comprises the amino acid sequence set forth in SEQ ID NOs:1, 5 or 6 i.e., all HDAC9 variants. In certain particularly preferred embodiments in this regard, the antibodies are highly selective

10

for human HDAC9 polypeptides or portions of human HDAC9 polypeptides.

In a further aspect, an antibody or fragment thereof is provided that binds to a fragment or portion of the amino acid sequence set forth in SEQ ID NOs:1, 5 or 6.

15

In another aspect, methods of treating a condition in a subject, wherein the condition is associated with abnormal HDAC9 gene expression, an increase or decrease in the presence of HDAC9 polypeptide in a subject, or an increase or decrease in the activity of HDAC 9 polypeptide, by the administration of an effective amount of an antibody that binds to a polypeptide with the amino acid sequence set out in SEQ ID NOs:1, 5 or 6., or a fragment or portion thereof to the subject are provided. Also provided are methods for the diagnosis of a disease or condition associated with abnormal HDAC9 gene expression or an increase or decrease in the presence of the HDAC9 in a subject, or an increase or decrease in the activity of HDAC 9 polypeptide, which comprises utilizing conventional methodologies, including, for example, the H4 histone assay that was previously described (Inokoshi, J., Katagiri, M., Arima, S., Tanaka, H., Hayashi, M., Kim, Y.-B., Furumai, R., Yoshida, M., Horinouchi, S., Omura, S. (1999) *Biochem. Biophys. Res. Com.* **256**, 372-376.).

20

25

In yet another aspect, the invention provides host cells which can be propagated in vitro, preferably vertebrate cells, in particular mammalian cells, or bacterial cells, which are capable upon growth in culture of producing a polypeptide that comprises the amino acid sequence set

30

WO 02/050285

9

PCT/EP01/14928

forth in SEQ ID NOs:1, 5 or 6 or fragments thereof, where the cells contain transcriptional control DNA sequences, where the transcriptional control sequences control transcription of RNA encoding a polypeptide with the amino acid sequence according to SEQ ID NOs:1, 5 or 6. or fragments thereof. This includes, but is not limited to, the propagation of HDAC9 in a plasmid and the production of DNA, RNA or protein in human or insect cells or bacteria using the endogenous HDAC9 promoter or any other transcriptional control sequence.

In yet another aspect of the present invention there are provided assay methods and kits comprising the components necessary to detect above-normal expression of polynucleotides encoding a polypeptide comprising an amino acid sequence as set forth in SEQ ID NOs:1, 5 or 6. , or polypeptides comprising an amino acid sequence set forth in SEQ ID NOs:1, 5 or 6. , or fragments thereof, in body tissue samples derived from a patient, such kits comprising e.g., antibodies that bind to a polypeptide comprising an amino acid sequence set forth in SEQ ID NOs:1, 5 or 6 or to fragments thereof, or oligonucleotide probes that hybridize with polynucleotides of the invention. In a preferred embodiment, such kits also comprise instructions detailing the procedures by which the kit components are to be used.

In another aspect, the invention is directed to use of a polypeptide comprising an amino acid sequence set forth in SEQ ID NOs:1, 5 or 6. or fragment thereof, polynucleotide encoding such a polypeptide or a fragment thereof, or antibody that binds to said polypeptide comprising an amino acid sequence set forth in SEQ ID NOs:1, 5 or 6. or a fragment thereof in the manufacture of a medicament to treat diseases associated with abnormal HDAC activity or gene expression.

Another aspect is directed to pharmaceutical compositions comprising a polypeptide comprising or consisting of an amino acid sequence set forth in SEQ ID NOs:1, 5 or 6. or fragment thereof, a polynucleotide encoding such a polypeptide or a fragment thereof, or antibody that binds to such a polypeptide or a fragment thereof, in conjunction with a suitable pharmaceutical carrier, excipient or diluent, for the treatment of diseases associated with abnormal HDAC activity or gene expression.

WO 02/050285

10

PCT/EP01/14928

In another aspect, the invention is directed to methods for the identification of molecules that can bind to a polypeptide comprising an amino acid sequence set forth in SEQ ID NOs:1, 5 or 6. and/or modulate the activity of a polypeptide comprising an amino acid sequence set forth in SEQ ID NOs:1, 5 or 6. or molecules that can bind to nucleic acid sequences that modulate the transcription or translation of a polynucleotide encoding a polypeptide comprising an amino acid sequence set forth in SEQ ID NOs:1, 5 or 6. Such methods are disclosed in, e.g., U.S. Patent Nos. 5,541,070; 5,567,317; 5,593,853; 5,670,326; 5,679,582; 5,856,083; 5,858,657; 5,866,341; 5,876,946; 5,989,814; 6,010,861; 6,020,141; 6,030,779; and 6,043,024, all of which are incorporated by reference herein in their entirety. Molecules identified by such methods also fall within the scope of the present invention.

In a related aspect, the invention is directed to use of the novel HDAC9 to identify associated proteins in HDAC biologically relevant complexes. At present, the proteins that associate with HDAC9 are not known. However, these may be characterized by determining whether HDAC9 associates with proteins that have been previously shown to interact with other HDACs (see Introduction). For example, components of HDAC9 complexes may be determined using conventional methods, including co-immunoprecipitation (see Example 9).

In yet another aspect, the invention is directed to methods for the introduction of nucleic acids of the invention into one or more tissues of a subject in need of treatment with the result that one or more proteins encoded by the nucleic acids are expressed and or secreted by cells within the tissue.

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

art from reading the following description and from reading the other parts of the present disclosure.

BRIEF DESCRIPTION OF THE DRAWINGS

- 5 Figure 1 shows the 1156 bp open reading frame that was identified using GENFAM (proprietary software) and used to search databases for the complete HDAC9 cDNA sequence. The respective ORF (SEQ ID NO:3) starts at nucleotide position no. 1 and ends at nucleotide position no. 1156.
- 10 Figures 2A and 2B show the full length cDNA sequence (SEQ ID NO:2) of HDAC9 and the amino acid sequence (SEQ ID NO:1), respectively. The full length cDNA sequence starts at nucleotide position no. 1 and ends at nucleotide position 2022.
- 15 Figure 3 shows the genomic DNA sequence in silico (AL022328) (SEQ ID NO:4), aligned with the sequence of clone 198929/HDAC9. The alignment was produced using proprietary software (Novartis Pharmaceuticals, Summit, NJ).
- 20 Figure 4 is a depiction of the alignment of HDAC9 predicted peptide and *S. pombe* Hda1 peptide. The query is HDAC9 peptide and the subject is *S. pombe* Hda1 peptide. The alignment was produced using Clustalw algorithm (Higgins, D.G., Thompson, J.D., Gibson, T.J. (1996) Using CLUSTAL for multiple sequence alignments. Methods Enzymol 266, 383-402).
- 25 Figure 5 shows the alignment of HDAC1 and HDAC9v1 and locations of the putative catalytic domain amino acids and Rb-binding domain. Catalytic domain amino acids are boxed and putative Rb domain amino acids are contained within crosshatched boxes. The alignment was produced using Clustalw algorithm (Higgins, D.G., Thompson, J.D., Gibson, T.J. (1996) Using CLUSTAL for multiple sequence alignments. Methods Enzymol 266, 383-402).

Figure 6 shows the alignment of HDACs 1-9v1. The alignment was produced using Clustalw algorithm (Higgins, D.G., Thompson, J.D., Gibson, T.J. (1996) Using CLUSTAL for multiple sequence alignments. *Methods Enzymol* 266, 383-402).

5 Figure 7 shows the Northern analysis of HDAC9. (A) Northern blot analysis of the distribution of HDAC9 in normal human tissues. GAPDH was hybridized to the same blot as a control for RNA loading. (B) Northern blot analysis of HDAC9 in matched tumor and normal tissues. GAPDH was hybridized to the same blot as a control for RNA loading.

10 Figure 8 shows Real Time PCR analysis of the distribution of HDAC9 in normal human tissues and cell lines relative to 18S ribosomal RNA. RNA from the human lung carcinoma cell line, A549 was used as an internal control.

15 Figure 9 shows the alignment of HDAC9v1 with class II HDACs (HDACs 4,5,6, 7). The alignment was produced using Clustalw algorithm (Higgins, D.G., Thompson, J.D., Gibson, T.J. (1996) Using CLUSTAL for multiple sequence alignments. *Methods Enzymol* 266, 383-402). Catalytic domain amino acids are boxed.

20 Figure 10 shows the alignment of HDAC9v1 with class I HDACs (HDACs 1,2,3,8). The alignment was produced using Clustalw algorithm (Higgins, D.G., Thompson, J.D., Gibson, T.J. (1996) Using CLUSTAL for multiple sequence alignments. *Methods Enzymol* 266, 383-402). Catalytic domain amino acids are boxed.

25 Figure 11 There are three HDAC9 sequence variants (HDAC9v1, HDAC9v2, and HDAC9v3). HDAC9v1 and HDAC9v2 were found by searching the human EST database and HDAC9v3 was found as a predicted transcript in the Celera Sequence database. (A) shows an alignment of the 3 HDAC9 variant peptide sequences. (B) shows a schematic of class I and class II HDAC peptide sequences. Catalytic domains are in filled boxes and putative LXCXE motifs are in open boxes (C) is a schematic of the genomic structures of HDAC9v1 and HDAC9v2.

WO 02/050285

13

PCT/EP01/14928

Exons are shown as filled boxes and introns are shown as lines between the filled boxes. Lengths of boxes and lines represent the lengths of exons and introns.

Figure 12 shows that HDAC9 is an enzymatically active histone deacetylase. (A) HDAC9 catalytic activity is comparable to the activity of HDAC3 and HDAC4. (B) shows that HDAC1 was more efficient than HDAC3, HDAC4, and HDAC9 at deacetylating the histone substrate in this assay.

Figure 13 shows that HDAC9 is a nuclear protein and shows that HDAC9-flag is in vitro translated.

Figure 14 shows DNA and peptide sequences for HDAC9v3 and HDAC9v2.

DETAILED DESCRIPTION OF THE INVENTION

All patent applications, patents and literature references cited herein are hereby incorporated by reference in their entirety.

In practicing the present invention, many conventional techniques in molecular biology, microbiology, and recombinant DNA are used. These techniques are well known and are explained in, for example, Current Protocols in Molecular Biology, Volumes I, II, and III, 1997 (F. M. Ausubel ed.); Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.; DNA Cloning: A Practical Approach, Volumes I and II, 1985 (D. N. Glover ed.); Oligonucleotide Synthesis, 1984 (M. L. Gait ed.); Nucleic Acid Hybridization, 1985, (Hames and Higgins); Transcription and Translation, 1984 (Hames and Higgins eds.); Animal Cell Culture, 1986 (R. I. Freshney ed.); Immobilized Cells and Enzymes, 1986 (IRL Press); Perbal, 1984, A Practical Guide to Molecular Cloning; the series, Methods in Enzymology (Academic Press, Inc.); Gene Transfer Vectors for Mammalian Cells, 1987 (J. H. Miller and M. P. Calos eds., Cold Spring Harbor Laboratory); and Methods in Enzymology Vol. 154 and Vol. 155 (Wu and Grossman, and Wu, eds., respectively).

SUBSTITUTE SHEET (RULE 26)

The following abbreviations used throughout the disclosure are listed herein below:
histone deacetylase (HDAC), histone deacetylase-like protein (HDLP)

In its broadest sense, the term "substantially similar", when used herein with respect to a
5 nucleotide sequence, means a nucleotide sequence corresponding to a reference nucleotide
sequence, wherein the corresponding sequence encodes a polypeptide having substantially the
same structure and function as the polypeptide encoded by the reference nucleotide sequence,
e.g. where only changes in amino acids not affecting the polypeptide function occur. Desirably
the substantially similar nucleotide sequence encodes the polypeptide encoded by the reference
10 nucleotide sequence. The percentage of identity between the substantially similar nucleotide
sequence and the reference nucleotide sequence desirably is at least 80%, more desirably at least
85%, preferably at least 90%, more preferably at least 95%, still more preferably at least 99%.
Sequence comparisons are carried out using Clustalw (see, for example, Higgins, D.G. et al.
Methods Enzymol. 266:383-402 (1996)). Clustalw alignments were performed using default
15 parameters.

A nucleotide sequence "substantially similar" to reference nucleotide sequence
hybridizes to the reference nucleotide sequence in 7% sodium dodecyl sulfate (SDS), 0.5 M
NaPO₄, 1 mM EDTA at 50°C with washing in 2X SSC, 0.1% SDS at 50°C, more desirably in
20 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄, 1 mM EDTA at 50°C with washing in 1X
SSC, 0.1% SDS at 50°C, more desirably still in 7% sodium dodecyl sulfate (SDS), 0.5 M
NaPO₄, 1 mM EDTA at 50°C with washing in 0.5X SSC, 0.1% SDS at 50°C, preferably in 7%
sodium dodecyl sulfate (SDS), 0.5 M NaPO₄, 1 mM EDTA at 50°C with washing in 0.1X SSC,
0.1% SDS at 50°C, more preferably in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄, 1 mM
25 EDTA at 50°C with washing in 0.1X SSC, 0.1% SDS at 65°C, yet still encodes a functionally
equivalent gene product.

"Elevated transcription of mRNA" refers to a greater amount of messenger RNA
transcribed from the natural endogenous human gene encoding the novel polypeptide of the
30 present invention present in an appropriate tissue or cell of an individual suffering from a

condition associated with abnormal HDAC9 activity than in a subject not suffering from such a disease or condition; in particular at least about twice, preferably at least about five times, more preferably at least about ten times, most preferably at least about 100 times the amount of mRNA found in corresponding tissues in humans who do not suffer from such a condition. Such elevated level of mRNA may eventually lead to increased levels of protein translated from such mRNA in an individual suffering from a condition associated with abnormal cellular proliferation as compared with a healthy individual. It is also understood that "elevated transcription of mRNA" may refer to a greater amount of messenger RNA transcribed from genes the expression of which is modulated by HDAC9 either alone or in combination with other molecules.

A "host cell," as used herein, refers to a prokaryotic or eukaryotic cell that contains heterologous DNA that has been introduced into the cell by any means, e.g., electroporation, calcium phosphate precipitation, microinjection, transformation, viral infection, and the like.

"Heterologous" as used herein means "of different natural origin" or represent a non-natural state. For example, if a host cell is transformed with a DNA or gene derived from another organism, particularly from another species, that gene is heterologous with respect to that host cell and also with respect to descendants of the host cell which carry that gene. Similarly, heterologous refers to a nucleotide sequence derived from and inserted into the same natural, original cell type, but which is present in a non-natural state, e.g. a different copy number, or under the control of different regulatory elements.

A "vector" molecule is a nucleic acid molecule into which heterologous nucleic acid may be inserted which can then be introduced into an appropriate host cell. Vectors preferably have one or more origin of replication, and one or more site into which the recombinant DNA can be inserted. Vectors often have convenient means by which cells with vectors can be selected from those without, e.g., they encode drug resistance genes. Common vectors include plasmids, viral genomes, and (primarily in yeast and bacteria) "artificial chromosomes."

"Plasmids" generally are designated herein by a lower case p preceded and/or followed by capital letters and/or numbers, in accordance with standard naming conventions that are familiar to those of skill in the art. Starting plasmids disclosed herein are either commercially available, publicly available on an unrestricted basis, or can be constructed from available

plasmids by routine application of well known, published procedures. Many plasmids and other cloning and expression vectors that can be used in accordance with the present invention are well known and readily available to those of skill in the art. Moreover, those of skill readily may construct any number of other plasmids suitable for use in the invention. The properties, construction and use of such plasmids, as well as other vectors, in the present invention will be readily apparent to those of skill from the present disclosure.

The term "isolated" means that the material is removed from its original environment (e.g., the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide or polypeptide present in a living animal is not isolated, but the same polynucleotide or polypeptide, separated from some or all of the coexisting materials in the natural system, is isolated, even if subsequently reintroduced into the natural system. Such polynucleotides could be part of a vector and/or such polynucleotides or polypeptides could be part of a composition, and still be isolated in that such vector or composition is not part of its natural environment.

As used herein, the term "transcriptional control sequence" refers to DNA sequences, such as initiator sequences, enhancer sequences, and promoter sequences, which induce, repress, or otherwise control the transcription of protein encoding nucleic acid sequences to which they are operably linked.

As used herein, "human transcriptional control sequences" are any of those transcriptional control sequences normally found associated with the human gene encoding the novel HDAC9 polypeptide of the present invention as it is found in the respective human chromosome. It is understood that the term may also refer to transcriptional control sequences normally found associated with human genes the expression of which is modulated by HDAC9 either alone or in combination with other molecules.

As used herein, "non-human transcriptional control sequence" is any transcriptional control sequence not found in the human genome.

The term "polypeptide" is used interchangeably herein with the terms "polypeptides" and "protein(s)".

As used herein, a "chemical derivative" of a polypeptide of the invention is a polypeptide of the invention that contains additional chemical moieties not normally a part of the molecule.

WO 02/050285

17

PCT/EP01/14928

Such moieties may improve the molecule's solubility, absorption, biological half life, etc. The moieties may alternatively decrease the toxicity of the molecule, eliminate or attenuate any undesirable side effect of the molecule, etc. Moieties capable of mediating such effects are disclosed, for example, in Remington's Pharmaceutical Sciences, 16th ed., Mack Publishing Co.,

5 Easton, Pa. (1980).

As used herein, "HDAC9" refers to the amino acid sequences of substantially purified HDAC9 obtained from any species, particularly mammalian, including bovine, ovine, porcine, murine, equine, and preferably human, from any source, whether natural, synthetic, semi-

10 synthetic, or recombinant.

As used herein, "HDAC activity", including "HDAC9 activity" refers to the ability of an HDAC polypeptide to deacetylate histone proteins, including ³H-labeled H4 histone peptide. Such activity may be measured according to conventional methods, for example as described in Inokoshi, J., Katagiri, M., Arima, S., Tanaka, H., Hayashi, M., Kim, Y.-B., Furumai, R.,

15 Yoshida, M., Horinouchi, S., and Omura, S. (1999) Biochem. Biophys. Res. Com. 256, 372-376.

A biologically "active" protein refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule.

The term "agonist", as used herein, refers to a molecule which when bound to HDAC9, causes a change in HDAC9 which modulates the activity of HDAC9.. Agonists may include

20 proteins, nucleic acids, carbohydrates, or any other molecules that bind to HDAC9.

The terms "antagonist" or "inhibitor" as used herein, refer to a molecule which when bound to HDAC9, blocks or modulates the biological activity of HDAC9. Antagonists and inhibitors may include proteins, nucleic acids, carbohydrates, or any other molecules, natural or

25 synthetic that bind to HDAC9.

HDAC9 was identified using proprietary computer software called GENFAM to search for new human sequences that are related to histone deacetylases in the Celera Human Genome Database, Incyte LIFESEQ® database and the public High Throughput Genomic database. An

30

1156 bp open reading frame (ORF) was identified and used to search a database of sequenced clones from pan-tissue and dorsal root ganglion cDNA libraries. Four clones were found to contain the ORF (M6, K10, P3, F23), two from each library. Of these clones, M6, from the pan-tissue library was determined to be the most complete cDNA as a result of sequence analysis and *in vitro* translation. BLAST (Altschul S.F. et al Nucleic Acid Res 25:3389-402 (1997)) was used to search the Genbank database using cDNA clone M6. Genomic sequence AL022328 was found to contain exons that were identical in sequence to the M6 cDNA. A Clustalw alignment of the antisense sequence of HDAC9 (2022 to 8) with genomic sequence AL022328 is shown in Figure 3. The first 7 bases of the HDAC9 predicted cDNA are not aligned, presumably because they occur following the next intron and this sequence was probably too short for the software to determine an alignment. The sequence of cDNA clone M6 was confirmed by automated DNA sequencing (ACGT, Inc., Northbrook, IL). Based upon the predicted cDNA sequence from genomic sequence AL022328, 44 bases were missing from the N-terminus of M6. This sequence was subsequently added by PCR.

The full length cDNA for HDAC9 predicts a protein of 673 amino acids. The HDAC9 cDNA sequence is 2022 base pairs in length. In order to determine the percent similarity of HDAC9 to other known HDACs, a Clustalw multiple sequence alignment was performed using complete peptide sequences for HDACs 1-9. HDAC9 is most similar in peptide sequence to human HDAC6 at 37%. The Clustalw alignment of HDAC9 with class II HDACs is shown in Figure 9. HDAC9 was also 40% similar to a yeast class II sequence hda1 from *S. pombe*. The Clustalw alignment of human HDAC9 and *S. pombe* is shown in Figure 4. HDAC9 was less similar to class I HDACs ($\leq 18\%$). The Clustalw alignment of HDAC9 to class I HDACs is shown in Figure 10. HDAC9 possesses a putative catalytic domain which encompasses approximately 317 aa (~6 to 323) based upon alignments of HDAC9 with the putative catalytic domains of all of the other known HDACs. To identify the catalytic domain of HDAC9, Clustalw alignments were performed separately using HDAC9 complete peptide and catalytic domain sequences from class I HDACs (1-3 and 8) or class II HDACs (4-7). 13 amino acids were previously shown to confer deacetylase activity, based upon inactivation by single amino acid mutations and the three dimensional structure formed by a complex of HDAC-like protein (HDLP), Zn²⁺ and HDAC inhibitors (Finnin, M. S., Donigian, J. R., Cohen, A., Richon, V. M.,

Rifkind, R. a., Marks, P. A., Breslow, R., and Pavletich, N. P. (1999) Structures of a histone deacetylase homologue bound to TSA and SAHA inhibitors. *Nature* 401, 188-193). These 13 amino acids include Pro 22, His 131, His 132, Gly 140, Phe 141, Asp 166, Asp 168, His 170, Asp 173, Phe 198, Asp 258, Leu 265, and Tyr 297. 12 out of 13 of these amino acids are
5 conserved in HDAC9. The amino acid that is not conserved is Leu 265. This hydrophobic residue forms part of the TS binding pocket and is replaced in HDAC9 with Glu at amino acid 272. Leu 265 is replaced with Met in HDAC8 and Lys in HDAC 6 domain 1. This suggests that this residue is not highly conserved and need not be identical to other HDACs. The second residue that differs from HDLP, HDAC1, and HDAC2, Asp 173 is substituted with Gln at
10 position 177 in HDAC9, a difference that is also present in the HDAC6 catalytic domain 1. Furthermore, Asp 173 is substituted with Asn in HDACs 4,5, 6 (domain 2), and 7. This evidence suggests that these Asp173 substitutions do not affect HDAC activity.

An amino acid sequence motif was previously found to be important for the binding of HDACs 1 and 2 to retinoblastoma protein (Rb). Complexes of HDACs 1 and 2 and Rb induce
15 repression of E2F responsive promoters (Brehm, A., Miska, E. A., McCance, D. J., Reid, J. L., Bannister, A. J., and Kouzarides, T. (1998) *Nature* 391, 597-601). An Rb-binding motif fits the sequence model LXCXE, where "X" can be any amino acid. The LXCXE domain has been found to be dispensible for growth suppression function of Rb, but is necessary for HDAC binding (Chen, T.-T. and Wang, J. Y. J. (2000) *Mol. Cell Biol.* 20, 5571-5580). The Rb-binding
20 domain that was previously determined in HDAC1 is located from amino acid 414 to amino acid 419 and is the sequence IACEE. So far, it has not been determined whether other HDACs are capable of binding to Rb. However, HDAC 9 contains a putative Rb-binding motif, LSCIL, that aligned with HDAC1 IACEE and is located between amino acids 560 and 564. Co-immunoprecipitation of HDAC9 with Rb is one strategy that may be used to validate the
25 function of this motif in HDAC9.

As a member of the HDAC family, HDAC9 could form biologically relevant complexes with proteins and display functions that have been described for other HDACs. For example, it is likely to be involved in the regulation of transcription as a component of complexes that are involved in transcriptional repression that is mediated through interactions of HDACs with
30 multi-protein complexes and which requires deacetylase activity. Thus, increased activity or

WO 02/050285

20

PCT/EP01/14928

expression of HDAC9 may be associated with numerous pathological conditions, including but not limited to, abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, host inflammatory or immune response, or psoriasis.

5 Thus, the DNA/amino acid sequence and predicted structure of HDAC9 will be useful for designing agents (e.g. antagonists or inhibitors) useful to ameliorate conditions associated with abnormal HDAC activity. These may include, for example, antiproliferative or antiinflammatory agents either through the use of small molecules or proteins (e.g. antibodies) directed against it or associated proteins in HDAC transcription repressor complexes. In
10 addition, protein derived from the HDAC9 sequence may also be used as a therapeutic to modify host cell proliferative or inflammatory responses.

To determine the expression pattern of the novel polypeptide, a panel of mRNAs from a variety of human tissues is subjected to Northern analysis. Data indicate that HDAC9 is
15 expressed in human tissues, being detectable in brain, colon, heart, kidney, liver, placenta, small intestine, spleen, stomach and testes. Thus, HDAC9 represents a transcribed gene.

Therefore, in one aspect, the present invention relates to a novel histone deacetylase (HDAC). As outlined above, HDAC9 is clearly a member of the HDAC family since it is highly similar to other HDAC proteins in the hda1 class II HDACs. It also shares many similarities
20 with the HDAC family.

The present invention relates to an isolated polypeptide comprising the amino acid sequence set forth in SEQ ID NO:1. For example, such a polypeptide may be a fusion protein including the amino acid sequence of the novel HDAC9. In another aspect the present invention relates to an isolated polypeptide consisting of the amino acid sequence set forth in SEQ ID
25 NO:1, which is, in particular, the novel HDAC9.

The invention includes nucleic acid or nucleotide molecules, preferably DNA molecules, in particular encoding the novel HDAC9. Preferably, an isolated nucleic acid molecule, preferably a DNA molecule, of the present invention encodes a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:1 SEQ ID NO 5 or SEQ ID NO 6. Likewise
30 preferred is an isolated nucleic acid molecule, preferably a DNA molecule, encoding a

WO 02/050285

21

PCT/EP01/14928

polypeptide consisting of the amino acid sequence set forth in SEQ ID NO:1, SEQ ID NO 5 or SEQ ID NO 6. Such a nucleic acid or nucleotide, in particular such a DNA molecule, preferably comprises a nucleotide sequence selected from the group consisting of (1) the nucleotide sequence as set forth in SEQ ID NO:2,, 7 or 8 which is the complete cDNA sequence encoding the polypeptide consisting of the amino acid sequence set forth in SEQ ID NO:1, 5 and 6, respectively, (2) the nucleotide sequence set forth in SEQ ID NO:3, which corresponds to the open reading frame of the cDNA sequence set forth in SEQ ID NO:2; (3) a nucleotide sequence capable of hybridizing under high stringency conditions to a nucleotide sequence set forth in SEQ ID NO:3; and (4) the nucleotide sequence set forth in SEQ ID NO:4, which corresponds to the endogenous genomic human DNA encoding the polypeptide consisting of the amino acid sequence set forth in SEQ ID NO:1. Such hybridization conditions may be highly stringent or less highly stringent, as described above. In instances wherein the nucleic acid molecules are deoxyoligonucleotides ("oligos"), highly stringent conditions may refer, e.g., to washing in 6X SSC/0.05% sodium pyrophosphate at 37 °C (for 14-base oligos), 48 °C (for 17-base oligos), 55 °C (for 20-base oligos), and 60 °C (for 23-base oligos). Suitable ranges of such stringency conditions for nucleic acids of varying compositions are described in Krause and Aaronson (1991), *Methods in Enzymology*, 200:546-556 in addition to Maniatis et al., cited above.

These nucleic acid molecules may act as target gene antisense molecules, useful, for example, in target gene regulation and/or as antisense primers in amplification reactions of target gene nucleic acid sequences. Further, such sequences may be used as part of ribozyme and/or triple helix sequences, also useful for target gene regulation. Still further, such molecules may be used as components of diagnostic methods whereby the presence of an allele causing a disease associated with abnormal HDAC9 expression or activity, for example, abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, host inflammatory or immune response, or psoriasis, may be detected.

The invention also encompasses (a) vectors that contain at least a fragment of any of the foregoing nucleotide sequences and/or their complements (i.e., antisense); (b) vector molecules, preferably vector molecules comprising transcriptional control sequences, in particular expression vectors, that contain any of the foregoing coding sequences operatively associated

with a regulatory element that directs the expression of the coding sequences; and (c) genetically engineered host cells that contain a vector molecule as mentioned herein or at least a fragment of any of the foregoing nucleotide sequences operatively associated with a regulatory element that directs the expression of the coding sequences in the host cell. As used herein, regulatory
5 elements include, but are not limited to, inducible and non-inducible promoters, enhancers, operators and other elements known to those skilled in the art that drive and regulate expression. Preferably, host cells can be vertebrate host cells, preferably mammalian host cells, such as human cells or rodent cells, such as CHO or BHK cells. Likewise preferred, host cells can be bacterial host cells, in particular *E.coli* cells.

10 Particularly preferred is a host cell, in particular of the above described type, which can be propagated in vitro and which is capable upon growth in culture of producing an HDAC9 polypeptide, in particular a polypeptide comprising or consisting of an amino acid sequence set forth in SEQ ID NO:1, wherein said cell contains some fragment or complete sequence of HDAC9 coding sequence in a construct that is controlled by one or more transcriptional control
15 sequences that is not a transcriptional control sequence of the natural endogeneous human gene encoding said polypeptide, wherein said one or more transcriptional control sequences control transcription of a DNA encoding said polypeptide. Possible transcriptional control sequences include, but are not limited to, bacterial or viral promoter sequences.

The invention includes the complete sequence of the gene as well as fragments of any of
20 the nucleic acid sequences disclosed herein. Fragments of the nucleic acid sequences encoding the novel HDAC9 polypeptide may be used as a hybridization probe for a cDNA library to isolate other genes which have a high sequence similarity to the HDAC9 gene or similar biological activity. Probes of this type preferably have at least about 30 bases and may contain, for example, from about 30 to about 50 bases, about 50 to about 100 bases, about 100 to about
25 200 bases, or more than 200 bases. The probe may also be used to identify a cDNA clone corresponding to a full length transcript and a genomic clone or clones that contain the complete HDAC9 gene including regulatory and promoter regions, exons, and introns. An example of a screen comprises isolating the coding region of the HDAC9 gene by using the known DNA sequence to synthesize an oligonucleotide probe. Labeled oligonucleotides having a sequence
30 complementary to that of the gene of the present invention may be used to screen a library of

human cDNA, genomic DNA or mRNA to determine which members of the library to which the probe hybridizes.

In addition to the gene sequences described above, homologs of such sequences, as may, for example, be present in other species, may be identified and may be readily isolated, without undue experimentation, by molecular biological techniques well known in the art. Further, there may exist genes at other genetic loci within the genome that encode proteins which have homology to one or more domains of such gene products. These genes may also be identified via similar techniques. For example, the isolated nucleotide sequence of the present invention encoding the novel HDAC9 polypeptide may be labeled and used to screen a cDNA library constructed from mRNA obtained from the organism of interest. Hybridization conditions will be of a lower stringency when the cDNA library is derived from an organism different from the type of organism from which the labeled sequence was derived. Alternatively, the labeled fragment may be used to screen a genomic library derived from the organism of interest, again, using appropriately stringent conditions. Such low stringency conditions will be well known to those of skill in the art, and will vary predictably depending on the specific organisms from which the library and the labeled sequences are derived. For guidance regarding such conditions see, for example, Sambrook et al. cited above.

Further, a previously unknown differentially expressed gene-type sequence may be isolated by performing PCR using two degenerate oligonucleotide primer pools designed on the basis of amino acid sequences within the gene of interest. The template for the reaction may be cDNA obtained by reverse transcription of mRNA prepared from human or non-human cell lines or tissue known or suspected to express a differentially expressed gene allele. The PCR product may be subcloned and sequenced to ensure that the amplified sequences represent the sequences of a differentially expressed gene-like nucleic acid sequence. The PCR fragment may then be used to isolate a complete cDNA clone by a variety of conventional methods. For example, the amplified fragment may be labeled and used to screen a bacteriophage cDNA library. Alternatively, the labeled fragment may be used to screen a genomic library.

PCR technology may also be utilized to isolate full length cDNA sequences. For example, RNA may be isolated, following standard procedures, from an appropriate cellular or tissue source. A reverse transcription reaction may be performed on the RNA using an

oligonucleotide primer specific for the most 5' end of the amplified fragment for the priming of first strand synthesis. The resulting RNA/DNA hybrid may then be "tailed" with guanines using a standard terminal transferase reaction, the hybrid may be digested with RNAase H, and second strand synthesis may then be primed with a poly-C primer. Thus, cDNA sequences upstream of the amplified fragment may easily be isolated. For a review of cloning strategies which may be used, see e.g., Sambrook et al., 1989, supra.

In cases where the gene identified is the normal, or wild type, gene, this gene may be used to isolate mutant alleles of the gene. Such an isolation is preferable in processes and disorders which are known or suspected to have a genetic basis. Mutant alleles may be isolated from individuals either known or suspected to have a genotype which contributes to disease symptoms related to abnormal HDAC activity, including, but not limited to, conditions such as abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, host inflammatory or immune response, or psoriasis. Mutant alleles and mutant allele products may then be utilized in the diagnostic assay systems described below.

A cDNA of the mutant gene may be isolated, for example, by using PCR, a technique which is well known to those of skill in the art. In this case, the first cDNA strand may be synthesized by hybridizing an oligo-dT oligonucleotide to mRNA isolated from tissue known or suspected to be expressed in an individual putatively carrying the mutant allele, and by extending the new strand with reverse transcriptase. The second strand of the cDNA is then synthesized using an oligonucleotide that hybridizes specifically to the 5' end of the normal gene. Using these two primers, the product is then amplified via PCR, cloned into a suitable vector, and subjected to DNA sequence analysis through methods well known to those of skill in the art. By comparing the DNA sequence of the mutant gene to that of the normal gene, the mutation(s) responsible for the loss or alteration of function of the mutant gene product can be ascertained.

Alternatively, a genomic or cDNA library can be constructed and screened using DNA or RNA, respectively, from a tissue known to or suspected of expressing the gene of interest in an individual suspected of or known to carry the mutant allele. The normal gene or any suitable fragment thereof may then be labeled and used as a probe to identify the corresponding mutant allele in the library. The clone containing this gene may then be purified through methods routinely practiced in the art, and subjected to sequence analysis as described above.

Additionally, an expression library can be constructed utilizing DNA isolated from or cDNA synthesized from a tissue known to or suspected of expressing the gene of interest in an individual suspected of or known to carry the mutant allele. In this manner, gene products made by the putatively mutant tissue may be expressed and screened using standard antibody screening techniques in conjunction with antibodies raised against the normal gene product, as described below. (For screening techniques, see, for example, Harlow, E. and Lane, eds., 1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor Press, Cold Spring Harbor.) In cases where the mutation results in an expressed gene product with altered function (e.g., as a result of a missense mutation), a polyclonal set of antibodies are likely to cross-react with the mutant gene product. Library clones detected via their reaction with such labeled antibodies can be purified and subjected to sequence analysis as described above.

The present invention includes those proteins encoded by nucleotide sequences set forth in any of SEQ ID NOs:2, 3, 4, 7 or 8 in particular, a polypeptide that is or includes the amino acid sequence set out in SEQ ID NO:1, 5 or 6 or fragments thereof.

Furthermore, the present invention includes proteins that represent functionally equivalent gene products. Such an equivalent differentially expressed gene product may contain deletions, additions or substitutions of amino acid residues within the amino acid sequence encoded by the differentially expressed gene sequences described, above, but which result in a silent change, thus producing a functionally equivalent differentially expressed gene product. 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 involved.

For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Functionally equivalent," as utilized herein, may refer to a protein or polypeptide capable of exhibiting a substantially similar *in vivo* or *in vitro* activity as the endogenous differentially expressed gene products encoded by the differentially expressed gene sequences described above. "Functionally equivalent" may also refer to proteins or polypeptides capable of interacting with other cellular or extracellular molecules in a manner

substantially similar to the way in which the corresponding portion of the endogenous differentially expressed gene product would. For example, a "functionally equivalent" peptide would be able, in an immunoassay, to diminish the binding of an antibody to the corresponding peptide (i.e., the peptide the amino acid sequence of which was modified to achieve the

5 "functionally equivalent" peptide) of the endogenous protein, or to the endogenous protein itself, where the antibody was raised against the corresponding peptide of the endogenous protein. An equimolar concentration of the functionally equivalent peptide will diminish the aforesaid binding of the corresponding peptide by at least about 5%, preferably between about 5% and 10%, more preferably between about 10% and 25%, even more preferably between about 25%

10 and 50%, and most preferably between about 40% and 50%.

The polypeptides of the present invention may be produced by recombinant DNA technology using techniques well known in the art. Therefore, there is provided a method of producing a polypeptide of the present invention, which method comprises culturing a host cell having incorporated therein an expression vector containing an exogenously-derived

15 polynucleotide encoding a polypeptide comprising an amino acid sequence as set forth in SEQ ID NOs: 1, 5 or 6 under conditions sufficient for expression of the polypeptide in the host cell, thereby causing the production of the expressed polypeptide. Optionally, said method further comprises recovering the polypeptide produced by said cell. In a preferred embodiment of such a method, said exogenously-derived polynucleotide encodes a polypeptide consisting of an amino

20 acid sequence set forth in SEQ ID NOs: 1, 5 or 6. Preferably, said exogenously-derived polynucleotide comprises the nucleotide sequence as set forth in any of SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO: 7 or SEQ ID NO:8. In case of using the nucleotide sequence set forth in SEQ ID NO:3, i.e. the open reading frame, the sequence, when inserted into a vector, may be followed by one or more appropriate translation stop codons, preferably by the

25 natural endogenous stop codon TGA beginning at nucleotide 2021 in the cDNA sequence.

Thus, methods for preparing the polypeptides and peptides of the invention by expressing nucleic acid encoding respective nucleotide sequences are described herein. Methods which are well known to those skilled in the art can be used to construct expression vectors containing protein coding sequences and appropriate transcriptional/translational control signals. These

30 methods include, for example, in vitro recombinant DNA techniques, synthetic techniques and in

vivo recombination/genetic recombination. See, for example, the techniques described in Sambrook et al., 1989, *supra*, and Ausubel et al., 1989, *supra*. Alternatively, RNA capable of encoding differentially expressed gene protein sequences may be chemically synthesized using, for example, synthesizers. See, for example, the techniques described in "Oligonucleotide Synthesis", 1984, Gait, M. J. ed., IRL Press, Oxford, which is incorporated by reference herein in its entirety.

A variety of host-expression vector systems may be utilized to express the HDAC9 gene coding sequences of the invention. Such host-expression systems represent vehicles by which the coding sequences of interest may be produced and subsequently purified, but also represent cells which may, when transformed or transfected with the appropriate nucleotide coding sequences, exhibit the HDAC9 gene protein of the invention *in situ*. These include but are not limited to microorganisms such as bacteria (e.g., *E. coli*, *B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing differentially expressed gene protein coding sequences; yeast (e.g. *Saccharomyces*, *Pichia*) transformed with recombinant yeast expression vectors containing the differentially expressed gene protein coding sequences; insect cell systems infected or transfected with recombinant virus expression vectors (e.g., baculovirus) containing the differentially expressed gene protein coding sequences; plant cell systems infected with recombinant virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant vectors, including plasmids, (e.g., Ti plasmid) containing protein coding sequences; or mammalian cell systems (e.g. COS, CHO, BHK, 293, 3T3) harboring recombinant expression constructs containing promoters derived from the genome of mammalian cells (e.g., metallothioneine promoter) or from mammalian viruses (e.g., the adenovirus late promoter; the vaccinia virus 7.5K promoter, or the CMV promoter).

Expression of the HDAC9 of the present invention by a cell from an HDAC9 encoding gene that is native to the cell can also be performed. Methods for such expression are detailed in, e.g., U.S. Patents 5,641,670; 5,733,761; 5,968,502; and 5,994,127, all of which are expressly incorporated by reference herein in their entirety. Cells that have been induced to express HDAC9 by the methods of any of U.S. Patents 5,641,670; 5,733,761; 5,968,502; and 5,994,127

can be implanted into a desired tissue in a living animal in order to increase the local concentration of HDAC9 in the tissue.

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the protein being expressed. For example, when a large quantity of such a protein is to be produced, for the generation of antibodies or to screen peptide libraries, for example, vectors which direct the expression of high levels of fusion protein products that are readily purified may be desirable. In this respect, fusion proteins comprising hexahistidine tags may be used, such as EpiTag vectors including pCDNA3.1/His (Invitrogen, Carlsbad, CA). Other vectors include, but are not limited, to the *E. coli* expression vector pUR278 (Ruther et al., 1983, EMBO J. 2:1791), in which the protein coding sequence may be ligated individually into the vector in frame with the lac Z coding region so that a fusion protein is produced; pIN vectors (Inouye & Inouye, 1985, Nucleic Acids Res. 13:3101-3109; Van Heeke & Schuster, 1989, J. Biol. Chem. 264:5503-5509); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. The pGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned target gene protein can be released from the GST moiety. Fusion proteins containing Flag tags, such as 3X Flag (Sigma, St. Louis, MO) or myc tags, for example pCDNA3.1/myc-His (Invitrogen, Carlsbad, CA) may be used. These fusions allow coimmunoprecipitation and Western detection of proteins for which antibodies are not yet available.

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

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

Among known bacterial promoters suitable for expression of polynucleotides and polypeptides in accordance with the present invention are the *E. coli* lacI and lacZ promoters, the T3 and T7 promoters, the T5 tac promoter, the lambda PR, PL promoters and the trp promoter.

Among known eukaryotic promoters suitable in this regard are the CMV immediate early promoter, the HSV thymidine kinase promoter, the early and late SV40 promoters, the promoters of retroviral LTRs, such as those of the Rous sarcoma virus ("RSV"), and metallothionein promoters, such as the mouse metallothionein-I promoter. For example, a plasmid construct could contain a HDAC9 transcriptional control sequence fused to a reporter transcription unit that encodes the coding region of β -Galactosidase, chloramphenicol acetyltransferase, green fluorescent protein or luciferase. This construct could be used to screen for small molecules that modulate HDAC9 transcription. Such molecules are potential therapeutics. Furthermore, an HDAC9 reporter gene could be used to examine the effects of an HDAC9 therapeutic in mammalian cells or xenografts using fluorescent reporters and imaging techniques, such as fluorescence microscopy or Biophotonic *in vivo* imaging, a technology that produces visual and quantitative measurements in real time (Xenogen, Palo Alto, CA). Changes in these reporters in normal, diseased or drug-treated tissue or cells would be indicators of changes in HDAC9 expression or activity.

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

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the coding sequence of interest may be ligated to an adenovirus transcription/translation control complex, e.g., the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by in vitro or in vivo recombination. Insertion in a non-essential region of the viral genome (e.g., region E1 or E3) will result in a recombinant virus that is viable and capable of expressing the desired protein in infected hosts (e.g., See Logan & Shenk, 1984, Proc. Natl. Acad. Sci. USA 81:3655-3659). Specific initiation signals may also be required for efficient translation of inserted gene coding sequences. These signals include the ATG initiation codon and adjacent sequences. In cases where an entire gene, including its own initiation codon and adjacent sequences, is inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only a portion of the gene coding sequence is inserted, exogenous translational control signals, including, perhaps, the ATG initiation codon, must be provided. Furthermore, the initiation codon must be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements, transcription terminators, etc. (see Bittner et al., 1987, Methods in Enzymol. 153:516-544). Other common systems are based on SV40, retrovirus or adeno-associated virus. Selection of appropriate vectors and promoters for expression in a host cell is a well known procedure and the requisite techniques for expression vector construction, introduction of the vector into the host and expression in the host per se are routine skills in the art. Generally, recombinant expression vectors will include origins of replication, a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence, and a selectable marker to permit isolation of vector containing cells after exposure to the vector.

In addition, a host cell strain may be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may be important for the function of the protein. Different host cells have characteristic and specific

mechanisms for the post-translational processing and modification of proteins. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells which possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. Such mammalian host cells include but are not limited to CHO, VERO, BHK, HeLa, COS, MDCK, 293, 3T3, WI38, etc. and are well known to one of skill in the art.

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

A number of selection systems may be used, including but not limited to, the herpes simplex virus thymidine kinase (Wigler, et al., 1977, Cell 11:223), hypoxanthine-guanine phosphoribosyltransferase (Szybalska & Szybalski, 1962, Proc. Natl. Acad. Sci. USA 48:2026), and adenine phosphoribosyltransferase (Lowy, et al., 1980, Cell 22:817) genes can be employed in tk⁻, hprt⁻ or aprt⁻ cells, respectively. Also, antimetabolite resistance can be used as the basis of selection for dhfr, which confers resistance to methotrexate (Wigler, et al., 1980, Natl. Acad. Sci. USA 77:3567; O'Hare, et al., 1981, Proc. Natl. Acad. Sci. USA 78:1527); gpt, which confers resistance to mycophenolic acid (Mulligan & Berg, 1981, Proc. Natl. Acad. Sci. USA 78:2072); neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin, et al., 1981, J.

Mol. Biol. 150:1); and hygromycin, which confers resistance to hygromycin (Santerre, et al., 1984, Gene 30:147) genes.

5 An alternative fusion protein system allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht, et al., 1991, Proc. Natl. Acad. Sci. USA 88: 8972-8976). In this system, the gene of interest is subcloned into a vaccinia recombination plasmid such that the gene's open reading frame is translationally fused to an amino-terminal tag consisting of six histidine residues. Extracts from cells infected with recombinant vaccinia virus are loaded onto Ni²⁺ nitriloacetic acid-agarose columns and histidine-tagged proteins are selectively eluted with imidazole-containing buffers.

10 When used as a component in assay systems such as those described below, a protein of the present invention may be labeled, either directly or indirectly, to facilitate detection of a complex formed between the protein and a test substance. Any of a variety of suitable labeling systems may be used including, but not limited to, radioisotopes such as ¹²⁵I; enzyme labeling systems that generate a detectable calorimetric signal or light when exposed to substrate; and
15 fluorescent labels.

Where recombinant DNA technology is used to produce a protein of the present invention for such assay systems, it may be advantageous to engineer fusion proteins that can facilitate labeling, immobilization, detection and/or isolation

Indirect labeling involves the use of a protein, such as a labeled antibody, which
20 specifically binds to a polypeptide of the present invention. Such antibodies include but are not limited to polyclonal, monoclonal, chimeric, single chain, Fab fragments and fragments produced by an Fab expression library.

In another embodiment, nucleic acids comprising a sequence encoding HDAC9 protein
25 or functional derivative thereof, may be administered to promote normal biological function, for example, normal transcriptional regulation, by way of gene therapy. Gene therapy refers to therapy performed by the administration of a nucleic acid to a subject. In this embodiment of the invention, the nucleic acid produces its encoded protein that mediates a therapeutic effect by promoting normal transcriptional regulation..

Any of the methods for gene therapy available in the art can be used according to the present invention. Exemplary methods are described below.

In a preferred aspect, the therapeutic comprises a HDAC9 nucleic acid that is part of an expression vector that expresses a HDAC9 protein or fragment or chimeric protein thereof in a suitable host. In particular, such a nucleic acid has a promoter operably linked to the HDAC9 coding region, said promoter being inducible or constitutive, and, optionally, tissue-specific. In another particular embodiment, a nucleic acid molecule is used in which the HDAC9 coding sequences and any other desired sequences are flanked by regions that promote homologous recombination at a desired site in the genome, thus providing for intrachromosomal expression of the HDAC9 nucleic acid (Koller and Smithies, 1989, Proc. Natl. Acad. Sci. USA 86:8932-8935; Zijlstra et al., 1989, Nature 342:435-438).

Delivery of the nucleic acid into a patient may be either direct, in which case the patient is directly exposed to the nucleic acid or nucleic acid-carrying vector, or indirect, in which case, cells are first transformed with the nucleic acid in vitro, then transplanted into the patient. These two approaches are known, respectively, as *in vivo* or *ex vivo* gene therapy.

In a specific embodiment, the nucleic acid is directly administered *in vivo*, where it is expressed to produce the encoded product. This can be accomplished by any of numerous methods known in the art, e.g., by constructing it as part of an appropriate nucleic acid expression vector and administering it so that it becomes intracellular, e.g., by infection using a defective or attenuated retroviral or other viral vector (see, e.g., U.S. Pat. No. 4,980,286 and others mentioned *infra*), or by direct injection of naked DNA, or by use of microparticle bombardment (e.g., a gene gun; Biolistic, Dupont), or coating with lipids or cell-surface receptors or transfecting agents, encapsulation in liposomes, microparticles, or microcapsules, or by administering it in linkage to a peptide which is known to enter the nucleus, by administering it in linkage to a ligand subject to receptor-mediated endocytosis (see e.g., U.S. Patents 5,166,320; 5,728,399; 5,874,297; and 6,030,954, all of which are incorporated by reference herein in their entirety) (which can be used to target cell types specifically expressing the receptors), etc. In another embodiment, a nucleic acid-ligand complex can be formed in which the ligand comprises a fusogenic viral peptide to disrupt endosomes, allowing the nucleic acid to avoid lysosomal degradation. In yet another embodiment, the nucleic acid can be targeted in vivo

WO 02/050285

34

PCT/EP01/14928

for cell specific uptake and expression, by targeting a specific receptor (see, e.g., PCT Publications WO 92/06180; WO 92/22635; WO92/20316; WO93/14188; and WO 93/20221). Alternatively, the nucleic acid can be introduced intracellularly and incorporated within host cell DNA for expression, by homologous recombination (see, e.g., U.S. Patents 5,413,923; 5,416,260; and 5,574,205; and Zijlstra et al., 1989, Nature 342:435-438).

In a specific embodiment, a viral vector that contains the HDAC9 nucleic acid is used. For example, a retroviral vector can be used (see, e.g., U.S. Patents 5,219,740; 5,604,090; and 5,834,182). These retroviral vectors have been modified to delete retroviral sequences that are not necessary for packaging of the viral genome and integration into host cell DNA. The HDAC9 nucleic acid to be used in gene therapy is cloned into the vector, which facilitates delivery of the gene into a patient.

Adenoviruses are other viral vectors that can be used in gene therapy. Adenoviruses are especially attractive vehicles for delivering genes to respiratory epithelia. Adenoviruses naturally infect respiratory epithelia where they cause a mild disease. Other targets for adenovirus-based delivery systems are liver, the central nervous system, endothelial cells, and muscle. Adenoviruses have the advantage of being capable of infecting non-dividing cells. Methods for conducting adenovirus-based gene therapy are described in, e.g., U.S. Patents 5,824,544; 5,868,040; 5,871,722; 5,880,102; 5,882,877; 5,885,808; 5,932,210; 5,981,225; 5,994,106; 5,994,132; 5,994,134; 6,001,557; and 6,033,8843, all of which are incorporated by reference herein in their entirety.

Adeno-associated virus (AAV) has also been proposed for use in gene therapy. Methods for producing and utilizing AAV are described, e.g., in U.S. Patents 5,173,414; 5,252,479; 5,552,311; 5,658,785; 5,763,416; 5,773,289; 5,843,742; 5,869,040; 5,942,496; and 5,948,675, all of which are incorporated by reference herein in their entirety.

Another approach to gene therapy involves transferring a gene to cells in tissue culture by such methods as electroporation, lipofection, calcium phosphate mediated transfection, or viral infection. Usually, the method of transfer includes the transfer of a selectable marker to the cells. The cells are then placed under selection to isolate those cells that have taken up and are expressing the transferred gene. Those cells are then delivered to a patient.

In this embodiment, the nucleic acid is introduced into a cell prior to administration in vivo of the resulting recombinant cell. Such introduction can be carried out by any method known in the art, including but not limited to transfection, electroporation, microinjection, infection with a viral or bacteriophage vector containing the nucleic acid sequences, cell fusion, chromosome-mediated gene transfer, microcell-mediated gene transfer, spheroplast fusion, etc. Numerous techniques are known in the art for the introduction of foreign genes into cells and may be used in accordance with the present invention, provided that the necessary developmental and physiological functions of the recipient cells are not disrupted. The technique should provide for the stable transfer of the nucleic acid to the cell, so that the nucleic acid is expressible by the cell and preferably heritable and expressible by its cell progeny.

The resulting recombinant cells can be delivered to a patient by various methods known in the art. In a preferred embodiment, epithelial cells are injected, e.g., subcutaneously. In another embodiment, recombinant skin cells may be applied as a skin graft onto the patient. Recombinant blood cells (e.g., hematopoietic stem or progenitor cells) are preferably administered intravenously. The amount of cells envisioned for use depends on the desired effect, patient state, etc., and can be determined by one skilled in the art.

Cells into which a nucleic acid can be introduced for purposes of gene therapy encompass any desired, available cell type, and include but are not limited to epithelial cells, endothelial cells, keratinocytes, fibroblasts, muscle cells, hepatocytes; blood cells such as T lymphocytes, B lymphocytes, monocytes, macrophages, neutrophils, eosinophils, megakaryocytes, granulocytes; various stem or progenitor cells, in particular hematopoietic stem or progenitor cells, e.g., as obtained from bone marrow, umbilical cord blood, peripheral blood, fetal liver, etc.

In a preferred embodiment, the cell used for gene therapy is autologous to the patient.

In an embodiment in which recombinant cells are used in gene therapy, a HDAC9 nucleic acid is introduced into the cells such that it is expressible by the cells or their progeny, and the recombinant cells are then administered in vivo for therapeutic effect. In a specific embodiment, stem or progenitor cells are used. Any stem-and/or progenitor cells that can be isolated and maintained in vitro can potentially be used in accordance with this embodiment of the present invention. Such stem cells include but are not limited to hematopoietic stem cells (HSC), stem cells of epithelial tissues such as the skin and the lining of the gut, embryonic heart muscle cells,

liver stem cells (see, e.g., WO 94/08598), and neural stem cells (Stemple and Anderson, 1992, Cell 71:973-985).

Epithelial stem cells (ESCs) or keratinocytes can be obtained from tissues such as the skin and the lining of the gut by known procedures (Rheinwald, 1980, Meth. Cell Bio. 21A:229).

5 In stratified epithelial tissue such as the skin, renewal occurs by mitosis of stem cells within the germinal layer, the layer closest to the basal lamina. Stem cells within the lining of the gut provide for a rapid renewal rate of this tissue. ESCs or keratinocytes obtained from the skin or lining of the gut of a patient or donor can be grown in tissue culture (Pittelkow and Scott, 1986, Mayo Clinic Proc. 61:771). If the ESCs are provided by a donor, a method for suppression of
10 host versus graft reactivity (e.g., irradiation, drug or antibody administration to promote moderate immunosuppression) can also be used.

With respect to hematopoietic stem cells (HSC), any technique which provides for the isolation, propagation, and maintenance in vitro of HSC can be used in this embodiment of the invention. Techniques by which this may be accomplished include (a) the isolation and
15 establishment of HSC cultures from bone marrow cells isolated from the future host, or a donor, or (b) the use of previously established long-term HSC cultures, which may be allogeneic or xenogeneic. Non-autologous HSC are used preferably in conjunction with a method of suppressing transplantation immune reactions of the future host/patient. In a particular embodiment of the present invention, human bone marrow cells can be obtained from the
20 posterior iliac crest by needle aspiration (see, e.g., Kodo et al., 1984, J. Clin. Invest. 73:1377-1384). In a preferred embodiment of the present invention, the HSCs can be made highly enriched or in substantially pure form. This enrichment can be accomplished before, during, or after long-term culturing, and can be done by any techniques known in the art. Long-term cultures of bone marrow cells can be established and maintained by using, for example, modified
25 Dexter cell culture techniques (Dexter et al., 1977, J. Cell Physiol. 91:335) or Witlock-Witte culture techniques (Witlock and Witte, 1982, Proc. Natl. Acad. Sci. USA 79:3608-3612).

In a specific embodiment, the nucleic acid to be introduced for purposes of gene therapy comprises an inducible promoter operably linked to the coding region, such that expression of
30 the nucleic acid is controllable by controlling the presence or absence of the appropriate inducer of transcription.

A further embodiment of the present invention relates to a purified antibody or a fragment thereof which specifically binds to a polypeptide that comprises the amino acid sequence set forth in SEQ ID NOs:1, 5 or 6 or to a fragment of said polypeptide. A preferred embodiment relates to a fragment of such an antibody, which fragment is an Fab or F(ab')₂ fragment. In particular, the antibody can be a polyclonal antibody or a monoclonal antibody.

Described herein are methods for the production of antibodies capable of specifically recognizing one or more differentially expressed gene epitopes. Such antibodies may include, but are not limited to polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab')₂ fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above. Such antibodies may be used, for example, in the detection of a fingerprint, target, gene in a biological sample, or, alternatively, as a method for the inhibition of abnormal target gene activity. Thus, such antibodies may be utilized as part of disease treatment methods, and/or may be used as part of diagnostic techniques whereby patients may be tested for abnormal levels of the HDAC9 polypeptide, or for the presence of abnormal forms of the HDAC9 polypeptide.

For the production of antibodies to the HDAC9 polypeptide, various host animals may be immunized by injection with the HDAC9 polypeptide, or a portion thereof. Such host animals may include but are not limited to rabbits, mice, and rats, to name but a few. Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluromic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and Corynebacterium parvum.

Polyclonal antibodies are heterogeneous populations of antibody molecules derived from the sera of animals immunized with an antigen, such as target gene product, or an antigenic functional derivative thereof. For the production of polyclonal antibodies, host animals such as those described above, may be immunized by injection with the HDAC9 polypeptide, or a portion thereof, supplemented with adjuvants as also described above.

Monoclonal antibodies, which are homogeneous populations of antibodies to a particular antigen, may be obtained by 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 of Kohler and Milstein, (1975, Nature 256:495-497; and U.S. Pat. No. 4,376,110), the human B-cell hybridoma technique (Kosbor et al., 1983, Immunology Today 4:72; Cole et al., 1983, Proc. Natl. Acad. Sci. USA 80:2026-2030), and the EBV-hybridoma technique (Cole et al., 1985, Monoclonal Antibodies And Cancer Therapy, Alan R. Liss, Inc., pp. 77-96). Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the mAb of this invention may be cultivated in vitro or in vivo. Production of high titers of mAbs in vivo makes this the presently preferred method of production.

In addition, techniques developed for the production of "chimeric antibodies" (Morrison et al., 1984, Proc. Natl. Acad. Sci., 81:6851-6855; Neuberger et al., 1984, Nature, 312:604-608; Takeda et al., 1985, Nature, 314:452-454) by splicing the genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable or hypervariable region derived from a murine mAb and a human immunoglobulin constant region.

Alternatively, techniques described for the production of single chain antibodies (U.S. Pat. No. 4,946,778; Bird, 1988, Science 242:423-426; Huston et al., 1988, Proc. Natl. Acad. Sci. USA 85:5879-5883; and Ward et al., 1989, Nature 334:544-546) can be adapted to produce differentially expressed gene-single chain antibodies. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide.

Most preferably, techniques useful for the production of "humanized antibodies" can be adapted to produce antibodies to the polypeptides, fragments, derivatives, and functional equivalents disclosed herein. Such techniques are disclosed in U.S. Patent Nos. 5,932, 448; 5,693,762; 5,693,761; 5,585,089; 5,530,101; 5,910,771; 5,569,825; 5,625,126; 5,633,425; 5,789,650; 5,545,580; 5,661,016; and 5,770,429, the disclosures of all of which are incorporated by reference herein in their entirety.

Antibody fragments that recognize specific epitopes may be generated by known techniques. For example, such fragments include but are not limited to: the F(ab')₂ fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the F(ab')₂ fragments. Alternatively, 5 Fab expression libraries may be constructed (Huse et al., 1989, Science, 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity.

An antibody of the present invention can be preferably used in a method for the diagnosis of a condition associated with abnormal HDAC9 expression or activity, for example, abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, host inflammatory or 10 immune response, or psoriasis, in a human which comprises: measuring the amount of a polypeptide comprising the amino acid sequence set forth in SEQ ID NOs:1, 5 or 6, or fragments thereof, in an appropriate tissue or cell from a human suffering from a condition associated with abnormal HDAC9 activity, wherein the presence of an elevated amount of said polypeptide or fragments thereof, relative to the amount of said polypeptide or fragments thereof in the 15 respective tissue from a human not suffering from a condition associated with abnormal HDAC9 activity is diagnostic of said human's suffering from such condition. Such a method forms a further embodiment of the present invention. Preferably, said detecting step comprises contacting said appropriate tissue or cell with an antibody which specifically binds to a polypeptide that comprises the amino acid sequence set forth in SEQ ID NOs:1, 5 or 6 or a fragment thereof and 20 detecting specific binding of said antibody with a polypeptide in said appropriate tissue or cell, wherein detection of specific binding to a polypeptide indicates the presence of a polypeptide that comprises the amino acid sequence set forth in SEQ ID NOs:1, 5 or 6 or a fragment thereof.

Particularly preferred, for ease of detection, is the sandwich assay, of which a number of variations exist, all of which are intended to be encompassed by the present invention.

25 For example, in a typical forward assay, unlabeled antibody is immobilized on a solid substrate and the sample to be tested brought into contact with the bound molecule. After a suitable period of incubation, for a period of time sufficient to allow formation of an antibody-antigen binary complex. At this point, a second antibody, labeled with a reporter molecule capable of inducing a detectable signal, is then added and incubated, allowing time sufficient for 30 the formation of a ternary complex of antibody-antigen-labeled antibody. Any unreacted material

is washed away, and the presence of the antigen is determined by observation of a signal, or may be quantitated by comparing with a control sample containing known amounts of antigen. Variations on the forward assay include the simultaneous assay, in which both sample and antibody are added simultaneously to the bound antibody, or a reverse assay in which the labeled
5 antibody and sample to be tested are first combined, incubated and added to the unlabeled surface bound antibody. These techniques are well known to those skilled in the art, and the possibility of minor variations will be readily apparent. As used herein, "sandwich assay" is intended to encompass all variations on the basic two-site technique. For the immunoassays of the present invention, the only limiting factor is that the labeled antibody be an antibody which is
10 specific for the HDAC9 polypeptide or a fragment thereof.

The most commonly used reporter molecules in this type of assay are either enzymes, fluorophore- or radionuclide-containing molecules. In the case of an enzyme immunoassay an enzyme is conjugated to the second antibody, usually by means of glutaraldehyde or periodate. As will be readily recognized, however, a wide variety of different ligation techniques exist,
15 which are well-known to the skilled artisan. Commonly used enzymes include horseradish peroxidase, glucose oxidase, beta-galactosidase and alkaline phosphatase, among others. The substrates to be used with the specific enzymes are generally chosen for the production, upon hydrolysis by the corresponding enzyme, of a detectable color change. For example, p-nitrophenyl phosphate is suitable for use with alkaline phosphatase conjugates; for peroxidase
20 conjugates, 1,2-phenylenediamine or toluidine are commonly used. It is also possible to employ fluorogenic substrates, which yield a fluorescent product rather than the chromogenic substrates noted above. A solution containing the appropriate substrate is then added to the tertiary complex. The substrate reacts with the enzyme linked to the second antibody, giving a qualitative visual signal, which may be further quantitated, usually spectrophotometrically, to give an
25 evaluation of the amount of HDAC9 which is present in the serum sample.

Alternately, fluorescent compounds, such as fluorescein and rhodamine, may be chemically coupled to antibodies without altering their binding capacity. When activated by illumination with light of a particular wavelength, the fluorochrome-labeled antibody absorbs the light energy, inducing a state of excitability in the molecule, followed by emission of the light at
30 a characteristic longer wavelength. The emission appears as a characteristic color visually

detectable with a light microscope. Immunofluorescence and EIA techniques are both very well established in the art and are particularly preferred for the present method. However, other reporter molecules, such as radioisotopes, chemiluminescent or bioluminescent molecules may also be employed. It will be readily apparent to the skilled artisan how to vary the procedure to
5 suit the required use.

This invention also relates to the use of polynucleotides of the present invention as diagnostic reagents. In particular, the invention relates to a method for the diagnosis of a condition associated with abnormal HDAC9 expression or activity, for example, abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, host inflammatory or immune
10 response, or psoriasis in a human which comprises: detecting elevated transcription of messenger RNA transcribed from the natural endogeneous human gene encoding the polypeptide consisting of an amino acid sequence set forth in SEQ ID NOs: 1, 5 or 6 in an appropriate tissue or cell from a human, wherein said elevated transcription is diagnostic of said human's suffering from the condition associated with abnormal HDAC9 expression or activity. In particular, said natural
15 endogeneous human gene comprises the nucleotide sequence set forth in SEQ ID NO: 4, 7 or 8. In a preferred embodiment such a method comprises contacting a sample of said appropriate tissue or cell or contacting an isolated RNA or DNA molecule derived from that tissue or cell with an isolated nucleotide sequence of at least about 20 nucleotides in length that hybridizes under high stringency conditions with the isolated nucleotide sequence encoding a polypeptide
20 consisting of an amino acid sequence set forth in SEQ ID NOs: 1, 5 or 6.

Detection of a mutated form of the gene characterized by the polynucleotide of SEQ ID NO: 4, 7 or 8 which is associated with a dysfunction will provide a diagnostic tool that can add to, or define, a diagnosis of a disease, or susceptibility to a disease, which results from under-expression, over-expression or altered spatial or temporal expression of the gene. Individuals
25 carrying mutations in the gene may be detected at the DNA level by a variety of techniques.

Nucleic acids, in particular mRNA, for diagnosis may be obtained from a subject's cells, such as from blood, urine, saliva, tissue biopsy or autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR or other amplification techniques prior to analysis. RNA or cDNA may also be used in similar fashion.
30 Deletions and insertions can be detected by a change in size of the amplified product in

comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to labeled nucleotide sequences encoding the HDAC9 polypeptide of the present invention. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase digestion or by differences in melting temperatures. DNA sequence differences may also be
5 detected by alterations in electrophoretic mobility of DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing (e.g., Myers et al., Science (1985) 230:1242). Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (see Cotton et al., Proc Natl Acad Sci USA (1985) 85: 4397-4401). In another embodiment, an array of oligonucleotides probes
10 comprising nucleotide sequence encoding the HDAC9 polypeptide of the present invention or fragments of such a nucleotide sequence can be constructed to conduct efficient screening of e.g., genetic mutations. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability (see for example: M. Chee et al., Science, Vol
15 274, pp 610-613 (1996)).

The diagnostic assays offer a process for diagnosing or determining a susceptibility to disease through detection of mutation in the HDAC9 gene by the methods described. In addition, such diseases may be diagnosed by methods comprising determining from a sample derived from a subject an abnormally decreased or increased level of polypeptide or mRNA. Decreased or
20 increased expression can be measured at the RNA level using any of the methods well known in the art for the quantitation of polynucleotides, such as, for example, nucleic acid amplification, for instance PCR, RT-PCR, RNase protection, Northern blotting and other hybridization methods. Assay techniques that can be used to determine levels of a protein, such as a polypeptide of the present invention, in a sample derived from a host are well-known to those of
25 skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays.

Thus in another aspect, the present invention relates to a diagnostic kit which comprises:

- (a) a polynucleotide of the present invention, preferably the nucleotide sequence of SEQ ID NO:2, 3, 4, 7 or 8 or a fragment thereof;
- 30 (b) a nucleotide sequence complementary to that of (a);

(c) a polypeptide of the present invention, preferably the polypeptide of SEQ ID NOs:1, 5 or 6 or a fragment thereof; or

(d) an antibody to a polypeptide of the present invention, preferably to the polypeptide of SEQ ID NOs:1, 5 or 6.

5 It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component. Such a kit will be of use in diagnosing a disease or susceptibility to a disease, particularly to a disease or condition associated with abnormal HDAC9 expression or activity, for example, abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, host inflammatory or immune response, or psoriasis.

10

The nucleotide sequences of the present invention are also valuable for chromosome localization. The sequence is specifically targeted to, and can hybridize with, a particular location on an individual human chromosome. The mapping of relevant sequences to chromosomes according to the present invention is an important first step in correlating those
15 sequences with gene associated disease. Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found in, for example, V. McKusick, Mendelian Inheritance in Man (available on-line through Johns Hopkins University Welch Medical Library). The relationship between genes and diseases that have been mapped to the same
20 chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes).

The differences in the cDNA or genomic sequence between affected and unaffected individuals can also be determined. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative
25 agent of the disease.

An additional embodiment of the invention relates to the administration of a pharmaceutical composition, in conjunction with a pharmaceutically acceptable carrier, excipient or diluent, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of HDAC9, antibodies to that polypeptide, mimetics, agonists, antagonists, or
30 inhibitors of HDAC9 function. The compositions may be administered alone or in combination

with at least one other agent, such as stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs or hormones.

5 In addition, any of the therapeutic proteins, antagonists, antibodies, agonists, antisense sequences or vectors described above 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
10 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. Antagonists and agonists of HDAC9 may be made using methods which are generally known in the art.

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

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which
20 facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, Pa.).

Pharmaceutical compositions for oral administration can be formulated using
25 pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combination of active
30 compounds with solid excipient, optionally grinding a resulting mixture, and processing the

mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums
5 including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel,
10 polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or
15 sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in
20 aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles
25 include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder which may contain any or all of the following: 1-50 mM histidine, 0.1%-2% sucrose, and 2-7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of the HDAC9, such labeling would include amount, frequency, and method of administration.

Pharmaceutical 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.

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, usually mice, rabbits, dogs, or pigs. The 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 HDAC9 or fragments thereof, antibodies of HDAC9, agonists, antagonists or inhibitors of HDAC9, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., ED₅₀ (the dose therapeutically effective in 50% of the population) and LD₅₀ (the dose lethal to 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index, and it can be expressed as the ratio, LD₅₀/ED₅₀. Pharmaceutical compositions which

exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, 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 that requires 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, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or once every two weeks depending on half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, 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. Pharmaceutical formulations suitable for oral administration of proteins are described, e.g., in U.S. Patents 5,008,114; 5,505,962; 5,641,515; 5,681,811; 5,700,486; 5,766,633; 5,792,451; 5,853,748; 5,972,387; 5,976,569; and 6,051,561.

The following Examples illustrate the present invention, without in any way limiting the scope thereof.

EXAMPLES

Example 1: Identification of a novel HDAC related human DNA sequence using bioinformatics
HDAC9 was identified using computer software for the identification of new members of gene families based on a strategy to find maximal evolutionary links among known HDAC family members by first searching the non-redundant amino acid database, followed by searching less

diverse databases such as the Celera Human Genome Database (CHGD), public High Throughput Genomic (HTG) database and the Incyte LIFESEQ™ database. Smith-Waterman (Pearson W. R. Comparison of methods for searching protein sequence databases. *Protein Sci* (1995) 4,1145-60) and Hidden Markov Models (probability models derived from diversity of amino acids at every position (Eddy S. R. Hidden Markov models. *Curr Opin Struct Biol* (1996) 6, 361-5) were performed. An 1156 bp open reading frame (ORF) was identified and used to search a database of sequenced clones from pan-tissue and dorsal root ganglion cDNA libraries.

Example 2: Construction of pan-tissue and dorsal root ganglion cDNA libraries

10 Pan-tissue and dorsal root ganglion cDNA libraries are prepared from polyA+ RNA. Total RNA was extracted from a pooled sample of 31 human tissues or dorsal root ganglia and isolated using TRIZOL reagent according to manufacturer's instructions (Life Technologies, Rockville, MD). mRNA is isolated using Polytract mRNA Isolation System III according to manufacturer's instructions (Promega, Madison, WI). Total RNA is hybridized to a biotinylated-oligo (dt) probe.

15 The oligo (dt)-mRNA hybrids are captured on streptavidin magnesphere particles and eluted in RNase-free H₂O. 3 ul of biotinylated-oligo(dt) probe (50 pmol/ul) and 13 ul of 20X SSC is added to 60-150 ug of RNA that is heated to 65°C in RNase free water. This mixture is incubated at room temperature until it is completely cooled. Streptavidin-paramagnetic particles (beads) are resuspended and washed 3 times in 0.5X SSC and then resuspended in 0.5X SSC. The RNA-oligo(dt) hybrids from the previous step are added to these beads. To release the poly-A RNA from the beads, the beads are resuspended in RNase-free water and magnetically captured and then the eluate from the beads is ethanol precipitated. First and second strand cDNA synthesis is performed using a modified procedure from Life Technologies (D'Alessio, J. M., Gruber, C.E., Cain, C. R., and Noon, M. C. (1990) *Focus* 12, 47). First strand synthesis is performed by

25 incubating 1-5 ug of RNA that is heated to 60°C in 1X 1st strand buffer (Life Technologies)/6 mM DTT/600 nM dNTPs/2 units anti-Rnase. This mixture is incubated at 40°C for 2 min, then Superscript II reverse transcriptase (RT) and 1 ul of Display Thermo RT terminator mix is added and the mixture is incubated at 40°C for 1 h, followed by incubation at 60°C for 10 min. Second strand synthesis is performed in 1x second strand buffer (Life Technologies) in DEPC-H₂O/66

30 nM/1 ul E.coli DNA ligase/4 ul E. coli DNA polymerase I/1 ul E. coli Rnase H. This mixture is

incubated at 10°C for 10 min and then at 16°C for 2h. To this mixture, 2 ul of T4 DNA polymerase is added and incubation is continued at 16°C for 5 min. The reaction is stopped with 10 ul of 0.5M EDTA, extracted with phenol/chloroform/isoamyl alcohol and then ethanol precipitated. Sal I and Not I adaptors are added to the 5' ends of the cDNAs by ligation for directional cloning using conventional methodology. The cDNAs are then passed through a size fractionation column to retrieve cDNAs that are >500 bp in length according to manufacturers instructions (Life Technologies, Rockville, MD). cDNAs are ligated to Sal I/Not I digested Gateway compatible pCMV-Sport6 vector (Life Technologies, Rockville, MD) using conventional methods. Competent DH10B cells (Life Technologies, Rockville, MD) are transformed with the resulting library using conventional methods. Semi-solid amplification of the libraries is performed according to the manufacturer's instructions (Life Technologies, Rockville, MD).

Example 3: Preparation of full length cDNA encoding the novel HDAC9 consisting of SEQ ID NO:1, 5 or 6. An 1156 base pair ORF was used to search a database of sequenced clones from pan-tissue and dorsal root ganglion cDNA libraries using BLAST. Four clones were found to contain the ORF (M6, K10, P3, F23), two from each library. Of these clones M6 from the pan-tissue library was determined to be the most complete, but missing approximately 44 bp from the N-terminus. A protein slightly smaller than that predicted for the complete cDNA was observed by *in vitro* translation. The result that proteins were observed by *in vitro* translation of the incomplete cDNA, suggests possibility of alternate translation initiation sites within HDAC9. Specifically, sequencing of HDAC9 in pCMVSPORT6 was performed using an automated ABI Sequencer (ACGT, Northbrook, IL). PCR was performed using conditions listed in the ABI Prism BigDye™ Terminator Cycle Sequencing Ready Reaction Kit manual and are as follows: denaturation at 96°C for 30 seconds, annealing at 50° C for 15 seconds, extension at 60°C for 4 minutes, for a total of 25 cycles. Each round of sequencing provided between 200 and 600 bp of sequence. PCR primers for 1st round sequencing were 5'-ATTAGGTGACACTATAG -3' (Sp6, sense) and 5'-TAATACGACTCACTATAGGG -3' (T7, antisense). Results of sequencing using Sp6 primer are as follows. Bolded sequence is pCMVSPORT6 vector sequence.

CTggTACCGTCCGGAATCCCGGGATATCGTCGACCCACGCGTCCG/GGCTGCT

CCGGGCCGAAGCCCCGAGTGCAGATCGAGCGTCCTGAGCGCCTGACCCGAGCCCT
 GGATCGCCTGCGGCAGCGCGGCTGGAACAGAGGTGTCTGCGGTTGTACAGCCGCG
 AGGCCTCGGAAGAGGAGCTGGGCCTGGTGCACAGCCAGAGTATGTATCCCTGGTC
 AGGGAGACCCAGGTCTAGGCAAGGAGGAGCTGCAGGCGCTGTCCGGACAGTTTGA
 5 CGCCATCTACTTCCACCCGAGTACCTTTCACTGCGCGCGGCTGGCCGCAGGGGCTGG
 ACTGCAGCTGGTGGACGCTGTGCTCACTGGAGCTGTGCA:AAATGGGCTTGCCCTGG
 TGAGGCCTCCCGGGCACCATGGCCAGAGGGCGGCTGCAACGGGTTCTGCGTGTTC
 ACAACGTGGCCATAGCAGCTGCACATGCCAAGCAGAAACACGGGCTACACAGGATC
 CTCGTCTGGACTGGGGGATGTGCACCATGGCAGGGGGATCCAGTATCTCTTGAAG
 10 GATGACCCAGCGTCTTTACTTCTCTGGCACCCTATGAGCATTGGGCGCCTTCT
 GGCCCTTCTGCGAGAGTCAGATGA_gACGCATGGGGGGCGGGGACAGGGCCTCGGC
 TTCCTGTCaACCTGCCCTGACCAAGT_gGGGGAATGGGAAACGCTGACTTACGTG
 GCTGGCCTTCTTGACCTTGTGTTCCa_gTGGCCTTTTGGAGTTTGACCTGA_gCTGG
 GTGCTTGGT_cTC_gGCAGGGATTGTGACT_cCa_gTi_gC_gGGACCCTGA_gGGGGCAAA. Results
 15 of sequencing using the T7 primer were:
 TCAAGCCACCAGGTGAGGATGGCACTGCAACATCTTCACTGAGGCTCCAGCTGCCC
 TCTCAGGTACATCAGGGCCTGGACGTCTCTGGGGAGGCCACAGAGGAAGGGCCTA
 GGCTAGGAGGTGCTCTCCATTACAGCACCCGGGCCAGGATCCCTGTAGCTGGGGTG
 TGGAGTTCTCTCCAGGAGGGGCCAGGACTCGGCCCCCTGCCAGCCCCGAAGCATTG
 20 CAGCCAGGAGTGCAGCGTGGGGGCCCTGCAGGCCATGGCCAGGCCACGCGCCACC
 AGCACCAGGTGAGGCTGGAAGCCATAGGCCAGGGGCAGCaCCAAGCCCAAGATGCA
 GCTCAGGAAACCACCGGTATCACTGGCAGTGGCGTGGAGACATGGAACATGGA[T
 AGGGCAGcCGCCTCCTTGCCCTGATGTTTACGCCACAGACTcCTCCCGTCATGGGCGA
 AGTCTGGAGGCCGTCCAgCTGT_{la}GGCCACGCACAG_{gt}CTCTGGGCTCC_{gt}GGGACA
 25 gGCCT:TTT_{la}GAAAGAtATT_{la}AGGGTGGGTTGTGA_{acagg}GCTGGAATGGCTGGTATAc
 AcTGT_{TT}AcCTGCCATT. 2nd and 3rd round sequencing primers are designed to prime sequence
 obtained from the previous round of sequencing. 2nd round primers are 5'-GTCATCA
 CTGGCAGTGGCGTG -3' (HUF7392, antisense) and 5'-TGGACTGCAGCTGGTGG -3' (DF-2,
 sense). Results of sequencing using the DF-2 primer were: CTGGcAAATGGGCTTGCCCTGG
 30 TGAGGCCTCCCGGGCACCATGGCCAGAGGGCGGCTGCCAACGGGTTCTGCGTGTTC

AACAACGTGGCCATAGCAGCTGCACATGCCAAGCAGAAACACGGGCTACACAGGAT
CCTGCTGCTGGACTGGGATGTGCACCATGGCCAGGGGATCCAGTATCTCTTTGAGGA
TGACCCAGCGTCTCTTACTTCTCTGGCACCGCTATGAGCATGGGCGCTTCTGGCCT
TTCTGCGAGAGTCAGATGCAGACGCACTGGGGCGGGACAGGGCCTCGGCTTAC
5 TGTC AACCTGCCCTGGAACCAGGTTGGGATGGGAAACGCTGACTACGTGGCTGCCTT
cCTGCACCTGCTGCTCCCACTGGCCTTTGAGTTTGACCTGAGCTGGTGCTGCTCG
GCAGGATTTGACTCAGCCATCGGGGACCTGAGGGGCAAATGCAGGCCACGCCAGA
GTGCTTCGCCACCTCACACAGCTGCTGCAGGTGCTGGCCGGCGGCGGCTGTGTGC
CGTGCTGGAGGGCGGCTACCACTGGAGTCACTGGCGGAGTCAGTGTGATGACAG
10 TACAGACGCTGCTGGGTGACCCGGcCCACCCCTGTCAAGGCCAATGGCGCCATGTC
AGAGTGCCTAgATTCATTCAgAGTGCCCGTGTGCCAGGcCCCGCACTGGAAAgAgG
CTTCagCAGCAAgATGTGACCGcTGTGCCGATGAACCCCA. Sequencing results for the
HUF7392 primer were: TGtaTAGGGcAGCCGCTCTTGCC
CCTGATGTTAGCCACAGACTCCTCCCGTCATGGGCGAGG
15 TCTGGAGGCCGTCCAGCTGTCCAGGGCCACGCACAGCAGCCTCTGGGCTCCGTG
GGACAGGCCTCTCCGAACAGCCACATCCAGGGTGGCTGCTGCAGCAGAGGCTGGAG
TGGCTGCTATACCACTGTTACCTGCCATCCAGCATCCCATCTAAGAGGTACAGGA
GCTTCCCAAGTGCACTGAGGGCTCCTCCCGGCCAGGGACTCGTGTGGCTGGCCC
AGGCTTCTGTCTCCTCCTCAGGGCTGACGCTTCTGTTGGATGACGTCAGGGGGCAG
20 AACCAATGTGATATCCGGCGTGTCAAGGGCAACAGCGGTGCCGACAGAGGTGCG
GGGCAGAGGCACgGCTGGTCCAgGAGGGAGCTCGGTGCAgATGCAGcTGCCTTACAC
ACTGgACCCCAAGGCAGCAGAGGTGGAGGCCTCCCTCTGGGGAGTG. 3rd round
sequencing primers were 5'-AACAGCGGTG C GGACAGA -3' (HUF2A, antisense) and 5'-
CTGGAGTCACTGGCGGAG -3' (DF3A, sense). Results of sequencing using DF3A primer
25 were: AgcaCAGA cGCTgCTGGGTGACCCGGCCACCCCTG
TCAGGGCCAATGGCGCCATGTCAAGTGCCTAGAGTCCATCCAGAGTGcCCGTGCT
GCCCAGGCCCCGCACTGGAAGAGCCTCCAGCAGCAAGATGTGACCGCTGTGCCGAT
GAGCCCCAGCAGCACTCCCCAGAGGGAGGCTCCACCTCTGCTGCTGGGGGTC
CAGTGTGTAAGGCAGCTGCATCTGCACCGAGCTCCCTCCTGGACCAGCCGTGCCTCT
30 GCCCGCACCCCTCTGTCCGCACCGCTGTTGCCCTGACAACGCCGGATATCACATTGG

WO 02/050285

52

PCT/EP01/14928

TTCTGCCCCCTGACGTCATCCAACAGGAAGCGTCAGCCCTGAGGGAGGAGACAGAA
GCCTGGGCCAGGCCACACGAGTCCCTGGCCCGGAGGAGGCCCTcACTGcACTTGGG
AAGCTCCTGTACCTcTIAgATGGGATGCTGGATGGGCAGGTGAACAgTGGTATA.

Results of sequencing using HUF2A primer were: TgcaCGGATGGTCCAGGAGGGAGCTCG

5 GTGCAAAATGCAGCTGCCTTACACACTGGACCCCGAGGCAGCAgAGGTGGAGGCCTC
CCCTcTGGGGAGTGGCTGCTGGGGCTCATCGGCACAGCGGTCACATCTTGTGCTGG
AGGCTCTTCCAGTGCGGGGCCCTGGGCAGCACGGGCACTCTGGATGGACTCTAGGGC
ACTCTGACATGGCGCCATTGGCCCTGACAGGGGTGGGGCCGGGTCACCCAGCAGCG
TCTGTACTGTATGCACACTGACTCCGCCAGTGACTCCAGGTGGTAGCCGCCCTCCA
10 GCACGGCACAgACCCGGCCGCCGGCCAGCACCTGCAGCAGCTGTGTGAGGTGGGcg
AAGCACTCTGGCGTGGCCTGCATTGCCCCCTCAGGGTCCCCGATGGCTTGAGTCAAA
TCCTGCCGAGACCAGCACCAGCTCAGGGTCAAACTCAAAGGCCAGTGGGAGCAGCA
GGTGAGGAAGGCCAGCCACgTATCAGCGTTTCCCATCCCAACCTGgTTCCAGGGGCA
GGTTGAACAGTGAAGCCGAGGGCCCTTGTCCCCgCCCCACCTTGCGTCTGCATcGA
15 CTCTCGCAGGAAAGGCCAAgAAGCgCCCATgCTATTTT. The overlapping sequence from
the combined sense and antisense sequencing was reconstructed to give the complete cDNA
sequence of HDAC9. See Figure 2A.

BLAST is used to search the Genbank database using cDNA clone M6 as the query to
identify a genomic sequence containing M6 cDNA sequence. The results of this search identified
20 a genomic sequence AL022328 that was found to contain exons that were identical in sequence
to the M6 cDNA. The sequence of cDNA clone M6 was confirmed by automated DNA
sequencing (ACGT, Inc. Northbrook, IL). See Figure 2A.

The remaining 44 bp of N-terminal sequence was added by PCR using the nested sense
strand primers 5'-GCGGTCGACGCCACCATGGGGACCGCGCTTGTGTACCATGAGGAC
25 ATG-3' and 5'-GTGTACCATGAGGACATGACGGCCACCCGGCTGCTCTGGGACGACC
CCGAGTGC-3' and the 3' primer 5'-GAACCAATGTGATATCCGGCGTTG-3'. The 5' primer
added a kozak sequence and a SalI site for cloning and the 3' primer sequence overlaps the
EcoRV site in HDAC9. PCR was performed using a step-cycle file for amplification using 1
cycle of 94°C for 30 seconds, 68°C for 30 seconds, and 72°C for 1 minute, followed by 20 cycles
30 of 94°C for 30 seconds and 72°C for 1 minute.

Example 3 HDAC9 sequence variants

Three variants of the HDAC9 sequence, HDAC9v1, HDAC9v2, and HDAC9v3 were found. HDAC9v1 is the original sequence found and described above. HDAC9v2 was found in the human dorsal root ganglion cDNA library and in AL022328 genomic sequence. HDAC9v3 is a predicted transcript that lacks a stop codon that was found in the Celera human genomic database. HDAC9v1 contains 20 exons and HDAC9v2 has 20 exons. Comparison of the peptide sequences of HDAC9 variants demonstrated that HDAC9v1 and HDAC9v2 were identical up to exon 17, but diverge after this exon. HDAC9v2 has an extended intron between exon 17 and 18 and an extended exon 18 that contains HDAC9v1 exon 19, but lacks 20, as a result of a single nucleotide insertion at nucleotide 446. This insertion frame shifts the sequence and shortens the peptide by 11 amino acids (Fig 11A). Compared to HDAC9v1 and HDAC9v2, HDAC9v3 has an internal deletion of amino acids 219 through 240 and diverges in its C-terminal beginning at amino acid 486. HDAC9 is the first HDAC enzyme for which sequence variants have been reported. HDAC9v1 is the sequence variant that is characterized, unless otherwise noted.

Example: 4 Identification of HDAC-associated sequence motifs.

The M6 clone was analyzed for the presence of motifs that would indicate an HDAC catalytic domain and a binding site for Rb and Rb-like proteins. HDACs are characterized by the presence of a catalytic domain with conserved amino acids. Most of the HDACs that have been identified to date have one catalytic domain, with the exception of HDAC6 that has two domains. N-terminal catalytic domains have been associated with class I HDACs, while C-terminal catalytic domains are associated with class II HDACs. An N-terminal catalytic domain was found in HDAC9 based upon PFAM prediction and alignment with the catalytic domains of

other HDACs. A set of conserved amino acids were previously shown to be critical for HDAC activity and provide the critical contacts for HDAC inhibitor, TSA, based upon single amino acid mutations in HDAC1 and the three dimensional structure formed by a complex of an HDAC-like protein (HDLP), Zn^{2+} and HDAC inhibitor TSA (Hassig CA, Tong JK, Fleischer TC, Owa T, Grable PG, Ayer DE, Schreiber SL. (1998) *Proc Natl Acad Sci U S A*. **95**, 3519-3524; Finnin, M. S., Donigian, J. R., Cohen, A., Richon, V. M., Rifkind, R. a., Marks, P. A., Breslow, R., and Pavletich, N. P. (1999) Structures of a histone deacetylase homologue bound to TSA and SAHA inhibitors. *Nature* **401**, 188-193). A bacterial protein with similarities in sequence and enzymatic activity to human HDACs and the only class I HDAC-like structure elucidated, HDLP was used as an HDAC template. Many of these conserved amino acids with a few exceptions were found in HDAC9 (Table 4). Alignments of HDAC peptide sequences indicated that the hydrophobic residue Leu 265 that forms part of the binding pocket in HDLP is replaced with Glu at amino acid 272 in HDAC9. Similarly, Leu 265 is also replaced with Met in HDAC8 and with Lys in HDAC6 domain 1. Furthermore, Asp 173 in HDLP is substituted with Gln at position 177 in HDAC9, a difference that was also found in the HDAC6 catalytic domain 1. This Asp is substituted with Asn in HDAC4, HDAC5, HDAC6 domain 2, and HDAC7. HDAC1-8 have been shown to be catalytically active, hence the amino acid substitutions in these proteins have no enzymatic consequences.

HDAC9 is similar in sequence to class I and class II HDACs. HDACs have been classified by their sequence similarity with yeast HDACs Rpd3, Hda1, and Sir2 and by catalytic domain location. Alignment of the peptide sequences of HDAC9, yeast HDACs Rpd3, Hda1, Hda1 subfamily member from fission yeast, cryptic loci regulator 3 (Clr3), and Sir2 determined

that HDAC9 had the highest sequence similarity with Ctr3 (Table 1). However, the sequence similarity is not high enough to categorize HDAC9.

Alignment of human HDACs 1-9 and Sir 1-7 peptide sequences demonstrated that HDAC9 was most similar to class II human HDAC6 (Table 2). Alignment of class I and class II HDAC catalytic domains with HDAC9 catalytic domains demonstrated that HDAC6 catalytic domain 1 has the most sequence similarity with HDAC9 (Table 3).

In order to compare the locations of catalytic domains in HDACs, PFAM predictions were made of the catalytic domains in HDAC peptides (Fig. 11B). The location of HDAC9 catalytic domain was at the N-terminus, similar to class I HDACs, and was estimated as spanning the amino acid sequence from amino acid 4 to 323. In addition, the average length of class I HDACs is 443 amino acids, while the average length of class II HDACs is 1069 amino acids. The 673 amino acid HDAC9 peptide is between the average sizes of class I and class II HDACs (Fig. 11B).

Table 1.

HDAC Class	HDAC Isoform	%Similarity to HDAC9
Class I	Rpd3	16
Class II	Hda1	18
	Ctr3	23
Class III	Sir2	5

Table 2.

HDAC Class	HDAC Isoform	% Similarity to HDAC9
Class I	HDAC1	14
	HDAC2	15
	HDAC3	15
	HDAC8	22
Class II	HDAC4	21
	HDAC5	19
	HDAC6	37
	HDAC7	20

Class III	Sir1	5
	Sir2	7
	Sir3	11
	Sir4	4
	Sir5	8
	Sir6	10
	Sir7	15

Table 3.

HDAC Class	HDAC Isoform	% Similarity to HDAC9
Class I	HDAC1	20
	HDAC2	20
	HDAC3	20
	HDAC8	19
Class II	HDAC4	39
	HDAC5	38
	HDAC6-1	55
	HDAC6-2	53
	HDAC7	40

- 5 The protein product of the retinoblastoma protein (Rb) gene is a transcriptional regulator that controls DNA synthesis, the cell cycle, differentiation and apoptosis and plays a tissue-specific role normal development. Rb complexes with the transcription factor E2F, an interaction that is regulated by phosphorylation. Mutations in Rb lead to a hereditary form of cancer of the retina, retinoblastoma. Mutations have also been found in a number of mesenchymal and
- 10 epithelial cancers. Mutations that affect regulators of Rb phosphorylation including, cyclin D1, cdk4, and p16 have been found in many cancers. Therefore, Rb function is thought to play a critical role in tumorigenesis (Sellers, W.R., Kaelin, W.G. Jr. (1997) *J. Clin. Oncol.* 15, 3301-3312, DiCiommo, D., Gallie, B.L., Bremner, R.(2000) *Semin. Cancer Biol.* 10, 255-269). An Rb-binding motif was previously defined as the amino acid sequence LXCXE, where "X" can be
- 15 any amino acid (Chen, T.-T. and Wang, J. Y. J. (2000) *Mol. Cell Biol.* 20, 5571-5580). The LXCXE domain in HDAC1 was found to be dispensible for growth suppression function of Rb, but necessary for HDAC binding to Rb. Two putative Rb-binding motifs were found in HDAC9 (Fig. 11A, green boxes). LLCVA is located between amino acids 510 and 515, and LSCIL located between amino acids 560 and 564. Both are present in HDAC9v1 and HDAC9v2.

WO 02/050285

57

PCT/EP01/14928

4-311000A/USN

HDAC Isoform		Amino acids in the catalytic domains of HDAC isoforms													
		Pro	Tyr	His	His	Gly	Phe	Asp	His	Asp	Phe	Asp	Leu	Tyr	
HDLP	22	91	131	132	140	141	166	168	170	173	198	258	265	297	
HDAC1	Pro	Glu	His	His	Gly	Phe	Asp	Asp	His	Asp	Phe	Asp	Leu	Tyr	
	29	98	139	140	148	149	174	176	178	181	205	264	271	303	
HDAC2	Pro	Glu	His	His	Gly	Phe	Asp	Asp	His	Asp	Phe	Asp	Leu	Tyr	
	30	99	140	141	149	150	175	177	179	182	206	265	272	304	
HDAC3	Pro	Asp	His	His	Gly	Phe	Asp	Asp	His	Asp	Phe	Asp	Leu	Tyr	
	23	92	134	135	143	144	167	168	171	174	199	259	266	298	
HDAC4	Pro	Trp	His	His	Gly	Phe	Asp	Asp	His	Asn	Phe	Asp	Leu	His	
	592	762	802	803	900	901	838	839	843	846	870	934	943	976	
HDAC5	Pro	Trp	His	His	Gly	Phe	Asp	Asp	His	Asn	Phe	Asp	Leu	His	
	705	793	832	833	841	842	868	869	873	876	900	964	973	1006	
HDAC6-1	Pro	Tyr	His	His	Gly	Tyr	Asp	Asp	His	Gln	Phe	Asp	Lys	Tyr	
	106	175	215	216	224	225	251	252	255	258	283	946	353	386	
HDAC6-2	Pro	Tyr	His	His	Gly	Phe	Asp	Asp	His	Asn	Phe	Asp	Leu	Tyr	
	501	570	593	594	602	603	647	648	651	654	679	742	749	782	
HDAC7	Pro	Tyr	His	His	Gly	Phe	Asp	Asp	His	Asn	Phe	Asp	Leu	His	
	502	589	629	630	638	639	668	669	673	676	700	764	773	806	
HDAC8	N/A	Tyr	His	His	Gly	Phe	Asp	Asp	His	Asp	Phe	Asp	Met	Tyr	
	100	141	142	150	151	176	177	180	183	208	267	274	306		
HDAC9	Pro	Tyr	His	His	Gly	Phe	Asp	Asp	His	Gln	Phe	Asp	Glu	Tyr	
	21	94	134	135	143	144	170	173	174	177	205	265	272	305	

Non-conserved amino acids (bold text). No alignment (N/A).

57

Example 5: mRNA distribution of HDAC9 in normal tissues

mRNA distribution of HDAC9 in normal tissues is investigated using Northern analysis.

Probes are prepared by ³²P-labeling a 750 bp EcorV/NotI HDAC9 fragment using Redi-Prime
5 random nucleotide labelling kit according to manufacturer's instructions (Amersham,
Piscataway, NJ). A Northern blot containing polyA+ RNA from 12 normal tissues (Origene
Technologies, Rockville, MD) and an array of matched tumor *versus* normal cDNAs (Clontech,
Palo Alto, CA) are probed with the [³²P]-labeled 750 bp EcorV/NotI HDAC9 fragment and
washed under high stringency conditions (68°C). Hybridized blots are washed two times for 15
10 min at 68°C in 2 X SSC /0.1% SDS followed by two 30 min washes in 0.1 X SSC/0.1% SDS at
68°C. The blot is exposed to film with an intensifying screen for 18 hr. Results indicate that an
approximately 3.0 Kb HDAC9 mRNA is detected in brain, colon, heart kidney, liver, lung,
placenta, small intestine, spleen, stomach and testes. HDAC9 message was not detected in
muscle, but GAPDH was also not detected. See Figure 7.

15 Analogous computer techniques using BLAST (Altschul, S.F. 1993, 1990 refs) are used to
search for identical or related molecules in nucleotide databases such as GenBank or the
LIFESEQ™ database. The basis of the search is the product score which is defined as:

$$\frac{\% \text{ sequence identity} \times \% \text{ maximum BLAST score}}{100}$$

20

The product score takes into account both the degree of similarity between two sequences and
the length of the sequence match. For example, with a product score of 40, the match will be
exact within a 1-2% error; and at 70, the match will be exact. Homologous molecules are
25 usually identified by selecting those which show product scores between 15 and 40, although
lower scores may identify related molecules.

The results of Northern analysis are reported as a list of libraries in which the transcript
encoding HDAC9 occurs. Abundance and percent abundance are also reported. Abundance
directly reflects the number of times a particular transcript is represented in a cDNA library, and

percent abundance is abundance divided by the total number of sequences examined in the cDNA library.

In this case, electronic Northern analysis of LIFESEQ™ database (Incyte Pharmaceuticals, Inc. Palo Alto, Calif) indicates tissue distribution of the HDAC9 sequence as seen in Table 5. These results are reported as a list of cDNA libraries in which the transcript encoding HDAC9 occurs. The presence of HDAC9 in 20 libraries from different tissue-specific and mixed tissue sources indicates that HDAC9, like other HDAC family members may be found as an expressed gene in a wide range of tissues. This result is supported by the Northern hybridization of an HDAC9 probe to mRNAs from 12 normal tissues (see Figure 7).

Table 5. Tissue distribution determined electronically from LIFESEQ™ database.

Tissue Category
Cardiovascular System
Connective Tissue
Digestive System
Embryonic Structures
Endocrine System
Exocrine Glands
Genitalia, Female
Genitalia, Male
Germ Cells
Hemic and Immune System
Liver
Musculoskeletal System
Nervous System
Pancreas
Respiratory System
Sense Organs
Skin
Stomatognathic System
Unclassified/Mixed
Urinary Tract

Example 6: Real time PCR survey of HDAC9 distribution in human normal tissues and cell lines.

Real Time PCR. Total RNA from cultured cell lines was isolated with the Rneasy 96 kit according to the manufacturers protocol (Qiagen, Valencia CA). RNA from human tissues was purchased (Clontech Inc, Palo Alto, CA) and the tissue sources are listed in table 6 below.

Table 6. Tissue sources of RNA for real time PCR analysis

Tissue	Sex of donor	Age range of donor (yrs.)	Number of samples pooled
Brain 1	M	57	1
Brain 2	F	16&36	2
Cerebellum	M	64	1
Spinal cord	M/F	17-72	31
Fetal brain	M/F	20-23 wks	8
Trachea	M/F	17-70	84
Liver 1	M	27	1
Liver 2	M/F	15&35	2
Fetal liver	?	15-24 wks	?
Stomach	M/F	23-61	15
Pancreas	M/F	17-69	18
Colon	M	35&50	2
Intestine	M/F	25&30	2
Kidney	M/F	24-55	8
Bone marrow	M/F	18-68	24
Spleen	M	22-60	7
Thymus	M	6-45	9
Thyroid	M/F	10-46	4
Adrenal gland	M	32-50	6
Salivary gland	M/F	13-78	43
Mammary gland	F	23-47	8
Skeletal muscle	M/F	23-56	10
Testis	M	28-64	25
Prostate 1	M	26-64	23
Prostate 2	M	14-60	10
Placenta	F	22-41	15

Numbers following tissues represent separate samples from the same tissue type: Male (M), Female (F)

- 5 Human cell lines, H1299 human lung carcinoma, T24 bladder carcinoma, SJRH30 muscle rhabdomyosarcoma, SJSA-1 osteosarcoma, human fibroblasts, and A549 human lung carcinoma, were obtained from American Type Tissue Culture Collection. Total RNA was isolated from

human cell lines using RNA easy kit according to the manufacturers instructions (Qiagen, Valencia, CA). RNAs were quantified using RT-PCR on an ABI Prism Sequence Detection System. The primers used for detection of HDAC9 were forward primer 5'-GGATCCAGTATCTCTT TGAGGATGAC-3', reverse primer 5'-AGAAGCGCCCATGCTCATA-3', and Taqman probe 5'-AGCGTCCTTACTTCTCCTGGCACCG-3'. The Taqman Reaction System (Eurogentec, Belgium) was used with 10 ng total RNA in a 25 µl reaction in the proportions indicated by the manufacturer but supplemented with 0.25 U/µl reverse transcriptase (MultiScribe ABI, Perkin Elmer, Branchburg NJ) and 0.08 U/µl RNaseOUT RNase inhibitor (Life Technologies, Gaithersburg, MD). The reverse reaction was initiated with a 5 min incubation at 48 °C for the reverse transcription of the mRNA followed by a 10 min incubation at 95 °C to inactivate the reverse transcriptase and simultaneously activate the 'hot-start' thermostable DNA polymerase. This was followed by 50 cycles of a two-step PCR reaction with alternating 15 sec at 95 °C and 60 sec at 60 °C. Computations were performed using ABI sequence detection software (version 1.6.3). The RT-PCR assays were standardized with cRNAs transcribed in vitro with the T7 RNA polymerase reaction using the Maxiscript kit (AMBION Inc., Austin, TX) according to the manufacturers protocol. The RT-PCR assays were standardized with a dilution series of total RNA isolated from A549 lung tumor cells. Parallel to the RT-PCR, the total amount of RNA in each reaction was quantitated in a fluorometric assay using the RiboGreen kit (Molecular Probes Inc., address) according to the manufacturers instructions, using mammalian ribosomal RNA provided with the kit as standard.

Real time PCR was also used to survey the distribution and levels of HDAC9 in tissues and tumor cell lines, relative to the levels of 18S ribosomal RNA . RNA from the human A549

lung carcinoma cell line was arbitrarily chosen as an internal control for the levels of total RNA in the samples. The levels of HDAC9 and 18S rRNA in A549 cells were set at 100 % and the levels of HDAC9 and 18S rRNA in other tissues and cell lines were measured as a percent of the level of these genes in A549 RNA. The levels of 18S ribosomal RNA ranged between 82% and 126% of the A549 internal control in all of the RNA samples, suggesting that there were similar amounts of RNA in the analyzed tissue samples. HDAC9 was detected at varying levels by real time PCR in a wide range of tissues (Fig. 8), confirming the Northern blot analysis (Fig. 7). In normal tissues, HDAC9 was detected at the highest levels in fetal brain (894%), cerebellum (538%), and thymus (589%). In tumor cell lines, HDAC9 was detected at the highest levels in SJRH30 cells (850%) (Fig. 8). These results suggest that HDAC9 is differentially expressed in some tissues at the RNA level.

Example 7: HDAC Enzyme Assay

Preparation of HDAC9-flag. A flag epitope tag sequence was added to the 3' end of HDAC9v1 by PCR. The PCR primers were 5'-ACGCCGGATATCACATTGGT TCTGC-3' and 5'-GCGGAATTCTTATTATTATCATCATCATCITTATAATCCCC GTGCACAGCCACCAGGTGAGGATGGCA -3'. The flag-tagged HDAC9v1 was reconstructed using the EcoRV site in the 1st primer and subcloned into the XbaI and EcoRI sites of human expression vector pCDNA3.1(-) (Invitrogen, Carlsbad, CA).

HDAC activity assay. HDAC activity assays are performed as previously described (Emiliani, S., Fischle, W., Van Lint, C., Al-Abed, Y., and Verdin, E. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2795-2800). 5×10^5 293 cells grown to 50% confluency in 100 mm dishes are transfected with 30 ug of C-terminally flag-tagged HDAC1, HDAC3, HDAC4, HDAC6, HDAC7, or HDAC9 using Geneporter transfection kit according to the manufacturers

WO 02/050285

63

PCT/EP01/14928

instructions. The cell culture medium is changed 5 h after transfection. 48 h after transfection cells are washed in cold PBS and scraped into 1 ml of IP buffer (50mM Tris-HCl pH 7.5, 120mM NaCl, 0.5mM EDTA, 0.5% NP-40) and incubated on a rocker for 20 min. Cellular debris is pelleted in a centrifuge at 14K for 20 min. The supernatant is precleared for 1 h with

5 protein G beads (Pharmacia Biotech) in IP buffer. Immunoprecipitations are performed by incubating the precleared supernatant with either α -FLAG M2 agarose affinity gel (Sigma) for 2 h at 4°C or anti-HDAC2 (Santa Cruz) for 1 h followed by incubation with protein G beads for 1 h at 4°C. The beads are then washed three times for 5 min in IP buffer and then washed three

10 times in high salt IP buffer (50mM Tris-HCl pH 7.5, 1000 mM NaCl, 0.5mM EDTA, 0.5% NP-40) at 4°C. IPs are then washed two times for 2 min in 1ml of HD-buffer (10mM Tris-HCl pH 8.0, 10mM NaCl, 10% glycerol). When trapoxin inhibition is determined Ips are incubated with 0.3, 3, 30 and 300 nM TPX in HD-buffer for 20 min. Supernatants are incubated with 100000 cpm substrate ($[^3\text{H}]\text{-Ac(H41-24) SGRGKGGKGLGKGAKRHRKVLRD}$, in vitro/chemically

15 acetylated using BOP-chemistry) in 30 ul HD-buffer or TPX in HD-buffer, resuspending the sepharose by gently tapping the tube and shaking in an Eppendorf 5436 Thermomixer at full speed at 37°C for 2h. 170 ul HD-buffer and 50ul stop-mix (1M HCl, 0.16M HAc) are added, vortexed for 15' min, 600ul ethylacetate is then added and vortexed for 45 minutes, then centrifuged at 14000g for 7 minutes. 540 ul of the organic (upper) phase is then counted in 5 ml scintillation liquid using conventional techniques.

20 **HDAC9 is catalytically active.** *In vitro* histone deacetylase assays using immunoprecipitated HDAC9 and an ^3H -acetylated histone H4 peptide as substrate were performed to determine whether HDAC9 was catalytically active and to compare the activity of HDAC9 to known catalytically active HDAC1, HDAC3, and HDAC4. An HDAC-related protein

that lacks catalytic activity, HDRP/MITR/HDACC was used as a negative control (Zhou, X., Richon, V.M., Rifkind, R.A., Marks, P.A. (2000) Identification of a transcriptional repressor related to the noncatalytic domain of histone deacetylases 4 and 5. *Proc Natl Acad Sci U S A* 97, 1056-61). These results demonstrated that HDAC9 could deacetylate the histone peptide
5 substrate at a level that was equivalent to HDAC3 and HDAC4 (Fig. 12A), while HDAC1 was more effective in this assay (Fig. 12B).

Example 8 HDAC9 expression and cellular localization

HDAC9 is expressed *in vitro* using 1 ug of the M6 clone, 2 ul of ³⁵S-Methionine and Sp6
TNT Quick Coupled Transcription/Translation System according to manufacturer instructions.
10 (Promega, Madison, WI). Proteins are electrophoresed on a SDS-PAGE gel according to conventional methods and visualized by a Storm phosphorimager. The complete HDAC9 sequence molecular weight is estimated in silico as 72 kda using VectorNTI Suite software (Informax, North Bethesda, MD). A doublet was observed on a 10% SDS-PAGE gel. Doublets have also been observed when HDAC1 is translated *in vitro*. These doublets suggest that there is
15 potentially a second translation initiation site. Furthermore, these results suggest that HDAC9 is an expressed gene. See Figure 13.

1X10⁵ Cos7 cells are plated onto chamber slides. Cells are transfected on the slides with 2 ug of flag epitope-tagged HDAC9 or a cytoplasmically expressed protein (Ena-flag) using
Genepor2 in serum free medium according to the manufacturers instructions. The cell culture
20 medium is changed 24 h after transfection. 48 h after transfection, cells are washed three times with PBS, fixed for 15 min. in 5% formaldehyde, washed two times in PBS, and blocked for 30 minutes at room temperature in 10% fetal calf serum (Sigma) in PBS with 0.5% Triton-X-100 to permeabilize the cells. The cells are washed again two times in PBS and then incubated with 25

mg/ml anti-Flag-FITC conjugate for 1 hour. The stained cells are washed with PBS and photographed using fluorescence microscopy.

HDAC9 is a nuclear protein. The translated HDAC9 peptide sequence predicts a 72 Kda protein and this was confirmed by *in vitro* translation (Fig. 13A). In order to determine the cellular localization of HDAC9, flag epitope-tagged HDAC9, Enabled (Ena) or pCMV4flag were transfected into Cos7 and 293 cells or cells were mock transfected without plasmid. The flag epitope was detected by fluorescence immunocytochemistry 48 h after transfection (Fig 13B). Ena is a cytoskeleton-associated cytoplasmic protein substrate of Abl tyrosine kinase that transduces the axon-repulsive function of the Roundabout receptor during axon guidance (Gertler FB, Comer AR, Juang JL, Ahern SM, Clark MJ, Liebl EC, Hoffmann FM. (1995) enabled, a dosage-sensitive suppressor of mutations in the Drosophila Abl tyrosine kinase, encodes an Abl substrate with SH3 domain-binding properties. Genes Dev. 9, 521-533. Bashaw GJ, Kidd T, Murray D, Pawson T, Goodman CS. (2000) Repulsive axon guidance: Abelson and Enabled play opposing roles downstream of the roundabout receptor. Cell. 101, 703-715). As expected, Ena was detected in the cytoplasm, whereas HDAC9 was detected in the nuclei of these cells. The detection of HDAC9 in the nuclei of both Cos7 and 293 cells suggested that HDAC9 was predominantly a nuclear protein.

Example 9: Identification of associated proteins in HDAC complexes

Transfection. 1×10^7 Cos7 cells are transfected with 10 ug of either C-terminally flag epitope-tagged HDAC1, HDAC2, HDAC3, HDAC4, HDAC6, HDAC7, or HDAC9 in pCDNA3.1 expression vector or Flag vector or buffer (Mock) as transfection controls. by electroporation using a Gene Pulser II instrument (Biorad, Hercules CA) set at 0.3Kv/ 500 uF.

WO 02/050285

66

PCT/EP01/14928

- Immunoprecipitation.** Immunoprecipitations are performed as described (Grozinger, C. M., Hassig, C. A., and Schreiber, S. L. 1999. Proc. Natl. Acad. Sci. USA 96, 4868-4873). Whole cell extracts are prepared 48h after transfection by scraping cells into JLB buffer (50 mM Tris-HCl, pH 8, 150 mM NaCl, 10% glycerol, 0.5% Triton-X-100) containing complete protease inhibitor cocktail (Boehringer-Mannheim). Lysis is continued at 4°C for 10 min. and then cellular debris is pelleted by centrifugation at 14K for 5 minutes. Supernatants are pre-cleared with Sepharose A/G-plus agarose beads (Santa Cruz). Recombinant proteins are immunoprecipitated from pre-cleared supernatant by incubation with α -FLAG M2 agarose affinity gel (Sigma) for 2 h at 4°C or anti-HDAC1 (Santa Cruz, Santa Cruz, CA) for 1 h at 4°C, followed by incubation with Sepharose A/G beads. For Western blot analysis, the beads are washed with MSWB buffer (50 mM Tris-HCl, pH 8, 150 mM NaCl, 1 mM EDTA, 0.1% NP-40) and the proteins are separated by SDS/PAGE. Western blots are probed with anti-flag M2 (Sigma), HDAC1 (Santa Cruz), anti-HDAC2 (Santa Cruz), anti-HDAC6 (Santa Cruz), anti-Rb (Pharmingen), or anti-mSin3A (Transduction Labs, Lexington, KY)
- HDAC9 associates with proteins in the mSin3A complex.** Class I HDACs, but not class II HDACs were previously found to be associated with the mSin3A complexes. The core HDAC1 complex consists of HDAC1, HDAC2, RbAp46, RbAp48. This core complex has been found to associate with an mSin3A complex that is involved in transcriptional repression through an Rb and E2F complex (Luo RX, Postigo AA, Dean DC.(1998) Rb interacts with histone deacetylase to repress transcription. Cell. 92, 463-473; Magnaghi-Jaulin L, Groisman R, Naguibneva I, Robin P, Lorain S, Le Villain JP, Troalen F, Trouche D, Harel-Bellan A. (1998) Retinoblastoma protein represses transcription by recruiting a histone deacetylase. Nature. 391, 601-605; Brehm A, Miska EA, McCance DJ, Reid JL, Bannister AJ, Kouzarides T. (1998)

Retinoblastoma protein recruits histone deacetylase to repress transcription. Nature. 391, 597-601). In order to determine whether HDAC9 was a part of this complex, endogenous HDAC1, HDAC2, Rb, and mSin3 proteins were co-immunoprecipitated from cells transfected with flag-epitope tagged HDAC1, HDAC3, HDAC4, HDAC6, HDAC7 or HDAC9. To assure that
5 transfected flag epitope-tagged HDACs could be detected in cells, the levels of HDAC expression were detected by immunoprecipitation and Western blotting with antiserum to the flag epitope. To determine which HDACs associated with components of the Sin3 complex, endogenous proteins in the Sin3 complex were immunoprecipitated and the associated HDACs were detected by Western blotting flag epitope-specific antibody HDAC9 was found to associate
10 with HDAC1, HDAC2, Rb, and mSin3A, suggesting that HDAC9 is a component of an mSin3A complex.

HDAC9 associates with SMRT and NCoR. Since corepressors SMRT and NCoR associate with the mSin3 core complex, experiments were performed to co-immunoprecipitate HDACs with NCoR and SMRT (Fig. 15). HDAC9 co-immunoprecipitated with both of these
15 proteins, suggesting that HDAC9 associates with SMRT, and NCoR. Western analysis of the flag-detected blots with anti-NCoR indicated that NCoR was immunoprecipitated. As previously reported, SMRT co-immunoprecipitated with HDAC4 and HDAC6, and HDAC6 and HDAC7 did not associate with the Sin3A complex.

HDAC9 associates with 14-3-3 and Erk proteins. HDAC4 was previously found to
20 associate with 14-3-3- β , 14-3-3- ϵ , CamK, Erk1, and Erk 2 proteins, which sequester HDAC4 in the cytoplasm and prevent phosphorylated HDAC4 and HDAC5 from entering the nucleus and repressing MEF2 activated transcription. In order to determine whether HDAC9 associate with these proteins, experiments were performed to co-immunoprecipitate HDACs with 14-3-3 and

Erk proteins. All of the HDACs tested associated with 14-3-3s and Erks. These results suggest that the association of HDACs with 14-3-3 and Erks might be a general mechanism of sequestering HDACs in the cytoplasm.

Classification of HDAC9. HDACs have been classified by sequence similarity to yeast HDACs, sequence length, location of catalytic domain, cellular localization, associating proteins, and sensitivity to HDAC inhibitors. The data in this study suggests that HDAC9 has characteristics of both class I and class II HDACs. HDAC9 had sequence similarity with class II yeast hda1 subfamily member Clr3 and HDAC6 catalytic domain 1. In addition, the 3 Kb HDAC9 transcript was only detected in kidney and testis, suggesting that it might have a limited tissue distribution like class II HDACs. HDAC9 was between class I and class II HDACs in length. Class I HDACs average 443 bp in length, whereas class II HDACs average 1069 bp in length. However, HDAC9 was found to have an N-terminal catalytic domain, as opposed to the C-terminal domains that have been found in class II HDACs. HDAC6 is an exception that has both N-terminal and C-terminal catalytic domains. Furthermore, class I HDACs are nuclear proteins, while class II HDACs are nucleocytoplasmic. Immunocytochemistry indicated that HDAC9 was predominantly nuclear and was detected in a different subcellular compartment in comparison to the Ena protein that is expressed in the cytoplasm. In contrast to the 3 Kb HDAC9 transcript that might be differentially expressed, a 3.5 Kb HDAC9 transcript that was identified by Northern analysis was expressed ubiquitously in normal tissues, tumor tissues and cell lines, similar to class I HDACs. In addition, HDAC9 was found to co-immunoprecipitate with proteins that were previously only associated with class I HDAC complexes, including HDAC1, HDAC2, mSin3A, and Rb. HDAC9 also has putative C-terminal LXCXE motifs that so far have only been

found in HDAC1. HDAC9 was also found to associate with NCoR and SMRT. This evidence suggests HDAC9 had characteristics that bridged those of class I and class II HDACs.

WO 02/050285

PCT/EP01/14928

What is claimed is:

- 5 1. An isolated polypeptide comprising the amino acid sequence set forth in
SEQ ID NO:1, SEQ ID NO 5 or SEQ ID NO 6.
2. An isolated polypeptide consisting of the amino acid sequence set forth in
SEQ ID NO:1, SEQ ID NO 5 or SEQ ID NO 6.
- 10 3. An isolated DNA comprising a nucleic acid sequence that encodes the
polypeptide of claim 1 or 2.
4. A vector molecule comprising at least a fragment of the isolated DNA
according to claim 3.
- 15 5. The vector molecule according to claim 4 comprising transcriptional control
sequences.
6. A host cell comprising the vector molecule according to claim 5.
- 20 7. The isolated DNA according to claim 3, comprising a nucleotide sequence
selected from the group consisting of (1) the nucleotide sequence set forth in SEQ ID NO:2, 7 or
8, being the complete cDNA sequence encoding the polypeptide as defined in claim 2; (2) the
nucleotide sequence set forth in SEQ ID NO:3, being the open reading frame of the cDNA
25 sequence encoding the polypeptide as defined in claim 2; (3) a nucleotide sequence capable of
hybridizing under high stringency conditions to a nucleotide sequence set forth in SEQ ID NO:3;
and (4) the nucleotide sequence set forth in SEQ ID NO:4, being the endogenous genomic
human DNA encoding the polypeptide as defined in claim 2.

WO 02/050285

PCT/EP01/14928

8. A vector molecule comprising at least a fragment of an isolated DNA molecule according to claim 7.

9. The vector molecule according to claim 8 comprising transcriptional control sequences.

10. A host cell comprising the vector molecule according to claim 9.

11. A host cell which can be propagated in vitro and which is capable upon growth in culture of producing a polypeptide according to claim 1 or 2, wherein said cell comprises at least one transcriptional control sequence that is not a transcriptional control sequence of the natural endogeneous human gene encoding the polypeptide of claim 2, wherein said one or more transcriptional control sequences control transcription of a DNA encoding a polypeptide according to claim 1 or 2.

12. A method for the diagnosis of a condition associated with abnormal regulation of gene expression which includes, abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, host inflammatory or immune response, or psoriasis in a human which comprises: detecting abnormal transcription of messenger RNA transcribed from the natural endogeneous human gene encoding the polypeptide as defined in claim 2 in an appropriate tissue or cell from a human, wherein said abnormal transcription is diagnostic of said condition.

13. The method of claim 12, wherein said natural endogeneous human gene comprises the nucleotide sequence set forth in SEQ ID NO:4, 7 or 8.

14. The method of claim 12, comprising contacting a sample of said appropriate tissue or cell or contacting an isolated RNA or DNA molecule derived from said tissue or cell with an isolated nucleotide sequence of at least about 15-20 nucleotides in length

WO 02/050285

PCT/EP01/14928

that hybridizes under high stringency conditions with the isolated nucleotide sequence as defined in claim 3.

15. A method for the diagnosis of a condition associated with abnormal
5 HDAC9 expression or activity in a human which comprises:
measuring the amount of a polypeptide comprising the amino acid
sequence set forth in SEQ ID NO:1, 5 or 6 or fragments thereof, in an appropriate tissue or cell
from a human suffering from said condition wherein the presence of an abnormal amount of said
polypeptide or fragments thereof, relative to the amount of said polypeptide or fragments thereof
10 in the respective tissue from a human not suffering from said condition associated with abnormal
HDAC9 expression or activity is diagnostic of said human's suffering from a condition

16. The method of claim 15, wherein said detecting step comprises contacting
said appropriate tissue or cell with an antibody which specifically binds to a polypeptide that
15 comprises the amino acid sequence set forth in SEQ ID NO:1, 5 or 6 or a fragment thereof and
detecting specific binding of said antibody with a polypeptide in said appropriate tissue or cell,
wherein detection of specific binding to a polypeptide indicates the presence of a polypeptide
that comprises the amino acid sequence set forth in SEQ ID NO:1, 5 or 6 or a fragment thereof.

- 20 17. An antibody or a fragment thereof which specifically binds to a
polypeptide that comprises the amino acid sequence set forth in SEQ ID NO:1, 5 or 6 or to a
fragment of said polypeptides.

- 25 18. An antibody fragment according to claim 17 which is an Fab or F(ab')₂
fragment.

19. An antibody according to claim 17 which is a polyclonal antibody.

- 30 20. An antibody according to claim 17 which is a monoclonal antibody.

WO 02/050285

PCT/EP01/14928

21. A method for producing a polypeptide as defined in claim 1 or 2, which method comprises:

5 culturing a host cell having incorporated therein an expression vector comprising an exogenously-derived polynucleotide encoding a polypeptide comprising an amino acid sequence as set forth in SEQ ID NO:1, 5 or 6 under conditions sufficient for expression of the polypeptide in the host cell, thereby causing the production of the expressed polypeptide.

22. The method according to claim 21, said method further comprising recovering the polypeptide produced by said cell.

10

23. The method according to claim 21, wherein said exogenously-derived polynucleotide encodes a polypeptide consisting of an amino acid sequence set forth in SEQ ID NO:1, 5 or 6.

15

24. The method according to claim 21, wherein said exogenously-derived polynucleotide comprises the nucleotide sequence as set forth in SEQ ID NO:2, 7 or 8.

25. The method according to claim 21, wherein said exogenously-derived polynucleotide comprises the nucleotide sequence as set forth in SEQ ID NO:3.

20

26. The method according to claim 21, wherein said exogenously-derived polynucleotide consists of the nucleotide sequence as set forth in SEQ ID NO:3.

27. The method according to claim 24, wherein said exogenously-derived polynucleotide comprises the nucleotide sequence as set forth in SEQ ID NO:4.

25

WO 02/050285

PCT/EP01/14928

Fig. 1

```
1  GGCGCCGAGG  CTTCTGCGTC  CGTCGTGGTT  CTTGGCTCCG
41  GGGCGGAGTT  CGCGATAGCG  ATCGGGGAGC  AGGACGCGGG
81  GCGTGGACCC  AGGTCCGAGG  CGAGGAAGCC  GTAACCCATG
121  CGCGGGGAGC  CTCCCCCTTC  GACTGCAGCC  TCGCTCCGTG
161  CCTTCTGCGC  GCCTGGGATC  CCGGAGCCTG  CCTAGGTTCT
201  GTGCGCTCCC  GCCCAGGCCG  GTGCCCCCGC  CCCGCCTGGG
241  CCCCAGGCAG  GTCCAGGCC  TCCGGCTGCT  CCCGGCCGAA
281  GCCCCGAGTG  CGAGATCGAG  CGTCCTGAGC  GCCTGACCCG
321  AGCCCTGGAT  CGCCTGCGGC  AGCGCGGCCT  GGAACAGAGG
351  TGTCTGCGGT  TGTCAAGCCG  CGAGGCCTCG  GAAGAGGAGC
391  TGGGCCTGGT  GCACAGAGTA  CCTTTCAGTG  CGCGCGGCTG
431  GCCGCAGGGG  CTGGACTGCA  GCTGGTGGAC  GCTGTGCTCA
471  CTGGAGCTGT  GCAAAATGGG  CTTGCCCTGG  TGAGGCCTCC
511  CGGGCACCAT  GGCCAGAGGG  CGGCTGCCAA  CGGGTTCTGT
551  GTGTTCAACA  ACGTGGCCAT  AGCAGCTGCA  CATGCCAAGC
601  AGAAACACGG  GCTACACAGG  ATCCTCGTCG  TGGACTGGGA
641  TGTGCACCAT  GGCCAGGGGA  TCCAGTATCT  CTTTGAGGAT
681  GACCCAGCG  TCCTTTACTT  CTCTGGCAC  CGCTATGAGC
721  ATGGGCGCTT  CTGGCCTTTC  CTGCGAGAGT  CAGATGCAGA
761  CGCAGTGGGG  CGGGGACAGG  GCCTCGGCTT  CACTGTCAAC
801  CTGCCCTGGA  ACCAGGTIGG  GATGGGAAAC  GCTGACTACG
841  TGGCTGCCCT  CCTGCACCTG  CTGCTCCAC  TGGCCTTIGA
881  GTTTGACCTT  GAGCTGGTGC  TGGTCTCGGC  AGGATTTGAC
921  TCAGCCATCG  GGGACCCTGA  GGGGCAAAAG  CAGGCCACGC
961  CAGAGTGCTT  CGCCACCTC  ACACAGCTGC  TGCAGGTGCT
1001  GGC CGCGCGC  CGGGTCTGTG  CCGTGCTGGA  GGGCGGCTAC
1041  CACCTGGAGT  CACTGGCGGA  GTCAGTGTGC  ATGACAGTAC
1081  AGACGCTGCT  GGGTGACCCG  GCCCCACCCC  TGTAGGGCC
1121  AATGGCGCC
```

WO 02/050285

PCT/EP01/14928

Fig. 2

A.

```

1 ATGGGGACCG CGCTTGTGTA CCATGAGGAC ATGACGGCCA CCGGGCTGCT
51 CTGSGACGAC CCCGAGTCGG AGATCGAGCG TCCTGAGCGC CTGACCGCAG
101 CCCTGGATCG CCTGCGGCAG CCGCGCCTGG AACAGAGGTC TCTGCGSTTG
151 TCAGCGCCGCG AGGCTTCGGA AGAGGAGCTG GCGCTGGTGC ACAGCCGAGA
201 GTATGATATCC CTGGTCAGGG AGACCGAGGT CCTAGGCAAG GAGGAGCTGC
251 AGGCGCTGTC CGACAGTTC GACGCCATCT ACTTCCACCC GAGTACCTTT
301 CACTGCGCGC GCTGCGCGCG AGGGGCTGGA CTGACGCTGG TGGACGCTGT
351 CCTCACTGGA GCTGTGCAAA ATGGGCTTGC CTGCTGAGGG CCTCCCGGGC
401 ACCATGGCCA GAGGGCGGCT GCCAACGGGT TCTGTGTGTT CACACAACGTG
451 GCCATAGCAG CTGCACATGC CAAGCAGAAA CACGGGTAC ACAGGATCCT
501 CGTCGTGGAC TGGGATGTGC ACCATGGCCA GGGGATCCAG TATCTCTTTG
1001 AGGATGACCC CAGCGTCCTT TACTTCTCTT GGCACCGCTA TGAGCATGGG
1051 CGCTTCTGGC GTTTCCTGCG AGAGTCAGAT GCAGACGCGT TGGGCGGGG
1101 ACAGGGCCTC GGCTTCACTG TCAACCTGCC CTGGAACCGG GTTGGGATGG
1151 GAAACGCTGA CTACGTGGCT GCCTTCTGCG ACCTGTGCTCT CCCACTGGCC
1201 TTTGAGTTTG ACCCTGAGCT GGTGCTGGTC TCGGCGAGAT TTGACTCAGC
1251 CATCGGGGAC CCTGAGGGGC AAATCGAGGC CACGCCAGAG TGCTTCGCCC
1301 ACCTCACACA GCTGCTGCAG GTGCTGGCG GCGGCCGGT CTGTCCGGTG
1351 CTGGAGGGCG GCTACCACCT GAGTCACTG GCGGAGTCA TGTGCATGAC
1401 AGTACAGACG CTGCTGGGTG ACCCGGCCCC ACCCCTGTCA GGGCCAATGG
1451 CGCCATGTCA GAGGTGCGAG GGGAGTGCCT TAGAGTCCAT CCAGAGTGCC
1501 CGTGTGCCCC AGGCCCCGCA CTGGAAGAGC CTCAGCAGC AAGATGTGAC
1551 CGCTGTGCCC ATGAGCCCCA GCAGCCATC CCCAGAGGG AGGCTCCAC
1601 CTCTGCTGCC TGGGGGTCCA GTGTGTAAAG CAGCTGCATC TGCACCGAGC
1651 TCCTCTCTGG ACCAGCGGTG CTTCTGCCCC GCACCTCTCT CCGCACCGC
1701 TGTTCGCCCTG ACAACGCGGG ATATCACATT GGTTCGTCCC CTGACCTCA
1751 TCCACAGGGA AGGCTCAGCC CTGAGGGAGG AGACAGAGGC CTGGGCCAGG
1801 CCACACAGAT CCCTGCCCCG GAGGAGGGCC CTCACTGCAC TTGGGAGGCT
1851 CCTGTACCTC TTAGATGGGA TGCTGGATGG GCAGGTGAAC AGTGGTATAG
1901 CAGCCACTCC AGCCTCTGCT GCAGCAGCCA CCCTGGATGT GGCTGTTCCG
2001 AGAGGCTCTGT CCCACGGAGC CCAGAGGCTG CTGTGCGTGC CCCTGGGACA
2051 GCTGGACCGG CCTCCAGACC TCGCCCATGA CCGGAGGAGT CTGTGGCTGA
2101 ACATCAGGGG CAGGAGGGCG GCTGCCCTAT CCATGTTCOA TGTCCTCACG
2151 CCACTGCCAG TGATGACCGG TGGTTTCCTG AGCTGCAICT TGGGCTTGGT
2201 GCTGCCCTGT GCCTATGGCT TCCAGCCTGA CCTGGTGTG GTGGGCTGTG
2251 GGCCTGGCCA TGGCCTGCAG GGCCTCCACG CTGCATCTCT GGCTGCAATG
2301 CTTGCGGGCG TGGCAGGGGG CCGAGTCCTG GCCCTCTTGG AGGAGGACTC
2351 CACACCCAG CTAGCAGGGA TCCTGGCCCG GGTGCTGAAT GAGAGGCAC
2401 CTCTAGCCT AGGCCCTTCC TCTGTGGCTT CCCAGAGGA CPTCCAGCC
2451 CTGATGTACC TGAGAGGGCA GCTGAGGCTT CAGTGAAGA TGTGCAAGTG
2501 CCATCTCTAC CTGGTGGCTT GA

```

B.

```

MGTAIVYHED MTATRLWDD PECEIERPER LTAALDLRLQ RGLERCLRLI SAREASEEEL
GLVHSPRYVS LVRETQVLGK EELQALSGQF DAIFYFHPSTF HCARIAAGAG LQIVDAVLIG
AVQNGALAVR PFCHHGQRAA AMFCVFNWV AIAAAHAKQK HGLHRIYVVD KDVHGGQIGI
YLFDDPSVL YFSWHRYEHG RFWFLEBSD ADAVGRGQGL GFTVNLPMNO VGMGNADYVA
AFILHLLPLA FEFDPELVLV SAGFDSAGD PRGQMATPE CFAHLTOLLQ VLAGSRVCAV
LEGGYHLESF AFSVCWTVQT LLKDPAPPLS GPMAPOQRC GSALESIQSA RAAQAPHWKS
LQQQDVTAVP MSPSSHSPG RPPPLLPQGP VCKAASAPF SLLDQPCICP APSVRTAVAL
TTPDITLVLP PDVIQREASA LREETEAHAR PHESLAREEA LTALGKLLYL LDGMIDQVNV
SGLAATPASA AAATLDVAVR RGLSHGAQRL LCVALGQLDR PPDLAHDGRS LWNIRGKEA
AALSHPHVST PLFVMTGGEL SCILGLVLPL AYGFQFDLVL VALGPHGLQ GPHAAALLAM
LRGLAGRVYL ALLEENSTPQ LAGILARVLN GRAPPSLQPS SVASPEDVQA IMYLRGQLSP
QWKMLQCHPH LVA

```

PCT/EP01/14928

AL022328	vs	HDAC9:	
AL022328		2	tcaggccaccagggtgaggatggcaactaca.....ctacactgcacacatc
			<<<<< 144 <<<<<
HDAC9		1	tcaggccaccagggtgaggatggca.....ctgcaacatct
AL022328		181	tccactgaggctccagctgcctctcaggtaacatcagggctcgagctcc
HDAC9		36	tccactgaggctccagctgcctctcaggtaacatcagggctcgagctcc
AL022328		231	tctggggaggccacagaggaagggcctatggtcaggagtgctctccatt
HDAC9		86	tctggggaggccacagaggaagggcctatggtcaggagtgctctccatt
AL022328		281	cagcaccocgggcaggatccctgctagctggggctgaggagtctcgaa....
HDAC9		136	cagcaccocgggcaggatccctgctagctggggctgaggagt.....<<<<< 82
AL022328		322	.cttacctctccaggaggccaggactcggccctctgcagcccccga
HDAC9		177	<<<<<.....ctctccaggaggccaggactcggccctctgcagcccccga
AL022328		449	gcattgcagccaggagtcagcgtgggggcctgcaggccatggccaggc
HDAC9		222	gcattgcagccaggagtcagcgtgggggcctgcaggccatggccaggc
AL022328		499	cccagcgcaccagcaccagggtcaggctggaagccataggccaggggcag
HDAC9		272	cccagcgcaccagcaccagggtcaggctggaagccataggccaggggcag
AL022328		549	caccagcccaagatgcagctcaggaaacaccggctcatctgtg.....c
HDAC9		322	caccagcccaagatgcagctcaggaaacaccggctcat.....<<<<< 204 <
AL022328		587	tcaccactggcagtgccgtggagacatggaacatggatagggcacggcc
HDAC9		360	<<<<<.....cactggcagtgccgtggagacatggaacatggatagggcacggcc
AL022328		838	tccttgcctctgatgttcagccacagactcctccttc.....cctacc
HDAC9		407	tccttgcctctgatgttcagccacagactcctc.....<<<<< 302 <<<<<cc
AL022328		1175	gtcatggggcaggtctggaggccggtccagctgtccaggggccacgcaca
HDAC9		442	gtcatggggcaggtctggaggccggtccagctgtccaggggccacgcaca
AL022328		1225	gcagccttga.....cttacctctgggctcgtgggacaggcctctcgga
HDAC9		492	gcagc.....<<<<< 139 <<<<<.....ctctgggctcgtgggacaggcctctcgga
AL022328		1399	acagccacatccagggtggctgctgcagcagaggctggagtggctgctat
HDAC9		527	acagccacatccagggtggctgctgcagcagaggctggagtggctgctat

WO 02/050285

PCT/EP01/14928

AL022328 1449 accactgttcacotgtg....cccacctgcccacccagcatcccatcta
 |||||||||||<<<< 725 <<<<|||||||||||||||||
 HDAC9 577 accactgttcac.....ctgcccacccagcatcccatcta

AL022328 2209 agaggtacaggagcttcccaagtgcagtgagggcctcctcccgggccagg
 |||||||||||<<<< 113 <<<<|||||||||||||||||
 HDAC9 612 agaggtacaggagcttcccaagtgcagtgagggcctcctcccgggccagg

AL022328 2259 gactcgtgtggcctgtg....cccacctggcccaggcttctgtctctc
 |||||||||||<<<< 113 <<<<|||||||||||||||||
 HDAC9 662 gactcgtgtggc.....ctggcccaggcttctgtctctc

AL022328 2407 cctcagggtgcagcttctcgttgatgagtcagggggcagaaccaatg
 |||||||||||<<<< 113 <<<<|||||||||||||||||
 HDAC9 697 cctcagggtgcagcttctcgttgatgagtcagggggcagaaccaatg

AL022328 2457 tgatatccggcgttgatcagggaacagcgggtgcggacagaggggcgggg
 |||||||||||<<<< 113 <<<<|||||||||||||||||
 HDAC9 747 tgatatccggcgttgatcagggaacagcgggtgcggacagaggggcgggg

AL022328 2507 cagaggcacggcgtggtccaggaggagctcgggtgcagatgcagctgcctt
 |||||||||||<<<< 113 <<<<|||||||||||||||||
 HDAC9 797 cagaggcacggcgtggtccaggaggagctcgggtgcagatgcagctgcctt

AL022328 2557 acacactggaccccaggcagcagaggtggaggcctccctctggggagt
 |||||||||||<<<< 113 <<<<|||||||||||||||||
 HDAC9 847 acacactggaccccaggcagcagaggtggaggcctccctctggggagt

AL022328 2607 ggtcgtcggggctcatcgccacagcggtcacatctagg...ctgacctt
 |||||||||||<<<< 113 <<<<|||||||||||||||||
 HDAC9 897 ggtcgtcggggctcatcgccacagcggtcacat.....ctt

AL022328 2722 gctgctggaggctcttcacgtgcggggcctgggcagcagggcactctgg
 |||||||||||<<<< 113 <<<<|||||||||||||||||
 HDAC9 933 gctgctggaggctcttcacgtgcggggcctgggcagcagggcactctgg

AL022328 2772 atggactctagggcactgtg...cctacctccctctgcacctctgacat
 |||||||||||<<<< 113 <<<<|||||||||||||||||
 HDAC9 983 atggactctagggcactgtg.....ctacctcgcacctctgacat

AL022328 2869 ggcgccattggccctgacaggggtggggccgggtcaccacagcagcgtctg
 |||||||||||<<<< 113 <<<<|||||||||||||||||
 HDAC9 1019 ggcgccattggccctgacaggggtggggccgggtcaccacagcagcgtctg

AL022328 2919 tactgtcatgcacactgactccgccagtgaactcagggtggtagccgccct
 |||||||||||<<<< 113 <<<<|||||||||||||||||
 HDAC9 1069 tactgtcatgcacactgactccgccagtgaactcagggtggtagccgccct

AL022328 2966 ggg....gtcacctccagcacggcagacacccggcggccggccagcacc
 <<< 189 <<<<|||||||||||||||||
 HDAC9 1116 gtccagcacggcagacacccggcggccggccagcacc

WO 02/050285

PCT/EP01/14928

AL022328 3193 tgcagcagctgtgtgaggtgggccaagcactctggcgtggcctgcatttg
HDAC9 1154 tgcagcagctgtgtgaggtgggccaagcactctggcgtggcctgcatttg
AL022328 3243 cccctgga...ctcacctcaggtccccgatggctgagtcacaaatcctgc
HDAC9 1204 ccc...ctcaggtccccgatggctgagtcacaaatcctgc
AL022328 3358 cgagaccagcaccagctcaggtcaaaactaca...gtcacctcaagg
HDAC9 1240 cgagaccagcaccagctcaggtcaaa...ctcaagg
AL022328 3605 ccagtgggagcagcaggtgcaggaaggcagccagtagtcagcgtttccc
HDAC9 1275 ccagtgggagcagcaggtgcaggaaggcagccagtagtcagcgtttccc
AL022328 3655 atcccaacctggc...ggcacctggttccaggcaggttgacagtga
HDAC9 1325 atcccaac...ctgggttccaggcaggttgacagtga
AL022328 3849 gcaggccctgtccccccccactgcgtctgcatctgactctgcagga
HDAC9 1360 gcaggccctgtccccccccactgcgtctgcatctgactctgcagga
AL022328 3899 aaggccagaagcggccatgctcatagcgggtgccaggagaagtaaggacg
HDAC9 1410 aaggccagaagcggccatgctcatagcgggtgccaggagaagtaaggacg
AL022328 3948 ctgcc...ctcacctggggtcatcctcaagagatactggatccccg
HDAC9 1459 <<<< 180 <<<<|ctgggtcatcctcaagagatactggatccccg
AL022328 4164 gccatggtgcacatcccagtcacacgagagatcctggg...cacacc
HDAC9 1495 gccatggtgcacatcccagtcacacgagagatc...c
AL022328 4355 tgtgtagcccggtttctgcttggcatgtgcagctgctatggccacgttg
HDAC9 1530 tgtgtagcccggtttctgcttggcatgtgcagctgctatggccacgttg
AL022328 4405 ttgaacacacagaaccogtggcagccgcccctctggccatggtgcccg
HDAC9 1580 ttgaacacacagaaccogtggcagccgcccctctggccatggtgcccg
AL022328 4455 aggcctacg...ctcacctcaccaggccaagccattttgcacagctcc
HDAC9 1630 aggc...ctcacaggccaagccattttgcacagctcc
AL022328 4589 agtgagcacagcgtccaccagctgcagtcagcccccctggggcagccg
HDAC9 1666 agtgagcacagcgtccaccagctgcagtcagcccccctggggcagccg

WO 02/050285

PCT/EP01/14928

```
AL022328 4639 cgcagtgaaagggtactctgtg....cgccacgggtggaagtagatggcg
      |||||||||||<<<< 266 <<<<|||||||||
HDAC9 1716 cgcagtgaaagggtact.....cggggtggaagtagatggcg

AL022328 4940 tcgaactgtccggacagcgctgcagctcctctgcctaggacotgggt
      |||||||||||
HDAC9 1751 tcgaactgtccggacagcgctgcagctcctctgcctaggacotgggt

AL022328 4990 ctccctgaccagggatcacatactctgggtgca....ctgacctgtgca
      |||||||||||<<<< 247 <<<<|||||
HDAC9 1801 ctccctgaccagggatcacatactctggg.....ctgtgca

AL022328 5272 ccaggcccagctcctcttcaggagcctcgcggtgacaaccgcagacac
      |||||||||||
HDAC9 1836 ccaggcccagctcctcttcaggagcctcgcggtgacaaccgcagacac

AL022328 5322 ctctgttccaggccgcgtgccgcaggcgatccagggtcgcggtcaggcg
      |||||||||||
HDAC9 1886 ctctgttccaggccgcgtgccgcaggcgatccagggtcgcggtcaggcg

AL022328 5372 ctcaggacgctogatctcgactcggggctggg....cttactcgtcca
      |||||||||||<<<< 68 <<<<|||||
HDAC9 1936 ctcaggacgctogatctcgactcgggg.....tcgtcca

AL022328 5476 gagcagccgggtggccgtcatgtcctcatggtacacaagcgcg
      |||||||||||
HDAC9 1972 gagcagccgggtggccgtcatgtcctcatggtacacaagcgcg
```

WO 02/050285

PCT/EP01/14928

Fig. 4

Score=267 bits (676), Expect=4e-71
Identities=143/354 (40%), Positives = 201/354 (56%) Gaps = 14/354 (3%)
Query:26 ERPRLTAALDRQRGLQRQ-----LRLSAREASEELGLVHSPFVSLVRETVL
E P R+ + +++ G LR+ AREA+ EEL VHS E V I+ +
Sbjct:79 EDPRLVLRVFAIKKAGYVSNVPSDVFLRIPAREATIEELLQVHSQENWYDRVINTKEM

Query:79 GKRELOALSGQFDATYFHPSTFHCARLAAGLQLDVAVLTGAVONGIALVRRPGHGQR
E+L L D+Y++ + CARLA G+ ++ AV+TG V+N A+VRPPGHH +
Sbjct:139SHEDLANLEKISDSLYNNESAFCARLAGGAIETCTAVVTGQVKNAPAVVRPPGHHAE

Query:139AAAGFCVFNNAVLAHAHAKKHG- LHRILVVDVHGHGQIQYLFEDDPSVLYFSMHR
GFC+FNNV++ A Q+ + R+L+VDND+HGG G Q F DDP+VLY S HR
Sbjct:199HKPGGFCLEFNNVSVTARSMLQRFDPDKIKRVLIVDNDLHNGTQMAFYDDPNVLYSLR

Query:197YEHGRFWEFLRESADAVRGCGGLGFTVNLFPNNQVGMGNADYVXXXXXXXXXXXXDPE
YE+GRF+P A+ G G GIG TVN+PW+ GMG+ DY+ DP+
Sbjct:259YENGRTYFGINYGCAENCCEGCGELGRTVNLFWSCAGCGEDGYIYAFQVVMVAVYEDPD

Query:257VILVVSAGFSAIGDEPGQMQATPECFALHTQLQLVLAGRVCAVLEGGYHLESLSVCM
LV+VS GFD+A GD GQ TP +AH+TQ+L LA G+V LEGY+L+S++ S
Sbjct:319LVVSCGPDAAAGDHIGQFLTPAAVAHWTQMLMGADGKVFISLEGGINLDSISTSALA

Query:317TVQTLGLDPAPPLSGPMAPQORCEGSALESIOSARAAQAPHWKSLOQDDVATP 370
Q+LIG P L A Q A+ +I Q+ +R+ ++ + A P
Sbjct:379VAQSLIGIPPGRLHTTYACQ-----AVATINHVTKIQSOYWRMCRPKHFDAMP 427

WO 02/050285

PCT/EP01/14928

Fig. 5

```

HDAC1 MAQTQGTTRKVCYYDGDVGNYYG--QHPMKPHRIEMTHNLLNLYGLYRKMSIYRPHK
HDAC9 -----MGTALVYHEDMTATRLWDDPECEIERPERLTAALDRLRQGLEQRCLLSARE
          :*: * . . : . : . :*: : : * : * : : . : :
HDAC1 ANAEEMTXHSDDYIKFLRSIRPDNMSEYSKQMRFNVGEDCPVFDGLFEFCQLSTQG--
HDAC9 ASEELGLLVHSPETVSLVRETQVLGKEELQALSGQFDA--IYFHPSTHCARLAAGAAGL
          * :*: : * :*: :* . : . :* . : . :*: :* .
HDAC1 SVASAVKLNKQQTDLAVNWAGGLHAKKSEASDFCYVNDIVLAILELLKYH--QRVLYID
HDAC9 QLVDAVLGTAVQNGALVRPPG-HHQRAAANDFCVFNNVAIAAAHAKQKHGLHRLLVVD
          . :*. ** . * . :*: . * ** :*: * ** . :*: :* . : * :*: * :*
HDAC1 IDHRRGDSVEEAFYTTDRVMTVSFHKYGEY---FFCTGLRDIGAGKGYAVNYPILRD
HDAC9 WDVHRRGDSIQILFSDDPVLYTFSWIRYEHGRFWDFLRSDADAVGRGQGLGFTVNLFWNQ
          * :*: :*: * * :* . :*: :* . :* :* :* :* :* :*
HDAC1 G-IDDESYEAIKFKVMSKVMEMFQPSAVVLQCGSDSLSGDRLGCFNLTIKGHAKCVEFVK
HDAC9 VGMGNADYVAFLHLLPLAFEFDPFLVLVSAGFDSAIGDPEGQMQATPECFALHTQLLQ
          : . : . * * * : : : * :* . :* :* * * * * * : * : * : . : : :
HDAC1 SFN-LFMMLGGGGCTTRNVARCPTYETAVALDTRIP-----
HDAC9 VLAGGRVCAVLGGGCHLESLESVCMTVQTLGDPAPPLSGFMAPCQRCGSALESIQSA
          : : : * * :* :* :* . . * . *
HDAC1 -----NELPYN-----
HDAC9 RAAQAPHWKSLLQQQDVTAVPMSPSHSPEGRPPPLFGGPVCKAAASAPSSLLDQPCLCP
          . :* .
HDAC1 -----DYFEYFGDPDKLHISPSNMTNQ-----TNEVLEKIKQRLFENLRMLPHA
HDAC9 APSVRTAVALTTPDITLVLPPDVIOQASALREETRAWARPHESLARREALTGKLLVL
          . * :* . * :* . : : : : : . : : : * * :
HDAC1 PGVQMQAIPEDAIPESGDEDEDPPDKRIS--ICSSDKRIACRE--FSDSEEEGEGGRE
HDAC9 LDGMLDGQVNSGIAATPASAAAAATLDVAVRRGLSHGAQKRLCVLALGQLDRPDLADHGRS
          . :* . :* . . . * : : . :* : * : . . : . : * .
HDAC1 NSSNFK-----
HDAC9 LMLNIRGKEAALSMFHVSTPLPVMTGGFLSCILGLVLPLAYGFQPDVLVLALGPGHGLQ
          * : :
HDAC1 -----KAKRVKTEDEKKKIDPEEK--KEVTRERKTKEEKPEAKGVKEEVKL
HDAC9 GPHAAALLAAMLRGLAGGRVLALLEENSTPQLAGILARVLNGEAPPSLGFSSVASPEDVQA
          . * : : * : : * : * . . * : . * : :
HDAC1 A-----
HDAC9 LMYLRGQLFQWKMQLQCHPHLVA

```

 Catalytic amino acids
 Potential RB-binding pocket residues

WO 02/050285

PCT/EP01/14928

Fig. 6

Sequence format is Pearson

Sequence 1: HDAC1	482 aa
Sequence 2: HDAC2	488 aa
Sequence 3: HDAC3	428 aa
Sequence 4: HDAC8	377 aa
Sequence 5: HDAC4	1084 aa
Sequence 6: HDAC5	1122 aa
Sequence 7: HDAC6	1122 aa
Sequence 8: HDAC7	855 aa
Sequence 9: HDAC9	673 aa

Start of Pairwise alignments

Aligning...

Sequences (1:2) Aligned. Score: 82
Sequences (1:3) Aligned. Score: 57
Sequences (1:4) Aligned. Score: 38
Sequences (1:5) Aligned. Score: 18
Sequences (1:6) Aligned. Score: 14
Sequences (1:7) Aligned. Score: 14
Sequences (1:8) Aligned. Score: 15
Sequences (1:9) Aligned. Score: 14
Sequences (2:3) Aligned. Score: 55
Sequences (2:4) Aligned. Score: 39
Sequences (2:5) Aligned. Score: 13
Sequences (2:6) Aligned. Score: 15
Sequences (2:7) Aligned. Score: 15
Sequences (2:8) Aligned. Score: 14
Sequences (2:9) Aligned. Score: 15
Sequences (3:4) Aligned. Score: 37
Sequences (3:5) Aligned. Score: 12
Sequences (3:6) Aligned. Score: 13
Sequences (3:7) Aligned. Score: 13
Sequences (3:8) Aligned. Score: 15
Sequences (3:9) Aligned. Score: 15
Sequences (4:5) Aligned. Score: 21
Sequences (4:6) Aligned. Score: 16
Sequences (4:7) Aligned. Score: 16
Sequences (4:8) Aligned. Score: 20
Sequences (4:9) Aligned. Score: 22
Sequences (5:6) Aligned. Score: 59
Sequences (5:7) Aligned. Score: 59
Sequences (5:8) Aligned. Score: 49
Sequences (5:9) Aligned. Score: 21
Sequences (6:7) Aligned. Score: 100
Sequences (6:8) Aligned. Score: 43
Sequences (6:9) Aligned. Score: 19
Sequences (7:8) Aligned. Score: 43

WO 02/050285

PCT/EP01/14928

Sequences (7:9) Aligned. Score: 19
Sequences (8:9) Aligned. Score: 20
Guide tree file created: [/biocinfv/software/biobenchsw/tmp/align/1478.dnd]
Start of Multiple Alignment
There are 8 groups
Aligning...
Group 1: Sequences: 2 Score:24259
Group 2: Sequences: 3 Score:18415
Group 3: Sequences: 4 Score:12882
Group 4: Delayed
Group 5: Sequences: 2 Score:9847
Group 6: Sequences: 3 Score:7569
Group 7: Sequences: 4 Score:5689
Group 8: Sequences: 8 Score:2841
Sequence:9 Score:3452
Alignment Score 36872
CLUSTAL-Alignment file created [/biocinfv/software/biobenchsw/tmp/align/1478.out]
CLUSTAL W (1.81) multiple sequence alignment

```
HDAC5      MNSPNESDGMSGREPSELEILPRTSLMSIPVTVEVKFVLFRAMPSSMGGGGGGSPFVEILR
HDAC6      MNSPNESDGMSGREPSELEILPRTSLMSIPVTVEVKFVLFRAMPSSMGGGGGGSPFVEILR
HDAC4      MSQSHPDGLSGKDPVELLNPAKVNHPSTVDVATALLQVAPSAVPMDIRLDHQFSLP
HDAC7      -----MDLRVGGKPFVEPPP-----
HDAC1      -----
HDAC2      -----
HDAC3      -----
HDAC8      -----
HDAC9      -----

HDAC5      GALVGSVDPTLREQQLQOELLALAKQQQLQKQLLFAEFCKQHDHLTRQHEVQLQKHKKQ
HDAC6      GALVGSVDPTLREQQLQOELLALAKQQQLQKQLLFAEFCKQHDHLTRQHEVQLQKHKKQ
HDAC4      -----VAEPALREQQLQOELLALAKQQQLQKQLLFAEFCKQHEQLSRQHEAQLNHNKQ
HDAC7      -----EPTLLALQRFQRLMHHLFLAGLQ-----
HDAC1      -----
HDAC2      -----
HDAC3      -----
HDAC8      -----
HDAC9      -----

HDAC5      QEMLAARKQQQEMLAARKQOELFQQFQREQQQOELFKQRLQOQLLILRNKESKESAIAS
HDAC6      QEMLAARKQQQEMLAARKQOELFQQFQREQQQOELFKQRLQOQLLILRNKESKESAIAS
HDAC4      QEMLAARKQOELLHQR-----KLERHQQ-----IQELEKQNRQQLQOLKNEKESAVAS
HDAC7      RSVEFNRHLSMDTP-----MPQLQVGPQOELRQLLHKDKSKESAVAS
HDAC1      -----
HDAC2      -----
HDAC3      -----
HDAC8      -----
HDAC9      -----

HDAC5      TEVALRLQEFLLSKSKEPTPGGLMHSLPQHPCWG--AHASLDQSSFPQSGPPGTPPBY
HDAC6      TEVALRLQEFLLSKSKEPTPGGLMHSLPQHPCWG--AHASLDQSSFPQSGPPGTPPBY
HDAC4      TEVALRLQEFVILNKKALAHNHLNMTSSDFATWYGRGTQSSLDQSSFPQSG--VETSY
HDAC7      EVVQRKLAEVILKKQQAALERTVHPHSPGIF-----YRTLEF-LETEGATRSHLS
HDAC1      -----
HDAC2      -----
```

WO 02/050285

PCT/EP01/14928

```

HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 KLFLFG-PYDSRDDFFLRKTASEFNLKVRSLAQKVAERRSSFLRRKDGTVISTFKKRA
HDAC6 KLFLFG-PYDSRDDFFLRKTASEFNLKVRSLAQKVAERRSSFLRRKDGTVISTFKKRA
HDAC4 NHFVLG-NYDAKDDFFLRKTASEFNLKVRSLAQKVAERRSSFLRRKDGTVISTFKKRA
HDAC7 LPPVSLPDDPPHFPLRKTASEFNLKVRSLAQKVAERRSSFLRRKDGTVISTFKKRA
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 VEITGAGPGASSEVCHSAPGSGFSEPN-ESHSTIAENGFTGSEVFNITPEMLPCHRALPLDS
HDAC6 VEITGAGPGASSEVCHSAPGSGFSEPN-ESHSTIAENGFTGSEVFNITPEMLPCHRALPLDS
HDAC4 LDVT-----DSACESAPGSGFSEPNSSGVSALNGIAPAVSIPAEISLANRLVAREQ
HDAC7 AETLG---DSFSSSSTPASGCSFVNSEHG-----
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 SPNCFSLTTSFSLPNISLGLQATVTVTHRLTASPLSTQCEAERQALQSLRQGGTLTK
HDAC6 SPNCFSLTTSFSLPNISLGLQATVTVTHRLTASPLSTQCEAERQALQSLRQGGTLTK
HDAC4 SAAPFLTTSFSLPNITLGLPATG-----PSATAGQCDTERLTLPALQQR--LE--
HDAC7 -----FNFLG-----DSRRKTHPTLGPRG-----
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 FMSTESI PGCLLGVALEGGSPGHASLLQHVLLLEQARQOSTLIA-----VPLHGQSP
HDAC6 FMSTESI PGCLLGVALEGGSPGHASLLQHVLLLEQARQOSTLIA-----VPLHGQSP
HDAC4 LPPGTHLTYLSTPLERDGG-AAHSFLLQHVLLLEQARQOSTLIA-----GALPLHAQS-
HDAC7 FILGSPHTPLFLHGLEPEAG-GTLPERLQFILLLDPGGERAPLLTYPGLGFLPFHFAQS
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

HDAC5 LVTGERVATSMRTVGKLPKHPRLERTQSFLPQSPQALQOLVMOOQHQQFLEKQKQ---
HDAC6 LVTGERVATSMRTVGKLPKHPRLERTQSFLPQSPQALQOLVMOOQHQQFLEKQKQ---
HDAC4 LVGADRVSPSIN---KLKQHPRLERTQSFLPQSPQALQOLVMOOQHQQFLEKQKQFQQ
HDAC7 LMTTERLS-----GSLHWFLEKTRSEPLPFSATAPPPGPMQFLEQLKTHVQ---
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC8 -----
HDAC9 -----

```

WO 02/050285

PCT/EP01/14928

```

HDAC5      CQLQLGKILTKTGILFQPTTHPEETTEELTQOQVLLCGALTMFREGSTESESTQEDL
HDAC6      CQLQLGKILTKTGILFQPTTHPEETTEELTQOQVLLCGALTMFREGSTESESTQEDL
HDAC4      CQLQMPRIIPRPFEPARQFESHPEETTEELRPHQ-ALLDFPYLDRLPGQREAHQAGVQV
HDAC7      ---VIERSAAPSEKPLRQIPSAEDLETGGGPGQVVDGSLHRELGRGQFPAQCPAPL
HDAC1      -----
HDAC2      -----
HDAC3      -----
HDAC6      -----
HDAC9      -----

HDAC5      FEEDEEDGEEFEDCIQVKDEEGESGAEFGPDLFEPGAGYKXLFSDAQPLQPLQVYQAPL
HDAC6      FEEDEEDGEEFEDCIQVKDEEGESGAEFGPDLFEPGAGYKXLFSDAQPLQPLQVYQAPL
HDAC4      XQEP1ESDEEEAE----PPREVEPGCRQ-PSQELLFRQDALLLEQCRHQLRNYQASH
HDAC7      QQHPQVLLNEQQR----LAGRLPRGSTGDTVLLFLACGGHRLSKAQ-----SSPA
HDAC1      -----
HDAC2      -----
HDAC3      -----
HDAC6      -----
HDAC9      -----

HDAC5      SLATVP-----HQALGRTOSSPAAPGGMKSFPDQPVKHL-FTTGYYVDTFMLKHQCMCGN
HDAC6      SLATVP-----HQALGRTOSSPAAPGGMKSFPDQPVKHL-FTTGYYVDTFMLKHQCMCGN
HDAC4      EAAG1FVSFGGHRPLSKAQSPASATFPVSVQLEPPTKFR-FTTGYYVDTFMLKHQCTCGS
HDAC7      APASLS-----APEPASQARVLSSEETPARTLPFTTGYYVDTFMLKHQCSGCD
HDAC1      -----HAQTGG-TRKVCYYVDGDLVGNYYTGQ
HDAC2      -----HAYSGGGKXKVCYYVDGDLVGNYYTGQ
HDAC3      -----HAKTVAYFYDDPVGNFHYGA
HDAC6      -----MELPEEPADSGSLVFFVITYSPEYVSMCD
HDAC9      -----MGATLVYHELMTATRLNDD

HDAC5      THVHPEHAGRIQSIWSRIQETGLLSKCEIRGRKATLDEIQTVHSEYHTLLYGTSPLNRO
HDAC6      THVHPEHAGRIQSIWSRIQETGLLSKCEIRGRKATLDEIQTVHSEYHTLLYGTSPLNRO
HDAC4      SSSHPHAGRIQSIWSRIQETGLRGRCECIRGRKATLEELQTVHSEARTLLYGTNPLNRO
HDAC7      NSKHPHAGRIQSIWSRIQETGLRGRCECIRGRKATLEELQTVHSEARTLLYGTNPLNRO
HDAC1      G--HPMKPHRIKMTDNLNLYGLYKMEIYRPHKATAEMTKYHSEDTYKFLRSIRPDNM
HDAC2      G--HPMKPHRIKMTDNLNLYGLYKMEIYRPHKATAEMTKYHSEDTYKFLRSIRPDNM
HDAC3      G--HPMKPHRIKMTDNLNLYGLYKMEIYRPHKATAEMTKYHSEDTYKFLRSIRPDNM
HDAC6      S--LAKIPKASHVHSLIENYALHQMRIYKPKVASHEEHATPHTDATLQHLQKVSQEGD
HDAC9      PECEIERPERLTAALDEIRQRGLTQRCLRLSAREASEEELGLVHSPETVSLVRETQVLGR
      . * . : . * : * . : :

HDAC5      KLDSEKLLGPISQRMAYVLPCCGIGVDSDTVWNEHSSAVSHAVGCLLELAFKYAAGEL
HDAC6      KLDSEKLLGPISQRMAYVLPCCGIGVDSDTVWNEHSSAVSHAVGCLLELAFKYAAGEL
HDAC4      KLDSEKLLGSLAS-VFRLPCCGIGVDSDTVWNEHSSAVSHAVGCLLELAFKYAAGEL
HDAC7      KLDSEKLLGSLAS-VFRLPCCGIGVDSDTVWNEHSSAVSHAVGCLLELAFKYAAGEL
HDAC1      SE-----YSKQMBYFNVGDCPVPFGLFPCQLSTGGSVASAVLNKQCTDIAYNM
HDAC2      SE-----YSKQMBYFNVGDCPVPFGLFPCQLSTGGSVASAVLNKQCTDIAYNM
HDAC3      QG-----FTKSLNMFNVGDCPVPFGLFPCQLSTGGSVASAVLNKQCTDIAYNM
HDAC6      DD-----HPSIE-YGLGYDCPATEGIFDYAAJGGATITAAQCLIDGWCKVAJNM
HDAC9      EE-----LQALSGQFDATYFHPSTFHCARLAAGAGLQLVAVLTGAV
      . : . :

HDAC5      KNGFAIIRPPGHAEESTAMGFCFFNSVAITAKLLQOK---LNVGKVLIVDWDIHGNGT
HDAC6      KNGFAIIRPPGHAEESTAMGFCFFNSVAITAKLLQOK---LNVGKVLIVDWDIHGNGT
HDAC4      KNGFAVVRPPGHAEESTAMGFCFFNSVAITAKLLQOK---LSVSKVLIVDWDIHGNGT

```

WO 02/050285

PCT/EP01/14928

```

HDAC7  KNGFAYVRPPGHEADHSTAMGFCFFNGVAIACROLOQOOSKASKSLITVDWVHRGNGT
HDAC1  AGG-----LHHAKKFEASGFCYVNDIVLAIILELLKY-----HORVLYIDIDIHSGDGV
HDAC2  AGG-----LHHAKKFEASGFCYVNDIVLAIILELLKY-----HORVLYIDIDIHSGDGV
HDAC3  AGG-----LHHAKKFEASGFCYVNDIVLAIILELLKY-----HORVLYIDIDIHSGDGV
HDAC8  SGG-----WHHAKKFEASGFCYVNDIVLAIILELLKY-----FERILYVDLDLHSGDGV
HDAC9  QNGLALVRPPGHEGQRAAANGFCVFNVAIAAAHAKQKHG-----LHRILVVDWVHRGNGT
..*      **...  . ** .*. .:      :      : * : * : * : *

HDAC5  QQAFYNDPFLVLYISLHRYDNGNFFPGS--GAPEEVGGGPGVGTNVNVAWTGGVDFPIGDV
HDAC6  QQAFYNDPFLVLYISLHRYDNGNFFPGS--GAPEEVGGGPGVGTNVNVAWTGGVDFPIGDV
HDAC7  QQAFYNDPFLVLYISLHRYDNGNFFPGS--GAPEEVGGGPGVGTNVNVAWTGGVDFPIGDV
HDAC1  QQTFFYQDFVLYISLHRYDNGNFFPGS--GAVDEVGAGSGEGFNVAWAGGLDFPFMGDA
HDAC2  EEAFTTTRVHTVSEIHYG--EYFPGT--GDLADI GAGKGYAVNYPFRDGDID----DE
HDAC3  EEAFTTTRVHTVSEIHYG--EYFPGT--GDLADI GAGKGYAVNYPFRDGDID----DE
HDAC8  QEAFTTTRVHTVSEIHYG--EYFPGT--GDMYVGAESGRTYCLNYPFRDGDID----DQ
HDAC9  EDAFTTTRVHTVSEIHYG--EYFPGT--GDVEVGLGKRTYCLNYPFRDGDID----DE
OYLFEEDPFLVLYISLHRYDNGNFFPGS--GAVDEVGAGSGEGFNVAWAGGLDFPFMGDA
: *      * : . * * : . 11* . 1* * : 1* . 1
:

HDAC5  EYLTAFTVVMPIAHEFSFDVVLVSAGFDVAGHLSPLGGYVTAACFGHLTRQIMTLAAG
HDAC6  EYLTAFTVVMPIAHEFSFDVVLVSAGFDVAGHLSPLGGYVTAACFGHLTRQIMTLAAG
HDAC4  EYLTAFTVVMPIAHEFSFDVVLVSAGFDVAGHLSPLGGYVTAACFGHLTRQIMTLAAG
HDAC7  EYLTAFTVVMPIAHEFSFDVVLVSAGFDVAGHLSPLGGYVTAACFGHLTRQIMTLAAG
HDAC1  SYEALFKPMKVMKVFQPSAVVLQCGSDELSGD--RLGCFNLTKGHAKCDEVVKSFNL
HDAC2  SYGQIKFKPIISKVMKVFQPSAVVLQCGSDELSGD--RLGCFNLTKGHAKCDEVVKSFNL
HDAC3  SYKHLFPQVINGVVDYFQPTCIVLQCGSDELSGD--RLGCFNLTKGHAKCDEVVKSFNL
HDAC8  KYQIQCEVILKEVYQAFNPKAVVLQCGSDELSGD--FMCSFNLTKGHAKCDEVVKSFNL
HDAC9  DTVAAFLHLLFLAFEFDFELVLVSAGFDVAGHLSPLGGYVTAACFGHLTRQIMTLAAG
..*      : : 1 1 * 111. * * : . 1 1 . . 1

HDAC5  GRVVLALGEGGDLTAJCDASEACVSAALLSEVLQ---PLDEAVLQQRPNINAVATLEKVI
HDAC6  GRVVLALGEGGDLTAJCDASEACVSAALLSEVLQ---PLDEAVLQQRPNINAVATLEKVI
HDAC4  GRVVLALGEGGDLTAJCDASEACVSAALLSEVLQ---PLDEAVLQQRPNINAVATLEKVI
HDAC7  GAVVLALGEGGDLTAJCDASEACVSAALLSEVLQ---PLDEAVLQQRPNINAVATLEKVI
HDAC1  PMLMLG-GGGYTIKNVACRTYETAVALDCEIENEL-FYNDYEFYFGPDFKLHISPSN-M
HDAC2  PMLMLG-GGGYTIKNVACRTYETAVALDCEIENEL-FYNDYEFYFGPDFKLHISPSN-M
HDAC3  PMLMLG-GGGYTIKNVACRTYETAVALDCEIENEL-FYNDYEFYFGPDFKLHISPSN-M
HDAC8  ATLILG-GGGYTIKNVACRTYETAVALDCEIENEL-FYNDYEFYFGPDFKLHISPSN-M
HDAC9  GRVVLALGEGGDLTAJCDASEACVSAALLSEVLQ---PLDEAVLQQRPNINAVATLEKVI
:      * : 1 . . . 1

HDAC5  EIQSNHNSCVQRFAGLGRSLREADAGETIELLTVSAMALLSVGAQOQAAAAAREHSFRP
HDAC6  EIQSNHNSCVQRFAGLGRSLREADAGETIELLTVSAMALLSVGAQOQAAAAAREHSFRP
HDAC4  EIQSNHNSCVQRFAGLGRSLREADAGETIELLTVSAMALLSVGAQOQAAAAAREHSFRP
HDAC7  EIQSNHNSCVQRFAGLGRSLREADAGETIELLTVSAMALLSVGAQOQAAAAAREHSFRP
HDAC1  TNONTNELYELKIKRDLFENLRMLPHAPGVQMQAIPEDAIYKESGDEDEDDEKRIISIRAS
HDAC2  TNONTNELYELKIKRDLFENLRMLPHAPGVQMQAIPEDAIYKESGDEDEDDEKRIISIRAS
HDAC3  TNONTNELYELKIKRDLFENLRMLPHAPGVQMQAIPEDAIYKESGDEDEDDEKRIISIRAS
HDAC8  FDRNEFRHQQIIMYIKGNLKHVV-----FDRNEFRHQQIIMYIKGNLKHVV-----
HDAC9  AFHNSLQCCQDVTAVMSPSSHSPEGRFPPLPGGFCVKAASAPSSLLDQFCLCPAPSV
AFHNSLQCCQDVTAVMSPSSHSPEGRFPPLPGGFCVKAASAPSSLLDQFCLCPAPSV

```

WO 02/050285

PCT/EP01/14928

```
HDAC6 -----
HDAC5 RTAVALTTFDITLVLPDVIOCEASALREETEMARPHESIAAREALTALGRLLYLIDGM

HDAC5 -----
HDAC6 -----
HDAC4 -----
HDAC7 -----
HDAC1 -EKPEAKGVKEEVKLA-----
HDAC2 GERTDTKGTKEQLSNP-----
HDAC3 -----KESDVEI-----
HDAC8 -----
HDAC9 LDGGVNSGIAATPASAJAATLDVAVRRGLSHGAQRLLCVALGQLDRPDLAHDGRSLWLN

HDAC5 -----
HDAC6 -----
HDAC4 -----
HDAC7 -----
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC6 -----
HDAC9 IIRKEAAALSMFHVSTPLFVMTGGFLSCILGLVPLAYGFQPELVLVALPGHGLQGPFA

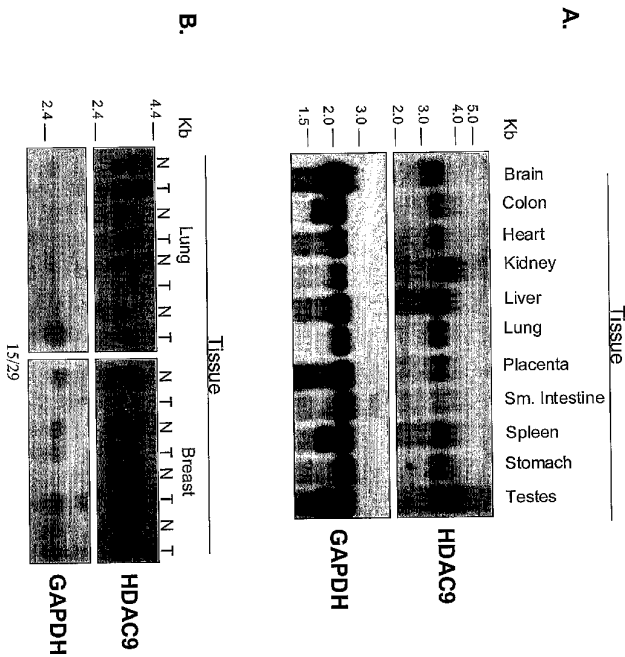
HDAC5 -----
HDAC6 -----
HDAC4 -----
HDAC7 -----
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC6 -----
HDAC9 ALLAAMLRGLAGGRVLALLEENSTPQLAGILARVLNGEAPFSLGPFSEVASPEDVQALMYL

HDAC5 -----
HDAC6 -----
HDAC4 -----
HDAC7 -----
HDAC1 -----
HDAC2 -----
HDAC3 -----
HDAC6 -----
HDAC9 RQGLEPQWKMLQCHPHLVA
```

WO 02/050285

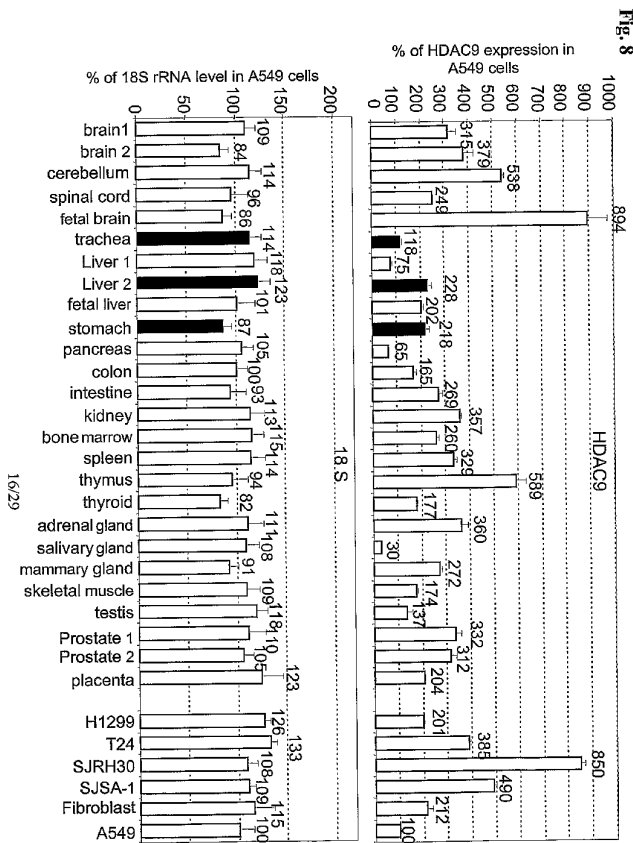
PCT/EP01/14928

Fig. 7



WO 02/050285

PCT/EP01/14928



WO 02/050285

PCT/EP01/14928

Fig. 9

```

Sequence format is Pearson
Sequence 1: HDAC6catalyticdomain      336 aa
Sequence 2: HDAC6catalyticdomain      329 aa
Sequence 3: HDAC6catalyticdomain1     302 aa
Sequence 4: HDAC6catalyticdomain2     401 aa
Sequence 5: HDAC7catalyticdomain      334 aa
Sequence 6: HDAC9completepeptide     673 aa
Start of Pairwise alignments
Aligning...
Sequences (1:2) Aligned. Score: 78
Sequences (1:3) Aligned. Score: 41
Sequences (1:4) Aligned. Score: 45
Sequences (1:5) Aligned. Score: 75
Sequences (1:6) Aligned. Score: 37
Sequences (2:3) Aligned. Score: 42
Sequences (2:4) Aligned. Score: 44
Sequences (2:5) Aligned. Score: 72
Sequences (2:6) Aligned. Score: 37
Sequences (3:4) Aligned. Score: 45
Sequences (3:5) Aligned. Score: 41
Sequences (3:6) Aligned. Score: 55
Sequences (4:5) Aligned. Score: 46
Sequences (4:6) Aligned. Score: 41
Sequences (5:6) Aligned. Score: 36
Guide tree      file created:  [/biocinv/software/biobenchw/tmp/align/
Start of Multiple Alignment      3664.out]
There are 5 groups
Aligning...
Group 1: Sequences: 2      Score:6517
Group 2: Sequences: 3      Score:6390
Group 3: Sequences: 4      Score:4801
Group 4: Sequences: 2      Score:5205
Group 5: Sequences: 6      Score:4795
Alignment Score: 15000
CLUSTAL-Alignment: file created: [/biocinv/software/biobenchw/tmp/align/
CLUSTAL W (1.81) multiple sequence alignment      3664.out]

```

PCT/EP01/14928

18/29

SUBSTITUTE SHEET (RULE 26)

WO 02/050285

PCT/EP01/14928

Fig. 10

```

Sequence format is: Pearson
Sequence 1: HDAC1catalyticdomain      310 aa
Sequence 2: HDAC2catalyticdomain      310 aa
Sequence 3: HDAC3catalyticdomain      310 aa
Sequence 4: HDAC6catalyticdomain      308 aa
Sequence 5: HDAC9completepeptide      673 aa
Start of Pairwise alignments
Aligning...
Sequences (1:2) Aligned, Score: 92
Sequences (1:3) Aligned, Score: 65
Sequences (1:4) Aligned, Score: 42
Sequences (1:5) Aligned, Score: 20
Sequences (2:3) Aligned, Score: 64
Sequences (2:4) Aligned, Score: 43
Sequences (2:5) Aligned, Score: 20
Sequences (3:4) Aligned, Score: 42
Sequences (3:5) Aligned, Score: 21
Sequences (4:5) Aligned, Score: 19
Guide tree      file created: [./bin/lnx/software/biobenchsw/tmp/align/
Start of Multiple Alignment      3650.out]
There are 4 groups
Aligning...
Group 1: Sequences: 2      Score: 6624
Group 2: Sequences: 3      Score: 5020
Group 3: Sequences: 4      Score: 4070
Group 4:      Delayed
Sequence 5:      Score: 2573
Alignment Score: 6332
CLUSTALW Alignment file created: [./bin/lnx/software/biobenchsw/tmp/align/
CLUSTAL W (1.61) multiple sequence alignment:      3650.out]

```

PCT/EP01/14928

21/29
SUBSTITUTE SHEET (RULE 26)

WO 02/050285

PCT/EP01/14928

NDAC1 catalytic domain -----
NDAC2 catalytic domain -----
NDAC3 catalytic domain -----
NDAC4 catalytic domain -----
NDAC5 complete peptide KDAALSDTHVSTPLPHHSGTLEGLGLVLPAYGFQPDLVVALGFPHGLQGPDAALL

NDAC1 catalytic domain -----
NDAC2 catalytic domain -----
NDAC3 catalytic domain -----
NDAC4 catalytic domain -----
NDAC5 complete peptide AAMLPLGLAGPHVLALLEENSTFQLAGCLAHVNLGEAPPSLGFSSVASPEDVQALMYLREQ

NDAC1 catalytic domain -----
NDAC2 catalytic domain -----
NDAC3 catalytic domain -----
NDAC4 catalytic domain -----
NDAC5 complete peptide LEPQVHMLQCHPHHVA

WO 02/050285

PCT/EP01/14928

Fig 11A

```

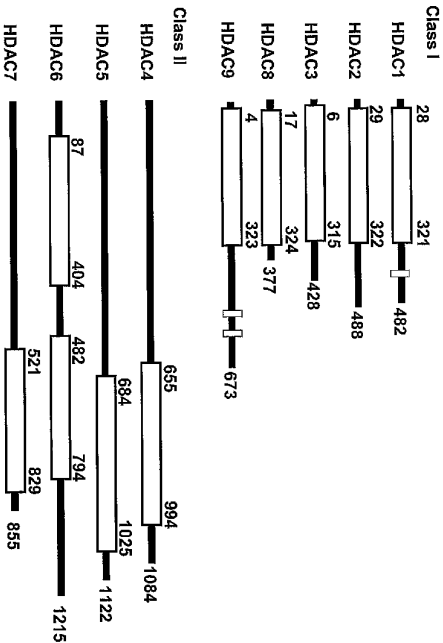
HDAC9v1 MGTALVYHEDMTATRLWDDPECEIERPERLTAALDRLRQGLEQRCILRSAREASEEEL
HDAC9v2 MGTALVYHEDMTATRLWDDPECEIERPERLTAALDRLRQGLEQRCILRSAREASEEEL
HDAC9v3 MGTALVYHEDMTATRLWDDPECEIERPERLTAALDRLRQGLEQRCILRSAREASEEEL
*****
HDAC9v1 GLVHSPFYVSLVRETQVLGKEELQALSGQFDIYFHFSTFHCARLAAGAGLQLVDAVLTG
HDAC9v2 GLVHSPFYVSLVRETQVLGKEELQALSGQFDIYFHFSTFHCARLAAGAGLQLVDAVLTG
HDAC9v3 GLVHSPFYVSLVRETQVLGKEELQALSGQFDIYFHFSTFHCARLAAGAGLQLVDAVLTG
*****
HDAC9v1 AVQNGLALVRPPGHHGQRAAANGFCVFNNVAIAAAHAKQKHGLHRLILVVDWDVHHGGGIQ
HDAC9v2 AVQNGLALVRPPGHHGQRAAANGFCVFNNVAIAAAHAKQKHGLHRLILVVDWDVHHGGGIQ
HDAC9v3 AVQNGLALVRPPGHHGQRAAANGFCVFNNVAIAAAHAKQKHGLHRLILVVDWDVHHGGGIQ
*****
HDAC9v1 YLFEDDPSVLYFSWHRYEHGRFWFFLRESADAVGRGQGLGFTVNLWNQVGMGNADYVA
HDAC9v2 YLFEDDPSVLYFSWHRYEHGRFWFFLRESADAVGRGQGLGFTVNLWNQVGMGNADYVA
HDAC9v3 YLFEDDPSVLYFSWHRYEHGRFWFFLRESADAVGRGQGLGFTVNLWN-----
*****
HDAC9v1 AFLHLLLPALAFEDPELVLSAGFDSAIGDFEGQMOATPECF AHLTQLLQVLAGGRVCAV
HDAC9v2 AFLHLLLPALAFEDPELVLSAGFDSAIGDFEGQMOATPECF AHLTQLLQVLAGGRVCAV
HDAC9v3 -----QFDPELVLSAGFDSAIGDFEGQMOATPECF AHLTQLLQVLAGGRVCAV
*****
HDAC9v1 LEGGYHLESIAESVCMVTQTLGDFAPFLSGPMAPCQRCESALESIQSARAAQAPHWKS
HDAC9v2 LEGGYHLESIAESVCMVTQTLGDFAPFLSGPMAPCQRCESALESIQSARAAQAPHWKS
HDAC9v3 LEGGYHLESIAESVCMVTQTLGDFAPFLSGPMAPCQRCESALESIQSARAAQAPHWKS
*****
HDAC9v1 LQQQDVTAVFMSPSHSPFGRFPFLPGGPFVCKAAASAPSSLLDQPCLCBAPSVRTAVAL
HDAC9v2 LQQQDVTAVFMSPSHSPFGRFPFLPGGPFVCKAAASAPSSLLDQPCLCBAPSVRTAVAL
HDAC9v3 LQQQDVTAVFMSPSHSPFGRFPFLPGGPFVCKAAASAPSSLLDQPCLCBAPSVRTAVAL
*****
HDAC9v1 TTPDITLVLPDPVIOQEA-----
HDAC9v2 TTPDITLVLPDPVIOQEAALREETEAWARPHESLAREEALTALGKLLYLLDGMLDGQVN
HDAC9v3 TTPDITLVLPDPVIOQEAALREETEAWARPHESLAREEALTALGKLLYLLDGMLDGQVN
*****
HDAC9v1 -----
HDAC9v2 SGIAATPASAAAATLDVAVRRGLSHGAQRLLCVALGQLDRPPDLAHDGRSLMLNIRGKEA
HDAC9v3 SGIAATPASAAAATLDVAVRRGLSHGAQSWGVGEGLEAMFGGSPAQRLLSHSTPAHGPFV
*****
HDAC9v1 -----CILGLVLPLAYGFQFDLVLVALGPGHGLQSPHAALLAAM
HDAC9v2 AALSMFHVSTFLPMTGGFLSCILGLVLPLAYGFQFDLVLVALGPGHGLQSPHAALLAAM
HDAC9v3 NALFPLFLRFLRMTGGFLSCILGLVLPLAYGFQFDLVLVALGPGHGCRAFTLHSHWLQC
*****
HDAC9v1 LRGLAGGRVLALLEENSTFQLAGILARVLNGEAPPSLGLSSVASPEDVQALMYLRGQLEF
HDAC9v2 LRGLAGGRVLALLEEVSAGWR--CCGVGRGKGF--VTASVFAGPELHTPASRDPGPGA
HDAC9v3 PGGWQC-----AESWPSNR-----RGRPGPYVPERAAGASVEDVAVPSSPGGLKSA
*****
HDAC9v1 QWKMLQCHPLVA
HDAC9v2 EWRGTS-----
HDAC9v3 K-----

```

WO 02/050285

PCT/EP01/14928

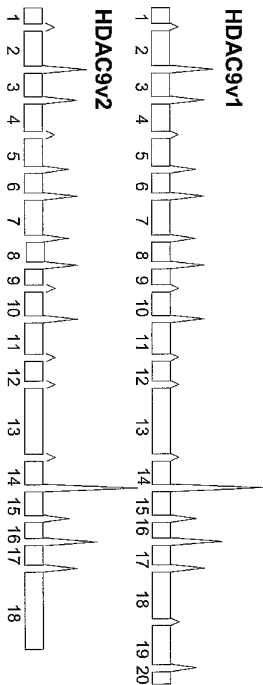
Fig. 11B



WO 02/050285

PCT/EP01/14928

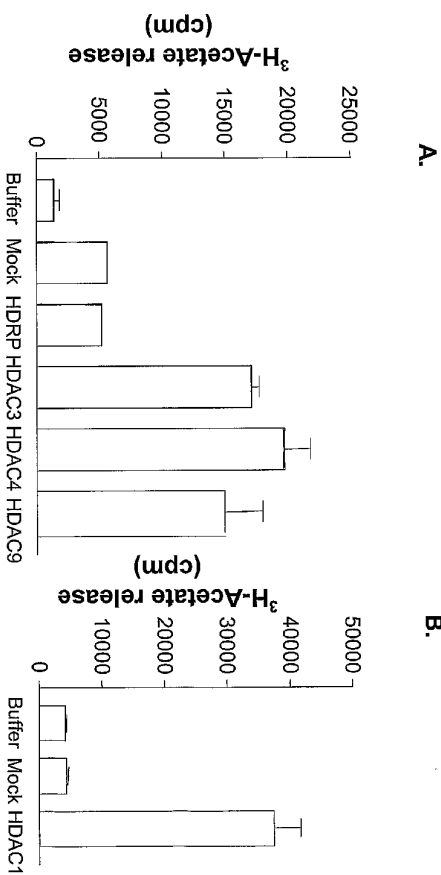
Fig. 11C



PCT/EP01/14928

WO 02/050285

Fig. 12



WO 02/050285

PCT/EP01/14928

Fig. 14.

SEQ ID NO:7

>HDAC9v2 DNA sequence

```

1  ATGGGGACCGCGCTTGTGTACCATGAGGACATGACGGCCACCCGGCTGCTCTGGGACGAC
61  CCCGAGTGGGAGATCGAGCGTCTCTGAGCGCTGACCGCAGCCCTGGATGCGCTGCGGCAG
121  CGCGGCTTGGAAACAGAGGTGTCTGCGGTGTCTGAGCCCGGAGGCTTCGGAAAGGAGGCTG
181  GGCGTGTGTGACAGCCAGAGTATGTATCCCTGGTTCAGGGAGACCCAGGTCTTAGGCAAG
241  GAGGAGCTGCAGGCGCTGTCCGGACAGTTCGACGCCATCTACTTCCACCCGAGTACCTTT
301  CACTGCGCGCGGCTGCGCGCAGGGCTGGACTGCACTGGTGGACGCTGTGCTCACTGGA
361  GCTGTGCAAAATGGGCTTGGCTTGGTGGGCTCCCGGGACCATGGCCAGAGGGCGGCT
421  GCCAACGGGTTCTGTGTGTCAACAACTGGCCATAGCAGCTGCATGCCAAGCAGAAA
481  CACGGCTACACAGGATCCTCGTGTGACTGGGATGTGACCATGGCCAGGGGATCCAG
541  TATCTCTTTGAGGATGACCCAGCGCTCTTTACTTCTCCTGGCACCGCTATGAGCATGGG
601  CGCTTCTGGCTTTCTCTCGAGAGTCAGATGCAGACGCACTGGGCGGGGACAGGGCCTC
661  GGCTTCACTGTCAACCTGCCCTGGAAACAGGTTGGGATGGGAAACGCTGACTACGTGGCT
721  GCCTTCTGCACCTGCTGCTCCCACTGGCTTGTGAGTTTGAACCTGAGCTGGTGTGGTC
781  TCGGACGGATTGACTCAGCCATCGGGGACCTGAGGGGCAATGCAGGCCACGCCAGAG
841  TGCTTCGCCACCTCACACAGCTGTGCAAGTGTGCGCGGGCGGGTCTGTGCGGTG
901  CTGAGGGCGGGCTACCACTGGAGTCACTGGCGGAGTCACTGTGCATGACAGTACAGACG
961  CTGTGGGTGACCGGCCCCACCCCTGTGAGGGCAATGGCGCATGTGAGAGGTGCGAG
1021  GGGAGTGCCTTAGAGTCCATCCAGATGCGCGGTGCTGCCAGGCCCGCATGTGAAAGAGC
1081  CTCGACGACAGATGATGACCGCTGTGCGGATGAGCCCGCAGCGCCACTCCCGAGGGG
1141  AGGCTTCACTCTGTGCTGGGCTTCCAGTGTGTAAAGCAGCTGCATCTGACCCGAGC
1201  TCCCTCTGGGACAGCGCTGCTCTGCGCCGACCCCTCTGTGCGCACCGCTGTGCGCTG
1261  ACAACGCGGATATCACATTGTTCTGCCCCCTGAGCTCATCAACAGGAAGCGTCAGCC
1321  CTGAGGAGGAGACAGAAAGCTGGGCCAGGCCACAGAGTCCCTGGCCCGGAGGAGGCC
1381  CTCCTGCACTTGGGAAGCTCTGTACTCTTAGATGGGATGTGGATGGGAGGTGAAC
1441  AGTGGTATAGCAGCCACTCCAGCCTCTGCTGACAGGCCACCCCTGGATGGCTGTTCGG
1501  AGAGGCTGTCCACGAGAGCCAGAGGCTGCTGTGCTGGCCCTGGGACAGCTGACCCGG
1561  CTTCCAGACCTCGCCCATGACGGGAGGAGTCTGTGGCTGAACATCAGGGGCAAGGAGCGC
1621  GCTGCCCTATCCATGTTCCATGTCCTCCACGCCACTGCCAGTGAACCGGTGTTCTCTG
1681  AGCTGCATCTTGGGCTTGGTGTGCTGCCCTGGCCTATGGCTTCCAGCCTGACCTGGTGTG
1741  GTGGGCTTGGGCTTGGCCTTGGCTGCGAGGCGCCCAAGCTGCACTCTGGCTGCAATG
1801  CTTGGGGGCTTGGGAGGGGCGAGTCTTGGCCCTCTGGAGGAGGTAGCTGGGAGGG
1861  TGGAGTGTCTGGGGGTGGGACGAGGGGAGGACCACTGACTGCTTCCGCTTCCGCCCCT
1921  GGTCCAGAACTCCACACCCAGCTAGCAGGATCTTGGCCGGGTGCTGAATGGAGAGGC
1981  ACTTCTAGCTAGGCTTTCTCTGTGCGCTCCCGAGGAGCTCCAGGCCCTGATGTA
2041  CTTGAGAGGGCAGCTGGAGCCTCAGTGGAGAGTGTGAGTGCCATCCTCACTGGTGGC
2101  TTGA

```

SEQ ID NO:5

>HDAC9v2 peptide sequence

```

1  MGTALVYHEDMTATRLNDDPECFIERFERLTAAALDRLRQRLBQRCRLRSAREASEEEL
61  GLVHSFRYVSLVRETVLQKEELQALSGQFDALYFHPSTFHCARLAAGAGLQLVDAVLTG
121  AVQNLALVLRPPHGHGQRAAANGFCVFNVAALAAAHAKQKHGLHRLVVDVHHGQGIQ
181  YLFEDDPVSLVFSWHRYEHRFWPFLRESADAVRGQGLFTVNLFPWNVGMGNADYVA
241  AFLHLLPLAFEDFELVLVSAGPDSALGDPGQMQATPECTFAHLTQLQLVLAGGRVCAV
301  LEGGYHLESLESVCMTVQLLGDPPAPPLSGPMAPCORCEGSALESIQSARAAQAPHWKS
361  LQQQDVIAVPMSPSSHSPEGRPPPLPGGPVCKAAASAPSSILDPCLCPAPSVRTAVAL
421  TTPDITLVLPFDVIQGEASALRETEAWARPHESLAREBALTALGKLLYLLDGMLDGQVN
481  SGLAATPASAAAATLDVAVRRGLSHGAQRLLCVALGQLDRPPDLAHDGRSLWNLIRGKEA

```

WO 02/050285

PCT/EP01/14928

541 AALSMFHVSTPLPVMITGGFLSCILGLVLEPLAYGFQFDLVLVALGPGHGLQGFHAALLAAM
601 LRLAGGRVLLALLEVSWAGWRCCGVGRGKGPVTASVTFAPGPELHTFASRDPGPGAERWG
661 TS

SEQ ID NO:8

>HDAC9v3 DNA sequence

1 ATGGGGACCGGCTTGTGTACCATGAGGACATGACGGCCACCCGGCTGCTCTGGGACGAC
61 CCCGAGTGGAGATCGAGCGTCTGAGCGCCTGACCGGACCCCTGGATCGCTGCGGCAG
121 CGCGGCTTGGAAACAGAGGTGTCTGGGGTGTGTCAGCCCGCGAGGCTTCGGAAGAGGAGCTG
181 GGCCTGGTGCACAGCCACAGATATGTATCCCTGGTCAGGGAGACCCAGGCTCTTAGGCAAG
241 GAGGAGCTGCAGGCGCTGTCCCGACAGTTCGACGCCATCTACTCCACCCGAGTACTTTT
301 CACTGCGCGCGGCTGGCCGCGAGGGCTGGACTGCAGCTGGTGGACCGCTGTGCTCACTGGG
361 GCTGTGCAAAATGCGCTTGGCTGGTGAGGCTCTCCCGGCACCAATGGCCAGAGGGCGGCT
421 GCCACCGGTTCTGCGTGTTCACCAACGTTGGCCATAGCAGCTGCACTGCCAAGCAGAAAA
481 CACGGGCTACACAGGATCCTCGTCGTGGACTGGGATGTGACCCATGGCCAGGGGATCCAG
541 TATCTCTTTGAGGATGACCCGAGCTCCTTTACTTCTCTGGCACCGCTATGAGCATGGG
601 CGCTTCTGGCCTTTCTGCGAGAGTCAGATGCAGACGAGTGGGGCGGGGACAGGGCCTC
661 GGCTTCACTGTCAACCTGCCCTGGAACAGTTTACCCCTGAGCTGGTGTCTGGTCTCGGCA
721 GGAATTGACTCAGCCATCGGGGACCTGAGGGGCAAAATGCAGGCCACGCCAGAGTGTCTT
781 GCCACCTCACAAGCTGTCTGAGGTGTGGCGGGCGGGGCTGTGTGCGCTGTCTGAG
841 GCGGGCTACCACTTGGAGTCACTGGCGGAGTCAGTGTGCATGACGTAAGAGCTGTGCTG
901 GGTGACCCGGCCCCACCCCTGTGAGGGCCAAATGGCGCCATGTCAAGAGTGCAGGGGAGT
961 GCCTAGAGTCCATCCAGAGTGCCTGTGCTGCCAGGCCCGCACTGGAGAGGCTCCAG
1021 CAGCAAGATGTGACCGCTGTGCCGATGAGCCCGAGCAGCCACTCCCGAGAGGGAGGCT
1081 CCACCTCTGCTGCTGGGGTCCAGTGTGTAAAGCAGCTGCATCTGACCGAGCTCCCTC
1141 CTGGACCTAGCGCTGCTCTGCCCGCACCTCTGTTCGCGACCGCTGTGCTCTGACAAAG
1201 CCGGATATCACATTTGGTTCTGCCCCCTGACGTCATCCACAGGAAGCGTCAAGCTGAGG
1261 GAGGAGACAGAAGCTCGGGCCAGGCCACAGAGTCCCTGGCCGGGAGGAGGCTCACT
1321 GCACTTGGGAAGCTCCTGTACTCTTAGATGGATGTGGATGGGCGAGGTGAACAGTGGT
1381 ATAGCAGCCACTCCAGCTCTGTCTGACAGCAGCCACCTGGATGGCTGTTCGGAGAGGC
1441 CTGTCCACAGGAGCCAGAGCTGGGGTGTGGGAGAAAGGCTGCTGGAGGCAATGCCAGT
1501 GGGTCTCAGACAGAGGCTCAGCAGTCAAGCAGCCCTGCCCATGGCCCGTGAATGCT
1561 CTTCACCTCTGCTCTGCGGTTTGGGCTCAGGAGATGACCGGTGGTTCTCTGAGCTGC
1621 ATCTTGGGCTTGGTGTCTGCCCTTGGCTTAGGCTTCCAGCTGACCTGGTGTGGTGGG
1681 CTGAGGCTTGGCCATGGCTGACAGGCCCCCAGCTGCACTCTGGCTGCAATGCTTCGGG
1741 GGCCTGGCAGGGGGCCGAGTCTGGCCCTCTGGAGGAGGAGCTCCAGGCCCTTAGTA
1801 CCTGAGAGGGCAGCTGGAGCTCAGTGAAGATGTTGAGTGCATCTCACTCGTGTGGC
1861 TTGAAATCGGCCAAG

SEQ ID NO:6

>HDAC9v3 peptide sequence

1 MGTALVYHEDMTATRLWDDPECEIERPERLTAALDRLRQRLBQRCLRLSAREASEEEL
61 GLVHSPYYVSLVRETQVLKEELQALSGQFDATYFHPSTFHCARLAAGAGLQLVDAVLTG
121 AVQNGALVRRPPGHHCQRAAANGFCVFNNVAIAAAHAKQKHGLHRLLVVDVHVGQGIQ
181 YLFEDDPVLYFVSWHYRHRFPWFPLRRSDADAVRGQGLGFTVNLFWNQDPFLVLVSA
241 GFDSAIGDPGQMQATFECPAHLVQLLQVLAGGRVCAVLEGGYHLESLAESVCMVTVQLL
301 GDPAPPLSGPMAPCQRCESALESIQSARAAQAPHWKSILQQQDVTAVPMSPSHSPGRRP
361 PPLPGGPVCKAAASAPSSLLDQPCLCAPAPSVRTAVALTTPDITLVLPDPV IQEASALR
421 ETEFAWARPHESLARBEAL TALGLLYLLDGMLDGQVNSGIAATPASAAAATLDVAVRRG
481 LSEGAQSWGVGRGLLEAMPGGSPAQRLLSHSTPAHGPVNALPFLPLRFGLRMTGGFLSC
541 ILGLVPLAYGFQPDVLVLVALGPGHGCRAPTLHSLWLCQFGNQGABSWSWRRGRPGPYV
601 PERAAGASVEDVAVPSSPGGLKSAK

29/29

SUBSTITUTE SHEET (RULE 26)

【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
27 June 2002 (27.06.2002)

PCT

(10) International Publication Number
WO 02/050285 A3

(51) International Patent Classification: C12N 15/55, 9/16, C07K 14/47 (74) Agent: BECKER, Konrad; Novartis AG, Corporate Intellectual Property, Patent and Trademark Department, CH-4002 Basel (CH).

(21) International Application Number: PCT/JP01/14928

(22) International Filing Date: 18 December 2001 (18.12.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data: 20 December 2000 (20.12.2000) US 60/256,827 23 May 2001 (23.05.2001) US 60/293,089 6 September 2001 (06.09.2001) US 60/317,984

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

(84) Designated States (regional): Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).

Published: with international search report before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

(71) Applicant (for all designated States except AT, US): NOVARTIS AG [CH/CH]; Lichtstrasse 35, CH-4056 Basel (CH).

(71) Applicant (for AT only): NOVARTIS PHARMA GMBH [AT/AT]; Brunner Strasse 59, A-1230 Vienna (AT).

(88) Date of publication of the international search report: 3 July 2003

(72) Inventors; and (75) Inventors/Applicants (for US only): COHEN, Dalia [FR/US]; 236 East Cedar Avenue, Livingston, NJ 07039 (US). BHATIA, Umesh [US/US]; 5212 Union Avenue, San Jose, CA 95124 (US). CAI, Richard, Lie [CN/US]; 367 Pond Road, Bridgewater, NJ 08807 (US). FISCHER, Denise, Dawn [US/US]; 17 Woodland Road, Bernardsville, NJ 07924 (US).

(15) Information about Correction: Previous Correction: see PCT Gazette No. 15/2003 of 10 April 2003, Section II

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

WO 02/050285 A3

(54) Title: HISTONE DEACETYLASE-RELATED GENE AND PROTEIN

(57) Abstract: Disclosed is an HDAC related genes and gene products. In particular, the invention relates to a protein and variants that is highly homologous to known HDACs and referred to herein as HDAC9, nucleic acid molecules that encode such a protein, antibodies that recognize the protein, and methods for diagnosing conditions related to abnormal HDAC9 activity or gene expression.

【国際調査報告】

INTERNATIONAL SEARCH REPORT		Inventor's Application No. PCT/EP 01/14928
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/55 C12N9/16 C07K14/47		
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		
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) BIOSIS, EMBL, EPO-Internal, WPI Data, PAJ, MEDLINE		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE NCBI 'Online! 12 December 1999 (1999-12-12) COVILLE G. : Database accession no. CAB63048 XP002227641 & VERDEL A., AND KHOCHBIN S. : "Identification of a new family of higher eukaryotic histone deacetylases" THE JOURNAL OF BIOLOGICAL CHEMISTRY, pages 2440-2445, XP002159646 page 2440, column 1, paragraph 1 ---	1-24
X	DATABASE EMBL 'Online! EBI; 1 April 1998 (1998-04-01) "Human sequence from clone RP3-402611 on chromosome 22q13" Database accession no. AL022328 XP002237050 ---	3-7
-/-		
☒ 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 "B" earlier document but published on or after the international filing date "C" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "D" document referring to an oral disclosure, use, exhibition or other means "E" document published prior to the international filing date but later than the priority date claimed "F" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "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 8 April 2003		Date of mailing of the international search report 28.04.03
Name and mailing address of the ISA European Patent Office, P.O. Box 5818 Patenteen 2 NL - 2280 HV Rijswijk Tel: (+31-70) 540-2040, Tx: 31 651 apo nl, Fax: (+31-70) 540-3016		Authorized officer Mabit, H

Form PCT/ISA/210 (second sheet) (July 1999)

INTERNATIONAL SEARCH REPORT		International Application No. PCT/EP 01/14928
C4(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WEIDLE U H ET AL: "INHIBITION OF HISTONE DEACETYLASES: A NEW STRATEGY TO TARGET EPIGENETIC MODIFICATIONS FOR ANTICANCER TREATMENT" ANTICANCER RESEARCH, HELENIC ANTICANCER INSTITUTE, ATHENS,, GR, vol. 20, May 2000 (2000-05), pages 1471-1485, XP001098720 ISSN: 0250-7005 abstract; figure 4 page 1474, column 2, paragraph 3 -page 1476, column 1, paragraph 4 ----	12-16
E	WO 02 30970 A (BAYER AG ;ZHU ZHIMIN (US)) 18 April 2002 (2002-04-18) claims 1-5,18-26,28,29 ----	1-11, 17-24
P,X	DATABASE EMBL 'Online! EBI; 14 November 2001 (2001-11-14) "Homo sapiens histone deacetylase 10 (HDAC10) mRNA" Database accession no. AF426160 XP002237051 ----	1-24
P,X	DATABASE EMBL 'Online! EBI; 20 August 2001 (2001-08-20) "Homo sapiens histone deacetylase 10 isoform alpha (HDAC10) mRNA" Database accession no. AF407273 XP002237052 ----	1-24
P,X	DATABASE EMBL 'Online! EBI; 6 August 2001 (2001-08-06) "homo sapiens histone deacetylase 10 isoform b mRNA" Database accession no. AF393962 XP002237053 ----	1-24
P,X	BERTOS NICHOLAS R ET AL: "Class II histone deacetylases: Structure, function, and regulation." BIOCHEMISTRY AND CELL BIOLOGY, vol. 79, no. 3, 2001, pages 243-252, XP009004311 22nd Annual West Coast Chromatin and Chromosomes Conference;Pacific Grove, California, USA; December 07-10, 2000 ISSN: 0829-8211 page 244, column 1, paragraph 2 page 244, column 2, paragraph 3 page 245, column 1, paragraph 1 page 249, paragraph 1 ----- -/-	1-24

Form PCT/ISA210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT		Int. Application No. PCT/EP 01/14928
C(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
T	FISCHER DENISE D ET AL: "Isolation and characterization of a novel class II histone deacetylase, HDAC10." JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 277, no. 8, 22 February 2002 (2002-02-22), pages 6656-6666, XP002227639 February 22, 2002 ISSN: 0021-9258	

Form PCT/ISA(210) (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT		international application No. PCT/EP 01/14928
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 claims 12-16 are directed to a diagnostic method practised on the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition. 2. ☒ Claims Nos.: 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: claims 25-27, and claims 7-10 and 13 partially 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.: 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/EP 01 /4928

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

Invention 1 (claims 1-27 partially)

Claims 1-27 related to SEQ ID N°1 comprising the amino acid sequence of SEQ ID N°1, an isolated DNA encoding such amino acid sequence, the DNA of SEQ ID N°2, a vector comprising said DNA, a host cell comprising said vector, a method for the diagnosis of a condition associated with abnormal regulation of gene expression by detecting abnormal transcription of the polypeptide of SEQ ID N°1, an antibody specific of SEQ ID N°1, and a method for producing a polypeptide of SEQ ID N°1.

Invention 2 (claims 1-27 partially)

Claims 1-27 related to SEQ ID N°5 comprising the amino acid sequence of SEQ ID N°5, an isolated DNA encoding such amino acid sequence, a DNA of SEQ ID N°7, a vector comprising said DNA, a host cell comprising said vector, a method for the diagnosis of a condition associated with abnormal regulation of gene expression by detecting abnormal transcription of the polypeptide of SEQ ID N°5, an antibody specific of SEQ ID N°5, and a method for producing a polypeptide of SEQ ID N°5.

Invention 3 (claims 1-27 partially)

Claims 1-27 related to SEQ ID N°6 comprising the amino acid sequence of SEQ ID N°6, an isolated DNA encoding such amino acid sequence, a DNA of SEQ ID N°8, a vector comprising said DNA, a host cell comprising said vector, a method for the diagnosis of a condition associated with abnormal regulation of gene expression by detecting abnormal transcription of the polypeptide of SEQ ID N°6, an antibody specific of SEQ ID N°6, and a method for producing a polypeptide of SEQ ID N°6.

International Application No. PCT/EP 01 /4928

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: claims 25-27, and claims 7-10 and 13 partially

The SEQ ID N°3 in claim 7 is defined as being the open reading frame of the cDNA sequence encoding the polypeptide as defined in claim 2. However, the SEQ ID N°3 in the sequence listing corresponds to a protein sequence. Moreover, it does not seem plausible that SEQ ID N°4 (1129 nt) corresponds to the endogenous genomic human DNA encoding the polypeptide as defined in claim 2 (673, 662 and 625 amino acids), as defined in claim 7. The inconsistencies render difficult, if not impossible, to determine the matter for which protection is sought to such an extent that a meaningful search is impossible. Consequently, the search has not been carried out for those parts of the application which do appear to be not clear, i.e. claims 25-27 and claims 7-10 and 13 partially.

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

Int'l Application No
PCT/EP 01/14928

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0230970	A	18-04-2002	
		AU 1501802 A	22-04-2002
		WO 0230970 A2	18-04-2002
		US 2002115177 A1	22-08-2002

フロントページの続き

(51)Int.Cl. ⁷	F I	テーマコード (参考)
C 1 2 N 5/10	C 1 2 N 9/10	
C 1 2 N 9/10	C 1 2 Q 1/68	A
C 1 2 Q 1/68	G 0 1 N 33/53	D
G 0 1 N 33/53	G 0 1 N 33/53	M
G 0 1 N 33/566	G 0 1 N 33/566	
// C 1 2 P 21/08	C 1 2 N 5/00	A
	C 1 2 P 21/08	

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

- (72)発明者 ダリア・コーエン
アメリカ合衆国07039ニュージャージー州リビングストン、イースト・シーダー・アベニュー
236番
- (72)発明者 ウメシュ・バティア
アメリカ合衆国95124カリフォルニア州サンノゼ、ユニオン・アベニュー5212番
- (72)発明者 カイ・リチャード・リー
アメリカ合衆国08807ニュージャージー州ブリッジウォーター、ポンド・ロード367番
- (72)発明者 デニーズ・ドーン・フィッシャー
アメリカ合衆国07924ニュージャージー州バーナーズビル、ウッドランド・ロード17番

专利名称(译)	<无法获取翻译>		
公开(公告)号	JP2004520027A5	公开(公告)日	2005-08-04
申请号	JP2002552162	申请日	2001-12-18
[标]申请(专利权)人(译)	瑞士商诺华公司		
申请(专利权)人(译)	诺华股份公司		
[标]发明人	ダリアコーエン ウメシュバティア カイリチャードリー デニース・ドーン・フィッシャー		
发明人	ダリア・コーエン ウメシュ・バティア カイ・リチャード・リー デニース・ドーン・フィッシャー		
IPC分类号	G01N33/53 C07K16/40 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N9/10 C12N9/16 C12N15/09 C12N15/55 C12P21/08 C12Q1/68 G01N33/566		
CPC分类号	C12N9/16		
FI分类号	C12N15/00.ZNA.A C07K16/40 C12N1/15 C12N1/19 C12N1/21 C12N9/10 C12Q1/68.A G01N33/53.D G01N33/53.M G01N33/566 C12N5/00.A C12P21/08		
代理人(译)	小島 一晃		
优先权	60/256827 2000-12-20 US 60/293089 2001-05-23 US 60/317984 2001-09-06 US		
其他公开文献	JP2004520027A		

摘要(译)

公开了HDAC相关基因和基因产物。特别地，本发明提供了蛋白质和变体称为HDAC9在高度同源的，并在本文中称为HDAC，核酸分子编码这样的蛋白质，抗体编码蛋白质，以及HDAC活性或基因表达对于与患者的异常有关的医疗状况的诊断方法。